

Finance, Risk and Accounting Perspectives Proceedings 2019

Editor: Dr. Othmar M. Lehner

ISBN: 978-3-9503518-5-9 (electronic)



Scientific Committee 2019

Prof. Othmar M. Lehner	Director ACRN Oxford Centre		
Prof. Minna Martikainen	Hanken School of Economics, Finland		
Prof. Bonnie Buchanan	University of Surrey, UK		
DDr. Julia Puaschunder	The New School NY, US		
Prof. James Ming Chen	Michigan State University, US		
Prof. Ahmed Marhfor	University of Quebec, Canada		
Dr. Michael Viehs	Oxford University and Hermes Investment Management, UK		
Dr. Juha Kansikas	University of Jyväskylä, Finland		
Prof. Heimo Losbichler	University of Applied Sciences Upper Austria		
Dr. Elisabeth Menschl	Johannes Kepler University Linz, Austria		
Prof. Steven C. Michael	University of Illinois, US		
Prof. Christiana Weber	University of Hanover. Germany		
Prof. Michael Dorff	Southwestern Law School, US		
Dr. Yifat Reuveni	Tel Aviv University, Israel		
Prof. Julia Mundy	University of Greenwich, UK		

Finance, Risk and Accounting Perspectives **Proceedings 2019** (2019), Edited by Prof. Dr. Othmar Lehner, ISBN: 978-3-9503518-5-9 (electronic).

TABLE OF CONTENTS

Foreword Othmar Lehner	1
The Determinants of Credit Cycle, its Forecast and Impact on the Credit Ratings <i>Natalya Dyachkova, Alexander M. Karmisky</i>	1
A Cognitive Load-theoretic Framework for Information Visualization Lisa Perkhofer	9
Dignity and Utility of Privacy and Information Sharing in the Digital Big Data Age Julia M. Puaschunder	26
Efficiency in the Highly Market-Segmented Chinese Banking Sector: A Meta-Frontier Non-Radial Directional Distance Function Approach	
Dong Xiang, Ning Zhang, Andrew C. Worthington	55
Factoring Transition Risks into Regulatory Stress-Tests Soline Ralite, Jakob Thomä	73
Early Stage Investing in Green SMEs: The Case of the UK Robyn Owen, Othmar Lehner, Fergus Lyon, Geraldine Brennan	88
The Problem of Heterogeneity within Risk Weights: Does Basel IV contain the Solution? <i>Christina Binder, Othmar Lehner</i>	108
How Social is Islamic Banking? M. Luthfi Hamidi, Andrew C. Worthington	136
Children's Toy or Grown-Ups' Gamble? LEGO Sets as an Alternative Investment Nikita Shimkus, Savva Shanaev, Satish Sharma, Binam Ghimire	150
Entrenchment, Equity Incentives and Workplace Safety Jesper Haga, Fredrik Huhtamäki, Dennis Sundvik	185
On Artificial Intelligence's Razor's Edge: On the Future of Democracy and Society in the Artificial Age <i>Julia M. Puaschunder</i>	203

Value at Looking Back: Towards an Empirical Validation of the Role of Reflexivity in Econo-Historic Backtesting: Economic Market Prediction Corrections Correlate with Future Market Performance <i>Julia M. Puaschunder</i>	224
How Blockchain Technology can Monetize New Music Ventures: An Examination of New Business Models <i>Marcus O`Dair, Robyn Owen</i>	253
Social Impact Investment for Housing Projects in Russian Regions Leonid A. Shafirov	271
The Influence of EU Non-Financial Reporting Directive on CSR Disclosure: Empirical Evidence from Finland Hanna Silvola, Eija Vinnari, Suvi Viljaranta	283
Designing visualizations to identify and assess correlations and trends: An experimental study based on price developments Lisa Perkhofer, Conny Walchshofer, Peter Hofer	294
Business Angel Ability to Accurately Identify Investment Opportunities: The Role of Mindfulness Catherine Crapsky, André Soares, Abdel Malik OLA	341

acrn A oxford Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings

homepage: http://www.acrn.eu/frap



Dear valued colleagues,

It is with great joy to welcome you to the 17th edition of the Finance, Risk and Accounting Perspectives Conference, running in parallel with the 3rd Social and Sustainable Finance and Impact Investing Conference. This year's motto is "The Fintech Revolution – Sustainability meets Technology in Finance, Risk and Accounting", and have we invited the global scholarly community to share their current research and to discuss the current innovations in technology and thinking in our fields together in a friendly and open, collaborative atmosphere.



Recent economic as well as technological developments have left us with much to learn and understand and I am personally convinced that further narrow-focused research inquiries - without accepting their complex interplay - will not lead to the right solutions! This is perhaps one of the reasons why this conference format has become so popular over the years – as it provides the opportunity to listen to a broad range of topics and perspectives and to engage in fruitful discussions. As always, we have carefully selected a maximum of 100 participants from various backgrounds in finance, risk and accounting and in addition borrow fascinating insights from other fields such as management and organisation theory or quantitative physics.

As academics, we are usually at the forefront of thinking how we can model and ultimately manage complex situations in our economy. Not only in our research but also our teaching we influence the way people will approach decision-making for the years to come and we certainly have a tremendous forming influence on the young students who will become ouf future leaders. However, in the last couple of years we seem to have lost our leading role and the young generation itself now takes over control in their own ways, demanding systemic change and more responsibility and sustainability for people and planet in the ways we conduct business. In order to contribute to their worthwhile endeavours we need to understand what really drives them and provide them with the tools to manoeuvre the complexity of our society, to look critical at purely ideological attempts of manipulation, and ultimately show them how a focus on profits and the financial side of things in general is equally important and needed - as long as those funds are being put to good use.

have the

Prof Dr Othmar M Lehner Director, ACRN Oxford Research Centre, Oxford, UK





Programme Sponsor



QR Code and Link to the Abstract Proceedings: https://tinyurl.com/frap2019-abstractproceedings



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings



homepage: http://www.acrn.eu/frap

The Determinants of Credit Cycle, its Forecast and Impact on the Credit Ratings

Natalya Dyachkova*, Alexander M. Karmisky

NRU Higher School of Economics, School of Finance, Moscow, Russia

Abstract

This paper presents a modern approach to the credit cycle theory and its impact on credit ratings that for several years remains as an important and modern financial problem.

Problem/Relevance - In our research, we study what macroeconomic factors drive and influence the credit cycle. This paper also describes the changes in the dynamics of credit cycles.

Research Objective/Questions - The study has the following structure – first section provides an overview of literature sources and the second part shows data and methods applied. The third part discusses the possibilities of modeling the relationship between credit ratings and credit cycles and includes empirical studies and the obtained results, demonstrating their economic analysis.

Methodology – We use statistical and econometric modelling for determination what economic factors influence credit cycle and how credit gap as indicator describes these changes.

Major Findings - The obtained empirical results from multiple choice models demonstrated the strong influence between credit gap and different macroeconomic factors.

Implications - Our results show the comparative analysis of credit cycles between different countries with various economic growth.

Introduction

The relationship between credit cycles and changes in credit ratings is relatively new problem in modern finance theory, but it plays an important role requiring some academic attention. In the dynamics, it is possible to consider that the existence of these relations between credit cycles and rating changes have a great influence on the level of creditworthiness of issuers. It also has an impact in the opposite direction, affecting the government spending.

In our research, we studied an empirical analysis of the influence of changes in credit ratings according to the dynamic of credit cycle with use of multiple choice models. The specification of our models have showed that different stages of the credit cycle have different impacts on behavior of credit ratings. Our results showed that from the forecasting basis of credit cycle, the changes that occur in credit ratings smoothly spread during credit cycle's transition from one "bottom" point to next "peak" point.

The term "credit cycle" is understood as the change in the volume of loans issued by banks in response to changes in market participants' expectations according to the future market conditions and, consequently, it also shows the value of market assets that can serve as a "sum of total loans" (Kiyotaki, Moore, 1997).

The authors of many studies note that cyclical changes in economic access to credit affect the dynamics of final demand and production costs. This creates a macroeconomic vulnerability, and determines the relationship between the credit cycle and the business cycle.

There are a lot of academic studies considering the evaluation of the relationship between credit ratings and the credit cycle. Most of them include consideration of macroeconomic factors and the extent to which they affect the level of credit ratings. According to the key-study of Amato, Furfine, (2003), we examined the relationship between the rating estimations and the parameters of the credit cycle, for example, by such indicators as the ratio of the volume of loans to GDP. This indicator is called credit gap. However, the allocation of credit cycle trends based on this indicator and its use for the analysis of the relevant dynamics is always effective. A study by Repullo, Saurina, (2011), devoted to the study of credit cycles and their quantitative determination, points to a shift in the cycle of fluctuations in the growth rates of issued bank loans relative to the cycle of fluctuations in GDP growth rates. When the GDP growth begins after the economic crisis, the volumes of issued bank loans continue to hold at minimal levels. At the beginning of the decline in GDP growth, the volumes of bank loans issued are still growing and reach their maximum volumes for the entire credit cycle. This is a peak of credit cycle. As a result, the dynamics of the credit gap indicator (the ratio of the volume of loans to GDP) shifts relative to the dynamics of the actual volume of banking loans.

Data & Methodology

The empirical sample used in this study includes 2,449 observations from 19 countries (including Russia) from 2000 to 2016, quarterly data used. The main sources for collecting information were the databases of the International Monetary Fund (IMF), the World Bank, and the Bloomberg information system.

The collected sample includes issuing companies with predominantly several ratings assigned by various foreign rating agencies, both investment and speculative. In order to fully take into account the whole set of rating categories, the account of the dynamics (changes) of grades, the rating data were grouped by issuing companies into clusters based on their forecast transitional "outlooks" (i.e., positive and negative rating forecasts).

In total, after cleaning the collected data, the sample included 1,683 issuers, of which 390 were commercial banks, the rest were large non-financial companies (see Figure 2). Throughout the entire 16-year period, each company has, as a minimum, two observations of a credit rating.

Credit ratings were transformed to a numerical scale. The number of credit ratings categories was about 25 grades from the highest level to the default.

Explaining variables included (see Table 1):

- Macroeconomic variables (4 variables);
- socio-demographic variables (2 variables);
- financial indicators (10 indicators).

|--|

Variable name		Scale	Obs.
Population growth / decrease	рор	in %	2 449
GDP growth	gdp	in %	2 394
Credit gap: loans-to-GDP ratio (net trend)	iy	in %	2 394
Inflation (CPI)	cpi	in %	2 449
Total loans, growth	са	in %	2 293
Import, growth	imports	in %	2 404
Export, growth	exports	in %	2 404
Credit spread ¹	tnarrowm	in %	2 384
Money supply, increase	money	in %	2 294
The growth rate of employed people	stir	in %	2 285
Loans to companies	tloans	in %	2 247
The volume of debt securities	stocks	in %	2 182
Loans-to-GDP ratio	debtgdp	in %	2 271
Exchange rate revaluation of a loan portfolio	xrusd	in %	144
denominated in foreign currency (only for Russia,			
dummy)			
Credit rating of the company	cr_rat	Category, from 1	2 449
		to 25	
The company's assets to GDP	assets	in %	2 135
1 Total observations: 2 449			
2 Countries (19). Australia Brazil Great Brita	in Germany India	Spain Italy Canada	China Mexico

2 Countries (19): Australia, Brazil, Great Britain, Germany, India, Spain, Italy, Canada, China, Mexico, Netherlands, Portugal, Russia, USA, Finland, France, Sweden, South Africa, Japan 3 Data sources: World Bank, IMF, Bloomberg financial database.

Table 1 contains descriptions of the main variables that were included in the selected sample and used in the creating of models. The total number of correctly selected (truly omitted) variables included in the final sample was 14. The sample included data from 19 countries, including Russia.

The trend component of a series of the ratio of loans to the economy to GDP (*trend_component*) was estimated in two ways: using the two-way Hodrick-Prescott filter (*indicator of the credit gap iy*), and using the band-pass filter (*indicator of the credit gap iy_bp*).

As an additional indicator, characterizing the credit cycle, the credit spread was used in modeling. The difference was calculated between the average market rate for bank loans and the risk-free rate.

To study the laws governing the formation of credit ratings of domestic issuers, a specific variable was used for Russian observations. This variable is based on the exchange currency rate of a portfolio of banking loans denominated in foreign currency (*xrusd*). The need to use this variable is due to the high share of foreign currency loans in the aggregate loan portfolio of Russian banks and companies. As a result, under conditions of fluctuations in the ruble exchange rate, the dynamics of such explanatory variables as the volume of loans to companies, the volume of all loans issued by Russian banks, the ratio of loans to GDP, and the credit gap may be distorted due to the impact of exchange currency rate. For example, in conditions of actual contraction of the credit market in the conditions of a double - banking and currency - crisis, a statistical illusion of a credit "boom" may arise. In order to prevent this from happening, statistical analysis and econometric modeling need to eliminate the distorting effect of the exchange rate re-evaluation of the foreign currency loan portfolio for Russian banks.

¹ The market rate for loans less than the risk-free rate

Empirical results

The empirical analysis made it possible to distinguish various stages of the credit cycle and to assess the changes in credit ratings, both at the beginning of the credit cycle (at the stage of recovery) and at its end (at the stage of the credit recession). This analysis showed that the onset of the crisis and the downgrading of credit ratings are out of sync. At the beginning of the crisis, credit ratings are relatively high, and only when the crisis starts and goes into recession, the ratings are beginning to be adjusted by rating agencies. This is typical for both developed and developing countries.

Based on the model estimates, one can clearly identify the end of one credit cycle and the beginning of the next one. Empirical estimates show that with the growth of the credit spread over one quarter more than 2.9 percent (%), the volume of issued bank loans begins to decrease, and changes in the dynamics of production and inflation are often observed simultaneously (see Table 2). Also, the turning points of the credit cycle can be predicted using the models of multiple choices, using variables from Tables 1 and 2.



Source: authors' calculations based on IMF data

Figure 1. The credit gap (changes in the loans-to- GDP ratio) in total for 19 countries, using a two-wayside Hodrick-Prescott filter (quarterly, in p.p.)

If the theory is true, the probability of a crisis arising may be controlled by the regulation of the amount of loans provided by banks and other financial institutions. It also can be controlled by government regulation of rates, for example, by changing the interbank lending rate. On the other hand, the regulation may concern the amount of reserves held by banks. These macroprudential regulation and monetary policy aim at reducing the risks that arise with the lending process, especially the effects of the default on the borrowers.

The proposition of the Basel Committee is that the amount of money reserved through creation of a counter-cyclical buffer should increase in time without crisis events, when there is an upward trend, and lower amount of reserves are made during downturns. This approach will provide an opportunity to accumulate a reasonable buffer that will be used in case of emergency or arising financial crises.



Source: authors' calculations based on Bloomberg data

Figure 2. The number of credit ratings issued by the international rating agency Moody's before the beginning of the credit recession (vertically - the number of companies in thousands, horizontally - the share in the sample, in p.p.).



Source: authors' calculations based on IMF data

Figure 3. The difference between 19 countries (in%) relative to the credit gap

The selection of macroeconomic and financial variables made it possible to form an optimal set of variables. When constructing a multiple choice model, the distributions of the coefficient estimates and, most importantly, the absolute probability and the factor estimation for the long-term period were obtained. In the given estimations of the models, all explanatory variables that were highly significant at the level of 1, 5, and 10 percent were highlighted. If the sign of the coefficient estimate contradicts the economic sense, the composition of the explanatory variables in the models changed and the models were re-evaluated.

Variable name	Probit models, specifications				
	M1 (OP)	M2 (OP), LS	M3 (MP)	M4 (M	1P), LS
Const	-0,446** (0,133)	-1,672** (0,271)			
GDP		-2,377* (1,641)	-7,948* (5,911)		-6,707** (1,521)
Assets	-1,627** (0,731)	-1,013* (1,228)	-0,621** (0,951)	-1,741* (0,925)	-2,315** (1,168)
IY	-0,185* (0,540)	-3,222*** (1,187)	-0,118** (0,577)	-3,055** (1,260)	-5,525** (3,856)
СРІ	-1,627** (1,348)	0,159* (0,179)	-1,137* (0,731)		0,194* (0,153)
Imports			-8,007** (4,647)		-3,117* (2,246)
Exports		-6,427** (9,577)	7,557** (4,631)		
Tnarrowm	8,038** (2.948)	2,207** (1,119)	-1,007* (7,108)	3,004* (1,037)	-1,717** (2,207)
Money	1,447** (8.878)	3,17** (2,145)			
Stir	0,003*** (0,126)	0,072*** (0,018)	0,036*** (0,099)	0,069** (0,185)	0,032*** (0,105)
Tloans			0,110* (0,677)	0,319** (0,875)	-0,139* (0,203)
Stocks		0,396** (0,745)	1,758** (1,218)	-1,042* (0,881)	0,419** (0,972)
Xrusd					-0,087*** (0,029)
Russia / dummy country			1	-	+ / with
Number of observations	1 682	1 767	2 282	2 138	2 282
R ²	0,618	0,623	0,811	0,818	0,809

According to the statistical analysis, the question of the endogeneity of the factors under consideration arose. For example, the credit spread, the credit gap and the unemployment rate can mutually influence each other. To minimize possible problems associated with endogenous parameters, LS estimation models were applied for two types of models. This approach allowed us to take into account the two-stage assessment of the categorical variable in the models under consideration, in which the endogenous variables X_{it} were explained by those variables that are not dependent on Z_{it} , taking into account time lags for one period ahead for X_{it} .

Conclusion

The empirical analysis made it possible to look through the different stages of the credit cycle and to calculate the changes in credit ratings, both at the beginning of the credit cycle (at the stage of "credit boom") and at its end (at the stage of the credit recession). This analysis showed that the onset of the crisis and the downgrading of credit ratings are not systemic. At the beginning of the credit cycle, credit ratings are relatively high, and only when a crisis starts and goes into next

stage near recession, the ratings are beginning to be downgraded by rating agencies. This is typical for both developed and developing countries.

The results of the analysis of the relationship between the credit cycle, identified by means of a credit spread and the dynamics of ratings, based on data from both Russian and foreign companies, are presented in Figure 2. Within the framework of joint trajectories of credit ratings and credit activity, the following stages have been singled out:

Step 1. At the beginning of the credit cycle, an increase in the volume of lending (banking loans) from the minimum levels against the low GDP growth and, moreover, continuing to decline credit ratings.

Step 2. Credit activity starts to increase; the level of credit ratings is also growing.

Step 3. Credit activity is no longer growing, but credit ratings are still emerging.

Step 4. Credit activity is beginning to decline - the market is on the verge of a credit recession. Credit ratings reach their maximum values.

Step 5. As a result of the credit recession, there is a parallel decrease in the volume of loans issued by banks, and the level of credit ratings.

In general, the results of modeling credit ratings using various probit models of multiply choices indicate high significance of factors such as GDP dynamics, employment, and the credit gap for rating formation. Some specifications are also characterized by the high importance of the credit spread. Judging by the "fixed effects" for models, no pronounced fundamental differences in the laws governing the formation of a rating between countries with developed and developing markets have been identified. A comparison of the results of evaluating the model of the M4 type with Russia and without Russia nevertheless shows that the role of country specificity may turn out to be significant.

The inclusion of a credit gap indicator calculated using a band-pass filter (iy_bp) instead of a two-way Hodrick-Prescott filter (iy) did not significantly improve the results. As can be seen from Table 2, the coefficient for this variable turned out to be insignificant.

In this case, it is necessary to take into account the fact that in the selected sample there were fewer issuers with a speculative rating than with an investment one. Perhaps this reduces the sensitivity of the dependent variable to the dynamics of the credit cycle.

It is also possible that the circumstance that the credit gap indicator has a symmetrical distribution plays a significant role, while other credit cycle metrics have an offset to the right (see Figure 3). As empirical analysis shows, credit ratings asymmetrically respond to the ascending and descending phases of the credit cycle, so using indicators with asymmetric distribution in statistical modelling can produce better results.

The presence of a right-handed bias and high investment grades of credit ratings also indicate a deterioration in macroeconomic parameters and a slowdown or absence of real GDP growth. In this case, all the ratings presented on the market tend to decrease in the long run and at the same time, the credit ratings of the investment class of large companies begin to gradually decrease and move to other categories of the lowest speculative grade.

Obviously, credit ratings are not only subject to cyclical fluctuations themselves (within the framework of the credit cycle), but they are also late in relation to the credit cycle. In the conditions of the beginning of the credit recession, the ratings are as high as possible and, moreover, can continue to grow, and in the conditions of the beginning of recovery credit ratings may continue to decline.

References

- Aikman D., Haldane A., Nelson B. (2010) Curbing the credit cycle. Presented at the Conference on Micro-Foundations for Modern Macroeconomics, New York, November 2010.
- Amato J.D., Furfine C. (2003) Are credit ratings procyclical? Bank for International Settlements. *Working Paper No. 129*.
- Blume M.E., Lim F., Mackinlay A.C. (1998). The declining credit quality of u.s. corporate debt: myth or reality? *The Journal of Finance, 53*.
- Cantor R., Mann C. (2003) Are corporate bond ratings procyclical? Special Comment, Moody's Investor Services.
- Claessens S., Kose M., Terrones M. (2012) How do business and financial cycles interact? Journal of International Economics, Elsevier, vol. 87(1), pages 178-190.
- Karminsky A., Polozov A. (2016). Handbook of ratings: approaches to ratings in the economy, sports, and society / Springer International Publishing AG.
- Kiff J., Kisser M., Schumacher L. (2013). Rating through-the-cycle: what does the concept imply for rating stability and accuracy? *IMF Working Paper*.
- Kiyotaki N., Moore J. (1997). Credit cycles // The Journal of Political Economy, v. 105, 2, April, 211-248.
- Lown C., Morgan D. (2006). The credit cycle and the business cycle: new findings using the loan officer opinion survey // Journal of Money, Credit, and Banking, Vol. 38, 6, September.

Nickell P., Perraudin W., Varotto S. (2000). Stability of rating transitions. *Journal of Banking and Finance, 24, 203-227*. Poghosyan T. (2015). How do public debt cycles interact with financial cycles? *IMF Working Paper, November 2015*.

Repullo R., Saurina J. (2011). The countercyclical capital buffer of Basel III: a critical assessment. *Working Papers, CEMFI, June 2011.*



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings



homepage: http://www.acrn.eu/frap

A Cognitive Load-theoretic Framework for Information Visualization

Lisa Perkhofer*

University of Applied Sciences Upper Austria

Abstract

This paper introduces cognitive load theory as a possible framework for information visualization in management accounting to enhance usability (effectiveness, efficiency, and satisfaction) by evaluating and improving design. Visualization is increasingly gaining importance as data samples are increasing in complexity because they enable decision-makers to extract larger amounts of information in a fast and easy manner. As opposed to traditional frameworks, which primarily focus on technical steps to transform raw data into the final layout, also called visualization, the cognitive perspective stresses the importance of humans and their cognitive capabilities to extract information from a particular and presented visualization. More precisely, it emphasizes the process of encoding data via the management accountant as well as decoding views via the respective decision-maker and the individual differences that may distort or alter interpretation. Cognitive load theory can explain a significant number of phenomena, including the mechanisms behind the frequently found problems with the adoption of new visualization options amongst accounting professionals as well as the striking influence of experience, which is responsible for a lot of contradicting results on visualization choice depending on the sample population or user group. Further, the framework contributes to the ongoing discussion on suitable evaluation methods. The proposed framework enables the information visualization community, which is responsible for the development of new visualization types, to evaluate and compare alternative visualization options outside of the traditional small-scale user studies or expert feedback; they can then shift the focus towards a more quantitative assessment based on both experimental research and an analysis within their area of application.

Index Terms— Cognitive Load, Information Processing, Information Visualization, Multi-Store Model of Memory, Visual Analytics

Introduction

The main purpose of management accounting is to provide decision-makers with relevant information for an easy, accurate and fast decision-making process (Appelbaum, Kogan, Vasarhelyi, & Yan, 2017; Dilla, Janvrin, & Raschke, 2010a; Perkhofer, Hofer, Walchshofer, Plank, & Jetter, 2019). Being able to fulfill this fundamental task is becoming more and more difficult as market dynamics increase (Eisl, Losbichler, Falschlunger, Fischer, & Hofer, 2012). The consequence is most likely a distortion of current working practice. Management accountants need to expand their scope from historical data reporting to real-time data processing, from using only in-house data to the inclusion of external data sources, from traditional paper based to interactive and online based reporting, and to shift the focus from reporting the past to predicting the future (Appelbaum et al., 2017; Goes, 2014). To achieve this shift, new tools and technical instruments such as algorithms, appropriate reporting software and especially interactive data visualization are necessary (Janvrin, Raschke, & Dilla, 2014; Perkhofer et al., 2019).

Interactive data visualization is especially helpful in contexts where problems (or tasks) are

high in complexity (Dilla et al., 2010a; Dilla & Raschke, 2015; Janvrin & Weidenmier Watson, 2017) or not sufficiently well-defined for computers to handle algorithmically, meaning human involvement and transparency is required (e.g., in fraud detection) (Dilla & Raschke, 2015; Kehrer & Hauser, 2013; Keim et al., 2008; Munzner, 2014). Visualization is supposed to boost information processing by relying on the human perceptual system, which is highly developed and allows for multiple processes to be executed simultaneously (Al-Kassab, Ouertani, Schiuma, & Neely, 2014; Mostyn, 2012). To optimally serve their purpose, different visualization types are created; this allows the decision-maker to decode information efficiently and effectively and consequently to enhance processing and insight (Chen & Jänicke, 2010).

With respect to Big Data, new and interactive forms of visualizations, which are generated by data scientists as well as information visualization researchers and are broadly available in software tools as well as free online libraries (e.g., D3.js), need to be taken into account. The pool of possible visualization options is extensive and ranges from bar, column and pie charts to parallel coordinate plots, force-directed graphs, and sunburst, Sankey and heatmap visualizations (Grammel, Tory, & Storey, 2010; Perkhofer et al., 2019; van Wijk, 2005). Unfortunately, the newer types of visualization created specifically to deal with Big Data have not yet found their way into either management accounting research or practice (Perkhofer et al., 2019). This can be attributed to a lack of knowledge on how to best use these types and how to attract potential users because there is currently no coherent and evidence based theoretical framework to guide the optimal visualization choice (Kopp, Riekert, & Utz, 2018; Sacha et al., 2017). Instead, small-scale user studies and expert interviews are used for evaluation (Plaisant, 2004; van Wijk, 2013), ignoring the benefits of experimental research to determine causal relations within their respective fields of application. Without such a theoretic framework, the ultimate goal of identifying and creating distinct images by emphasizing particular data or task characteristics in order to facilitate understanding and comprehension while at the same time accounting for individual differences of the decision-maker (Anderson et al., 2011; Lurie & Mason, 2007) cannot be achieved.

This paper therefore contributes to closing this gap by introducing a possible framework from a related research discipline, also proposed by Pike et al., namely cognitive science in order to focus on the decision-maker and his or her capabilities and needs (Pike, Stasko, Chang, & O'Connell, 2009). Although cognitive science has long been identified as an essential part and a pillar for visualization design and visual analytics, limited research can be found proposing it to be a conceptual or theoretical framework. In particular, a lot of explanatory power on differences between decision-makers, users, or students in information processing and learning is provided by focusing on cognitive load, one of the most frequently mentioned possibilities for contradicting results in previous findings (Chandler & Sweller, 1991; Chang, Ziemkiewicz, Green, & Ribarsky, 2009; Pollock, Chandler, & Sweller, 2002; van Merriënboer & Sweller, 2005). This framework will not only help explain some of the phenomena observed within the InfoVis and management accounting discipline (e.g., data-ink ratio, superiority of dashboards, superiority of user experience, user-specific differences), but will also allow for the inclusion of established methods from experimental research. By doing so, a shift in focus from being mainly qualitative towards an increased use of quantitative methods, which is necessary for data driven theory development, is possible.

In summary, the contributions are twofold: first, by relying on cognitive load theory (CLT) some of the user-related heterogeneity within the results in the discipline can be explained. The focus needs to be shifted from a rather theoretical and mechanical perspective of generating visualizations (Chen & Jänicke, 2010) towards the human factor and how visualizations are being used in their respective fields of application. The paper therefore adds to our understanding of individual users' needs within management accounting and increases our understanding of interactive data visualization. Second, because cognitive load is based on experimental psychology the frequently used qualitative research methods in the context of interactive data

visualization use (Isenberg, Isenberg, Chen, Sedlmair, & Möller, 2013) are extended by quantitative assessment possibilities. With the introduction of new measurement instruments, it is possible to mitigate some of the long-standing criticisms in evaluation practices (Plaisant, 2004) (e.g., lack of generalizability, unreliable or unreplaceable results, bias towards specific user groups). Instruments of CLT used in other disciplines (e.g., self-assessment questionnaires, eye tracking, EEG, fMRI, efficacy measures) will be analyzed in terms of their applicability in InfoVis research within the discipline of management accounting.

The remainder of this paper is structured as follows. First, related work concerning the InfoVis community, current practices and related advantages as well as disadvantages will be discussed. Finally, CLT and its three components will be introduced before discussing instruments, possible new insights and a research agenda for the management accounting community.

Related Work

The discipline of information visualization (InfoVis) is "concerned with the design, development, and application of computer-generated interactive graphical representations of information" (Al-Kassab et al., 2014, p. 409). Researchers have developed a huge pool of possible visualization options depending on different tasks and different contexts, which are widely available for practitioners to use (although some are incorporated in software tools such as tableau, PowerBI or QlikView, most are available on free libraries such as D3.js or D3plus which require the user to have a basic coding-knowledge). The major focus of the discipline is directed towards the development of new visualization options or the enhancement of existing ones by using technology that allows for interaction to be introduced (Chen & Jänicke, 2010). However, scientists are way ahead of their business counterparts (Grammel et al., 2010) and current research agendas are increasingly directed towards understanding the ever-increasing gap between practice and academia (Amar & Stasko, 2005; Chen, 2005; Perkhofer et al., 2019; Plaisant, 2004). It is largely unknown why visualizations, which are developed to increase understanding in a specific context and which are accessible to their respective users, are scarcely used.

The Visualization Reference Model

Multiple models outlining the process of constructing visualizations under working conditions (e.g., Card, Mackinlay, & Shneiderman, 1999; Chi & Riedl, 1998; Ware, 2012) have been introduced in order to identify routines, but also to improve existing visualization applications. Analyzing practices and behaviors of users has consequently helped in finding tool-support features as well as strengths and weaknesses of existing visualization options. The most frequently cited process model in this context is the one by Card et al. (1999), which is presented in Figure 1.



Figure 1: Visualization Reference Model by Card et al. (1999)

According to the authors, four steps are necessary to transform raw data into views (visualizations). First, raw data needs to undergo data transformation processes in order to form data tables. These data tables can then be altered by using filtering techniques, by adding information or calculations, and by merging columns or deleting them. This modified data table is then used and further altered into visual structures by applying visual mapping strategies. These visual structures include conventional charts (*such as the line chart, the bar chart or the scatterplot*) as well as enhanced visualization options used in data-rich contexts, for example:

- multidimensional visualizations (such as the parallel coordinates plot, the star plot or the *heatmap visualization*),
- hierarchical visualizations (such as tree visualizations, sunburst visualization or the Sankey chart),
- network visualizations (*such as the node-link chart, the chord diagram, or the dependency graph*)
- or geographical visualizations (such as the Choropleth or proportional symbol maps).

According to Card et al. (1999), the visual structure used depends on both the user themselves and on the inherent task that needs to be performed. The task and the user therefore determine the visualization option applied, which can but must not correspond to the task-dependent recommendations made by domain experts or visualization researchers (Grammel et al., 2010). After the definition of the visual structure, the last step calls for a rendering of the data table according to the visual properties in order to generate the view (Card et al., 1999; Grammel et al., 2010). Although the process seems to be straightforward, its applicability needs to be questioned. The next section will discuss some of the problems associated with current practice, including relying on Card's process model.

Disadvantages of Current InfoVis Practices

First, although experiments should be the pillar for design recommendations (Pike et al., 2009; Plaisant, Fekete, & Grinstein, 2008), existing approaches in the domain of information visualization seem to lack quantitative evaluation methods (Elmqvist & Soon Yi, 2013; Chen, 2005). Several authors in this field of research have raised the question concerning a quantitative method that can help with the unique needs of this research area (Dilla et al., 2010, Peck et al., 2012, Eisl et al., 2013). In their paper on current evaluation techniques for visual analytics, Plaisant and her colleagues present and distinguish three levels of evaluation and metrics which are displayed in the following figure. Most of the presented evaluation methods are used within small-scale usability studies (discussed in the next bullet point) and present objective measures, which place a high focus on time and error.



Figure 2: Three Evaluation Levels for Visual Analytics (Plaisant et al., 2008)

In order to be able to use the presented metrics, evaluation requires a roll out or at least a finished prototype and a comparison with a suitable alternative. Current practice not only binds many resources but also encourages mostly unfair comparisons and a naïve favoritism of the visualization introduced by the respective designer.

Another problem is that the higher the functionalities (interaction) and the higher the complexity of the data, the more urgent is the need for measures that go beyond traditional time and error, not least because in visual analytics knowing the task beforehand is becoming increasingly difficult as hypotheses are generated during the process of exploration with the visual structure (Brehmer & Munzner, 2013). Furthermore, the tasks used in many studies artificially reduce complexity and therefore might not be applicable in real life situations. According to Elmqvist and Soo Yi (2013), empirical evaluation for visualization beyond time and error is challenging due to difficulties in choosing the right task as well as finding correct measures that also take into account the large differences in perceptions among different people.

Second, currently available recommendations for choosing and designing an optimal visualization either rely on expert interviews and small-scale user studies in an ad-hoc manner or represent "awesome" examples without empirical proof (Chen, 2005; Isenberg et al., 2013). These methods are useful in terms of enhancing design and obtaining initial insights into their applicability, however, they are not comprehensive or comparable and therefore lack generalizability (Isenberg et al., 2013):

- With respect to expert interviews during the design stage of visualizations, researchers direct their focus towards a limited number of users with high domain knowledge. However, how experts approach a visualization is very different compared to novices and consequently their design and content needs also differ. By considering the cognitive limitations of novices, it is possible to enlarge the current user base considerably (Falschlunger, Lehner, Treiblmaier, & Eisl, 2016b).
- Small-scale user studies are also helpful in gaining initial insights but relying on them solely as a research discipline is dangerous. Evaluation is often based on a task-performance comparison between two or more alternatives, while the majority changes

not one but multiple attributes at the same time, making it extremely difficult to identify positive or negative features and distinguish their different impacts. This also means that "findings can hardly be broken down into a cause-effect relation and questions such as what has caused the performance increase remain unclear and speculative" (Isenberg et al., 2013, p. 2834).

The discipline is therefore criticized for not having sufficient theoretical depth or as a consequence for not having theoretical models that allow for a guided quantitative assessment to take place. As a result, no stable recommendation and no clear indications on the influencing factors affecting decision-making based on various visualization types and designs are available (Brehmer & Munzner, 2013). Without such clear guidelines, it is hard even for experts to choose the optimal visualization in a specific situation.

Third, during the four-step approach based on Card et al. (1999) there is significant room for error, especially for non-experts confronted with newer and data-rich visualizations. This is of high importance as we are confronted with a situation where the new visualizations are used scarcely due to a lack of technological support and/or a lack of coding skills amongst their intended users, but also due to low previous exposure (Perkhofer et al., 2019). Grammel et al. (2010) identified that three out of these four steps are especially challenging for novices, namely translating tasks into data attributes, constructing views/visualizations, and interpreting visualizations. The most significant observation was that a participant's choice of using a particular visualization was highly biased towards already known ones (bar – 34%, line – 23%, and pie – 13% charts). Their lack of experience with newer but more appropriate visualizations led to non-optimal choices, affecting user experience and decision-making outcomes. What's more, the researchers also observed a change in behavior as frustration levels increased and motivation sank (Grammel et al., 2010). Having limited exposure to these newer forms of visualization has also been identified as a barrier for adaption (Perkhofer et al., 2019).

Fourth, for most InfoVis applications (e.g., reporting of results in a scientific or managerial context) a separation of this four-step process into at least two parts depending on roles involved is necessary, a fact not sufficiently addressed in current literature (see Figure 3). According to Pike et al. (2009), very rarely does only one individual undertake the entire analysis process on their own as the necessary qualifications (skills in visual analytics and coding) and authority (to reach and implement decisions) are scarcely combined in a single user. Thus, in many situations, at least two users are involved, each having their own task within the chain of decision-making. While the *designer* (or more recently called the data scientist) is mostly concerned with the generation of the views and the preparation of the data for said views, the *decision-maker* or *end-user* is concerned with the interpretations of a predetermined visualization and the derivation of information-supported actions (see Figure 2).



Figure 3: Reference Model including Designer and User-Roles

The model therefore should also account for communication error and interpretation space as not only one but two users are involved. More precisely, depending on the expert level of the designer, it is more or less likely that the user or decision-maker will receive a task-optimized visualization. However, even if a visualization expert creates the visualization and therefore creates an optimal solution (based on current evaluation methods), the lack of involvement of the user in the visualization generation process (Chen, 2005) and the lack of experience in terms of interpreting the ordinal visualizations can have a serious effect on the results (Falschlunger et al., 2016b). This is the case because information processing is based on individual capabilities (for further explanation see Chapter Cognitive Load Theory) and adding a second person to this process increases the error span as encoding and decoding of data into visual structures is separated (Dilla, Janvrin, & Raschke, 2010b; Speier, 2006). For correct interpretation, the designer and the decision-maker need to work closely together and balance out each other's weaknesses. However, there is definitely a necessity to place a higher focus on the end-user and their needs for optimal information processing as he or she is the one using the presented information for decision-making.

And *fifth*, results of an extensive literature review conducted by Isenberg et al. in 2013 indicate that the average number of participant is only 23.8, supporting the claim that studies are lacking generalizability and statistical power. Experiments, on the other hand, would allow for cause-effect relationships as well as generalizable finding to be deduced. However, they are not among the top 20% of studies because they are "time- and resource-intensive to design, conduct, and analyze" (Isenberg et al., 2013, p. 2822). Nonetheless, experiments contribute a great deal in explaining variability in error rates and response times when exploring other disciplines such as managerial decision-making or experimental psychology. Other factors like task-, data-, and/or user-characteristics influence results significantly (Falschlunger, Lehner, & Treiblmaier, 2016a; Speier, 2006; Ziemkiewicz & Kosara, 2009) and need to be considered in order to draw the right conclusions.

Looking at current practice it can be concluded that results of user studies and expert interviews are often not replicable or reliable and contradicting findings are the consequence (Dilla et al., 2010b). Further, relying on the visualization reference model might oversimplify reality. This practice limits the discipline's full potential and hinders its further development (Chen, 2005). The next subsection discusses whether CLT can be used as a framework to overcome some of the mentioned problems within the discipline.

Cognitive Load Theory

CLT, predominantly used in education and learning research, was introduced by John Sweller and Paul Chandler (1991) and is based on the multi-store model of memory by Atkinson and Shiffring (1968). CLT investigates in detail the interplay between working memory and long-term memory, and how to overcome processing constraints (Chandler & Sweller, 1991; Chandler & Sweller, 1991). In order to understand its implications on information visualization, this chapter starts by introducing CLT as well as its fundamental basis – the multi-store model of memory.

The core components of the theory are schemas, their use during processing as well as their construction during learning. Only if schemas are available for information processing can humans' full potential be reached. In the context of InfoVis and in managerial decision-making, these schemas are often called "internal representation". An internal representation is the image or respectively the visualization option that pops up in the user's head in order to fulfill the required task. Only if a match between the internal and the external representation (the one handed to the user) occurs is the cognitive burden of processing minimized and generating insight

possible (Speier, 2006; Vessey & Galletta, 1991). To explain the importance of schemas in more detail, a closer look on the multi-store model of memory (Atkinson & Shiffring, 1968) (see Figure 4) is necessary. The model provides insights into the fundamental mechanisms behind information processing and helps to understand individual processing limitations and their corresponding behavioral reactions (e.g., anxiety and stress under a state of information overload (Bawden & Robinson, 2009)).

The multi-store model of memory consists of three essential stores: sensory store, short-term or working memory, and long-term memory (Atkinson & Shiffring, 1968). Their interplay and how information runs though these three stores is depicted in Figure 4 and briefly described in the following paragraphs.



Figure 4: Multi-store Model of Memory by Atkinson and Shiffring (1968)

Sensory memory is responsible for the immediate registration of a stimulus as well as for the assignment of incoming information to the appropriate sensory dimension. Sensory memory is not consciously controlled and most of the information is forgotten immediately. This is caused by the fact that sensory memory has a decay of only a few hundred milliseconds; it only gets passed on to working memory, where the actual processing takes place, if the user consciously focuses on specific information (Atkinson & Shiffring, 1968). Research on sensory memory is concerned with attracting attention and for visual information there needs to be a special focus on iconic memory (Dick, 1974). CLT, however, is concerned with schema construction and automation. Therefore, attention is a prerequisite but has not been conceptualized or analyzed in detail.

The opposite is true for working memory as cognitive load is described as the burden placed on working memory by the material or task (Huang, Eades, & Hong, 2009). In other words, cognitive load deals with information processing in working memory and how to do so in an efficient and effective manner (Chandler & Sweller, 1991). Looking at its function within the multi-store model of memory, this focus is inherently logical. All information that needs to be processed needs to run through working memory (Baddeley & Hitch, 1974). It is necessary to put emphasis on efficient and effective processes because it represents the bottleneck of the human cognitive architecture (Atkinson & Shiffring, 1968). Thereby, working memory is restricted in its capacity and duration (Atkinson & Shiffring, 1968; Sweller, van Merrienboer, & Paas, 1998), a fact that has been known since 1956 under the catchphrase "magical number seven" introduced by George A. Miller (1956). To be more specific, Miller's study states that only seven (plus or minus two) chunks of information can be processed simultaneously before the overloading of the user takes place. Miller's chunks correspond to single data points if no schemas are available in long-term memory, however, chunks can also be bundles of information if such schemas are available. The amount of information constructing one schema is unlimited, but one schema still functions as only one chunk of information irrespective of its size. By 'chunking', only a fraction of the initial load is placed on working memory and higher order processing can take place (Baddeley, 2000).

Long-term memory is therefore responsible for storing schemas that can be retrieved if needed

for processing (Atkinson & Shiffring, 1968; Miller, 1956). In comparison to short-term memory, long-term memory is assumed to have no capacity restrictions. This can explain the enhanced processing capacity of experienced students, users or decision-makers (Sweller, 1988), and also the need for experience (training and rehearsal). However, for information to be stored in long-term memory, greater amounts of time, cognitive effort, and rehearsal are needed (Atkinson & Shiffring, 1968).

According to the theoretical foundation provided by the multi-store model of memory, even highly complex information can be processed as one chunk in working memory while imposing no noticeable limits on information processing if information is stored in schemas (van Merriënboer & Sweller, 2005). This can also mean that one visualization option or even one dashboard consisting of multiple visualizations could be seen as one information chunk if stored in the long-term memory. Therefore, it is particularly important that schemes for information processing exist. This also means that the user needs to have some kind of experience with the task as well as with the visualization option. If no scheme is available, either limited or false interpretation is the consequence.

Important questions in this context arise: What does the process of schema creation look like? How can the process be influenced? These questions can be answered by taking a look at the three components of CLT (Sweller et al., 1998).

Components of CLT

Firstly, there is *intrinsic cognitive load*, which represents the effort required simply because of the complexity of the material or task. In this regard, complexity is determined by the element interactivity which exists (Brücken, Plass Jan L., & Leutner Detlev, 2003). A task has low element-interactivity when parts of a task can be learned in isolation, serially rather than simultaneously, because they do not interact. It is therefore possible to understand and learn the elements of such a task without holding too much information in working memory at any given time. To overcome the processing problems associated with high element-interactivity, expertise (stored schemas) is necessary. "Once a schema has been constructed, the interacting elements are incorporated within the schema and do not need to be considered individually within working memory" (Sweller et al., 1998, p. 261).

Secondly, *extraneous cognitive load* represents the component the theory was originally designed to elucidate – the load imposed on working memory by the instructional design and delivery method (Sweller, 2010). Extraneous cognitive load is believed to be under the direct control of instructional designers (Sweller et al., 1998) and can be interpreted as a kind of overhead which does not contribute to the understanding of the material (Brücken et al., 2003). The effects of good instructional design, however, may only be evident if the intrinsic cognitive load is high, since otherwise sufficient cognitive resources may exist anyway to successfully work with badly designed instructions (van Merriënboer & Sweller, 2005).

Thirdly, *germane cognitive load* refers the load incurred by the facilitating of knowledge or the creation of schemas in the course of practice and repetition (Mostyn, 2012; Sweller et al., 1998). Free processing resources, which arise via the reduction of extraneous and intrinsic cognitive load, are used in order to store information in long-term memory. However, users have to be motivated to redirect these freed resources towards learning. Although a high variability of schemata use during learning increases germane cognitive load because it requires mindful engagement, ultimately it yields more appropriate schema construction (van Merriënboer & Sweller, 2005).

Overall, according to CLT, learning (schema construction or extension) will be difficult or even hindered if the load imposed by the task and/or the material at hand is too high. Having said this,

the theory leaves us with an almost circular problem in relation to highly complex tasks: on the one hand, resources cannot be freed to create schemata; on the other, without schemata it is not possible to process the material in order to successfully perform the task. However, in real life there seem to be ways to overcome this circularity as even highly complex problems can be solved. This inconsistency between observation and the theory has been tackled in more current research, resulting in a change in the construct of intrinsic cognitive load; this load is now considered to be alterable (Sweller, 2010). One possible resolution in this context is to split highly complex problems into smaller, less complex part-problems, which can then be easily learned or solved. As another possibility, an artificially reduced version of the main problem can be provided as an intermediate goal in order to generate schemas which can subsequently be used for higher order processing and understanding (Pollock et al., 2002).

Explaining Phenomena in InfoVis

This chapter tries to match the results from InfoVis within the field of management accounting with those from CLT to demonstrate congruency. By closely monitoring cognitive load and its components for schema construction and automation, a lot of helpful insights for the InfoVis can be deduced.

Design Principles

Starting with *extraneous cognitive load* guiding principles of instructional design can be found, which correlate with and help explain information visualization design guidelines. The core goal, if possible, is to decrease the demand on working memory in order to allow novices to create schemes and for experienced users to effectively use stored ones. Two main effects can be found in this context:

- *The split-attention effect*: Information that needs to be processed together to facilitate understanding should be presented in juxtaposition; no working memory resources should be wasted in searching for information or mentally integrating the various information sources (Brücken et al., 2003; Chandler & Sweller, 1991). In an InfoVis context, this results in an integration of relevant labels or numbers within one visualization option (Falschlunger, Eisl, Losbichler, & Greil, 2014); however, it also means constructing one-page dashboards in order to increase understanding regarding correlations and relationships within the data (Yigitbasioglu & Velcu, 2012).
- *The redundancy effect*: This effect deals with reducing information that is irrelevant to the task or that is presented twice but in different forms (Sweller et al., 1998). In InfoVis this effect has established under a different name by Edward D. Tufte (1983), but with the same meaning: data-ink-ration no ink should be wasted on data that is not information.

With respect to *intrinsic cognitive load*, element interactivity needs to be investigated in more detail. Unfortunately, InfoVis largely ignores the *quantity, quality, and dimensionality* of the data being visualized. This leads not only to impaired decision-making but also to inconsistent findings. It has been shown that a visual representation with too much or too complex information can negatively affect the ability of the user to locate or extract relevant information because it overburdens working memory (Falschlunger et al., 2016a; Parsons & Sedig, 2014). On the other hand, the amount presented needs to be sufficiently comprehensive to enable the performance of the task (Huang et al., 2009). Choosing the right amount of data is a balancing act and experts should be capable of giving advice depending on the user's experience level. Second, it has been shown that the dimensionality of the data must be considered as well. The higher the

dimensionality (or complexity) of the data, the lower the overall amount of information displayed should be (Falschlunger et al., 2016a). Ignoring these facts, as has often been the case in previous research, might be the reason for some of the contradictions and unfair comparisons mentioned earlier in this paper.

It is also possible that two alternative visualization can result in no difference in the traditional metrics such as efficiency and effectiveness even though one visualization is obviously worse than the other (Huang et al., 2009). This can the case if intrinsic load or element interactivity is low and the additional effort of extraneous load can be compensated. The effects of good instructional design may only be evident if the intrinsic cognitive load is high since otherwise sufficient cognitive resources may exist anyway to successfully work with badly designed instructions (van Merriënboer & Sweller, 2005).

Novices vs. Experts

Differentiating between novices and experts is of high importance as having schemas for information processing is essential. If a schema or an appropriate internal representation exists, information processing and decision-making quality can be drastically enhanced. This corresponds neatly with findings in InfoVis stating that the construction of the internal representation is not only determined by the task itself, but more importantly by the prior experience and domain expertise of the decision-maker (Dilla et al., 2010b; Falschlunger et al., 2016b). With respect to novel visualization types in the context of Big Data, this also explains the barrier observed by Perkhofer and her colleagues (2019). It can be assumed that the user or decision-maker is familiar with charts and graphics encountered in everyday life (e.g., bar charts, column, chats, pie charts), however, limited knowledge can be expected for newer forms of multidimensional and hierarchical visualizations (Perkhofer et al., 2019). Moreover, prior experience and domain expertise could be key elements which foster valid and timely decision-making outcomes.

In this context, a phenomenon called *the expert reversal effect* also needs to be mentioned. While novices need more context and a reduced problem span in order to perform their tasks and classify information, experts might be more efficient with a reduced view concentrating on the core message, which can be higher in complexity (Pollock et al., 2002; van Merriënboer & Sweller, 2005). This behavior is highly visible when working with dashboards. In the initial phase a reduced information span and a higher amount of contextual information (explanations on relationships within the data, explanation on interaction with the dashboard, etc.) is recommended, while with repetition and training data density can be increased and contextual information eliminated without negatively influencing decision-making outcomes (Yigitbasioglu & Velcu, 2012).

Individual User Differences

For InfoVis, it is important to understand that depending on individual and situational influences (experience, culture, personality, dataset and contextual tasks) the optimal visualization choice differs (Falschlunger, Treiblmaier, & Lehner, 2015; Perkhofer & Lehner, 2019). A closer look at the process of decoding views is required as well as a better understanding of the designer and how data is being encoded. This is of importance as there seems to be a growing discrepancy between designers and users and a bridge between these two actors needs to be built (Perkhofer et al., 2019). However, individual differences and most importantly differences in experience (lack of schemas in long-term memory) explain why the involvement of a second person in the view generation process has such a big influence. This internal representation might be completely

different between designer and user (Grammel et al., 2010). Nonetheless, the goal of the InfoVis community should always be to best support the end-user.

Recent research demonstrates that an individual's cognitive load affects decision-making outcomes (Falschlunger et al., 2016a; Perkhofer & Lehner, 2019). More precisely, decision-making outcomes are posited to be impaired as soon as the decision-maker's cognitive state reaches its capacity limits (Al-Kassab et al., 2014; Tortosa-Edo, López-Navarro, Llorens-Monzonís, & Rodríguez-Artola, 2014). CLT therefore serves as a solid base to transform serendipity into intentional design by increasing usability (in terms of efficiency, effectivity, and satisfaction) based on the individual needs of the user. It is important to monitor the user's cognitive state during information processing when working with visualizations in order to enhance their user experience and decision-making outcomes.

Evaluation Methods of Cognitive Load

A measure that could discern differences between visualization tools and design options beyond traditional time and error measurements and enhance understanding is cognitive load or cognitive effort (Huang et al., 2009; Perkhofer & Lehner, 2019). Cognitive load is a latent construct and therefore is not directly observable (Brücken et al., 2003); however, it can be measured in various different ways in experimental research by monitoring the demand it places on working memory (Anderson et al., 2011). There are also multiple methods do which are classified and summarized in the following table.

Table 1: Classification of methods for measuring cognitive load by Brücken et al. (2003, p. 55)

	Causal Relationships		
Objectivity	Indirect		Direct
Subjective	✓ Self-reported invested mental effort	\checkmark	Self-reported stress level
		\checkmark	Self-reported difficulty of the materials
Objective	✓ Physiological measures	\checkmark	Brain activity measures (e.g., fMRI)
	✓ Behavioral measures	\checkmark	Dual-task performance
	✓ Learning outcome measures		

Hossain and Yeasin (2014) summarize different measurement models for information overload, namely subjective measurement via self-reports or objective measurement via physiological, behavioral, outcome-oriented (e.g., decision quality) or brain activity measures. Nonetheless, subjective measures and measures based on outcomes have a reputation for causing spurious correlations as, firstly, decision-making is influenced by many factors, not only by cognitive load, and secondly, participants are very often not aware of their cognitive state and therefore impair self-assessment. Dimoka et al. (2011) point out that measuring activities related to decision-making directly rather than asking individuals can open the "black box" and introduce new insights within IS related research.

Using physiological measures in InfoVis shows promising results. This can be attributed to the fact that these measures help place the user back in the center of attention and build visualizations exactly according to their individual needs. In this context, InfoVis and visual analytics have already started to investigate how users scan data with their eyes in order to accommodate their physical and mental limitations (Perkhofer & Lehner, 2019; Pike et al., 2009). Consequently, one measurement method which is increasingly gaining importance is eye tracking (Riche, 2009; Wang et al., 2014).

Research Agenda

Based on the critical evaluation of current practice and the promising framework of cognitive load, three explicit areas for future research are presented in the following sections.

Using Strategies From CLT to Introduce New Visualization Types

It seems only logical to use identified strategies for schema creation and integrate them into the design and introduction of new visualization options to accounting professionals. Possibilities therefore can be deduced from effects observed in experimental psychology by relying on CLT (Brücken et al., 2003; Chandler & Sweller, 1991; Sweller et al., 1998; van Merriënboer & Sweller, 2005). However, their use needs to be tested in the context of InfoVis.

- *Worked examples*: Help learners to focus on problem states and provide them with useful solution steps which can be utilized for more efficient schema creation.
- *Completion problem effect*: Presenting parts of the solution already found, leaving the learner with a reduced problem space; free resources can be used to focus on problem states and solution steps.
- *Modality effect*: Presenting information in a multimodal fashion (e.g., visual and audio); this ensures processing with both slave systems of working memory (van Merriënboer & Sweller, 2005), enhancing overall capacity because they can work simultaneously.

Eye Tracking to Measure Cognitive Load

As discussed in Chapter Evaluation Methods of Cognitive Load, eye tracking is gaining importance in the context of InfoVis. This is the case as eye movements are said to be cognitively controlled (Cole et al., 2013) and the data provides detailed information about the interaction with the visual stimulus. When analyzing eye tracking data, fixations are of particular importance. Fixations are dwells where the eye stops and information is processed. Conversely, during eye movement, called saccade, one is virtually blind. Long and many fixations indicate thinking processes and can be associated with a high visual complexity and cognitive load (Goldberg and Helfman, 2014; Toker et al., 2013; Wang et al., 2014). Additionally, data on pupil diameter over time can be used for analysis (Granholm et al., 1996). If a state of high mental load occurs, the pupil diameter increases (Kang and Wheatley, 2015).

Despite the fact that all measures in isolation present rather weak results, in their study on dashboard use in a managerial context Perkhofer and Lehner (2019) identified a combined measure which promises to reliably predict decision-making outcomes by including cognitive load as a complementary mediator to material complexity (measured by element interactivity and data density). However, their measure was only introduced and tested with static stimuli and as a result no interaction was used for information acquisition by the decision-maker. More research is needed to see how the metric works and/or needs to adapt with respect to more complex and interactive visualization types and dashboards.

Higher Order Statistical Methods to Identify Individual Influences

Including physiological measures, making use of latent variables, and testing multiple influences at the same time requires a change in statistical analysis. Commonly used analysis based on ANOVA (Cairns, 2007; Kay, Haroz, Guha, & Dragicevic, 2016) needs to be enriched by using

methods that allow for the identification of multiple factors to really identify all possible user- and situation-related influences. Higher order statistical methods such as structural equation modelling (CB-SEM or PLS-SEM) enable the evaluation of complex models and test whether or not the empirical data fits the conceptual or theoretical model. Making use of SEM might finally allow researchers to take issues surrounding possible influences such as gender or heritage out of the equation by simply including them as control variables.

Conclusion

For an efficient and effective use of visualization within management accounting, a deeper understanding of the underlying mechanisms on view generation and information decoding by the decision-maker is needed (Grammel et al., 2010). In this paper, the author has reported theoretic findings on whether CLT can become such a theoretic framework of visualization for management accounting. The focus of CLT lies in efficient and effective schema construction, creation, and automation in order to reduce the demand on working memory. A key aspect in achieving the goal of enhancing processing capabilities is to consider the decision-maker's cognitive capacities and constraints. Only if the load imposed on the decision-maker is appropriate can information processing and learning take place (Chandler & Sweller, 1991). CLT thereby differentiates between the load imposed by the instructional design (extraneous load), the inherent load of the material and task (intrinsic load), and the load directed towards learning (germane load) (Sweller, 2010). This corresponds very well with the often-used definitions of the internal and external problem representation used in InfoVis, and the influence of the task and the user on the effectiveness and efficiency of visualizations (Vessey & Galletta, 1991). Through evaluation of phenomena in InfoVis in the context of management accounting more closely, CLT provides even further explanations for many of them, such as the increasing gap between research and practice (van Wijk, 2005), the problems associated with adaption (Plaisant, 2004), or the loadreducing influence of experience (Grammel et al., 2010).

Introducing cognitive load as a possible framework not only increases understanding but also means introducing new measures and measurement instruments already established in experimental psychology. Expert feedback or small-scale user studies, which are the most common methods used (Isenberg et al., 2013), can be enriched by experimental studies that go beyond time and error measures to understand influences in information processing on a more fundamental level. This could be a step towards overcoming current deficits in evaluation practice, which are rather unreliable and result in heterogeneous outcomes (Chen, 2005). This shift increases internal consistency and allows for multiple influences (such as experience, gender, cultural background, etc.) to be measured, tested or controlled. Therefore, using CLT as a framework puts the focus back on the individual user and thereby answers the call for placing a stronger emphasis on the importance of the human factor (Herman, Melancon, & Marshall, 2000) and on finding empirical evidence (Chen, 2005).

From the authors' perspective, for all the reasons mentioned, CLT provides a solid framework for information visualization and therefore a solid base for further research. In particular, three areas were found to be of high relevance, namely:

- (1) Using strategies well established in education for learning to foster schema construction for new visualization types;
- (2) Introducing physiological measures and the construct cognitive load to predict decisionmaking outcomes; and
- (3) Using higher order statistical methods such as SEM over ANOVA in order to identify all user-dependent and therefore individual influencing factors.

Unfortunately, up to this point, cognitive science and its (mainly quantitative) measurement

approaches have rarely been used to evaluate visualizations (Herman et al., 2000; Huang et al., 2009). This paper therefore tries to contribute to this gap while at the same time answering the research calls of Chen and Jänicke (2010) and of Chen (2005) for a theoretical framework for information visualization based on human cognition.

References

- Al-Kassab, J., Ouertani, Z. M., Schiuma, G., & Neely, A. (2014). Information visualization to support management decisions. *International Journal of Information Technology & Decision Making*, 13(2), 407–428.
- Amar, R. A., & Stasko, J. T. (2005). Knowledge precepts for design and evaluation of information visualizations. *Visualization and Computer Graphics, IEEE Transactions on, 11*(4), 432–442.
- Anderson, E. W., Potter, K. C., Matzen, L. E., Shepherd, J. F., Preston, G. A., & Silva, C. T. (2011). A user study of visualization effectiveness using EEG and cognitive load. *Computer Graphics Forum*, 30(3), 791–800.
- Appelbaum, D., Kogan, A., Vasarhelyi, M., & Yan, Z. (2017). Impact of business analytics and enterprise systems on managerial accounting. *International Journal of Accounting Information Systems*, 25, 29–44.
- Atkinson, R.C., & Shiffring, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J.T. Spence (Eds.), *The psychology of learning and motivation* (pp. 89–195). New York: Academic Press.
- Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423.
- Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (pp. 47–89). Academic Press.
- Bawden, D., & Robinson, L. (2009). The dark side of information: Overload, anxiety and other paradoxes and pathologies. *Journal of Information Science*, 35(2), 180–191.
- Brehmer, M., & Munzner, T. (2013). A multi-level typology of abstract visualization tasks. *IEEE transactions on visualization and computer graphics*, 19(12), 2376–2385.
- Brücken, R., Plass Jan L., & Leutner Detlev (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 53–61.
- Cairns, P. (2007). HCI...not as it should be: inferential statistics in HCI research. In M. B. Rosson & D. Gilmore (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 195–201). 1240624: ACM.
- Card, S., Mackinlay, J., & Shneiderman, B. (1999). *Readings in information visualization: Using vision to think.* San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293–332.
- Chang, R., Ziemkiewicz, C., Green, T. M., & Ribarsky, W. (2009). Defining insight for visual analytics. *IEEE Computer Graphics & Applications, 29*(2), 14–17.
- Chen, C. (2005). Top 10 unsolved information visualization problems. *IEEE Computer Graphics & Applications, 25*(4), 12–16.
- Chen, M., & Jänicke, H. (2010). An information-theoretic framework for flow visualization. *IEEE transactions on visualization and computer graphics*, 16(6), 1216–1224.
- Chi, E. H.-H., & Riedl, J. T. (1998). An operator interaction framework for visualization systems. In *Proceedings IEEE* Symposium on Information Visualization (Cat. No.98TB100258) (pp. 63–70). IEEE Comput. Soc.
- Dick, A. O. (1974). Iconic memory and its relation to perceptual processing and other memory mechanisms. *Perception & Psychophysics, 16*(3), 575–596.
- Dilla, W., Janvrin, D. J., & Raschke, R. (2010a). Interactive Data Visualization: New Directions for Accounting Information Systems Research. *Journal of Information Systems*, 24(2), 1–37.
- Dilla, W., Janvrin, D. J., & Raschke, R. (2010b). Interactive data visualization: new directions for accounting information systems research. *Journal of Information Systems*, 24(2), 1–37.
- Dilla, W. N., & Raschke, R. L. (2015). Data visualization for fraud detection: Practice implications and a call for future research. *International Journal of Accounting Information Systems, 16*, 1–22.
- Eisl, C., Losbichler, H., Falschlunger, L., Fischer, B., & Hofer, P. (2012). *Reporting Design: Status quo und neue Wege in der Gestaltung des internen und externen Berichtswesens.* FH Oberösterreich, KPMG Advisory AG, pmOne AG.
- Falschlunger, L., Eisl, C., Losbichler, H., & Greil, A. (2014). Improving information perception of graphical displays: An experimental study on the display of column graphs. In V. Skala (Ed.), *Proceedings of the 22nd WSCG* (pp. 19–26).
- Falschlunger, L., Lehner, O., & Treiblmaier, H. (2016a). InfoVis: The Impact of Information Overload on Decision Making Outcome in High Complexity Settings. In : Special Interest Group on Human-Computer Interaction, Proceedings of the 15th annual Pre-ICIS Workshop on HCI Research in MIS (1–6, Paper 3). AIS Electronic Library: Association for Information Systems.

- Falschlunger, L., Lehner, O., Treiblmaier, H., & Eisl, C. (2016b). Visual Representation of Information as an Antecedent of Perceptive Efficiency: The Effect of Experience. In *Proceedings of the 49th Hawaii International Conference on System Sciences (HICSS)* (pp. 668–676). IEEE.
- Falschlunger, L., Treiblmaier, H., & Lehner, O. (2015). Cognitive differences and their impact on information processing: An empirical study comparing survey and eye tracking data. In F. D. Davis, R. Riedl, J. Vom Brocke, P.-M. Léger, A. B. Randolph, B. Weber, et al. (Eds.): *Lecture Notes in Information Systems and Organization, Information Systems and Neuroscience* (pp. 137–144). Springer International Publishing.

Goes, P. B. (2014). Big data and IS research. MIS Quarterly, 38(3), 3-8.

- Grammel, L., Tory, M., & Storey, M. A. (2010). How information visualization novices construct visualizations. *IEEE transactions on visualization and computer graphics, 16*(6), 943–952.
- Herman, I., Melancon, G., & Marshall, S. M. (2000). Graph visualization and navigation in information visualization: A survey. *IEEE transactions on visualization and computer graphics, 6*(1).
- Huang, W., Eades, P., & Hong, S. H. (2009). Measuring effectiveness of graph visualizations: A cognitive load perspective. *Information Visualization*, 8(3), 139–152.
- Isenberg, T., Isenberg, P., Chen, J., Sedlmair, M., & Möller, T. (2013). A systematic review on the practice of evaluating visualization. *IEEE transactions on visualization and computer graphics*, 19(12), 2818–2827.
- Janvrin, D. J., Raschke, R. L., & Dilla, W. N. (2014). Making sense of complex data using interactive data visualization. *Journal of Accounting Education*, 32(4), 31–48.
- Janvrin, D. J., & Weidenmier Watson, M. (2017). "Big Data": A new twist to accounting. *Journal of Accounting Education, 38*, 3–8.
- Kay, M., Haroz, S., Guha, S., & Dragicevic, P. (2016). Special Interest Group on Transparent Statistics in HCI. In J. Kaye, A. Druin, C. Lampe, D. Morris, & J. P. Hourcade (Eds.), *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '16* (pp. 1081–1084). New York, New York, USA: ACM Press.
- Kehrer, J., & Hauser, H. (2013). Visualization and visual analysis of multifaceted scientific data: A survey. *IEEE transactions on visualization and computer graphics*, 19(3), 495–513.
- Keim, D. A., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., & Melancon, G. (2008). Visual analytics: Definition, process, and challenges. In A. Kerren, J. T. Stasko, J. D. Fekete, & C. North (Eds.), *Lecture Notes in Computer Science: Vol. 4950. Information Visualization* (pp. 154–175). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Kopp, T., Riekert, M., & Utz, S. (2018). When cognitive fit outweighs cognitive load: Redundant data labels in charts increase accuracy and speed of information extraction. *Computers in Human Behavior, 86*, 367–376.
- Lurie, N. H., & Mason, C. H. (2007). Visual representation: implications for decision making. *Journal of Marketing*, 71(1), 160–177.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 101(2), 343–352.
- Mostyn, G. R. (2012). Cognitive load theory: what it is, why it's important for accounting instruction and research. *Issues in Accounting Education*, 27(1), 227–245.
- Munzner, T. (2014). Visualization analysis and design (1st Edition). AK Peters Visualization Series: CRC Press.
- Parsons, P., & Sedig, K. (2014). Adjustable properties of visual representations: improving the quality of humaninformation interaction. *Journal of the Association for Information Science & Technology*, 65(3), 455–482.
- Perkhofer, L., & Lehner, O. (2019). Using gaze behavior to measure cognitive load. In F. Davis, R. Riedl, J. Vom Brocke, P.-M. Léger, & A. Randolph (Eds.), *Lecture Notes in Information Systems and Organisation: NeuroIS Retreat* 2018. Information Systems and Neuroscience. NeuroIS Retreat 2018 (1st ed., pp. 73–83). Springer International Publishing.
- Perkhofer, L. M., Hofer, P., Walchshofer, C., Plank, T., & Jetter, H.-C. (2019). Interactive visualization of big data in the field of accounting. *Journal of Applied Accounting Research*, 5(1), 78.
- Pike, W. A., Stasko, J., Chang, R., & O'Connell, T. A. (2009). The Science of Interaction. *Information Visualization*, 8(4), 263–274.
- Plaisant, C. (2004). The challenge of information visualization evaluation. In M. F. Costabile (Ed.), *Proceedings of the working conference on Advanced Visual Interfaces 2004 (AVI)* (pp. 109–116). New York, NY: ACM.
- Plaisant, C., Fekete, J.-D., & Grinstein, G. (2008). Promoting insight-based evaluation of visualizations: From contest to benchmark repository. *IEEE transactions on visualization and computer graphics*, 14(1), 120–134.
- Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction*, 12(1), 61–86.
- Sacha, D., Zhang, L., Sedlmair, M., Lee, J. A., Peltonen, J., Weiskopf, D., et al. (2017). Visual Interaction with Dimensionality Reduction: A Structured Literature Analysis. *IEEE transactions on visualization and computer* graphics, 23(1), 241–250.
- Speier, C. (2006). The influence of information presentation formats on complex task decision-making performance. *International Journal of Human-Computer Studies, 64*(11), 1115–1131.
- Sweller, J. (1988). Cognitive load during problem solving: effects on learning. Cognitive Science, 12, 257-285.

- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane gognitive load. *Educational Psychology Review*, 22(2), 123–138.
- Sweller, J., van Merrienboer, Jeroen J. G., & Paas, Fred G. W. C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, 10(3), 251–296.
- Tortosa-Edo, V., López-Navarro, M. A., Llorens-Monzonís, J., & Rodríguez-Artola, R. M. (2014). The antecedent role of personal environmental values in the relationships among trust in companies, information processing and risk perception. *Journal of Risk Research*, *17*(8), 1019–1035.

Tufte, E. R. (1983). The visual display of quantitative information (1st). Connecticut: Graphics Press.

- van Merriënboer, Jeroen J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17(2), 147–177.
- van Wijk, J. J. (2013). Evaluation: A challenge for visual analytics. Computer, 46(7), 56-60.
- van Wijk, J. J. (2005). The value of visualization. In Proceedings of the 2005 IEEE VIS (pp. 79-86).
- Vessey, I., & Galletta, D. (1991). Cognitive fit: an empirical study of information acquisition. Information Systems Research, 2(1), 63-84.
- Ware, C. (2012). Information Visualization: Perception for design (3rd). Oxford: Elsevier Ltd.
- Yigitbasioglu, O. M., & Velcu, O. (2012). A review of dashboards in performance management: implications for design and research. *International Journal of Accounting Information Systems*, 13(1), 41–59.
- Ziemkiewicz, C., & Kosara, R. (2009). Preconceptions and individual differences in understanding visual metaphors. *Computer Graphics Forum, 28*(3), 911–918.



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings



homepage: http://www.acrn.eu/frap

Dignity and Utility of Privacy and Information Sharing in the Digital Big Data Age

Julia M. Puaschunder*

The New School, Columbia University, Princeton University

Financial support of the Academy of Behavioral Finance and Economics, Eugene Lang College of The New School, Fee Board, Fritz Thyssen Foundation, the Janeway Center Fellowship, New School for Social Research, Prize Fellowship, the Science and Technology Global Consortium, the University of Vienna, Systema B, and Vernon Arts and Sciences is gratefully acknowledged. The author thanks Professor Andrea Prat for a most interesting lecture on 'Industrial Organization' at Columbia University during Fall 2017 and Professor Anwar Shaikh for most valuable input and benevolent guidance. All omissions, errors and misunderstandings in this piece are solely the author's.

Abstract

Today enormous data storage capacities and computational power in the e-big data era have created unforeseen opportunities for big data hoarding corporations to reap hidden benefits from individual's information sharing, which occurs bit by bit in small tranches over time. This paper presents underlying dignity and utility considerations when individual decision makers face the privacy versus information sharing predicament. Thereby the article unravels the legal foundations of dignity in privacy but also the behavioral economics of utility in communication and information sharing. For Human Resources managers the question arises whether to uphold human dignity in privacy or derive benefit from utility of information sharing. From legal and governance perspectives, the outlined ideas may stimulate the e-privacy discourse in the age of digitalization but also serving the greater goals of democratisation of information and upheld humane dignity in the realm of e-ethics in the big data era.

Keywords: Behavioral Economics, Behavioral Political Economy, Democratisation of information, Dignity, Education, Exchange value, Governance, Preferences, Social media, Utility, Values

Introduction

Although communication and non-communication are day-to-day decisions of individuals; to this day, there is no stringently tested utility theory of information sharing and privacy. We lack a coherent decision science framework about when people choose to share information and when they rather want to stay silent for the sake of privacy. From the economic perspective, information sharing may impose temporal irreversible lock-ins or tipping points. The point of information sharing may be a reference point, in which one bit of more communication gives less utility than one bit of less information shared, hence one bit of more privacy, grants more utility in the sense of Kahneman & Tversky's (1979) behavioral decision science finding 'losses loom larger than gains.' There may also be a marginal decreasing utility derived from one bit more information shared but an exponential marginal utility gain from one more unit of information received given the fact that information can be put into context and an exponential marginal utility increase, as the more information one holds, the more complex connections one can make and use.

In the past, communication was depicted to decentralize organizations (Crémer, Garicano & Prat, 2007). Media was initially promoted to offer means of information transfer, political participation and protection against political abuse (Delli Carpini & Keeter, 1989; Neuman, Just & Crigler, 1992; Norris & Sanders, 2003; Prat & Strömberg, 2005; Snyder & Strömberg, 2010). Evidence suggests that media coverage increases voter information, which increases the responsiveness of votes to policy, which increases the effort and selection of politicians, thus producing better policies (Prat & Strömberg, 2013). Media thus traditionally was portrayed as helping to keep politicians accountable (Prat & Strömberg, 2013). Media coverage was found to improve selection and incentives of politicians alongside voting responsiveness (Iyengar & Kinder, 1987; Snyder & Strömberg, 2010).

Critical studies in this regard show that there are negative downsides of transparency (Prat, 2005). Mass media can also erode social capital, as they potentially isolate people from realworld experiences (Olken, 2009; Putnam, 2000). A positive relation between federal funds per capita allocation to areas where the media covered political parties in power was found (Snyder & Strömberg, 2010). Research has been done on the effect of conventional media on politics including a nomenclature of biases that impose problems – especially against minority opinions (Prat & Strömberg, 2013). Ideological biases are found in conventional media and media effects captured on vote choice (Prat & Strömberg, 2013). While the negative facets of information on elections and the role of social media on voting outcomes has been widely discussed recently; yet to this day no stringent theoretical or empirical framework for the utility of privacy and information sharing on social media exits.

In the digital age, to study the trade-off between information sharing and privacy has leveraged into unprecedented importance. Social media revolutionized human communication around the globe. As never before in the history of humankind, information about individuals can be stored and put in context over time and logically placed within society thanks to unprecedented data conservation and computational powers. The big data era, however, also opened gates to unprecedentedly reap benefits from information sharing and big data generation (Puaschunder, 2017). The so-called nudgital society was recently introduced, shedding light onto the undescribed hidden social class division between social media users and social media users share private information in their wish to interact with friends and communicate to public. The social media big data holder can then reap surplus value from the information shared by selling it to marketers, who can draw inferences about consumer choices. The big data can also be used for governance control purposes, for instance border protection and tax compliance control.

Drawing from the economic foundations of utility theory, this paper seeks the introduce the first application of utility theory to a preference-values predicament between communication and privacy in the new media era. Behavioral economics insights are advanced in shedding novel light on the conflict between the humane wish to communicate now versus combined information held by unknown big data compilers in the future. An exponential loss of privacy and hyperhyperbolic risks in the future for the information sharer are introduced as behavioral economic decision-making fallibilities. For the overconfident information sharer, it remains largely unforeseeable what the sum of the individual information sharing tranches can lead to over time and what information its Gestalt holds for those who have big data insights over time, which can also be analyzed in relation to the general population. Governance gains a critical stance on new media use for guiding on public concerns regarding privacy and information sharing in the digital age (Puaschunder, 2017). While there is some literature on the history of media on politics (Prat & Strömberg, 2013), the wide societal implications of fake news and discounting misinformation has widely been overlooked in contemporary behavioral economics research and the externalities literature. Social sciences literature on privacy and information sharing has to be reconsidered in the age of social media.

The article is structured as follows: An introduction of the theory of utility and communication and information sharing is followed by an outline of the impetus of the digital big data age on privacy. The first utility theory of information sharing and privacy will be theoretically introduced. Hyperbolic decision making fallibility will become the basis of argumentations around hyper-hyperbolic discounting – the novel argument that information sharing in tranches may lead to an underestimation of the privacy infringements when these bits of information can be put together over time and are compared to big data in order to infer about the individual in relation to the general population.

The subjective additive utility of information-shared tranche by tranche may underestimate the big data holder's advantage to reap benefits from information shared. The discussion introduces problems of the contemporary nudgital society (Puaschunder, 2017), in which big data compilers can reap a surplus value from selling compiled information (The New York Times, November 14, 2017)¹ or manipulate vulnerable population segments based on their previously shared information (The Economist, November 4, 2017).² Implications lead to open questions about ethics in the information age and recommendations for a reclaiming of the common good of shared knowledge in education about information sharing in the digital age as well as the democratization of information. Challenging contemporary behavioral insights theory aims at fostering a more informed, self-determined and protected digital society in the wish to uphold e-ethics in the 21st century big data social media era.

Utility theory

Economic theory is built upon the idea of utility, which captures people's preferences or values (Fishburn, 1968). Human are believed to strive to maximize utility on a constant basis by weighting their preferences and values on the pleasure they would receive from different options. In neoclassic economics, utility theory primarily focuses on prescriptive utility maximization giving recommendations how individuals should behave to maximize their utility. Prescriptive utility maximization theory serves as normative guide in helping the decision maker codify preferences. If preferences would violate rational preference choices, the theory suggests strategies so the informed decision maker can revise their rational reference choices and judgments to eliminate preference inconsistency. Using utility theory, preferences are constantly transformed into corresponding numerical utility data that is portrayed to maximize the individual's pursuit of happiness. Utility theory provides a powerful set to determine how to compare actual alternatives. It enables the decision maker's optimal preferences to be transformed into a numerical utility structure guided by an optimization algorithm.

In doing so, standard microeconomic utility theory has been of aid to explain how to maximize individual outcomes in very many different domains ranging from marketing research (Greenberg & Collins, 1966; Marquardt, Makens & Larzelere, 1965; Stafford, 1966), food industry quality control of products and corporate strategies (Read, 1964; Stillson, 1954) and production (Aumann & Kruskal, 1958, 1959; Suzuki, 1957).

¹ https://www.nytimes.com/2017/11/14/business/dealbook/taxing-companies-for-using-our-personal-

 $data.html?rref=collection\%2Fsectioncollection\%2Fbusiness\&action=click\&contentCollection=business\®ion=stream\&modu\ le=stream_unit\&version=latest\&contentPlacement=8\&pgtype=sectionfront$

² https://www.economist.com/news/leaders/21730871-facebook-google-and-twitter-were-supposed-save-politics-good-information-drove-out

Expected utilities

Utility maximization lies at the heart of common sense rational decision making. It provides a powerful tool to resolve multidimensional value maximization by weighting the pros and cons of alternatives in additive utility theory. Uncertainty can be coped with by comparing expectations as to what might result from alternative choices in expected utility theory. The so-called stochastic utility theory makes assumptions in terms of probabilities of choice rather than capturing actual preferences (Debreu, 1958, 1960; Fishburn, 1968; Suppes, 1961). Expected utility theory therefore introduces a first temporal discussion of expectations of utility rather than the actual utility derived from a choice later (Alchian, 1953; Marschak, 1950; Strotz, 1953). Given future uncertainty, the utility of any P can be computed as the weighted sum of the utilities of the x in X, the weights being the probabilities assigned by P. The expected utility equation reads (Bernoulli, 1954):

$$u(P) = P(x^{1})u(x^{1}) + P(x^{2})u(x^{2}) + \dots + P(x^{m})u(x^{m}),$$
 (Equation 1)

whereby u stands for utility, P denotes probability of the u of the separate alternatives x^1 , $x^2 \dots x^m$. The weighting of different alternatives in relation to each other is exhibited in indifference curves (Edgeworth, 1881; Pareto, 1906/2014).

Indifference curves

According to utility theory, individuals are constantly evaluating competing choice options. Given a consumer's budget constraints and prices of goods, commodity bundle arise, in which consumers weight alternative options based on their expected utility derived. Indifference curves connect points on a graph representing different quantities of two goods, between which a consumer is indifferent. For instance, if a consumer weights whether or not to buy good x or good y, the indifference curves would outline how much of good x and good y can be consumed to end with the same utility given the budget of the consumer. This is how much the utility of one good must be increased to offset a decrease in the utility of another good.

Graph 1 represents standard indifference curves for two competing goods, good x and good y. The indifference curve outlines that the consumer has no preference for one combination or bundle of goods over a different combination of the same curve. All points on the curve hold the same utility for the consumer. The indifference curve is therefore the locus of various points of different combinations of two goods providing equal utility to the consumer. Indifference curves represent observable demand patterns for individual consumers over commodity bundles (Edgeworth, 1881; Pareto, 1906/2014).³

In classical economics, an individual is believed to always being able to rank any consumption bundles by order of preference (Jevons, 1871).⁴ The curves are negatively sloped curves – as the quantity consumed of one good x increases, total satisfaction increases if not offset by a decrease in the quantity of good y. Indifference curves fulfill the axioms of completeness, transitivity and strictly convex preferences. Preferences are complete, meaning the consumer ranks all available alternative combinations of commodities in terms of satisfaction provided. Preference rankings are reflexive, transitive, continuous and strongly monotonous. There is a strongly holding substitution assumption, meaning that consumers are believed to willingly give up or trade-off some of one good to get more of another. The fundamental assertion is that there is a maximum amount that a consumer is willing to give up of one commodity to get another commodity, which

³ https://archive.org/details/mathematicalpsy01goog

⁴ http://www.econlib.org/library/YPDBooks/Jevons/jvnPE.html

will leave the consumer indifferent between the new and the old situation (Silberberg, 2000). Indifference curves are determined by the prices of goods and budget constraints of consumers.



Graph 1: Indifference curve (blue curve) for two goods, good *x* and good *y* given their budget constraint (red line)

Indifference curves for substitute goods are straight lines – see the straight line in green in graph 2 representing a budget constraint for substitute goods.



Graph 2: Indifference curve (green line) for substitute goods, good x and good y

The slope of an indifference curve in absolute value is the marginal rate of substitution, which explains the rate at which consumers are willing to give up one good in exchange for more of the other good. For most goods the marginal rate of substitution is not constant so their indifference curves are curved.

Dignity

Dignity is the right to be valued and respected for one's own sake and to be treated ethically. Everyone has a right to respect for their dignity.⁵ As an inherent, inalienable right, dignity is a core concept in fields such as morality, ethics, law and politics. Often connected to identity and respect for integrity and other fundamental freedoms and rights, dignity is often used to uphold the ethical considerations around oppressed and vulnerable groups, who do not have insights about the consequences of their actions. Individuals derive self-worth from dignity. While dignity itself seems to be a vague concept, it is often used as a boundary condition of what is right, just and fair to argue for the improvement of conditions for discriminated, vulnerable and targeted. Violations of dignity are felt as humiliation, instrumentalization or objectification, degradation and dehumanization. Privacy infringements have been argued to hold concerns regarding dignity. In the age of digital media and big data, when individual decision makers may have hyper-hyperbolic discounting fallibility regarding their share of data, dignity infringements may happen mainly unnoticed. Individuals may be endowed with reason and conscience but their decision making capabilities may not be able to discount the worth of their information in the future and in relation to other individuals' data forming big data insights. The freedom of expression may hold a shadow of the future. Dignity based protection of medical patients and in biology settings may serve as dignity-based privacy beacons in the age of big data. Dignity has become the legal-ethical foundation of new reproductive and genetic technologies, medicine and genetic ethics research on humans, life and health sciences, ethics around cloning, medical integrity, bioethics, but also against war cruelty, criminal punishment, imprisonment, terrorism, weapons, abortion, sex work and defamation. The core idea of dignity is prevalent in cultures of the world and has been extended onto animals and the environment. Dignity has been associated with privacy, yet utility and the rational choice theories are in direct contention with a dignitybased ontology. For Human Resources managers the question arises whether to uphold human dignity in privacy or derive benefit from utility of information sharing.

Utility from information sharing or dignity in privacy

The wish for communication is inherent in human beings as a distinct feature of humanity. Leaving a written legacy that can inform many generations to come is a humane-unique advancement of society. At the same time, however, privacy is a core human value. People choose what information to share with whom and like to protect some parts of their selves in secrecy. Protecting people's privacy is a codified virtue around the world to uphold the individual's dignity. Yet to this day, no utility theory exists to describe the conflict arising from the individual preference to communicate and the value of privacy.

The humane preference for communication

The act of conveying intended meanings from one entity or group to another through the use of mutually understood signs and semiotic rules is the act of communication. Communication is a key feature of humans, animals and even plants (Witzany, 2012). Steps inherent to all human communication are the formation of communicative motivation and reason, message composition as further internal or technical elaboration on what exactly to express, message encoding, transmission of the encoded message as a sequence of signals using a specific channel or medium, noise sources influencing the quality of signals propagating from the sender to one or

⁵ United Nations 1998 UNESCO Declaration on the Human Genome and Human Rights. At Article 2
more receivers, reception of signals and reassembling of the encoded message from a sequence of received signals, decoding of the reassembled encoded message and interpretation or sense making of the presumed original message (Shannon, 1948). Information sharing implying giving up privacy is at the core of communication. Communication can be verbal and non-verbal. Comprising very many different domains ranging from business, politics, interpersonal, social to mass media; communication is a humane-imbued wish and center core of every functioning society.

In society, language is used to exchange ideas and embody theories of reality. Language is the driver of social progress (Orwell, 1949). Linguists find discourse and information sharing inseparable from socio-economic societal advancement (Fowler, Hodge, Kress & Trew, 1979). Language and communication modes are implicit determinations of social strata (Orwell, 1949). Different institutions and media sources have different varieties of language and information sharing styles. Access to information is related to social status and market power. Social visibility is a powerful and cheap incentive to make people contribute more to public goods and charities and be less likely to lie, cheat, pollute or be insensitive and antisocial (Ali & Benabou, 2016). Information receipt is an implicit determinant to classify and rank people to assert institutional or personal status in society (Fowler et al., 1979). Mass communication echoes in economic cycles in the creation of booms and busts (Puaschunder, 2018). Media is also a hallmark of propaganda and political control (Besley & Prat, 2006; Prat & Strömberg, 2013). At the same time, privacy is a human virtue around the world.

Privacy as a human virtue

Privacy is the ability of an individual or group to seclude themselves, or information about themselves, and thereby share information about themselves selectively. The right to privacy grants the ability to choose which information about parts of the self can be accessed by others and to control the extent, manner and timing of the use of those parts we choose to disclose. Privacy comprises of the right to be let alone, the option to limit the access others have to one's personal information and secrecy as the option to conceal any information about oneself (Solove, 2008).

The degree of privacy varies in autonomy levels throughout individualistic and collectivism cultures. While the boundaries and contents protected and what is considered as private differ widely among cultures and individuals, the common sense in the world is that some parts of the self should be protected as private.

Privacy has a valued feature of being something inherently special or sensitive to a person, which can create value and specialty if shared with only a selected person or group. The domain of privacy partially overlaps with security, confidentiality and secrecy, which are codified and legally protected throughout the world, mainly in privacy laws but also in natural laws of virtues of integrity and dignity. Privacy is seen as a collective core human value and fundamental human right, which is upheld in constitutions around the world⁶ (Johnson, 2009; Warren & Brandeis, 1890).

In personal relations, privacy can be voluntarily sacrificed, normally in exchange for reciprocity and perceived benefits. Sharing private information can breed trust and bestow meaningfulness to social relations. Giving up privacy holds risks of uncertainty and losses, which are undescribed in economics and in particular the behavioral economics literature on intertemporal decision-making (Gaudeul & Giannetti, 2017). People tend to be more willing to

⁶ E.g., Asian-Pacific Economic Cooperation, Australia, Brazil, Canada, China, European Union, Italy, Japan, Korea,

Organisation for Economic Co-operation and Development, South Africa, United Kingdom, United Nations, United States, Universal Declaration of Human Rights – to name a few.

voluntarily sacrifice privacy if the data gatherer is seen to be transparent as to what information is gathered and how the information will be used (Oulasvirta, Suomalainen, Hamari, Lampinen & Karvonen, 2014). Privacy as a prerequisite for the development of a sense of self-identity is a core of humanness (Altman, 1975). Privacy is often protected to avoid discrimination, manipulation, exploitation, embarrassment and risks of reputational losses, for instance, in the domains of body parts, home and property, general information of private financial situations, medical records, political affiliation, religious denomination, thoughts, feelings and identity.

Technological shocks have a history of challenging privacy standards (Warren & Brandeis, 1890). The age of instant messaging and big data, however, has leveraged the idea of privacy to another dimension. The concept of information privacy has become more significant as more systems controlling big data appear in the digital age. With advances in big data, face recognition, automated licence-plate readers and other tracking technologies, the upholding privacy and anonymity has become increasingly expensive and the cost is more opaque than ever before (Ali & Benabou, 2016).

Privacy in the digital big data era

The amount of big data stored each second has reached an all time high in the digital era. Internet privacy is the ability to determine what information one reveals or withholds about oneself over the internet, who has access to personal information and for what purpose one's information may be used. Privacy laws in many countries have started to adapt to changes in technology in order to cope with unprecedented constant information surveillance possibilities, big data storage opportunities and computational power peaks. For instance, Microsoft reports that 75 percent of U.S. recruiters and human-resource professionals use online data about candidates, often using information provided by search engines, social-network sites, photo and video sharing tools, personal web appearances like websites and blogs, as well as Twitter.

Social media tools have become large-scale factories with unpaid labor (Puaschunder, 2017). For instance, Facebook accounts for the largest social-network site with nearly 1,490 million members, who upload over 4.75 billion pieces of content about their lives and that of others daily. The accuracy of this information also appears questionable, with about 83.09 million accounts assumed to be fake. Aside from directly observable information, social media sites can also easily track browsing logs and patterns, search queries or secondary information giving inferences about sexual orientation, political and religious views, race, substance use, intelligence and overall personality, mental status, individual views and preferences (Kosinski, Bachrach, Stillwell, Kohli & Graepel, in press; Kosinski, Stillwell & Graepel, 2013).

As for the unprecedented possibilities to collect data, store big data and aggregate information that can be compared to big data Gestalt over time and society, privacy has leveraged into one of the most fragile areas of concern in the electronic age, demanding for legal protection, regulatory control and e-ethics (Flaherty, 1989). Today, the existing global privacy rights framework in the digital age has been criticized to be incoherent, inefficient and in need for revision. Global privacy protection shields are demanded to be established. Yet to this day there is no economic framework on information sharing and privacy control. While – for instance – Posner (1981) criticizes privacy for concealing information, which reduces market efficiency; Lessig (2006) advocates for regulated online privacy. As of now we lack a behavioral decision making frame to explain the privacy paradox of the individual predicament between the humane-imbued preference to communicate and information share versus value of privacy. We have no behavioral economics description of inconsistencies and moderator variables in the decision between online information sharing behavior and retroactive preference reversal preferences in the eye of privacy concerns in the digital big data era.

A utility theory of information sharing and privacy

Building on classical utility theory, individuals are constantly evaluating competing choice options. Individuals weight alternative options based on their expected utility derived. Indifference curves would then connect points on a graph representing different quantities of two goods, between which an individual is indifferent.

In the case of the privacy paradox of information sharing preferences and privacy values, a person would weights whether or not to share information s or choose the information to remain private p. The respective indifference curves would outline how much of information sharing s and privacy p can be enabled to end with the same utility given the budget of overall information held by the decision maker.

Graph 3 represents the respective indifference curves for information sharing s and privacy p. That is, the individual has no preference for one combination or bundle of information sharing or privacy over a different combination of the same curve. All points on the curve hold the same utility for the individual. The indifference curve is therefore the locus of various points of different combinations of privacy and information sharing providing equal utility to her or him. Indifference curves are thereby seen to represent potentially observable behavioral patterns for individuals over information bundles. The indifference curve for information sharing s and privacy p is subject to communication and information constraints, hence all information budgets and communication opportunities. There is only a finite amount of information and there may be environmental conditions determining if people can exchange and share information. As exhibited in graph 3, the indifference curve for information sharing s and privacy p is a straight line given the assumption that information sharing or privacy are substitutes.

Graph 3: Indifference curve (blue line) for information sharing s and privacy p given the total information and communication constraint





While in classical economics, an individual was believed to always being able to rank any consumption bundles by order of preference (Jevons, 1871);⁷ the indifference curve for information sharing *s* and privacy *p* subject to communication and information constraints may feature a hyper-hyperbolic element or temporal dimension. The information share moment may

⁷ http://www.econlib.org/library/YPDBooks/Jevons/jvnPE.html

thereby be a reference point. At the moment of the information sharing decision, it may not be foreseeable what the future implication of the information sharing is.

In general, the costs and benefits of communication are assumed as linear subtraction of positive benefits of communication b_c minus the negative consequences of communication c_c . The nature of the problem is intertemporal as information sharers cannot foresee the future implications of their information sharing divided by variance σ (Prat, 2017).

$$\frac{b_c - c_c}{\sigma}$$
 (Equation 2)

However, the digital social media era has heralded a hyper-hyperbolic discounting fallibility. Individuals have lost oversight of the consequences of their individual information sharing given big data hoarding capabilities, which also allow drawing inferences about the individual in relation to others.

In the digital big data era, information share online may hold unforeseen risks of privacy merchants or social media capitalists that commercialize information reaping hidden benefits from the information provided (Etzioni, 2012; Puaschunder, 2017; The Economist, November 4, 2017).8 The subjective additive utility of information shared tranche by tranche may underestimate the big data holder's advantage to reap benefits from information shared given unprecedented data storage and big data computation power advantages of the big data era. Unprecedented computational power and storage opportunities have created the possibility to hoard information over time and put it in context with the rest of the population in order to draw inferences about the information sharer (The New York Times, November 14, 2017).⁹ The digital age and era of instant information sharing have therefore heralded problems of individuals who give in their basic humane need for information communication to become vulnerable over time. The big data information holder may thereby benefit from the history of information and the relation of the individual's information in comparison to the general population to an unknown degree given missing e-literacy and transparency. Comparison to the general public may lead to an implicit underrepresentation and hence discrimination of vulnerable groups. For instance, certain groups that may not be represented online will therefore likely face an under-advocacy of their rights and needs.

While regular hyperbolic discounting captures a game theoretical predicament of the self now versus the self later, the information offering more of a Gestalt in the eyes of the big data holder, leverages hyperbolic discounting to a game theory against uncertainty on the end of the big data holder. The hyper-hyperbolic discounting fallibility therefore may describe that at the moment of information sharing, the individual has hardly any grasp what is implied in the giving up of privacy. The individual only focuses on the current moment trade-off between information sharing and privacy upholding, but hardly has any insights what the compiled information over time holds for big data moguls. As for holding computational and storage advantages, the social media big data moguls can form a *Gestalt* which is more than the sheer sum of the individual information shared, also in comparison to the general populace's data. The shared information can also be resold to companies (Etzioni, 2012; The New York Times, November 14, 2017).¹⁰ In relation to other people's information, the big data moguls can make predictions about their

⁸ https://www.economist.com/news/leaders/21730871-facebook-google-and-twitter-were-supposed-save-politics-good-information-drove-out

⁹ https://www.nytimes.com/2017/11/14/business/dealbook/taxing-companies-for-using-our-personal-

 $data.html?rref=collection\%2Fsectioncollection\%2Fbusiness\&action=click\&contentCollection=business\®ion=stream\&modu\ le=stream_unit\&version=latest\&contentPlacement=8\&pgtype=sectionfront$

¹⁰ https://www.nytimes.com/2017/11/14/business/dealbook/taxing-companies-for-using-our-personal-

 $data.html?rref=collection\%2Fsectioncollection\%2Fbusiness\&action=click\&contentCollection=business\®ion=stream\&modu\ le=stream_unit\&version=latest\&contentPlacement=8\&pgtype=sectionfront$

choices and behaviors.¹¹ Information can also be used for governance purposes, for instance tax compliance and border control mechanisms (Puaschunder, 2017). Some governments have recently used big data to check the accuracy of tax reports but also to detect people's political views when crossing borders (Puaschunder, 2017). Lastly, the use of big data inferences also implies hidden persuasion means – nudging can be turned against innocent information sharers who have no long-term and computational advantage to foresee the impact of the information share (The Economist, November 4, 2017; Puaschunder, 2017).¹²

While behavioral economics hyperbolic discounting theory introduces the idea of timeinconsistency of preferences between an individual now and the same individual in the future; hyper-hyperbolic discounting underlines that in the case of information sharing preferences this fallibility is exacerbated since individuals lose control over their data and big data moguls can reap surplus value from the social media consumer-workers' information sharing and derive information complied over time and in relation to the general norm to draw inferences about the innocent information sharer. With the modern digital era, all these features open an information sharer versus information reaper divide in the big data age (Puaschunder, 2017).

From the social media big data capitalist view, the information gain of one more person sharing information is exponentially rising. Hence, the marginal utility derived from one more person providing information is increasing exponentially and disproportionally to the marginally declining costs arising from one more person being added to the already existing social media platform. Communication costs and benefits are assumed to not be additive and separable.

Expected utility and subjective probability in the digital big data era

In accordance with neo-classical utility theory, decision makers weight alternatives based on the resulting consequences dependent on uncertain aspects of the environment. But in the digital big data era, individuals simply lack an oversight of the consequences of information sharing. Assumptions on preferences of information sharing are skewed leading to an underestimation of the consequences of amalgamated information and private information evaluated in relation to other's data. Assignment of utilities to the consequences are underinflated. The utility of information sharing is thus the underweighted sum of the utilities of the consequences.

Time preferences

Following the standard neo-classical nomenclature of time preferences among the population, an information sharing preference over time is introduced. Multi-period decision-making addresses that for each time period, another set of preferences for the same options can be expected. The different perspectives on privacy can inform management practice and policy as outlined in the following. The populace may therefore be theoretically categorizes into

(1) *Extreme impatience*: extreme information sharing as the individual values immediate pleasure of information sharing, information is shared without hesitation, impression management may play a role in this. These customers may not demand privacy protection from corporations.

(2) *Impatience*: discounting the future impact of information, uninformed information sharing nature. This is the case if an individual shares information although having a hunch that this information sharing may create problems in the future, called the privacy paradox. Consumers

¹¹ https://www.nytimes.com/2017/11/14/business/dealbook/taxing-companies-for-using-our-personal-

data.html?rref=collection%2Fsectioncollection%2Fbusiness&action=click&contentCollection=business®ion=stream&modu le=stream_unit&version=latest&contentPlacement=8&pgtype=sectionfront

¹² https://www.economist.com/news/leaders/21730871-facebook-google-and-twitter-were-supposed-save-politics-good-information-drove-out

with impatience will also likely continue to use social media, despite the social backlash recently when revealing data leakages.

(3) *Eventual impatience*: discounting the future impact of information at some point in the future leads to controlled information sharing, very likely choosing what categories to expose to public. These customers should be targeted and convinced that social media executives try to ensure privacy protection and adapt to a changing climate quickly and efficiently to uphold dignity in privacy of users.

(4) *Time perspective*: related to hyper-hyperbolic discounting awareness, individuals may control information sharing, for instance, these individuals may participate in social media only to reap information from others but not contribute additional information beyond what is required. This type has a controlled privacy and is engaged in social media solely to reap benefits of other's information from social media networks. These customers should also be specially treated and convinced that social media executives try to ensure privacy protection and adapt to a changing climate quickly and efficiently to uphold dignity in privacy of users. A second line of social media account that may cost but ensures privacy protection and secure data handling should be considered by social media providers.

(5) *No time preference*: at the present time the individual neither discounts nor overcounts the future with respect to the present, which may be true for individuals who do not at all participate in social media communication and are blasé about information sharing and gaining information on social media. No change in portfolio composition and communication strategy is required for this kind of preference segment.

(6) *Persistence*: consistent preference structure regarding information sharing may result in informed information sharing with no regrets. Also for this consumer segment, no changes are required in the status quo.

(7) *Variety*: consistently varying preference structure regarding information sharing, likely dependent on the content of information shared, may result in information sharing with regrets afterwards. These individuals have no stringent position towards information sharing or privacy preference, likely have categories for what to share and what not. This type varies in information preferences over time and by subject category. These customers are hard to read and should be studied in further research to detect a pattern of information sharing and privacy concerns.

This nomenclature addresses the problem of capturing valid preference structures over time that allow stringent predictions of choice behavior and strategies for future planning (Fishburn, 1968). The nomenclature also highlights that the selection of information sharing or privacy has an impact on our later choices. Addressing this predicament, Klein and Meckling (1958) suggest the best strategy given future uncertainties is to concentrate attention on immediate decisions that lead toward the main objective while preserving a reasonable degree of freedom in future choices. For management practitioners that means that means to remain flexible and have several option strategies at hand. While Strotz (1957) considers the maximization of utility in an additive, discounted form over a continuous-time future this could be helpful for managerial contexts as A powerful strategy to alleviate hyperbolic discounting has been unraveled in prewell. commitment and consistent planning as means to curb harmful decision-making fallibility. Joint decision making of putting two time alternatives next to each other has been proven as another powerful strategy to curb hyperbolic discounting in the public policy domain (Puaschunder & Schwarz, 2012). KI Yet the age of social media generated big data may impose novel hyperbolic discounting fallibility onto the information-sharing individual (Behears, Choi, Laibson, Madrian & Sakong, 2011; Chabris et al., 2008; Koopmans, 1964). Future research may test the reliability and validity of the nomenclature and unravel moderator variables and variances between different populations, e.g., such as age, cultural heritage, gender, etc.

Expected utility and subjective probability

In accordance with neo-classical utility theory, alternatives are weighted based on the resulting consequences dependent on uncertain aspects of the environment. Assumptions on preferences between such alternatives lead to an assignment of utilities to the consequences and to the alternatives plus an assignment of subjective probabilities to the possible states of the environment. The utility of an alternative can therefore be written as a weighted sum of the utilities of the consequences. The weight for any alternative-consequence pair is the subjective probability associated with the states of the environment that yield the given consequence when the given alternative is used (Fishburn, 1968).

Regarding expected utility, the overall expected utility equation for information sharing and privacy reads

$$u = \sum w * u_s + w * u_p, \tag{Equation 3}$$

whereby w stands for weight, u_s is the utility of information sharing and u_p the utility of privacy. The weighted expected utility equation reads

$$u(P) = P(x^{1})u_{s}(x^{1}) + P(x^{2})u_{p}(x^{2}),$$
 (Equation 4)

whereby $P(x^1)u_s(x^1)$ is the probability of information sharing utility and $P(x^2)u_p(x^2)$ the probability of privacy.

In the digital age, the utility of privacy is expected to have a marginal exponential value given the exponential rise of utility for the big data holder to reap benefits from data. The more data is held, the more complex relations can be unraveled by the big data holder. Information can be put into context of time and population correlates. While there is a marginally declining cost of an additional social media user using an established social network, there is an disproportionally large social network gain with another person joining for the social network provider, who can reap an excessive exponentially increasing marginal utility of another person joining and sharing another piece of information. Given the absence of any taxation of this gain,¹³ social media has leveraged into an IT monopoly (Soros, 2018).

Discussion: Reclaiming the commons of knowledge: A theory of self determination in the digital big data age

In the age of instant communication and social media big data; the need for understanding people's tradeoff between communication and privacy has leveraged to unprecedented momentum. For one, enormous data storage capacities and computational power in the e-big data era have created unforeseen opportunities for big data hoarding corporations to reap hidden benefits from individual's information sharing.

In the 21st century, the turnover of information and the aggregation of social informational capital has revolutionized the world. In the wake of the emergence of new social media communication and interaction methods, a facilitation of the extraction of surplus value in shared information has begun. Computational procedures for data collection, storage and access in the large-scale data processing have been refined for real-time and historical data analysis, spatial and

¹³ https://www.nytimes.com/2017/11/14/business/dealbook/taxing-companies-for-using-our-personal-

 $data.html?rref=collection\%2Fsectioncollection\%2Fbusiness\&action=click\&contentCollection=business\®ion=stream\&modu\ le=stream_unit\&version=latest\&contentPlacement=8\&pgtype=sectionfront$

temporal results as well as forecasting and now casting throughout recent decades. All these advancements have offered a multitude of in-depth information on human biases and imperfections as well as social representations and collective economic trends (Minsky, 1977; Moscovici, 1988; Puaschunder, 2015; Wagner & Hayes, 2005; Wagner, Lorenzi-Cioldi, Mankova & Rose, 1999).

The digital age has brought about unprecedented opportunities to amalgamate big data information that can directly be used to derive inferences about people's preferences in order to nudge and wink them in the nudgitalist's favor. In today's nudgital society, information has become a source of competitive advantage. Technological advancement and the social media revolution have increased the production of surplus value through access to combined information. Human decisions to voluntarily share information with others in the search for the humane pleasure derived from communication is objectified in human economic relations. Unprecedented data storage possibilities and computational power in the digital age, have leveraged information sharing and personal data into an exclusive asset that divides society in those who have behavioral insights derived from a large amount of data (the nudgers) and those whose will is manipulated (the nudged).

The implicit institutional configuration of a hidden hierarchy of the nudgital society is structured as follows: Different actors engage in concerted action in the social media marketplace. The nudgital-brokers are owners and buyers of social media space, which becomes the implicit means of the production. In the age of instant global information transfer, the so-called social media industrialist-capitalist provides the social media platform, on which the social media consumer-workers get to share information about their life and express their opinion online for free. In their zest for a creation of a digital identity on social media platforms, a 'commodification of the self' occurs. Social media consumer-producer-worker are sharing information and expressing themselves, which contributes to the creation of social media experience (Puaschunder, 2017).

The hidden power in the nudgitalist society is distributed unevenly, whereby the social media consumer-workers are slaves, who receive no wages in return for their labor, falling for their own human nature to express themselves and communicate with one-another. Social media consumer-workers also engage in social media expression as for their social status striving in the social media platforms, where they can promote themselves. By posing to others in search for social status enhancement and likes, they engage in voluntary obedience to the social media capitalist-industrialist who sells their labor power product of aggregated information to either capitalists or technocrats. The social media consumer-worker's use value is inherent in their intrinsic motivation to satisfy a human need or want to communicate and gain respect from their community. The use value of the commodity is a social use value, which has a generally accepted use-value derived from others' attention and respect in the wake of information sharing in society. The social media provider gives the use value an outlet or frame, which allows the social media consumer-worker to express information, compare oneself to others and gain information about the social relation to others. The consumer-laborer thereby becomes the producer of information, releasing it to the wider audience and the social media industrialist. This use value only becomes a reality by the use or consumption of the social media and constitute the substance of consumption. The tool becomes an encyclopedic knowledge and joy source derived from the commodity.

But the use of social media is not an end in itself but a means for gathering more information that can then be amalgamated by the social media capitalist-industrialist, who harvests its use value to aid nudgers (Marx, 1867/1995). It is a social form of wealth, in form of social status and access to knowledge about others that the use value materializes on the side of the industrialist in the exchange value. For the social media industrialist, who is engaged in economic and governmental relations, the exchange value of the information provided by his or her social media

consumer-laborers is the information released and consumption patterns studied. In exchange this allows to derive knowledge about purchasing and consumption patterns of the populace and therefore creates opportunities to better nudge consumers and control the populace. With the amalgamated information, the social media industrialist-capitalist can gain information about common trends that can aid governmental officials and technocrats in ensuring security and governance purposes. Further, the social media platform can be used for marketing and governmental information disclaimers as media influences politics (Calvo-Armengol, de Marti & Prat, 2015; Prat, 2017; Prat & Strömberg, 2013).

Exchange value is a social process of self-interested economic actors taking advantage of information sharing based on utility derived from consuming the social media. The social media industrialist-capitalist can negotiate a price based on the access to the social media consumer-worker's attention and sell promotion space to marketers. The exchange value of the commodity of information share also derives from the subjective perception of the value of amalgamated data. Exchanged information can be amalgamated by the social media industrialist-capitalist and traded to other market actors. Exchange value is derived from integrating everything the worker is and does, so both in his creative potential and how he or she relates to others. Exchange value also stems from the exchange of the commodity of amalgamated information that enables an elite to nudge the general populace. The amalgam of information as a premium signals the average opinion and how the majority reacts to changing environments, which allows inferences about current trends and predictions how to react to market changes.

Underlying motives may be the humane desire for prestige and distinction on both sides – the industrialist-capitalist's and the consumer-worker's. From the industrialist-capitalist's perspective, monetary motives may play a role in the materialization of information; on the consumer-worker's side it is the prestige gained from likes, hence respect for an online identity created (Ali & Benabou, 2016). Individuals may experience a warm glow from contributing to the public good of common knowledge (Ali & Benabou, 2016). The benefits of the superior class are the power to nudge, grounded on the people's striving for prestige and image boosts. Impression management and emotions may play a vital role in seducing people to share information about themselves and derive pleasure for sharing (Evans & Krueger, 2009; Horberg, Oveis & Keltner, 2011; Lerner, Small & Loewenstein, 2004). Social norms and herding behavior may be additional information sharing drivers (Paluck, 2009). The realization of prestige stems from creating a favorable image of oneself online, which signs up the workers in a psychological quasi-contract to provide more and more information online and in a self-expanding value. Prestige is also gained in the materialization of information as asset by the capitalist-industrialist, who reaps the surplus value of the commodification of the self of the consumer-worker based on socio-psychological addiction to social media (Marx, 1867/1995; Soros, 2018). In the wake of an addiction to social media, users get distracted from profitability for their own terms and experience a loss of autonomy bit by bit. The social media capitalist-industrialist therefore increases their capital based on the social media consumer-worker's innocent private information share. The social media capitalist-industrialist also accumulated nudgital, the power to nudge.

This information sharing opens a gate for the social media provider to reap surplus value from the information gathered on social platforms and to nudge the social media consumer-producers or resell their amalgamated information to nudgers. Crucial to the idea of exploitation is the wealth or power of information in the digital age. While classical economic literature finds value in organizational hierarchy to economize transaction costs, the age of big data has opened a gate to reap disproportional benefits from individual data and information sharing. Surplus of information can be used to nudge in markets and by the force of governments. To acknowledge social media consumers as producers leads to the conclusion of them being underpaid workers in a direct wage labor exploitation. Surplus gravitates towards the social media owning class. Information becomes a commodity and commodification a social product by the nature of communication. Commodification of information occurs through the trade of information about the consumer-worker and by gaining access to nudge consumer-workers on social platforms. The transformation of a labor-product into a commodity occurs if information is used for marketing or governance purposes to nudge people. In the contemporary big data society, the nudged social media user therefore end up in a situation where they are unwaged laborers, providing the content of entertainment within social media, whereas the social media industrialist-capitalist, who only offers the information brokerage platform and is not subject to tax per information share, reaps extraordinary benefits from the amalgamated information shared. Not just labor power but the whole person becomes the exchange value, so one could even define the consumer-worker as utility-slave.

The technological complexity of digital media indicates how interrelated social, use and exchange value creation are. All commodities are social products of labor, created and exchanged by a community, with each commodity producer contributing his or her time to the societal division of labor. Use value is derived by the consumer-worker being socially related insofar as private consumption becomes collective. The use value thereby becomes the object of satisfaction of the human need for social care and want for social interaction. The use value becomes modified by the modern relations of production in the social media space as the consumer-worker intervenes to modify information. What the consumer-worker says on social media, for the sake of communication and expression but also in search for social feedback, is confined by the social media industrialist-capitalist, who transforms the use value into exchange value by materializing the voluntary information share by summing it up and presenting it to nudgers, who then derive from the information marketability and nudgitability of the consumer-workers. All information sharing has value, or labor value, the abstract labor time needed to produce it. The commodification of a good and service often involves a considerable practical accomplishment in trade. Exchange value manifests itself totally independent of use value. Exchange means the quantification of data, hence putting it into monetary units. In absolute terms, exchange value can be measured in the monetary prices social media industrialist-capitalists gain from selling advertisement space to nudging marketers but also to public and private actors who want to learn about consumer behavior in the digital market arena and influence consumers and the populace. The exchange value can also be quantified in the average consumption-labor hours of the consumers-workers. While in the practical sense, prices are usually referred to in labor hours, as units of account, there are hidden costs and risks that have to factored into the equation, such as, for instance, missing governmental oversight and taxing of exchange value.

Overall, there is a decisive social role difference between the new media capitalist-industrialist and the social media consumer-worker. The social media provider is an industrialist and social connection owner, who lends out a tool for people to connect and engage with. As the innovative entrepreneur who offers a new media tool, the industrialist also becomes the wholesale merchant in selling market space to advertisement and trading information of his customers or workers, who are actively and voluntarily engaging in media tools (Schumpeter, 1949). The social media consumers turn into workers, or even slaves if considering the missing direct monetary remuneration for their information share and since being engaged in the new media tool rather than selling their labor power for money in the market place holds opportunity costs of foregone labor. While selling their commodity labor power, the social media consumer-workers are also consumers of the new media tool laden information, which can be infiltrated with advertisement. The social media capitalist-industrialist not only reaps exchange value benefits through access to people's attention through selling advertisement space, but also grants means to nudge the consumers into purchasing acts or wink the populace for governance authorities (Marx, 1867/1995). The social media capitalist-industrialist thereby engages in conversion of surplus value through information sharing into profit as well as selling attention space access and private data of the consumer-workers.

When the new media consumer-workers' amalgam of provided information gets added up to big data sets, it can be used by capitalists and governance specialists. Over time the nudgital society emerges, as the nudging social media industrialist-capitalists form a Gestalt of several bits and pieces put together about the nudged social media consumer-producer-worker-slaves. Information gets systematically added up providing invaluable behavioral insights. Information in its raw form and in amalgamated consistency then gets channeled from the broad working body on social media into the hands of a restricted group or societal class. This circulation of information and the distribution into those who provide a medium of information exchange and those who exchange information that then forms a society in those who nudge and those who are nudged implying an inherent social class divide.

In the nature of exchange, nudgital becomes an abstract social power, a property claim to surplus value through information. Value can be expropriated through the exchange of information between the industrialist-capitalist and the nudgitalist. Exchange value has an inherent nature of implicit class division. Exchange value represents the nudgitalists' purchasing power expressed in his ability to gain labor time that is required for information sharing as a result of the labor done to produce it and the ability to engage in privacy infringements. The social media industrialist-capitalists implicitly commands labor to produce more of data through social nudging and tapping into humane needs to communicate and express themselves, whereby he or her use a reacting army of labor encouraging information share through social gratification in the form of likes and emoticons (Posner, 2000). The reacting army of labor is comprised of social media users, who degrade into hidden laborers that are not directly compensated for their information share and cheerleading others to do the same. The nudgital society's paradox is that information sharing in the social compound gets pitted against privacy protecting alienation.

In all these features apparent becomes the rise of the monopolistic power of giant IT platform companies (Soros, 2018). For instance, Facebook and Google are believed to control over half of all internet advertising revenue (Soros, 2018). While these companies initially played an important innovative and liberating role, by now it has become apparent that they also exploit the social environment (Soros, 2018). Social media companies know how people think and influence them to behave in a certain way without their users having insights or being aware of the hidden influence (Soros, 2018). As George Soros points out at the World Economic Forum 2018, this has far-reaching adverse consequences on the functioning of democracy, particularly on the integrity of elections. It is believed that social media can prime how people evaluate politicians consciously and unconsciously based on the available content (Iyengar & Kinder, 1987). The profitability of these corporations is based on the absence of direct payments for the information shared to the social media users or taxation being imposed on the IT giants (Soros, 2018).¹⁴ While these platforms were initially set up to make the world more flat, by now they have turned to monopoly distributors of the public good knowledge. Acknowledging these monopolistic IT giants as public utilities will help making them more accountable and subject to stringent regulations, aimed at preserving competition, innovation and fair and open universal access to information (Soros, 2018).

The nudgitalist exploitation also holds when technocrats use heuristics and nudges to create selfish outcomes or undermine democracy. Ethical abysses of the nudgital society open when the social media is used for public opinion building and public discourse restructuring. Social media not only allows to estimate target audience's preferences and societal trends but also imposes direct and indirect influence onto society by shaping the public opinion with real and alternative facts. Government officials' gaining information about the populace that can be used to interfere in the democratic voting process, for instance in regards to curbing voting behavior or

¹⁴ https://www.nytimes.com/2017/11/14/business/dealbook/taxing-companies-for-using-our-personal-

 $data.html?rref=collection\%2F section collection\%2F business \& action=click \& content Collection=business \& region=stream \& modu le=stream_unit \& version=latest \& content Placement=8 \& pgtype=section front$

misinformation leading people astray from their own will and wishes. The social intertwining of the media platform and the democratic act of voting has been outlined in recent votes that were accused to have been compromised by availability heuristic biases and fake news. Data can also be turned against the social media consumer-worker by governance technocrats for the sake of security and protection purposes, for instance, social media information can be linked together tax verification purposes.

Governments have been transformed under the impact of the digital revolution. Instant information flow, computational power and visualization techniques, sophisticated computer technologies and unprecedented analytical tools allow policy makers to interact with citizens more efficiently and make well-informed decisions based on personal data. New media technologies equip individuals with constant information flows about informal networks and personal data. Novel outreach channels have created innovative ways to participate in public decision making processes with a partially unknown societal impact at a larger scale, scope and faster pace than ever before. Big data analytics and the internet of things automate many public outreach activities and services in the 21st century. Not only do we benefit from the greatly increasing efficiency of information transfer, but there may also be potential costs and risks of ubiquitous surveillance and implicit persuasion means that may threaten democracy. The digital era governance and democracy features data-driven security in central and local governments through algorithmic surveillance that can be used for corporate and governmental purposes. Open source data movements can become a governance regulation tool. In the sharing economy public opinion and participation in the democratic process has become dependent on data literacy. Research on the nudgital society holds key necessary information about capacitybuilding and knowledge sharing within government with respect for certain inalienable rights of privacy protection. The nudgital society's paradox that information sharing in the social compound gets pitted against privacy protecting alienation requires an ideological superstructure to sustain and tolerate hidden exploitation.

All these features are modern times phenomena as the technology and big data creating computational power is currently emerging. The transferability of the commodity of information itself, hence the big data amalgamation over time and space to store, package, preserve and transport information from one owner to another appears critical. The legal leeway to allow private information sharing implicitly leads to individuals losing their private ownership rights to the commodity of information upon release on social media and the right to trade information. The transferability of these private rights from one owner to another may infringe on privacy protection, human rights and humane dignity upholding mandates.

Not only pointing at the ethical downfalls of the nudgital society, also defining social media users as workers is of monumental significance to understand the construction of the nudgital society and bestow upon us social media consumer-workers labor rights. The technical relationship between the different economic actors are completely voluntary and based on trust (Puaschunder, 2016). The creation of use value is outsourced to the community (e.g., in likes) and the share of information about the workers from the social media capitalist to the market or nudgitalists remains without a clear work contract and without protection of a labor union. The worker-employer relationship needs to be protected and a minimum wage should settle for the market value of the worker producing during the working day. Wages would be needed to maintain their labor power of the workers minus the costs of the production. Unpaid laborers should not only be compensated for their opportunity costs of time but should enjoy the workers' privilege of right to privacy, prevention of misuse of the information they share and have the right to access to accurate information but also protection from nudging in the establishment of the right to voluntary fail.

The nature of making profit from information in exchange value is questionable. Information exchange of the industrialist-capitalist is different than neoclassical goods and services trade

insofar since for the capitalist-industrialist making money off privacy and the consumer-workers share of information without knowledge and/or control over the recipient of the amalgamated mass of privacy released. Workers are never indifferent to their use value and their inputs may also produce unfavorable outcomes for them. The exchange value will sell for an adequate profit and is legally permitted, yet it can destroy the reputation and standing as well as potentially the access of the individual to country entrance if considering the proposed social media information release mandate at border controls. Care must be taken for privacy infringement and the product of amalgamated big data and how useful it is for society.

By shedding light on these risks of the social media age and the implicit dynamism of capitalism forming around information, a social formation of social media workers' right can be heralded. Social media user-workers should be defined to hold inalienable rights to privacy and be forgotten (part 4.1), to be protected from data misuse of information they share (part 4.2), they should be granted the right to access of accurate information (part 4.3) and – in light of the nudgitalist audacity – the right to fail (part 4.4).

People's right to privacy and to be forgotten

The transformation of a use value into a social use value and into a commodity has technical, social and political preconditions. Information gets traded and ownership of privacy transferred in information sharing. Upon sharing information on social media, the consumer-worker bestows the social media capitalist-industrialist with access to previously private information. The social media capitalists then transforms the information into use value by offering and selling the bundled information to nudgitalists, who then can draw inferences about certain consumer group's preferences and guide their choices.

Overall, the nudgital society leads to a dangerous infringement upon the independence of individuals in their freedom of choice and a social stratification into those who have access to the amalgamated information of social media consumer-workers. There is a trade-off between communication and privacy in an implicit contract of the use of personal data. Power is exercised through the accumulation of information, including the quality of insatiability of social media consumer-workers to constantly upload information and the social media capitalist-industrialist reaping profits from selling it.

Social media thereby reveals to hold a sticky memory that allows storage of information in the international arena eternally. Privacy and information share regulations depend on national governments. For instance, in the commodification of privacy, the EU is much more beneficial to consumers than that of the US. Data protection and commercial privacy are considered as fundamental human rights to be safeguarded in Europe. Europe appears in a better position, since it does not have any IT platform monopolistic giants of its own (Soros, 2018). Not only does Europe have much stronger privacy and data protection laws than America. EU law also prohibits the abuse of monopoly power irrespective of how it is achieved (Soros, 2018). US law measures monopoly by the inflated price paid by customers for a service received, which is impossible to prove when the services are free and there is no utility theory of privacy and information sharing that captures the value and price of information (Soros, 2018). In contrast, the US approach towards commercial privacy focuses on only protection the economic interests of consumers. Current privacy regulations are considered as not sufficient in targeting actions that cause non-economic and other kinds of harm to consumers.

Privacy and information sharing guidelines appear to be culturally-dependent phenomena. Information about privacy boundary conditions can be obtained from the transatlantic dialog between the US and Europe on privacy protection. While in Europe health care data is public, in Canada, there is a public interest to make the data more public. The EU's privacy approach is based on Article 7 and 8 of the Charter of the Fundamental Rights of the EU, which grants individuals rights to protection, access and request of data concerning him or herself. European privacy is oriented around consumer consent. The 2016 EU General Data Protection Regulation (GDPR) ruled the right to be forgotten under certain circumstances. Consumer consent and dealing with incomplete, outdated and irrelevant information is legally regulated. GDPR establishes regulatory fines for non-complying companies applicable to foreign companies whose data processing actions are related to 'good and services' that they provide to data subjects in the EU, so also including US companies operating in the virtual space accessible by European citizens. The EU privacy approach offers member states flexibility in data management for national security and other exceptional circumstances but also protects civilians from common potential circumstances of a companies' country-of-origin or operational locations. The EU's privacy approach has higher regulatory costs, is not specified by sectors and the right to be forgotten still needs enforcement validity.

The US approach to privacy is sector specific. Commercial privacy is pitted against economic interests and neither seen as civil liberty nor as constitutional right. US privacy is regulated by the Federal Communications Commission (FCC) and the Federal Trade Commission (FTC). Overall in the US, the general definitions of unfair and deceptive give the FTC a wider scope for monitoring and restricting corporate privacy infringements. The FTC has a wide variety of tools for data protection, yet the responsibility is split between the FTC and the FCC, which increases bureaucratic and regulatory costs and limits industry oversight.

So while the EU framework treats commercial privacy as a basic human right leading to a more extensive protection of individual's privacy including data collection, use and share; the EU framework is also non-sectoral and allows sovereign nation states to overrule common data management policies for the sake of national security and protection. The US framework lacks a centralized privacy regulation approach, yet is sector but split regarding oversight in the domains of the FCC and FTC.

People's right to prevent misuse of information they share

By US standards, social media is required by the FTC to ask users for permission if it wants to alter its privacy practices. Under Section 5 of the FTC Act that states that (1) unfair practices are causes or is likely to cause substantial injury to consumers or cannot reasonably be avoided by consumers; and (2) deceptive practices are practices that likely are misleading or actually misleading the consumer.

In August 2016, the decision of WhatsApp to share more user data – especially user phone numbers – with Facebook in order to track customer-workers' use metrics and refine targeted user advertising also opened a gate to discriminatory pricing. This decision faced a huge backlash in the EU, where data sharing was ordered to be halted and Germany deemed these practices as illegal. In the US the Federal Trade Commission (FTC) began reviewing joint complaints from consumer privacy groups. The recent WhatsApp data sharing is a possible violation of this requirement since it only allowed consumers to opt out of most of the data sharing while lacking clarity and specificity. WhatsApp's restrictive opt out option and incomplete data sharing restrictions were argued to be perceived as unfair and deceptive (Tse, in speech, March 25).

People's right to access to accurate information

Traditional media studies advocate for independence of the media. Commercial motives have ever since raised doubts about reputation and credibility of outlets (Prat & Strömberg, 2013). Technological shocks have always created new opportunities but also opened gates to novel downfalls in the communication realm. Novel technologies for information sharing but also monitoring of communication are prone to significant change in the nature of communication. In such technological leaps, attention to privacy is recommended (Ali & Benabou, 2016).

In the nudgital society, profits appear in the circuit of information and take on different forms in the new media age. The possibility of trading information and reaping benefits from information sharing of others determines the unequal position of people in society. The possession of knowledge stems from the surplus derived from the activity of production, hence the information share of social media consumer-producers. This confrontation of labor and consumption is not apparent in the modern marketplace. The class division remains quite invisible in the implicit workings of the system.

The nudgitalist act becomes problematic when being coupled with infiltration with fake news and alternative facts that curb democratic acts, e.g. manipulating voting behavior. Ethical questions arise if it there is a transparency about the capitalist's share of information and a fair social value benefits distribution among the capitalist and the worker. In addition, under the cloak of security and protection, privacy infringements by sharing information with the nudgitalist is questionable. In the political domain, knowledge has been acknowledged as a public good. Voters who spend resources on obtaining information to keep their government accountable produce a positive externality for their fellow citizens (Prat & Strömberg, 2013).

By outlining the nudgital market procedures and acknowledging knowledge as a public good, fairness in the distribution of gains should be accomplished and privacy infringing information sharing limited, curbed by taxation or guided by the legal oversight. Access to information about the storage, preservation, packaging and transportation of data is non-existent, demanding for more information about behind-the-scenes' social media conduct. Transforming private information from use value to exchange value is an undisclosed and therefore potentially problem-fraught process that holds implicit inequality within itself. From a societal standpoint, also the missing wealth production in the social media economy appears striking. Thereby the dangers of information release and transfer and the hidden exchange value accrued on the side of the media innovator are left unspoken. The importance of shedding light on such, though, is blatant as for stripping the populace from inalienable rights of privacy while reaping benefits at the expense of their susceptibility. Nudges in combination with misinformation and power abuse in the shadow of subliminal manipulation can strip the populace from democratic rights to choose and voluntary fail (Benabou & Laroque, 1992).

As a policy response to the negative implications of the nudgital society, taxing IT giants may enable to raise revenue for reducing cost and noise in collecting political information. For instance, by making news freely available without commercial interruptions. A mixture strategy could be introduced, in which consumers are given the choice to either choose a free account that releases information or pay for a private account, which restricts third party use of their data.

Facebook has recently acknowledged the rise of fake news having an impact on voting behavior and therefore roles out a bottom-up accuracy check mechanism.¹⁵ Truthfulness appears hard to quantify on social media since truth is not easily verifiable and integrity of information embodied in prices is missing through the free information exchange on social media. Reputation and social self-determination mechanism appear as alternate sources of information accuracy checks in the absence of classical price mechanisms (Benabou & Laroque, 1992).

¹⁵ http://www.telegraph.co.uk/news/2018/01/19/facebook-start-trust-ratings-media-outlets-fights-back-against/

People's right to choose and fail

In the personal information sharing age and nudgital society, attention must be given to privacy and human dignity. The nudgital society opens a gate to gain information about consumer choices and voting preferences. The uneven distribution of key information about people's choices opens a gate to tricking people into choices. The so-called nudging attempt though raises ethical questions about human dignity and the audacity of some to know better what is better for society as a whole. Because governance is a historical process, no one person can control or direct it, thereby creating a global complex of governance connections that precedes the individual administration. Structural contradictions describe the class struggle between the nudged in opposition to the nudgers in the nudgital society. Since societal actors who involuntarily are nudged are separated from an active reflection process when being nudged, the moral weight is placed on the nudger. Though democratically elected and put into charge, the nudgers checksand-balances of power seem concentrated and under disguise through the middle man of social media capitalist-industrialists who collect information. Rather focusing on how to trick people into involuntary choices, the revelations should guide us to demand to educate people on a broad scale about their fallibility in choice behavior.

In a self-enlightened society, people have a right to voluntary fail. Nudging implies a loss of degrees of freedom and disrespect of human dignity, hence the nudgital society will lead to structural contradictions. Their rational thinking and voluntary engagement in governmentally-enforced action becomes divorced from rational reflection. No one entity should decide to control or direct other's choices, thereby creating a global complex of social connections among the governed for the sake of efficiency for the common good. The economic formation of human decision making in society should never precede the human voluntary decision.

There is an inherent inequality of social positions, manifested primarily in the respective capacities of reaping benefit from amalgamated information, which leads to a disparity of social position. The distribution of power leads to a natural order of human activity, in which the nudgers are in charge of nudging the populace. Moral value is separated from economic value and hence placing the fate of the populace into the arms of the behavioral economists raises problems of lack of oversight and concentration of objective economic value rule in the nudgital society.

Overall, with the communication on the nudgital society just having started, it remains on us to redesign the apparatus of production in ways that make the infringement on private information through the natural tendency to share information, care about others and express oneself. Governance crises are rooted in the contradictory character of the value creation through big data. The formation of value is a complex determination and we still need more research to understand the deep structures of market behavior in the digital age.

Conclusion and future prospects

The article presented a first theoretical introduction of a trade-off faced by today's corporate decisions makers between serving their customers and workforce to ensure either dignity of privacy or reap benefits in the utility of information sharing. Potential limitations are that some communication may not be integrated in the framework, such as nonverbal communication or emotional responses. In general information exchange is very heterogeneous and vast international differences are assumed to exist. In addition, in what time online communication and under what circumstances decisions regarding communication and privacy are made, remains a completely undiscussed topic.

As a next research step, a stringent hypotheses-testing of the presented predicament is recommended. For instance, future research projects featuring a multi-methodological approach will help gain invaluable information about the actual performance and behavior regarding utility from information sharing and dignity in privacy upholding. Interaction of individuals on social media should be scrutinized in order to derive real-world relevant economic insights for legal and policy making purposes alongside advancing an upcoming scientific field.

Following empirical investigations should employ a critical survey of the intersection of analytic and behavioral perspectives to decision making in information sharing or privacy upholding. Literature discussion featuring a critical analysis how to improve e-literacy should be coupled with e-education and enhancement of e-ethicality. Research should be directed towards a critical analysis of the application of behavioral economics on hyper-hyperbolic discounting in the digital age. In the behavioral economics domain, both approaches, studying the negative implications of information sharing and decision making to uphold privacy but also finding ways how to train new media users wiser decisions should be explored. Interdisciplinary viewpoints and multi-method research approaches should be covered in the heterodox economics readings but also in a variety of independent individual research projects. Research support and guidance should be targeted at nurturing interdisciplinary research interests on privacy and information sharing in the fields of behavioral economics and public affairs.

More concretely, future studies should define the value that data has to individuals and data sovereignty in the international context. When people share information, they should be informed to consider what the benefit and value from information sharing is for them and what the benefit for social media industrialists-capitalists is. The sovereignty of data and the human dignity of privacy should become debated as civic virtual virtue in the 21st century. Individuals should be informed that sharing data is a personal security risk, if considered to be asked for social media information upon entry of a country.

Future studies should describe what companies and institutions constitute the complex system that helps establishing the nudgital society and the influence that social media has. The implicit underlying social structure of the nudgital society based on a complicated information gathering machinery should become subject to scrutiny and how, in particular, the nudgital class division is supported by a comprehensive social network data processing method. How social media advertising space can be used to specialize on targeted propaganda and misleading information to nudge the populace in an unfavorable way should be unraveled. The role of politicians' use of various channels and instruments to manipulate the populace with targeted communication should be scrutinized.

In the recent US election the profit and value of detailed market information has been found to have gained unprecedented impetus. Future research should also draw a line between the results of the 2016 US presidential election, and the study of heuristics to elucidate that heuristics played a key role in Trump's election as they made people less likely to vote logically. This would be key as it would help explain how people chose to vote, and why they do not always make the most logical choice when voting. This line of research could help to more accurately promote future elections' candidates, how to better predict election outcomes and how to improve democracy.

In addition, nudging through means of visual merchandising, marketing and advertising should be captured in order to uphold ethical standards in social media. Nudging's role in selling products, maximizing profits but also creating political trends should be uncovered. While there is knowledge on the visual merchandising in stores and window displays, little appears to be known how online appearances can nudge people into making certain choices. In particular, the familiarity heuristic, anchoring and the availability heuristic may play a role in implicitly guide people's choices and discreetly persuade consumers and the populace. Not to mention advancements of online shopping integrity and e-commerce ethics, the prospective insights gained

will aid uphold moral standards in economic market places and hopefully improve democratic outcomes of voting choices.

Contemporary studies could also address if the age of instant messaging has led to a loss of knowledge in information sharing. Future research should also investigate how search engines can be manipulated to make favorable sources more relevant and how artificial intelligence and social networks can become dangerous data manipulation means. The role of data processing companies may be studied in relation to the idea of data monopoly advantages – hence situations in which data processing companies may utilize data flows for their own purposes to support sponsored causes or their own ideals. Due to the specific time period of the digital age not extrapolations to past time periods is possible but the results appear useful in determining future behaviors.

The current research in this area lacks empirical evidence, demanding for further investigations on how nudges can directly impact individual's choices and new media can become a governance manipulation tool. What social instruments are employed on social media and what prospects data processing has in the light of privacy infringement lawsuits should be uncovered. How social media is utilized to create more favorable social personas for political candidates should be explored. How internet online presences allow to gain as much attraction as possible for the presence of political candidates is another question of concern. Another area of concern is how selective representations influence the voting population and what institutions and online providers are enabling repetitiveness and selectivity. How gathered individual information is used to parse data to manipulate social internet behavior and subsequent action is another topic to be investigated. Future research goals will include determining what this means for the future political landscape and how internet users should react to political appearances online. Information should be gathered how we choose what media to watch and if political views play a role in media selection and retention. Does distrust in the media further political polarization and partisanship, needs to be clarified. Future studies should also look into the relationship between individual's political ideologies and how they use and interact on social media, especially with a focus on the concept of fake news and alternative facts. Where do these trends come from and who is more susceptible to these negative impacts of the digital society? Has social media become a tool to further polarize political camps, is needed to be asked. All these endeavors will help outlining the existence of social media's influence in governance and data processing to aid political campaigning in order to derive inferences about democracy and political ethicality in the digital age.

How social media tools nudge people to not give everything at once but put it together in a novel way that it creates surplus, should be analyzed. In small bits and pieces individuals give up their privacy tranche be tranche. Small amounts of time are spent time for time. People, especially young people, may have a miscalibration about the value of information released about them. Based on hyperbolic discounting myopia, they may underestimate the total future consequences of their share of privacy.

The time spent on social media should become closer subject to scrutiny and the impact on opportunity costs onto the labor market. For instance, countries that ban social media, such as China, or restrict internet, like slowing it down or censoring certain media, could become valuable sources of variance to compare to. Network theories for e-blasting information should become another area of interest to be studied in relation to hyper-hyperbolic discounting fallibilities. Emotional reactions and emotional externalities of communication could be another area of behavioral economics research in the privacy and information sharing predicament domain. The role of attention should be addressed as another moderator variable that is quite unstudied in the digital media era (Prat, in speech, November 2017). Thereby interesting new questions arise, such as how to measure attention – is it the time allocation or the emotional arousal information bestows individuals with (Wouter & Prat, forthcoming)?

The preliminary results may be generalized for other user-generated web contents such as blogs, wikis, discussion forums, posts, chats, tweets, podcastings, pins, digital images, videos, audio files, advertisements but also search engine data gathered or electronic devices (e.g., wearable technologies, mobile devices, Internet of Things). Certain features of the nudgital society may also hold for tracking data, including GPS, geolocation data, traffic and other transport sensor data and CCTV images or even satellite and aerial imagery. All preliminary results should be taken into consideration for future studies in different countries to examine other cultural influences and their effects on social class and heuristics.

Innovative means should be found to restore trust in media information and overcome obstacles such as the availability heuristic leading to disproportionate competitive advantages of media controlling parties. As remedies, consumer education should target at educating social media users about their rights and responsibilities on how to guard their own and other's privacy. E-ethicality trainings could target at strengthening the moral impetus of big data and artificial ethicality in the digital age. Moral trade-offs between privacy infringements and security should also become subject to scrutiny.

Promoting governance through algorism offers novel contributions to the broader data science and policy discussion (Roberts, 2010). Future studies should also be concerned with data governance and collection as well as data storage and curation in the access and distribution of online databases and data streams of instant communication. The human decision-making and behavior of data sharing in regards to ownership should become subject to scrutiny in psychology. Ownership in the wake of voluntary personal information sharing and data provenance and expiration in the private and public sectors has to be legally justified (Donahue & Zeckhauser, 2011). In the future, institutional forms and regulatory tools for data governance should be legally clarified. Open, commercial, personal and proprietary sources of information that gets amalgamated for administrative purposes should be studied and their role in shaping the democracy. In the future we also need a clearer understanding of the human interaction with data and their social networks and clustering for communication results. The guarantee of safety of the information and the guarantee of the replacement or service, should a social media fail its function to uphold privacy law as intended, is another area of blatant future research demand. Novel qualitative and quantitative mixed methods featuring secondary data analysis, web mining and predictive models should be tested for holding for the outlined features of the new economy alongside advancing randomized controlled trials, sentiment analysis and smart contract technologies. Ethical considerations of machine learning and biologically inspired models should be considered in theory and practice. Mobile applications of user communities should be scrutinized.

As for consumer-worker conditions, unionization of the social media workers could help uphold legal rights and ethical imperatives of privacy, security and personal data protection. Data and algorithms should be studied by legal experts on licensing and ownership in the use of personal and proprietary data. Transparency, accountability and participation in data processing should also become freed from social discrimination. Fairness-awareness programs in data mining and machine learning coupled with privacy-enhancing technologies should be introduced in security studies of the public sector. Public rights of free speech online in the dialogue based on trust should be emphasized in future educational programs. Policy implications of the presented ideas range from security to human rights and law to civic empowerment. Citizen empowerment should feature community efforts to protect data and information sharing to be free of ethical downfalls. Social media use education should be ingrained in standard curricula and children should be raised with an honest awareness of their act of engagement on social media in the nudgital society of the digital century.

Future research may also delve into moderator variables of the utility derived from information sharing and privacy. For instance, extraversion and introversion could be moderating the overall

pleasure derived from communication or silence. Future research may also address prescriptive recommendations how to educate individuals about the risks and dangers of information sharing in the digital age. Attention must also be paid to how to uphold accuracy in times of fake news and self-created social information. Certain societal segments that are not represented strongly online should somehow be integrated into big data in order to democratize the information, which is considered as big data 'norm,' or standard by which the social media user is measured on. At the same time, psychologically guide studies could unravel a predictive approach and validate the outlined ideas' validity by testing the proposed theoretical assumptions in laboratory and field study settings. In particular, the proposed nomenclature's validity could be studied and the percentage of information sharing types captured in the population. The moderator variable age could be phased in as it appears to be conundrum why younger people, who have more to lose given a longer time ahead to live are in particular prone to use new social media and lavishly share their lives in e-blasts to public. Regarding direct implications, a tax may be used to offset problems of the costs and risks of social media privacy infringements in the big data era.¹⁶ Drawing from utility usually measured by the willingness to pay different amounts of money for different options, laboratory experiments may operationalize the value of privacy by measuring how much money people would be willing to pay for repurchasing their data or having a social media account that can only be viewed but no personal data can be resold or put in context to others. These attempts could also serve as a guideline for policy regulations and free market solutions. Social media could offer services of having accounts that are private in that sense that no surplus value can be reaped by reselling information or big data storage and computation can occur. This may serve as an indicator of revealed preferences of social media privacy. The privacy paradox may be scrutinized in behavioral economics laboratory and field experiments. Potential individual influencing factors such as gender, age, trust and personality differences may be tested for in order to retrieve information on how to educate the social media user and regulate the social media provider. Regarding regulation, splitting social media power cartels may be one solution to decrease the big data social media user disadvantage. Taxation for information sharing may create another incentive to slow down unreflected information share. The tax revenues could be used to offset some of the societal costs of privacy infringement. In addition, fines for privacy infringement could help to uphold e-ethics in the digital age. From the economics perspective, interesting moderator variables for future studies is the distinction between active and passive communication. Further, model robustness checks could follow and learning effects depicted. Access to information what happens with data and how big data is used appears crucial for learning people a well-calibration of their relation to their information. Communication costs and benefits are assumed to not be additive and separable, leaving an interesting field for future studies in this domain. The communication patterns could be classified in different types of communication in the future, e.g., certain node specificities detected, such as communication within a family, with friends and in hierarchical situations like at work. The absolute and relative influence of information sharers could become part of a network description approach as well. Impact factor measurements could be based on status, search engine rank and connections to capture global influence. Complexity of information would need to be controlled based on information processing times and time allocation preferences to information, hence attention. Communication costs should in the future to be separated in economic models in fixed and variable communication costs and a potential separation between fixed communication costs for social media providers and a variable communication costs for social media users be depicted (Prat, 2017).

¹⁶ https://www.nytimes.com/2017/11/14/business/dealbook/taxing-companies-for-using-our-personal-

<u>data.html?rref=collection%2Fsectioncollection%2Fbusiness&action=click&contentCollection=business®ion=stream&modu</u> <u>le=stream_unit&version=latest&contentPlacement=8&pgtype=sectionfront</u>

Overall, the presented piece can also serve as a first step towards advocating for education about information sharing in order to curb harmful information sharing discounting fallibility. From legal and governance perspectives, the outlined ideas may stimulate the e-privacy infringement regulations discourse in the pursuit of the greater goals of democratization of information, equality of communication surplus and upheld humane dignity and e-ethics in the big data era.

References

Alchian, A.A. (1953). The meaning of utility measurement. American Economic Review, 32, 26-50.

- Ali, N.S. & Benabou, R. (2017). Image versus information: Changing societal norms and optimal privacy. Cambridge, MA: National Bureau of Economic Research working paper 22203 retrieved from https://scholar.princeton.edu/sites/default/files/rbenabou/files/imagevsinfo-21_002.pdf
- Altman, I. (1975). *The environment and social behavior: Privacy, personal space, territory, and crowding*. Monterey: Brooks. Aumann, R.J. & Kruskal, J.B. (1958). The coefficients in an allocation problem. *Naval Research Logistics Quarterly*, *5*,
- 111-123.
- Aumann, R.J. & Kruskal, J.B. (1959). Assigning quantitative values to qualitative factors in the naval electronics problem. *Naval Research Logistics Quarterly*, *6*, 1-16.
- Benabou, R. & Laroque, G. (1992). Using privileged information to manipulate markets: Insiders, gurus, and credibility. *The Quarterly Journal of Economics*, 107, 3, 921-958.
- Beshears, J.L., Choi, J.J., Laibson, D.I., Madrian, B.C. & Sakong, J. (2011). Self-control and liquidity: How to design a commitment contract. RAND Working Paper Series, WR- 895-SSA.
- Besley, T. & Prat, A. (2006). Handcuffs for the grabbing hand? Media capture and government accountability. *American Economic Review*, 96, 3, 720-736.
- Bernoulli, D. (1954). Specimen theories novae de mensura sortis. Commentarii Academiae Scientiarum Imperialis Petropolitanae, 5, 175-192.
- Calvo-Armengol, A., de Marti, J. & Prat, A. (2015). Communication and influence. *Theoretical Economics*, *10*, 2, 649-690.
- Chabris, Ch., Laibson, D. & Schuldt, J. (2008). Intertemporal choice. New York: Palgrave Dictionary of Economics.
- Crémer, J., Garicano, L. & Prat, A. (2007). Language and the theory of the firm. *The Quarterly Journal of Economics*, *122*, 1, 373-407.
- Debreu, G. (1958). Stochastic choice and cardinal utility. Econometrica, 26, 440-444.
- Debreu, G. (1960). Topological methods in cardinal utility theory. In K.J. Arrow, S. Karlin and S. Suppes (Eds.), *Mathematical models in the social sciences*, pp. 16-26. Stanford, CA: Stanford University press.
- Delli Carpini, M.X. & Keeter, S. (1989). *What Americans know about politics and why it matters*. New Haven: Yale University Press.
- Donahue, J.D. & Zeckhauser, R.J. (2011). *Collaborative governance: Private sector roles for public goals in turbulent times*. Princeton, NJ: Princeton University Press.
- Edgeworth, F.Y. (1881). *Mathematical physics: An essay on the application of mathematics to the moral sciences*. London: Kegan. Retrieved from https://archive.org/details/mathematicalpsy01goog
- Etzioni, A. (2012). The privacy merchants: What is to be done? The Journal of Constitutional Law, 14, 4, 929-951.
- Evans, A.M. & Krueger, J.I. (2009). The psychology and economics of trust. *Social and Personality Psychology Compass*, *3*, 6, 1003-1017.
- Fishburn, P.C. (1968). Utility theory. Management Science, 14, 5, 335-378.
- Flaherty, D. (1989). Protecting privacy in surveillance societies: The federal republic of Germany, Sweden, France, Canada, and the United States. Chapel Hill, NC: The University of North Carolina Press.
- Fowler, R., Hodge, B., Kress, G. & Trew, T. (1979). Language and control. London: Routledge.
- Gaudeul, A. & Giannetti, C. (2017). The effect of privacy concerns and social network formation. *Journal of Economic Behavior & Organization*, 141, 233-253.
- Greenberg, A. & Collins, S. (1966). Paired comparison taste tests: Some food for thought. *Journal of Marketing Research*, *3*, 76-80.
- Horberg, E.J., Oveis, C. & Keltner, D. (2011). Emotions as moral amplifiers: An appraisal tendency approach to the influences of distinct emotions upon moral judgments. *Emotion Review*, *3*, 237-244.
- Iyengar, S. & Kinder, D.R. (1987). News that matter. London: University of Chicago Press.
- Jevons, W.St. (1871). *The theory of political economy*. London: Macmillan. Retrieved from http://www.econlib.org/library/YPDBooks/Jevons/jvnPECover.html
- Johnson, D. (2009). Ethical theory and business. Upper Saddle River, NJ: Pearson Prentice Hall.

- Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 2, 263-291.
- Koopmans, T.C. (1964). On flexibility of future preferences. In M.W. Shelly and G.L. Bryan (Eds.), *Human Judgments and Optimality*. New York: Wiley.
- Kosinski, M., Bachrach, Y., Stillwell, D.J., Kohli, P. & Graepel, T. (in press). Manifestations of user personality in website choice and behavior on online social networks. *Machine Learning Journal*.
- Kosinski, M., Stillwell, D. & Graepel, T. (2013). Private traits and attributes are predicable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, *110*, 15, 5802-5805.
- Lerner, J.S., Small, D.A. & Loewenstein, G. (2004). Heart strings and purse strings: Carryover effects of emotions on economic transactions. *Psychological Science*, *15*, 5, 337-341.
- Lessig, L. (2006). Code: Version 2.0. New York, NY: Basic.
- Marschak, J. (1950). Rational behavior, uncertain prospects, and measurable utility. Econometrica, 18, 111-141.
- Marquardt, R., Makens, J. & Larzelere, H. (1965). Measuring the utility added by branding and grading. *Journal of Marketing Research*, 2, 45-50.
- Marx, K. (1867/1995). Capital: A critique of political economy. Moscow: Progress.
- Minsky, H.P. (1977). Banking and a fragile financial environment. The Journal of Portfolio Management, 3, 4, 16-22.
- Moscovici, S. (1988). Notes towards a description of social representations. *Journal of European Social Psychology*, *18*, 3, 211-250.
- Neuman, W., Russel, MR.J. & Crigler, A.N. (1992). Common knowledge. Chicago: Chicago University Press.
- Norris, P. & Sanders, D. (2003). Message or medium? Campaign learning during the 2001 British General Election. *Political Communication*, 20, 233-262.
- Olken, B.A. (2009). Do television and radio destroy social capital? Evidence from Indonesian villages. *American Economic Journal: Applied Economics*, 1, 4, 1-33.
- Orwell, G. (1949). 1984. New York: Harcourt Brace.
- Oulasvirta, A., Suomalainen, T., Hamari, J., Lampinen, A. & Karvonen, K. (2014). Transparency of intentions decreases privacy concerns in ubiquitous surveillance. *Cyberpsychology, Behavior, and Social Networking*, *17*, 10, 633-638.
- Paluck, E.L. (2009). What's in a norm? Sources and processes of norm change. Journal of Personality and Social Psychology, 96, 594-600.
- Pareto, V. (1906/2014). Manual of political economy. Oxford, UK: Oxford University Press.
- Posner, R.A. (1981). The economic of privacy. The American Economic Review, 71, 2, 405-409.
- Posner, R.A. (2000). Law and social norms. Cambridge, MA: Harvard University Press.
- Prat. A. (2005). The wrong kind of transparency. American Economic Review, 95, 3, 862-877.
- Prat, A. (2017). Lecture notes 'Industrial Organization.' New York: Columbia University.
- Prat, A. & Strömberg, D. (2005). Commercial television and voter information. CEPR Discussion Paper 4989, April 2005.
- Prat, A. & Strömberg, D. (2013). *The political economy of mass media*. New York: Columbia University working paper. Retrievable at http://www.columbia.edu/~ap3116/papers/mediasurvey11.pdf
- Puaschunder, J.M. (2015). On the social representations of intergenerational equity. Oxford Journal of Finance and Risk Perspectives, 4, 4, 78-99.
- Puaschunder, J.M. (2016). Trust and reciprocity drive common goods allocation norms. Proceedings of the Cambridge Business & Economics Conference. Cambridge, UK: Cambridge University. Proceedings of the 2015 6th International Conference of the Association of Global Management Studies at Alfred Lerner Hall of Columbia University, New York: The Association of Global Management Studies.
- Puaschunder, J.M. (2017). Nudgital: Critique of Behavioral Political Economy. *Archives of Business Research*, *5*, 9, 54-76.
- Puaschunder, J.M. (2018). On the collective soul of booms and busts. Retrievable at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2799646
- Puaschunder, J.M. & Schwarz, G. (2012). *The future is now: How joint decision making curbs hyperbolic discounting but blurs social responsibility in the intergenerational equity public policy domain.* Harvard University Situationist Law and Mind Sciences Working Paper.
- Read, D.R. (1964). A quantitative approach to the comparative assessment of taste quality in the confectionery industry. *Biometrics*, 20, 143-155.
- Roberts, A. S. (2010). *The logic of discipline: Global capitalism and the architecture of government*. New York: Oxford University Press.
- Salanié, F. & Treich, N. (2005). Over-savings and hyperbolic discounting. *European Economic Review*, 50, 6, 1557-1570.
- Schumpeter, J. A. (1949). Economic theory and entrepreneurial history. Cambridge, MA: Harvard University Press.
- Shannon, C.E. (1948). A mathematical theory of communication. The Bell System Technical Journal, 27, 379-423.
- Silberberg, S. (2000). The structure of economics: A mathematical analysis. Boston: McGraw-Hill.

- Snyder, J.M. & Strömberg, D. (2010). Press coverage and political accountability. *Journal of Political Economy*, *118*, 2, 355-408.
- Soros, G. (2018). *Remarks delivered at the World Economic Forum in Davos*. Retrievable at https://www.georgesoros.com/2018/01/25/remarks-delivered-at-the-world-economic-forum/
- Stafford, J.E. (1966). Effects of group influences on consumer brand preferences. *Journal of Marketing Research*, *3*, 68-75.
- Stillson, P. (1954). A method for defect evaluation. Industrial Quality Control, 11, 9-12.
- Strotz, R.H. (1957). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, 23, 165-180.
- Suppes, P. (1961). Behavioristic foundations of utility. *Econometrica*, 29, 186-202.
- Suzuki, G. (1957). Procurement and allocation of naval electronic equipments. *Naval Research Logistics Quarterly*, 4, 1-7.
- Wagner W., Lorenzi-Cioldi, F., Mankova, I. & Rose, D. (1999). Theory and method of social representations. *Asian Journal of Social Psychology*, *2*, 1, 95-125.
- Warren, S. & Brandeis, L. (1890). The right to privacy. Harvard Law Review, 4, 5, 193-220.
- Witzany, G. (2012). Biocommunication of fungi. Dordrecht: Springer.
- Wouter, D. & Prat, A. (forthcoming). Attention in organizations. In Y. Bramoulle, A. Galeotti & B. Rogers, *The Oxford handbook of network economics*. Oxford: Oxford University Press.



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings



journal homepage: http://www.acrn-journals.eu/

Efficiency in the Highly Market-Segmented Chinese Banking Sector: A Meta-Frontier Non-Radial Directional Distance Function Approach

Dong Xiang¹, Ning Zhang², Andrew C. Worthington^{3,*}

¹Qilu University of Technology, China ²Jiangxi University of Finance and Economics, China ³Griffith University, Australia

Abstract

We apply a meta-frontier non-radial directional distance function approach to the efficiency analysis of 143 Chinese banks. Relative to the group frontier in their own market segment, the Big-5 Chinese banks are highly efficient. However, their low meta-technology ratios indicate that they underperform other banks. As befits their ownership structure, the joint-stock commercial banks also display high levels of efficiency, whereas the low efficiency scores and high meta-technology ratios of the city commercial banks evidence their historical role as channels for implementing local government policy in China. Foreign banks display the highest meta-technology ratios, suggesting they possess best-practice operating environments.

JEL classification: C23, D24, G21

Keywords: Meta-frontier analysis, Non-radial directional distance function, Bank efficiency, China

Introduction

Following more than three decades of sustained economic growth, China is now the world's second-largest economy after the US, with the Chinese economy now thoroughly integrated with most other global economies. However, China's financial system remains arguably unsound, and certainly less well developed than are those in the developed world. This is problematic because China will be unable to maintain stable and sustainable growth without a sound financial system. In addition, in light of the recent global financial crisis (GFC), it is widely recognized that banking unsoundness can trigger financial instability elsewhere. The key is stability, which is largely contingent on efficiency (Schaeck et al., 2006) such that banks that are more efficient are less likely to become problematic (Wheelock and Wilson 2000). Efficient banks will also perform their key financial intermediation function in a least costly way, thereby fostering economic growth and development. In particular, recent analysis suggests that an inefficient banking sector combined with poor financial infrastructure may already be restraining financial and economic growth and development in China (Berger et al., 2009).

Consequently, China's banking sector has attracted increasing attention from practitioners and academics alike both inside and outside China. As an example, using a sample of 38 banks over the period 1994–2003, Berger et al. (2009) suggested that while most foreign banks in China had relatively high efficiencies, the Big-5 state-owned banks were among the least-efficient banks because of their poor revenue performance and high levels of nonperforming loans. In another

study based on a somewhat larger sample of 88 banks from 1996 to 2006, Berger et al. (2010) found that the foreign ownership of banks in China brought with it more advanced monitoring and managerial expertise, and was often associated with better networks, partnerships, and even direct or indirect affiliation with international conglomerates.

In a similar vein, Ariff and Can (2008) concluded that joint-stock banks in China were both more cost- and profit-efficient than were their state-owned counterparts, while Kumbhakar and Wang (2007) found that private ownership in Chinese joint-equity banks improved technical efficiency. Elsewhere, Jiang et al. (2009) and Fu et al. (2015) suggested that state-owned banks in China were both relatively efficient, and that foreign banks were consistently least efficient. Finally, Fu and Heffernan (2009) argued that the relative efficiency (RE) hypothesis appeared to prevail over the structure-conduct-performance (SCP) hypothesis in the Chinese banking sector, implying that larger banks were generally more efficient.

Given these rather mixed results, an obvious way forward is further efficiency analysis of the Chinese banking sector. However, conventional efficiency analysis, including data envelopment analysis (DEA) and stochastic frontier analysis (SFA) such as used above, assumes that all observed firms share a common production technology, and so construct a single production frontier to measure efficiency. This is clearly an erroneous assumption in many of the world's banking sectors, not least in China. In response, a few studies apply so-called meta-frontier methods to the efficiency analysis of banking sectors, including some in China. For example, Fu et al. (2016) compared the efficiency of Taiwanese and Chinese banks, and Zhu et al. (2015) did the same for Chinese state-owned commercial banks (SOCBs), joint stock commercial banks (JSCB) and city commercial banks (CCBs).

We make a novel contribution to this important literature in the following ways. First, in light of the strong market segmentation prevailing in the Chinese banking sector, we separate China's banks into five groups. These comprise the five very large state-owned banks (Big-5)¹, the joint stock commercial banks (JSCBs), the city commercial banks (CCBs), the rural commercial banks (RCBs), and foreign banks (FB), a segmentation also recognized by China's monetary authorities and information agencies (e.g. CSMAR). This allows us to compare efficiency better between both banks operating in very similar environments (or segments) and across banks as a whole.

Second, asset quality in Chinese banks has been long been a concern. Given that the statecontrolled banks arguably lack power over loan quality, the sheer volume of potential problem loans in Chinese banks is an ongoing problem. Bearing this in mind, we apply a non-radial directional distance function approach, where loan quality in the form of nonperforming loans (NPLs) serves as one of the outputs. In this approach, a better performing bank is one that that can both maximize loans and/or income while minimizing NPLs.

Third, we examine the efficiency of 143 banks operating in China over the period of 2006–13, accounting for some 96% of all banking assets in China. We compare the efficiency of the Big-5 banks, JSCBs, CCBs, RCBs, and FBs. Fortunately, this period includes the GFC. Thus, we can examine efficiency trends in Chinese banks before, during, and after the most recent financial crisis. Lastly, we employ a bootstrap truncated model, which enables further detailed investigation of the firm-level determinants of Chinese bank efficiency.

Our results reveal that relative to the group frontier for their segment, the Big-5 banks and the JSCBs display high levels of technical efficiency. However, the low meta-frontier ratio for the Big-5 may indicate that they have more restricted conditions such that they are unable to make full use of their technological potential to achieve higher-order economic goals, such as cost-cutting and optimal profit-maximization.

¹ Some comparable analyses exclude the China Bank of Communications (CBOC), thereby concentrating on only the Big-4 Chinese banks. However, given its key role in the evolution of the Chinese banking sector discussed in Section Key developments in the Chinese Banking Sector, we follow Berger et al. (2009) and others and include the CBOC in our analysis.

At the other end of the spectrum, the worst performers are the CCBs in terms of efficiency scores. Largely owned by city governments or local authorities, the CCBs often serve as the major implementation channels of local government policy, particularly through fixed investment. At the same time, the CCBs attempt to dominate selected local markets, especially in small business finance, and for which they receive some market protection via local government policy. Elsewhere, the FBs do not exhibit good efficiency, evidencing that they remain immature in the Chinese market following their recent entry in 2007.

Therefore, their banking business awaits further expansion in terms of both the types of business and the operating locations in which they compete in China. However, their high metatechnology ratios reveal that the FBs have generally the best bank-operating environment, which they derive from better management, superior market discipline, and developed corporate cultures. Finally, over time, the technical efficiency of Chinese banks displays an upward trend, with only a very slight setback during the 2008 GFC.

The remainder of the paper is organized as follows. The Section 2 Key developments in the Chinese Banking Sector provides a brief overview of the Chinese banking sector with a focus on recent historical developments potentially affecting efficiency and thereby justifying a segmented market approach. The Sections Theoretical Framework and Data discuss the methodology and data. The final Sections present the empirical results and the concluding remarks.

Key developments in the Chinese Banking Sector

Prior to 1978, a single bank, the People's Bank of China (PBOC), dominated the Chinese banking sector. Following a series of fundamental economic and financial reforms after 1978, we can categorize the evolution of the Chinese banking system into four main periods.

1978–1984

During this period, the single-bank system dominated by the PBOC transformed into an effectively two-tier banking system (Jiang et al., 2009). Commercial banking operations were detached from the PBOC, resulting in it emerging as China's financial sector regulatory and supervisory authority. Eventually, the PBOC's commercial banking operations were distributed among four specialized state-owned banks: the Agricultural Bank of China (ABC), the China Construction Bank (CCOB), the Industrial and Commercial Bank of China (ICBC), the Bank of China (BOC) (together the Big-4).

1984–2001

The early part of this period saw the establishment of three policy-related banks to assume the policy-lending activities of the Big-4 banks, funded mainly through the issue of state bonds and loans by the PBOC. In 1986, the first of 12 joint-equity state-control dominated banks, the China Bank of Communications (CBOC), helped form the Big-5 banks. Urban and rural credit cooperatives started to merge and form city and rural commercial banks, the central bank lifted the restriction on foreign banks operating representative offices in China, and some regulatory allowance was made for foreign investors to hold minority stakes in domestic banks.

2001–2006.

Following China's entry into the World Trade Organization (WTO) in 2001, the Chinese government accelerated its banking reforms. In 2003, the China Banking Regulatory Commission

(CBRC) took on the role of banking supervisory authority from the PBOC, with the PBOC repurposed for overall financial stability and systemic liquidity. In terms of capital adequacy, a minimum capital adequacy ratio, as defined in Basel I, was introduced, along with a risk-based framework, with guidelines for credit, market, and operational risk. A crucial milestone in the financial liberalization process was the conclusion of the negotiations for China's accession to the WTO in late 2001. The commitments agreed to under WTO implied the full opening up of the Chinese banking system to foreign affiliates by the end of 2006.

2006–2013

The segmented commercial banking sector formed, now comprising the Big-5 banks, JSCBs, CCBs, RCBs, and FBs, along with urban and rural credit cooperatives, finance, trust and investment, and financial leasing companies and postal savings institutions, etc. Following the joint-stock conversions of the largest four state-owned banks, all of the Big-5 now operated as commercial banks.

Despite these reforms, a number of challenges remain for the Chinese banking sector. For example, foreign ownership remains relatively small, as does foreign involvement in governance. The banking system remains dominated by largely state-owned enterprises, and the availability of funds, not borrower quality, continues to drive most bank lending. Chinese banks also continue to be largely constrained by government intervention at different levels and are subject to substantial political influence (Bailey et al., 2011). These conditions clearly support the need for an analysis that considers strong market segmentation, which we undertake in the following section.

Theoretical Framework

Meta- and group-frontiers

Assume that there are j=1,...,N banks and that each bank uses input vector $x \in \mathfrak{R}^M_+$ to jointly produce outputs vector $y \in \mathfrak{R}^S_+$ and undesirable outputs $b \in \mathfrak{R}^J_+$. Here the multi-output production technology with undesirable output can be expressed as

 $T = \{(x, y, b) : x \text{ can produce } (y, b)\},\$

(1)

where T is often assumed to satisfy the standard axioms of production theory (Färe and Grosskopf, 2005). For instance, inactivity is always possible, and finite amounts of inputs can produce only finite amounts of outputs. In addition, we often assume inputs and desirable outputs are strongly or freely disposable. For a reasonable model of joint-production technologies, as described in Färe et al. (1989), we need to impose weak disposability and null-jointness assumptions on T. We express the two assumptions as follows:

- (i) If $(x, y, b) \in T$ and $0 \le \theta \le 1$, then $(x, \theta y, \theta b) \in T$;
- (ii) If $(x, y, b) \in T$ and b=0, then y=0.

The weak-disposability assumption implies that reducing undesirable outputs, such as NPLs, in a bank are costly in terms of the proportional reduction in total loans. The null-jointness assumption states that NPLs are not avoidable in bank loans and that the only feasible way to remove NPLs is to cease loans.

After specifying the bank joint-production technology T, it is possible to use parametric techniques or nonparametric data envelopment analysis (DEA) to construct the bank production

technology. In this paper, we employ a nonparametric DEA piecewise linear production frontier to construct the bank joint production technology. Then we can express T for N generators exhibiting constant returns to scale (CRS) as

$$T = \{ (x, y, b) : \sum_{n=1}^{N} z_n x_{mn} \le x_m, m = 1, \dots, M, \\ \sum_{n=1}^{N} z_n y_{sn} \ge y_s, s = 1, \dots, S, \\ \sum_{n=1}^{N} z_n b_{jn} = b_j, j = 1, \dots, J, \\ z_n \ge 0, n = 1, \Lambda, N. \}$$
(2)

Generalized directional distance functions

The directional distance function, developed by Chambers et al. (1996), is a relatively new method for performance measurement. According to Fukuyama and Weber (2009), the conventional directional distance function as a radial measure of efficiency (inefficiency) may overestimate efficiency when there are non-zero slacks. The literature mostly supports non-radial measures of efficiency for measuring bank efficiency because of their several advantages (e.g. Fukuyama and Weber, 2009, Barros et al., 2012). For this reason, Barros et al. (2012) and Zhou et al. (2012) provide a formal definition of the non-radial directional distance function, which is a generalized form of basic directional distance function. Following Zhou et al. (2012), we define the generalized directional distance function as follows:

$$\overset{\mathcal{V}}{D}(x, y, b; g) = \sup\{\mathbf{w}^{T}\boldsymbol{\beta}: ((x, y, b) + g \cdot diag(\boldsymbol{\beta})) \in T\}$$
(3)

where $\mathbf{w} = (w_m^x, w_s^y, w_j^b)^T$ denotes a normalized weight vector relevant to numbers of inputs and output, $g = (-g_x, g_y, -g_b)$ is an explicit directional vector, and $\beta = (\beta_m^x, \beta_s^y, \beta_j^b)^T \ge 0$ denotes the vector of scaling factors. We compute the value of D(x, y, b; g) by solving the following DEAtype model:

We may set the directional vector g in different ways based on specific bank goals. If $\stackrel{P}{D}(x, y, b; g) = 0$, then the bank to be evaluated is located on the frontier of best practices in the direction of g.

The overall technical efficiency (TE) for the bank is the average efficiency of each factor. Suppose that β_x^* , β_y^* , and β_b^* represent the optimal solutions to Equation (4), then the TE can be formulated as:

$$TE = 1 - \overset{\mathbf{r}}{D}(x, y, b; g) = 1 - \left[\frac{1}{M + S + J}\left(\sum_{m=1}^{M} \beta_{xm}^{*} + \sum_{s=1}^{S} \beta_{ys}^{*} + \sum_{j=1}^{J} \beta_{bj}^{*}\right)\right].$$
 (5)

Meta-frontier generalized directional distance function

We now combine the concept of meta-frontier production technologies in O'Donnell et al. (2008) with that of the generalized distance function to develop our meta-frontier generalized directional distance function and investigate the group heterogeneity of bank joint-production in China. For this, we define the group- and meta-frontier technologies.

Suppose that *H* groups show some technological heterogeneity, such that human resources, management knowhow, and other specific constraints may prevent banks in some groups from accessing the management technologies in other groups. Following Battese et al. (2004) and O'Donnell et al. (2008), we define the group-frontier technology of group *h* as $T_h = \{(x, y, b) : x \text{ can produce } (y, b)\}, h=1,2,...H$. Assume that T_h is specified as the nonparametric joint production technology described in equation (3). Then we define the generalized directional distance function for group *h* as

$$\overset{\mathcal{V}}{D}{}^{g}(x, y, b; g) = \sup\{\mathbf{w}^{T}\boldsymbol{\beta} : ((x, y, b) + g \cdot diag(\boldsymbol{\beta})) \in T_{h}\}, h = 1, 2...H$$
(6)

Unlike a group-frontier technology, we can construct a meta-frontier technology from all observations for all groups by enveloping all group-frontier technologies. Therefore, we can consider the following definition: $T_m = \{T_1 \cup T_2 \cup ... \cup T_H\}$. Here we formulate a generalized function based on a meta-frontier technology as follows:

$$\overset{\mathcal{V}}{D}^{m}(K,L,F,E,C;g) = \sup\{\mathbf{w}^{T}\boldsymbol{\beta}: ((x,y,b) + g \cdot diag(\boldsymbol{\beta})) \in T_{m}\}$$
(7)

In equation (7), we calculate $D^{g}(\cdot)$ by solving model (4) with use the of input and output data from group *h*. Suppose N_h observations for group *h*. We compute $D^{m}(\cdot)$ by solving the following:

D. Xiang, N. Zhang, A.C. Worthington / Finance, Risk and Accounting Perspectives Proceedings (2019) 55-72

where z_n^h represents the intensity variables for constructing the meta-frontier technologies. Equation (8) indicates that we need data on all the banks to construct the meta-frontier. The meta-frontier joint-production technologies cover all group-frontier technologies. After solving equations (6) and (7), we obtain the optimal solutions for all variables under group- and meta-frontier technologies, respectively.

O'Donnell et al. (2008) point out that we can decompose the technical efficiency from metafrontier technologies (*MTE*) into within-group technical efficiency (*GTE*) and the metatechnology ratio (*MTR*). *GTE* measures the relative efficiency of observations under specific group-frontier technologies, whereas *MTR*, also referred to as the technology gap ratio, measures how close a group-frontier technology is to a meta-frontier technology. O'Donnell et al. (2008) show that *MTE* does not exceed *GTE*, such that *MTR* is no greater than unity. The higher value the *MTR*, the closer the group-frontier technology is to the meta-frontier technology. An *MTR* value equal to unity implies no gap between the two technologies, that is, a complete overlap between the group- and meta-technologies.



Fig. 1: A theoretical meta-frontier

Following O'Donnell et al. (2008), we decompose MTE into within-group technical efficiency (GEC) and a meta-technology ratio (MTR) index, respectively. Here we consider a simple example in Figure 1 to illustrate the decomposition of MTE. Suppose that the input is x and the output is y and that the bank under analysis belong to two heterogeneous groups. We compute two group frontiers, XX' and YY'. Based on a specific bank a operating at XX', its performances under the group frontier GTE and the MTE according to M are formulated as:

$$GTE = \frac{y_2 - y_1}{x_2 - x_1}, \quad MTE = \frac{y_3 - y_1}{x_3 - x_1}$$
(9)

Therefore, *MTRT* captures the gap between the group frontier and the meta-frontier as: if $y_2=y_3$ and $x_2=x_3$ which means no technology gap, then the MTR=1.

$$MTR = \frac{MTE}{GTE} = \left[\frac{y_3 - y_1}{y_2 - y_1}\right] * \left[\frac{x_2 - x_1}{x_3 - x_1}\right]$$
(10)

Drawing on the intermediation approach to bank efficiency estimation, we develop three different input–output specifications to examine the technical efficiency of Chinese banks, as shown in Table 1. In Model A, the inputs are personnel expenses (PEX), total deposits (DEP) and physical capital (FIX), and the outputs are total loans (LOA) (including loans, advances, and other receivables), non-interest income (NON) and the non-performing loan ratio (NPL) (the ratio of non-performing loans to total loans). We use Model B to investigate efficiency after incorporating input prices, as proxied by personnel expenses divided by total assets (PLB), interest expenses divided by total deposits (PDP), and operating expenses less labor expenses divided by fixed assets (PPC). The outputs are normalized total loans (LOA), non-interest income (NON) and the non-performing loan ratio (NPL). Finally, we use Model C to examine the efficiency of the banks when we include additional outputs in the form of normalized pre-tax profits (PRO) and the non-performing loan ratio (NPL), but the inputs are the same as Model B.

Table	1	Model	specification
-------	---	-------	---------------

	Model A	Model B	Model C
Inputs	Personnel expenses (PEX)	Personnel expenses divided	Personnel expenses divided
		by total assets (PLB)	by total assets (PLB)
	Total deposits (DEP)	Interest expenses divided by	Interest expenses divided by
		total deposits (PDP)	total deposits (PDP)
	Physical capital (FIX)	Operating expenses less	Operating expenses less
		labor expenses divided by	labor expenses divided by
		fixed assets (PPC)	fixed assets (PPC)
Outputs	Total loans (LOA)	Total loans normalized by	Pre-tax profits (PRE)
		equity (LOA)	normalized by equity
	Non-interest income	Non-interest income	Non-performing loans to
	(NON)	normalized by equity (NON)	total loans ratio (NPL)
	Non-performing loans to	Non-performing loans to	
	total loans ratio (NPL)	total loans ratio (NPL)	

Bootstrap truncated regression

In early research in this area, Tobit regression was typical in the second-stage regressions used to examine the determinants of DEA efficiency analysis. However, more recently, bootstrap truncated regression has proposed. One disadvantage with Tobit and OLS regression is that the efficiency scores obtained through empirical DEA computation and not observed directly. Thus, OLS models (including the Tobit model), which assume independently distributed error terms,

are invalid. In addition, empirical estimates of the efficiency frontier draw on only a select sample of banks, which eliminates some efficiency production possibilities not observed in the sample. Further, the two-stage regression model depends on explanatory variables that not directly observed but estimated in the first stage. This implies that the error term correlates with the second-stage explanatory variables.

To overcome these disadvantages, the truncated bootstrap regression approach of Simar and Wilson (2007) enables consistent inferences within models, explaining efficiency scores while simultaneously producing standard errors and their confidence intervals. The truncated bootstrap model is as follows:

$$\phi^{\wedge} = z_i \beta + \varepsilon_i \tag{11}$$

Our aim is to recognize the relationship between MTE scores ϕ^{\wedge} and explanatory variables z_i , which refer to the vector of parameters with some statistical noise ε_i in Equation (11). Previous studies have suggested some estimation procedures based on the OLS or Tobit model. However, because of the biased estimation as mentioned above, our approach follows the following steps:

1. Calculate the MTE score ϕ^{\wedge} for each bank by using the meta-frontier distance approach model according to (11).

2. Conduct the truncated regression of ϕ^{\wedge} and z_i by using the maximum likelihood function to estimate β^{\wedge} and $\sigma_{\varepsilon}^{\wedge}$ of β and σ_{ε} , respectively.

3. Repeat the following steps *B* times (*B* = 2000) to obtain a set of bootstrap estimates $\{\phi_{i,b}^*, b=1,...B\}$.

a. Draw ε_i from the N (0, $\hat{\sigma}^2$) distribution with left truncation at (1- $\beta^2 z_i$);

b. Calculate
$$\phi_i^* = z_i \beta^{\wedge} + \varepsilon_i$$
.

c. Generate a pseudo data set $(x_i^* y_i^*)$, where $x_i^* = x_i$ and $y_i^* = y_i \phi_i^{\wedge} / \phi_i^*$;

d. Substitute a new meta-frontier estimate θ_i^* with the set of pseudo data $(x_i^* y_i^*)$

4. For each DMU, calculate the bias-corrected estimate $\phi^{\uparrow}_{i} = \phi^{\uparrow}_{i} - bia^{\uparrow}s_{i}$, where $bia^{\uparrow}s_{i}$ is the bootstrap estimator of bias defined as $bia^{\uparrow}s_{i} = 1/B\sum_{i=1}^{B}\phi^{\uparrow*}_{i,b} - \phi^{\uparrow}_{i}$

5. Conduct the truncated regression of $\phi^{\uparrow}{}_{i}$ on z_{i} to obtain the estimates $(\beta^{\uparrow}, \sigma^{\uparrow})$ of (β, σ) .

6. Repeat the following three steps *B2 (B2=2000)* times to obtain a set of bootstrap estimates $\{\beta_{b}^{\wedge^{*}}, \sigma^{\wedge^{*}}, b=1,...B2\}$.

- a. For i=1,...,n, ε_i is drawn from $N(0, \sigma^{\wedge})$ with left truncation at $(1-\beta^{\wedge} z_i)$;
- b. For i=1,...,n, execute $\phi^{**} = \beta^{\wedge} z_i + \varepsilon_i$;
- c. Again, conduct the truncated regression of ϕ^{**}_{i} on z_i to obtain the estimates of $(\beta^{\wedge^*}, \sigma^{\wedge^*})$.

See Simar and Wilson (2007) for the estimation algorithm. Some researchers have conducted empirical analyses to verify that in the two-stage DEA model, truncated bootstrap regression can provide robust results than the Tobit model and result in smaller standard errors and less variance.

Correlates with bank efficiency

The firm-specific variables in Table 2 include intangible assets (INT), debt to equity ratio (DER), the ratio of cash and dues from banks to assets (CASH), the ratio of loans to deposit (LDR), the net interest margins (NETI), the ratio of investment in marketable securities to loans (INV) and the capital adequacy ratio (CAR). Intangible assets (INT) (in logarithms) represent the accounting value of goodwill, patents, copyrights, trademarks, formulae, organizational costs, customer lists, franchises and licenses, purchased servicing rights, and capitalized value of software development, advertising costs and servicing rights. We interpret an intangible asset as an indicator of future growth opportunities (Ozkan, 2001). Thus, we expect a bank with substantial intangible investments to be more efficient. Though a number of studies examine the value relevance of intangible assets (Oliveira et al., 2010; Dahmash et al., 2009), there has been previously no attention to the role of intangible assets as a determinant of bank efficiency.

A higher debt to equity ratio (DER) may reduce agency costs through pressure to generate cash flow to pay interest expenses and through the threat of liquidation, which can cause personal losses to managers' salaries, reputations, etc. In this respect, DER may positively affect bank efficiency. Conversely, further increases in DER may incur significant agency costs, because a relatively high DER makes bankruptcy or financial distress more likely because of risk shifting or a reduced effort to control risk (Berger and Patti, 2006). In this case, DER has a negative effect on bank efficiency. The ratio of cash and dues from banks to assets (CASH) represents liquidity risk. We expect CASH to affect bank efficiency negatively due to opportunity cost of holding liquid assets (Kwan, 2003). The ratio of loans to deposit (LDR) represents insolvency risk. LDR reflects a bank's ability to covert deposits to loans (Dietsch and Lozano-Vivas, 2000), and thus exerts a positive influence on efficiency.

We justify net interest margins (NETI) as follows. Over the past few decades, interest margins on loans have declined significantly and banks have become more involved in non-loan services, such as fund managements, financial planning, and insurance services to improve their bottom line. In addition, the ratio of investment in marketable securities to loans (INV) is positively associated with efficiency because investment securities are more efficient than loans in generating revenues. Finally, we specify the capital adequacy ratio (CAR) as a measure of regulatory compliance.

Variable	Definition	Mean	Min.	Max.
AST	Total assets (RMB bil.)	711.218	2.464	18917.752
EQC	Total equity (RMB bil.)	43.719	0.195	1278.463
TC	Total costs (RMB bil.)	22.899	0.043	577.786
Р	Profits (RMB mil.)	10345.424	-57.386	338537.000
PEX	PERSONNEL EXPENSES (RMB mil.)	1598.142	0.009	47697.000
DEP	Total deposits (RMB bil.)	543.064	0.133	14620.825
FIX	Fixed assets (RMB bil.)	5.574	0.001	158.968
PLB	Personnel price (%)	2.779	0.001	22.232
PPC	Physical capital price (%)	300.816	0.809	3562.606
PDP	Deposit price(%)	2.607	0.188	28.800
LOA	Loans (RMB bil.)	360.886	0.015	9922.374
NII	Non-interest income (RMB mil.)	4078.043	-93.548	137628.000
CIT	Equals one if the bank is city commercial	0.611	0	1
	bank; zero otherwise			
RUR	Equals one if the bank is rural commercial	0.152	0	1
	bank; zero otherwise			
FOR	Equals one if the bank is foreign bank; zero	0.079	0	1
	otherwise			
LST	Equals one if the bank is listed; zero	0.164	0	1
	otherwise			
NAT	Equals one if the bank is national; zero	0.157	0	1
	otherwise			
INT	Intangible assets (logarithm)	15.691	0.001	24.004
NPL	Non-performing loans/total loans	0.014	0.000	0.226
LDR	Loans/deposits	0.652	0.082	3.946
DER	Debt/equity	14.725	1.368	55.738
CASH	Cash and dues from banks/total assets	0.165	0.000	0.424
NETI	Net margin (net interest income/revenue) (%)	0.526	0.198	0.954
INV	Investments/total loans	0.016	0.000	0.708
CAR	Capital adequacy	0.143	0.012	0.667

Table 2 Variable definitions and statistics

Data

We investigate the efficiency of Chinese banks over the period 2006 to 2013, which as discussed earlier, is the fourth period in the long history of Chinese banking reforms. During this time, all the major banks had completed transition in their ownership structure and most city or rural commercial banks had transformed from other financial institutions into banks. In this sense, we minimize the likelihood of bias in our estimates in terms of the violation of homogeneity. The sample is an unbalanced panel comprising 143 Chinese banks over 8 years, with 682 bank-year observations for the Big-5 banks, 11 JSCBs, 85 CCBs and 24 RCBs. The data is from the China Stock Market Accounting Research database (CSMAR) and Bankscope. Table 2 provides summary statistics for the variables.

Results and Discussion

Efficiency estimates

We base the estimates of the meta-frontier non-radial directional distance function (MNDDF) models on the CRS free-oriented slack-based DEA models given by Eq. 11. In the analysis, we use GTE, GCE and GPE to refer to the DEA estimates of Models A, B and C relative to the group frontiers; MTE, MCE, and MPE to the DEA estimates of Models A, B and C relative to

the meta-frontiers, respectively; and TER, CER, and PER to the DEA estimates of the meta-technology ratio.

	Null hypothesis (Ho)	Chi-Square	p-value	
Technical eff	iciency with respect to the meta-frontier			
MTE	MTE (group1=group5)	276.39	0.000	
MCE	MCE (group1=group5)	14.17	0.007	
MPE	MPE (group1=group5)	16.04	0.003	
Technical efficiency with respect to the group frontiers				
GTE	MTE (group1=group5)	387.19	0.000	
GCE	MCE (group1=group5)	275.69	0.000	
GPE	MPE (group1=group5)	452.28	0.000	

Table 3 Kruskal–Wallis tests of efficiency differences

To investigate whether there are statistically significant differences between the group- and meta-frontiers, we apply a Kruskal–Wallis test to the differences in the efficiency scores between the five bank segments. As shown in Table 3, all the null hypotheses are rejected at the 1% level, evidencing that the banks in the different segments are operating under the different boundaries of restricted technology sets, where the restrictions can derive from the differences in business environments such as regulations, restrictions, policies and/or business markets, etc. In other words, the segments lead to the different technological frontiers to which the performance of the sample banks is to measured relative.

Table 4 Efficiencies by group

Segmen	Model	Mean	Std.	Min.	Max.	
Technical efficiency with group frontiers						
Big-5	А	0.916	0.097	0.713	1.000	
	В	0.918	0.111	0.628	1.000	
	С	0.726	0.171	0.417	1.000	
JSCB	А	0.938	0.073	0.518	1.000	
	В	0.662	0.217	0.288	1.000	
	С	0.560	0.201	0.100	1.000	
ССВ	А	0.634	0.216	0.189	1.000	
	В	0.494	0.233	0.164	1.000	
	С	0.328	0.221	0.093	1.000	
RCB	А	0.849	0.128	0.536	1.000	
	В	0.753	0.216	0.241	1.000	
	С	0.459	0.283	0.021	1.000	
FB	А	0.821	0.194	0.378	1.000	
	В	0.697	0.271	0.122	1.000	
	С	0.549	0.349	0.150	1.000	
Meta-technology ratio						
Big-5	А	0.948	0.046	0.783	1.000	
	В	0.480	0.120	0.345	1.000	
	С	0.372	0.084	0.195	0.550	
JSCB	А	0.889	0.133	0.352	1.000	
	В	0.824	0.294	0.386	1.000	
	С	0.446	0.134	0.288	1.000	
ССВ	А	0.700	0.128	0.460	1.000	
	В	0.924	0.380	0.271	1.000	
	С	0.955	0.077	0.327	1.000	
RCB	А	0.434	0.137	0.201	0.760	

Segmen	Model	Mean	Std.	Min.	Max.
	В	0.553	0.214	0.143	1.000
	С	0.564	0.118	0.319	1.000
FB	А	0.981	0.031	0.840	1.000
	В	0.753	0.424	0.166	1.000
	С	0.722	0.249	0.251	1.000
Technical	efficiency	with res	pect to the	meta-from	ntier
Big-5	А	0.868	0.106	0.700	1.000
	В	0.441	0.128	0.313	1.000
	С	0.264	0.074	0.156	0.550
JSCB	А	0.835	0.145	0.330	1.000
	В	0.535	0.215	0.183	1.000
	С	0.259	0.177	0.035	1.000
CCB	А	0.450	0.200	0.130	1.000
	В	0.456	0.235	0.054	1.000
	С	0.313	0.210	0.038	1.000
RCB	А	0.376	0.151	0.140	0.760
	В	0.412	0.189	0.104	1.000
	С	0.254	0.169	0.014	1.000
FB	А	0.805	0.194	0.370	1.000
	В	0.498	0.308	0.064	1.000
	С	0.411	0.319	0.041	1.000

Table 4 provides descriptive statistics for the estimates for all the selected banks across the five segments. For example, in the first segment, we calculate the technical efficiency scores (GTE) of the Big-5 banks with respect to the group frontier to vary between 0.713 and 1.000 over the 8-year period, with an average of 0.916 and a standard deviation of 0.097. The efficiency scores (GCE) for the same segment range between 0.628 and 1.000, with an average GCE of 0.918, while the profit efficiency scores lie between 0.417 and 1.000 with an average GPE of 0.726. This means that given their business environment, the Big-5 banks are approximately 91.6% efficient in minimizing inputs and bad outputs i.e. non-performing loans, and maximizing good outputs i.e. loans and non-interest income. When we consider input prices or/and profits, the efficiency levels are 91.8% and 72.6%, respectively. In other words, the existing business environment allows for an 8.4%, 8.2%, and 27.4% improvement management in minimizing inputs and bad outputs and maximizing good outputs and maximizing good outputs including profits.

As shown in the upper panel in Table 4, relative to the group frontier in each segment, the Big-5 banks exhibit the best performance in GCE and GPE, whereas, the JSCBs have the best performance in GTE with average GTE scores of 93.8%. The good performance of the Big-5 in terms of GCE and GPE may well result from the well-organized business operating systems in these very large banks, while the better performance of JSCBs in TE may reflect the ownership advantages of the joint-stock firms in terms of governance, which is consistent with Jiang et al. (2009).

The worst performers in terms of efficiency are the CCBs, with average scores of 63.4%, 49.4%, and 32.8% for GTE, GCE, and GPE, respectively. On one hand, Chinese local governments normally use CCBs as the major channels for implementation of their policies, particularly those related to fixed investment. On the other hand, CCBs attempt to dominate local financial markets, particularly retail finance, deriving protection from competition from local governments. The low level of efficiency relative to the group frontiers suggests that the CCBs have greater management potential to improve under the local banking environment.

With respect to the FBs, relative to their boundaries in the restricted technology sets, they do not exhibit good performance in their management efficiency. We interpret this in two ways.
First, FBs largely remain at the startup/emerging growth stage of banking operations in China. Therefore, their banking business awaits further expansion in terms of both the types of business and operating locations in China. Second, there is a great potential for FBs to expand their banking business in China given the boundaries of the technology sets, as indicated by the meta-technology ratio.

As discussed, the meta-technological ratio measures the gap between the group (here, segment) frontiers and the meta-frontier. The second panel in Table 4 reveals that FBs with the highest levels of technology at 98.1% display the smallest gap between the group frontier and the meta-frontier. That is, the FBs have the best environment for banking business, which they could derive from their better management, greater market discipline, and/or established corporate cultures (Berger et al., 2009). There is no doubt that the opening-up of banking markets starting from the end of 2006 under WTO entry agreements greatly improved the performance potential of FBs in China's banking markets. However, as in the first panel, FBs still have room to realize this potential in China. In the meantime, CCBs have the highest levels of meta-technology ratio in Models B and C, which means that depending on their sizes, the CCBs can realize high levels of saving costs and/or maximizing profits when operating locally. As discussed, they also benefit from the protection of local governments in local financial markets.

Interestingly, the Big-5 have the lowest meta-frontier ratios, which indicates that the advanced technology they possess is not readily able to be fully converted into economic advantages like cost-saving and profit-maximizing. In other words, the Big-5 banks appear to have more restricted conditions. These restricted conditions may derive from the regulations, policies or the functions they carry out. As discussed in Section Key developments in the Chinese Banking Sector, the Big-5 has long undertaken the functions of implementing government fiscal policy. Therefore, it is unsurprising that the focus of the Big-5 departs from the conventional business goals of profit maximization and cost minimization, as reflected in their low meta-technology ratios (MR). The low MR levels of the Big-5 also provide good evidence that the Chinese banking sector is far from fully competitive.

In the second panel in Table 4, all groups and models have the maximum value of one, except the Big-5 with a maximum value of 0.55 for Model C (the RCBs have a maximum value of 0.76 for Model A. This means that, over the sample period of 2006–13, there must have been at least one year when the combination of inputs and outputs placed one bank (i.e. a Big-5 bank or a rural commercial bank) on their respective group frontier, but well below the meta-frontier.

Finally, we provide the efficiency scores relative to the meta-frontier, which is the product of the efficiency scores relative to the group frontiers and the meta-technology ratios, in the third panel of Table 4. When business environment is ignored, that is, all banks in the five segments are measured against a common technological frontier i.e. the meta-frontier, all efficiency scores are lower than are those relative to the group frontiers. For example, the average efficiency score of the Big-5 is 86.8%, which is much lower than the 91.6% in the first panel. The difference arises from the meta-technology ratios in the second panel. In a similar vein, the RCBs exhibits low levels of efficiency relative to the meta-frontier, which is largely attributable to the large technological gap between the group frontier and meta-frontier. We interpret this as being the result of the very tough restrictions imposed on the banking business of the RCBs.

Firm-level determinants of efficiency

In the second stage, following suggestions by Simar and Wilson (2007), we employ a bootstrap truncated regression model to analyze the firm-level factors behind the various measures of bank performance. As we pooled the bank data to find the firm-level determinants of bank efficiency, we specify the efficiency scores relative to meta-frontier as the dependent variables. As a simple

test of the potential for multicollinearity, we use a correlation matrix (results not shown). As the maximum correlation coefficient is only about 0.5, we believe multicollinearity is not a significant concern.

Variables	Coefficient	Boot. SD	z-statistic	p-value
Dependent variab	ole MTE			
CAR	0.928***	0.138	6.729	< 0.001
CASH	-0.751***	0.228	-3.296	0.001
DER	0.013***	0.002	7.192	< 0.001
INT	0.005***	0.002	2.722	0.007
LDR	0.914***	0.077	11.889	< 0.001
NETI	-0.665^{***}	0.098	-6.796	< 0.001
Obs.	1,690			
Log-likelihood	160.705***			
Dependent variab	ole MCE			
CAR	0.163*	0.084	1.940	0.052
CASH	0.586***	0.157	3.729	< 0.001
DER	0.017***	0.001	12.998	< 0.001
INT	0.001	0.001	-0.250	0.803
LDR	0.043	0.029	1.491	0.136
NETI	0.066	0.063	1.055	0.291
Obs.	1,690			
Log-likelihood	243.697***			
Dependent variab	ole MPE			
CAR	0.247^{*}	0.127	1.947	0.052
CASH	0.464**	0.230	2.024	0.043
DER	0.001	0.002	0.381	0.703
INT	-0.001	0.002	-0.610	0.542
LDR	-0.058	0.050	-1.156	0.248
NETI	0.269***	0.097	2.770	0.006
Obs.	1,690			
Log-likelihood	276.738***			

Table 5 Firm-level efficiency determinants

Table 5 provides the maximum likelihood estimates. As shown, an increase in the capital adequacy ratio (CAR) improves all three measures of technical efficiency. This finding is consistent with the notion that banks will be less likely to incur loans losses and are better placed in attracting deposits in a banking sector with a high capital adequacy requirement (Kasman and Yildirim, 2006; Grigorian and Manole, 2006). Intangible assets (INT) have a significant and positive effect on technical efficiency, indicating that it may actually be a proxy for advanced technology or management techniques, which can improve the technical efficiency of the banks. Financial leverage (DER) is significantly and positively associated with the first and second (including input prices) measures of efficiency, but has no significant effect on the third measure (including profits) of efficiency. The positive effect on efficiency of DER is consistent with the benefit of debt tax shields. This is also in line with the agency cost hypothesis, which implicitly indicates a positive relationship between DER and efficiency (Margaritis and Psillaki, 2007). In a similar vein, the loan to deposit ratio (LDR) also shows a significant and positive effects on efficiency. As expected, the ability to convert deposits into loans is in line with the improved efficiency of banks. This result is consistent with a number of studies (e.g., Barros, et al., 2007; Valverde et al., 2007).

As an indicator of liquidity risk, the ratio of cash and dues from banks to assets (CASH) exhibits a negative effect on TE but a positive effect on CE and PE. The negative effects of CASH on TE are due to the opportunity cost of holding liquid assets (Kwan, 2003). The positive effects of CASH on CE and PE are then perhaps because liquid assets makes bankruptcy or financial distress less likely due to liquidity risk and thus improves bank revenues. Interest margins (NETI), as expected, have positive effects on PE, but negative effects of NETI provide some evidence that increasing competition in Chinese banking sector can help improve bank efficiency. Thus, banks may still depend on lending services to achieve significant improvements in cost and profit efficiencies despite the fact loans are generally more costly than securities.

Concluding Remarks

The purpose of this analysis was to examine Chinese bank performance, as measured by technical, cost and profit efficiency, across a wide variety of banks and a relatively long sample period. First, taking account of the heterogeneity arising from the high degree of market segmentation in Chinese banking sector, we applied the so-called meta-frontier analysis approach to the analysis of the performance of Chinese banks. Given that the frontier analysis techniques draw on the distance of the performance of assumed relatively homogeneous banks to better performing banks, the meta-frontier analysis approach makes efficiency comparisons across subgroups of banks operating in different market segments. Under this approach, a common meta-frontier and group frontiers are the boundary of an unrestricted technology set and the boundaries of restricted technology sets, respectively, and the so-called meta-technology ratio measures the distance between the group frontier and the meta-frontier.

Second, given that state-dominated banks arguably lack control on loan qualities, we applied a non-radial directional distance function approach, where loan quality (as measured by nonperforming loans) serve as one output. In this approach, a better performer is a bank that can maximize loans and/or incomes, but minimize NPLs. Third, we specified three models to examine the performance of Chinese banks over the period post-2006 which include both GFC and post-GFC periods. Finally, we employed a bootstrap truncated model to investigate the effects of risk management on bank efficiency, including liquidity, credit, and insolvency risk, etc.

Our results revealed significant differences in the group frontiers between the different segments. We found that relative to the group frontier in each segment, the Big-5 banks exhibited the best performance in GCE and GPE, whereas the joint stock commercial banks displayed the best performance in GTE. We argued the good performance of the Big-5 in terms of GCE and GPE might have benefited from the well-organized business operating systems of these very large banks, whereas the better performance of the JSCBs in TE may reflect the ownership advantages of joint stock firms in terms of governance. However, the low levels of the meta-frontier ratio for the Big-5 may indicate that they have more restricted conditions where they are unable to make full use of their technological potentials to achieve higher economic goals such as cost cutting and profit maximization. These restricted conditions may also derive from the additional regulations, policies, or functions they bear. The low MR levels of the Big-5 evidence that the Chinese banking sector is far from fully competitive.

At the other end of the spectrum, the worst performers are the city commercial banks largely owned by city governments or their authorities in terms of efficiency scores. This can be a benefit for the banks in that the local governments use the CCBs as a conduit for the implementation of local public policy. The CCBs also benefit from the protection given by the local governments from non-local banks. Therefore, the CCBs have the highest meta-technology ratios in terms of cost and profit-augmented technical efficiency. The foreign banks (FBs) did not display good performance in their management efficiency, which indicates that FBs remain at a startup/early growth stage of banking operation in China. Therefore, their banking business model awaits further expansion in terms of both the types of business and the operating locations in China. However, the high meta-technology ratios for the FBs reveals that they have the best banking business environment, we argue derives from their better management, greater market discipline, and/or developed corporate cultures. Finally, the rural commercial banks (RCBs) exhibit low levels of efficiency over the period, which is largely attributable to the large technological gaps between their group frontier and meta-frontier. We interpret this as resulting from the very tough government regulations imposed on their banking business.

Over the sample period of 2006–13, the technical efficiency of the Chinese banks exhibit a stable and increasing trend, except for a very slight and expected decline during the 2008 GFC. By comparison, bank costs and profits fluctuate widely, particularly after the GFC. With respect to the firm-level determinants of bank efficiency, the estimates of a bootstrap truncated regression model reveals that the capital adequacy ratio (CAR), financial leverage (DER), intangible assets (INT), and the loan to deposit ratio (LDR) have a significant and positive effect on technical efficiency. In contrast, the ratio of cash and dues from banks to assets (CASH) and interest margins (NETI) exert negative effects on technical efficiency. In addition, CAR, CASH, and DER are positively associated with the cost-augmented technical efficiency, while CAR, CASH, and NETI are positively associated with profit-augmented technical efficiency.

References

- Aigner, D., Lovell, C.A.K., Schmidt, P., 1977. Formulation and estimation of stochastic frontier production models. Journal of Econometrics 6, 21–37.
- Ariff, M., Can,L., 2008. Cost and profit efficiency of Chinese banks: A non-parametric analysis. China Economic Review 19(2), 260–273.
- Bailey, W., Huang, W., Yang, Z., 2011. Bank loans with Chinese characteristics: Some evidence on inside debt in a state-controlled banking system, Journal of Financial and Quantitative Analysis 46(6), 1795–1830.
- Barros, C.P., Ferreira, C., and Williams, J., 2007. Analysing the determinants of performance of best and worst European banks: a mixed logit approach. Journal of Banking and Finance 31, 2189–2203.
- Barros, C.P., Managi, S., Matousek, R., 2012. The technical efficiency of the Japanese banks: non-radial directional performance measurement with undesirable outputs. Omega 40, 1–8.
- Battese, G.E., Rao, D.S.P., O'Donnell, C.J., 2004. A metafrontier production function for estimation of technical efficiencies and technology gaps for firms operating under different technologies. Journal of Productivity Analysis 21, 91–103.
- Berger, A. N., Hasan, I., Zhou, M., 2010. The effects of focus versus diversification on bank performance: Evidence from Chinese banks, Journal of Banking and Finance 34, 1417–1435.
- Berger, A. N., Hasan, I., Zhou, M., 2009. Bank ownership and efficiency in China: What will happen in the world largest nation?, Journal of Banking and Finance 33, 113–130.
- Berger, A.A., Patti, E.B.D., 2006. Capital structure and firm performance: A new approach to testing agency theory and an application to the banking industry. Journal of Banking and Finance 30, 1065–1102.
- García-Herrero, A. Gavilá, S. and Santabárbara, D. 2009. What explains the low profitability of Chinese banks? Journal of Banking and Finance 33 (11). 2080–2092.
- Chambers, R.G., Chung, Y., F., & Färe, R., 1996. Benefit and distance functions. Journal of Economic Theory 70,407–419.
- Chiu, C.R., Chiu, Y. H., Chen, Y. C., & Fang, C. L. 2016. Exploring the source of metafrontier inefficiency for various bank types in the two-stage network system with undesirable output. Pacific-Basin Finance Journal 36, 1-13.
- Dahmash, F., Durand, R., & Watson, J. (2009). The value relevance of reliability of reported goodwill and identifiable intangible assets. British Accounting Review 41, 120–137.
- Dietsch, M., Lozano-Vivas, A., 2000. How the environment determines banking efficiency: a comparison between French and Spanish industries. Journal of Banking and Finance 24, 985–1004.
- Färe R, Grosskopf S, Noh D-W, Weber W (2005) Characteristics of a polluting technology: theory

and practice. Journal of Econometrics, 126(2):469-492.

- Färe, R., Grosskopf, S., Lovell, C.A.K., & Pasurka, C., 1989. Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. Review of Economics and Statistics 71,90–98.
- Fu, T.T., Juo, J. C., Chiang, H. C., Yu, M. M., & Huang, M. Y. 2016. Risk-based decompositions of the meta profit efficiency of Taiwanese and Chinese Banks. Omega 62, 34-46.
- Fu, X., and Heffernan S., 2009. The effects of reform on China's bank structure and performance, Journal of Banking and Finance 33, 39–52.
- Fukuyama, H., Weber, W.L., 2009. A directional slacks-based measure of technical efficiency. Socio-Economic Planning Sciences 43, 274–287.
- Grigorian, D, V., & Manole, V. 2006. Determinants of Commercial Bank Performance in Transition: An Application of Data Envelopment Analysis, Comparative Economic Studies, Palgrave Macmillan, 48(3), 497–522.
- Jiang, C., Yao, S., Zhang, Z., 2009. The effects of governance changes on bank efficiency in China: A stochastic distance function approach, China Economic Review 20, 717–731.
- Kasman A., Yildirim, C., 2006. Cost and profit efficiencies in transition banking: the case of new EU members. Applied Economics 38,1079–1090.
- Kumbhakar, S.C., Wang, D., 2007. Economic reforms, efficiency, and productivity in Chinese banking, Journal of Regulatory Economics 32, 105–129.
- Kwan, S.H., 2003. The x-efficiency of commercial banks in Hong Kong. Journal of Banking and Finance, 30, 1127–1147.
- Margaritis, D., Psillaki, M., 2007. Capital structure and firm efficiency. Journal of Business Finance and Accounting 34, 1447–1469.
- Martin, M. (2012). China's banking system: Issues for congress. Congressional Research Service, 20.
- O'Donnell, C.J., Prasada Rao, D.S., Battese, G.E., 2008. Metafrontier frameworks for the study of firm-level efficiencies and technology ratios. Empirical Economics 34, 231–255.
- Oliveira, L., Rodrigues, L.L., Craig, R., 2010. Intangible assets and value relevance: evidence from the Portuguese stock exchange. The British Accounting Review 42, 241–252.
- Ozkan, A., 2001. Determinants of capital structure and adjustment to long run target: evidence from UK company panel data. Journal of Business Finance and Accounting 28, 175–198.
- Schaeck, K., Martin, C., Simon, W., 2006. Are competitive banking systems more stable? Journal of Money, Credit and Banking 41(4), 711–734.
- Simar, L. and Wilson. P.W. (2007). Estimation and inference in two-stage semi-parametric models of production processes. Journal of Econometrics, 136: 31-64.
- Valverde, S.C., Humphrey, D.B., Paso, R.L.D., 2007. Do cross-country differences in bank efficiency support a policy of "national champions"? Journal of Banking and Finance 31, 2173–2188.
- Wheelock, D. C., Wilson, P. W., 2000. Why do banks disappear? The determinants of U. S. bank failures and acquisitions, The Review of Economics and Statistics 82, 127–138.
- Zhou, P., Ang, B.W., Wang, H., 2012. Energy and CO2 emission performance in electricity generation: a non-radial directional distance function approach. European Journal of Operational Research 221,625–635.
- Zhu, N., Wang, B., & Wu, Y. 2015. Productivity, efficiency, and non-performing loans in the Chinese banking industry. The Social Science Journal 52, 468-480



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings



homepage: http://www.acrn/frap

Factoring Transition Risks into Regulatory Stress-Tests

Soline Ralite*, Jakob Thomä

2°Investing Initiative

Abstract

A debate has recently emerged as to whether climate risks may be material for financial stability, driven by a solid body of evidence that climate risks may create value destruction for key industrial sectors that are prominently represented in financial markets. As a result, financial supervisory authorities are starting to explore how these risks can be integrated into existing stress-testing frameworks. This paper proposes a methodology that financial supervisors could follow to build "late & sudden" transition scenario that could be used as input into either traditional or climate-specific stress-tests of regulated entities. The methodology specifically focuses on equity and corporate bonds tied to climate sensitive sectors (fossil fuels, power, steel, cement, automotive and aviation).

Keywords: climate stress-test, climate transition risks, scenario analysis

Introduction

In order for global temperatures to stay below +1.5° above pre-industrial era by the end of the century, estimates suggest that \$2.4 trillion would have to be invested annually until 2035 for the transition of the industrial, energy, agricultural, residential and transport sectors (IPCC, 2018). Similarly, a significant amount of capital will need to be moved out of current high- carbon investments in a range of sectors, including fossil-fuel mining, utilities, certain types of highcarbon manufacturing, and transport infrastructure. These necessary shifts in global financing flows will give rise to a new set of financial risks associated with the transition to a low-carbon economy. Failing to anticipate these so called "energy transition" risks might lead to large-scale mispricing of carbon-intensive assets (De Greiff et al., 2018), inevitably followed by sudden repricing when the market finally realizes the depth of the transition to come. In the speech that he gave to the European Commission on March 21th 2019, Mark Carney, the governor of the Bank of England, highlighted the need for financial supervisors to conduct climate stress-tests to assess the resilience of their regulated entities to such risks, and specifically to consider the eventuality of a "Climate Minsky moment", i.e. a sudden materialization of climate risks. Similarly, the European Systemic Risk Board (ESRB) recommended exploring how transition risks could be integrated into mainstream banking stress-testing frameworks (ESRB, 2016), and described in its scientific advisory board's report "Too Late, too sudden?" (2016) how a "late & sudden" transition scenario could impact overall financial stability.

The aim of this paper is to provide financial supervisors with a methodology that they could follow to build such a "late & sudden" transition scenario to input either in traditional or climatespecific stress-tests of regulated entities. The methodology specifically focuses on equity and corporate bonds tied to climate sensitive sectors (fossil fuels, power, steel, cement, automotive and aviation). The Literature review of this paper describes the scope of traditional regulatory stress-tests and reviews past climate stress-testing initiatives, The next section details the methodology that financial regulators could follow to estimate the impact of a "late & sudden" scenario on equity & bonds in climate-sensitive sectors, and the Results Section presents illustrative results obtained by applying this methodology.

Literature review

The scope of traditional regulatory stress-tests

A stress-test can be defined as a "what-if analysis that examines the effect of scenarios or sensitivities on the financial position of a bank, or a group of banks." (Cass 2013). Stress-tests are conducted internally by financial institutions as part of their risk management strategy, by regulators as part of the macroprudential policy framework, or by outside actors providing external analysis.

Stress-tests usually consist of three main parts: (i) a description of several disruptive economic scenarios and how they could propagate to the financial sector, (ii) a list of macroeconomic and sectoral parameters, as well as the values that they would take under each above-mentioned scenario, and (iii) impact indicators reflecting how each scenario impacts the financial sector. The time horizon of the scenarios is usually three years, and the scenario parameters and impact indicators displayed are often limited to a dozen. Table 1 displays the main characteristics of the stress-test conducted yearly by the Fed reserve in the United States, and by the ESRB in Europe.

	Fed reserve stress-test (USA)	ESRB stress-test (EU)
Risk scenarios considered	 Severe global recession accompanied by a global aversion for long-term fixed income assets, leading to a broad-based and deep correction in asset prices Weakening economic activities across all sectors, accompanied by rapid declines in long-term rates and flattening yield curves 	 Abrupt and sizeable repricing of risk premia in global financial markets – triggered e.g. by a policy expectation shock Adverse feedback loop between weak bank profitability and low nominal growth Public and private debt sustainability concerns; Liquidity risks in the nonbank financial sector
Time horizon Macroeconomic & sectoral indicators displayed	3 years GDP growth rate, unemployment rate, National House Price Index	3 years GDP growth rate, unemployment, HICP, Real estate prices
Impact indicators displayed	Dow Jones Stock Market Index, U.S. BBB corporate yield, Projected loan/revenues/net income losses & capital ratios for participating institutions	Stock prices, Long-term rates, Exchange rates, Foreign demand & commodity prices, SWAP rates, Credit spread indices

 Table 1 Main characteristics of US & EU regulatory stress-tests

Sources: Authors, based on ESRB (2018) & Board of the Governors of the Federal Reserve System (2018)

A growing interest from regulators for climate stress-testing

There is a growing debate as to whether climate risks may be material for financial stability. The claim is driven by a solid body of evidence that climate risks may create value destruction for key industrial sectors that are prominently represented in financial markets (e.g. energy, utilities). Indeed, according to Moody's analysis, \$9 trillion of their rated debt may be at immediate or elevated risk of downgrade in response to environmental risks (2Dii, 2017). Around \$15-20 trillion of market capitalization in stock markets is tied up with companies that are covered in the decarbonization scenarios of the International Energy Agency (2Dii, 2017). As a result, financial supervisory authorities are starting to explore how climate risks – and especially transition risks can be integrated into existing stress-testing frameworks. Associated recommendations around integrating such risks have been put forward by the UNEP Inquiry (2Dii / UNEP Inquiry 2015), as well as a number of leading think tanks (Bruegel 2016). The European Systemic Risk Board (ESRB) also recommended exploring how transition risks could be integrated into mainstream banking stress-testing frameworks (ESRB 2016). Research initiatives along these lines have been launched by the financial supervisory authorities in Sweden, the Netherlands, United Kingdom, and France (2Dii, 2015). We detail hereafter some of the most notable initiatives.

In 2017, Battiston et al. assessed the exposure of European financial institutions to fossil fuel production sectors and energy intensive sectors and analyzed the losses that these institutions would bear assuming an arbitrary 100% shock in the market capitalization of the fossil fuel sector. They concluded that such a shock to the equity value of the fossil fuel sector wouldn't threaten European financial stability, although specific banks could be significantly impacted. In 2017, the Dutch Central Bank assessed the potential impact of floods on credit losses and quantified the exposure of Dutch financial institutions to transition risks ("Waterproof? An exploration of climate-related risks for the Dutch financial sector", 2017). This report was followed by another, more in-depth analysis of transition risks and their impact on financial institutions' expected losses ("An energy transition risk stress test for the financial system of the Netherlands", 2018). This latter analysis was conducted using a CGE model, whose production functions were modified to reflect the consequences of several transition scenarios, and the macroeconomic impact was then distributed across sectors based on their emission levels. In 2018, the California Insurance Commissioner's Office conducted, in partnership with the 2°II, a climate scenario analysis project on insurance companies operating in California, aiming at quantifying the current and future exposure of these institutions to transitions risks, and physical risks to a lesser extent. However, the impact of these risks, were they to materialize, wasn't quantified. The Bank of England also included the impact of climate change and of a delayed transition in its UK insurers stress tests in 2019, partly based on the methodology presented in this paper. Finally, building on Battiston's 2017 paper, Battiston & Monasterolo published in 2019 a stress-testing methodology aiming at pricing transition risks in today's value of equity and corporate bonds in the energy & power sector, as well as in sovereign bonds' value ("A carbon risk assessment of central banks' portfolios under 2° aligned climate scenarios", 2019).

As highlighted above, methodologies already exist to quantify the exposure of financial institutions' portfolios to climate risks, and first steps have been taken to assess the potential losses that these institutions would incur, were these risks to materialize. However, all existing studies consider a "smooth" transition scenario, and the consequences of a proper "late & sudden" scenario haven't yet been assessed. The aim of this paper is to bridge this gap, by developing a "too late, too sudden" transition scenario that can be used by financial supervisors as inputs into either traditional or climate-specific stress-tests of regulated entities, including not only assets tied to the energy and power sector, but also key industrial (steel & cement) and transportation (aviation & automotive) sectors.

Methodology

Choice of the transition scenario

Many uncertainties remain as to the form that a low-carbon transition would take. The ESRB scientific advisory board's report "Too Late, too sudden?" (2016) identified two types of scenario outcomes, a "gradual", smooth ambitious scenario and a late & sudden one. This concept has been further developed by the UN PRI in 2018 operating under the premise of an "Inevitable Policy Response" (UN PRI, 2018). In addition to the two more ambitious scenarios, transition outcomes could also of course involve a "do nothing" approach or a limited climate transition ambition, which are of little interest to assess the materiality of transition risks.

Considering that the purpose of a stress-testing exercise is to assess the impact of a "worstcase" scenario on the financial system, a "too late, too sudden" scenario is more suited than a "smooth transition" one. Such a "late & sudden" scenario assumes that no climate action is taken for a few years, and that ambitious action is taken afterwards to stay below the 2°C

threshold by the end of the century. This approach also allows for a "sentiment" shock at the moment climate action is taken, leading to a sudden repricing of financial assets.

However, these types of scenarios haven't yet been explored by macroeconomic or energyeconomy models, and few information is therefore available to quantify there economic implications. The following pages aim at bridging this gap by providing a methodology to estimate the impact of such trajectories on sectoral profits and financial assets' value (equity and corporate bonds).

Estimating the impact of a "too late, too sudden" transition on sectoral profits

Before estimating the impact of the energy transition on financial assets' value, the first step is to understand how the profits of companies issuing these securities would be affected by an energy transition, and in particular by a "too late, too sudden" transition. To do so, we could have either used sectoral value-added estimates produced by a CGE (Computable General Equilibrium) energy-economy model, or model these changes in sectoral profits using a bottom- up approach. We chose the second approach as there isn't, to our knowledge, any energy- economy model assessing the consequences of a truly "late & sudden" scenario on sectoral value-added. Although we focus on changes in aggregated sectoral profits in this paper, a bottom-up approach also allows for a corporate-level analysis, and choosing this approach thus leaves room for future refinements (See Section Results for more details). The first step was then to choose the key indicators needed to model sectoral profits, and next step was to estimate the value that these indicators would take under a "delayed action" scenario, based on their value under a "smooth" transition scenario.

Figure 1 below illustrates how transition risks, and particularly those stemming from a "too late, too sudden" transition, would impact carbon-intensive industries' profits across the entire "profit value chain", and details the indicators needed to quantify each of these impacts.

-

Net j	a b b b c <th>T + Interests) Indicators needed to quantify the impact</th>	T + Interests) Indicators needed to quantify the impact
1	Increased cost of emitting CO2 : Under a transition scenario, the implementation of a carbon tax will cut the margin of carbon intensive industries proportionally to their emissions. Under a "too late, too sudden" scenario, carbon prices would need to be higher than under a "smooth" transition scenario, in order to foster a quick decrease in emissions.	 Production Carbon intensity of production Carbon tax
2	Increased cost of production inputs : During a low carbon transition, carbon intensive goods will increase in prices due to pass-through of direct emissions costs. Industries using such carbon intensive goods as production inputs will thus be impacted.	- Prices of production inputs
3	Additional depreciation costs and R&D expenditures: Under a transition scenario, significant capital expenditures in low-carbon technologies will increase companies' annual depreciation costs (included in Operating Expenses). Under a "too late, too sudden" scenario, the depreciation costs of "brown" capital stocks build up before the transition starts add up to these green expenses. R&D expenditures will also likely increase.	- CAPEX - R&D expenditures - All other OPEX
4	Changes in revenues : Companies' revenues will be affected through a change in prices and consumer demand: As they become increasingly costly to produce, prices of carbon intensive goods will likely increase, and consumers will, in turn, decrease their demand for such goods. A delayed transition, as it would increase the costs bared by carbon-intensive industries, would likely deepen this effect.	- Production - Prices

Figure 1 Impact of transition on sectoral profits & indicators needed to quantify the impact

A range of initiatives have already sought to quantify the sectoral impacts of a "smooth" energy transition, and provide some indicators allowing to quantify its impact on the profits' determinants detailed in Figure 1. Two relevant initiatives in this regard are the EU H2020-funded ET Risk project and UNEP FI's working group on transition risks (UNEP FI, 2018). To the knowledge of the authors, no research has however yet been conducted to understand the impact on sectoral profits of a delayed transition scenario, although initiatives looking at this issue are under way (notably led by UN Principles for Responsible Investment as part of their "Inevitable Policy Response" work) (PRI, 2018).

To bridge this gap, we developed "too late, too sudden" scenarios including the indicators needed to quantify the impact of a delayed transition on sectoral profits, based on the work done in the ET RISK project. Table 2 below details the indicators, mainly taken from the ET Risk project and IEA's scenarios, that we used to build our "too late, too sudden" scenarios.

Sector	Target companies	Geography	Indicators used for profits calculation	Source of the data for the BaU & smooth transition
Oil	Upstream oil	Europe, North	Production, Prices	- Production data taken from
Coal	Coal mining	America, South &		WEO2018 SDS & NPS
Natural gas	Upstream natural gas	Central America, Middle East, Africa, Asia Pacific, Eurasia		- Prices data taken from ETP2017 B2DS & RTS
Power	Power generators (Coal, Gas, Solar, Wind)	Europe, USA, Latin America	Production, Prices, Levelized Cost of Electricity, Subsidies	- Production data taken from WEO2018 SDS & NPS - Electricity prices, LCOE & Subsidies taken from ET Risk
Steel	Crude steel producers	Brazil, USA, Mexico, Erence	Production, Prices, Carbon	Production, Prices, carbon prices & carbon
Cement	Cement producers	Germany, Italy	Carbon intensity	from ET Risk
Automotive	Car producers	World average	Production, Net margin by powertrain type	Production data & net margin derived from ETP2017 & BNEF
Aviation	Airlines (international)		Demand, Fuel efficiency, Fuel prices	Demand taken from ETP2017 B2DS & RTS, fuel prices & fuel efficiency taken from ET Rick

Table 2 Sectors included in the analysis and indicators used for profits calculation

To estimate the values that the indicators mentioned in Table 2 would take under a delayed transition scenario (i.e. starting in 2025, as explain above), we followed the following principles:

Production & efficiency:

• For each sector, the **additional emissions occurring before the start of the transition** under a delayed action scenario compared to a smooth transition scenario (date at which the transition starts) **have to be offset by 2040**, assuming a climate lag of 60 years (the temperature of 2100 is determined by the GHG emitted 60 years before).

• Additional emissions occurring before 2025 can be **offset through either a drop in production or a surge in energy efficiency**, depending on the sector considered. For example, cement being

an essential material to build the infrastructures needed for the 10 billion humans expected by 2050, assuming a major drop in production wouldn't make sense (as confirmed by the IEA in ETP 2017), a surge in energy efficiency due to sudden R&D efforts seems more realistic (or the development of a substitution product, but we didn't consider this possibility in the paper).

• This can either be done at global level, i.e. considering the emissions occurring in each sector before 2025 in the IEA New Policy Scenario (or any other global Business as Usual scenario of this kind) or done at the portfolio level, using tools such as the Paris Agreement Capital Transition Assessment (PACTA) Transition Monitor¹. The tool **quantifies the current exposure of investment portfolios to "climate-relevant"** sectors and technologies (Fossil fuels – Oil, coal, natural gas; Power – Coal, gas, renewables; and Automotive – Electric vehicles, hybrid vehicles, ICE vehicles) and **provides a forward-looking assessment of its alignment with 2°C scenarios** (based on the production and investment plans of the companies financed by the portfolio). The emissions occurring before the start of the transition, which have to be offset afterwards, will thus reflect the production that is currently funded by the portfolio are calibrated to its current **trajectory**.

The results displayed in this paper are based on the "global market" approach (i.e. the production in each sector before the start of the transition follows the NPS scenario).



Figure 2 below illustrate these principles for the coal sector:

Figure 2. Illustrative Trajectory of the coal production financed by an investment portfolio under a delayed transition scenario starting in 2025

Prices and profit margins:

• Fossil fuel prices under a delayed transition scenario evolve proportionally to demand; prices for other sectors slowly reach the levels of a "smooth" transition once the "late & sudden" transition starts.

¹ See https://www.transitionmonitor.com/ for more details

• No impact on gross or operating margins is assumed for building material industries (Steel & Cement), as the authors didn't find any reasonable way to estimate this under a delayed transition scenario. The impact that a delayed transition would have on profits for these sectors might thus be underestimated.

• In line with literature,² carbon prices are assumed to be 1.5 times higher in 2040 under a "too late, too sudden" scenario compared to a "smooth" transition scenario, to foster quicker energy efficiency improvements once the late & sudden transition has started.

Although the approach developed above represents a valuable first step in the development of "too late, too sudden" transition scenarios including all the indicators needed for financial analysis, there are several caveats to bear in mind. First, the approach overlooks possible interactions between sectors (in reality, emissions may decrease less than needed in an industry and more than needed in another) – although it takes into account risk propagation across industries (e.g. an increase in oil prices impacts airlines expenses). Second, in the absence of alternative solutions, it features a very simplistic price dynamic. Finally, in the absence of alternative solutions, it neglects changes in net margins for some sectors.

Estimating equity value under a "too late, too sudden" transition scenario

As explained above, the energy transition will impact companies' revenues and charges, with the amplitude of the effect varying depending on the sector they operate in. These changes in the companies' profits will then impact their market value, as the demand for shares issued by weaken companies will decrease.

To estimate changes in share prices under a "too late, too sudden" scenario, we rely on Gordon's formulation of future dividends' flows (Gordon 1959). The equity market price VE at time t0 is given by:

 $V_{E,t_0=\frac{D_1}{r-g}}$ with D1 being the expected dividends for the next year, r being the cost risk of capital

for the company, and g being the dividend's growth rate.

Assuming that dividends for a given year are proportional to the net profits of the company for this year, and explicitly modeling the future evolution of profits, we derive the following formula:

$$V_{E,t_0=\alpha\sum_{t=t_0}^{t*}\frac{P_t}{(1+r)^t}*(1+x)}$$

With Pt being the profits made by the company in year t (modelled as explained in Section Choice of the transition scenario), t* the date until which we explicitly model cash-flows, x the percentage of modelled value in the terminal value, and α the proportionality coefficient between net profits and dividends.

In simple words, the value of equity for a given company is assumed to equate the Net Present Value of its future cash-flows.

² See Advance_2020_Med2C ("smooth" transition scenario) and Advance_2030_Med2C (slightly delayed transition scenario) on IAMC's 1.5° online scenario database

We set r = 5%; $t_0 = 2025$ (i.e. we assume a sudden repricing of equity in 2025, date at which the TLTS transition starts, due to a market sentiment shock), $t^* = 2040$, x = 10% and $\alpha = 1$ for all scenarios.

The difference between $V_{E,t0}$ under the Business as Usual and the "too late, too sudden" scenarios is the equity value put at risk by the transition.

Estimating corporate bonds' value under a "too late, too sudden" transition scenario

The most influential factors that affect a bond's value are its yield, prevailing interest rates (as they affect the discount rate of the bond's cash flows) and the bond's probability of default. As one cannot anticipate how a "late & sudden" transition would affect inflation, and thus long- term interest rates, we focus in this Section on default-risk as the sole driver of bond value changes under a transition scenario, and discount rates are kept constant across all scenarios.

Estimating the probability of default under a transition scenario

Studies have shown that bonds' probabilities of default are heavily correlated with the main financial ratios of their issuers (Tang & Yan, 2010). Among the financial ratios contributing most to the prediction of default, NI/TA (net income over total assets) is likely to be the most strongly affected by companies' profit losses due to the transition.

We assume that a X% change in net income translates into a X% change in NI/TA ratio (i.e. we assume that total assets are not affected by the transition) and, based on Zmijewski's bankruptcy model (Zmijewski, 1984), we then derive the marginal effect that a change in NI/TA would have on the probability of the bond defaulting over the course of the upcoming year.

Zmijewski's probit model is as follows:

$$PD = \varphi(-4.336 - 4.513\frac{NI}{TA} + 5.679\frac{TL}{TA} + 0.004\frac{CA}{CL})$$
(1)

Where PD is the 1-year probability of default, ϕ the standard normal cumulative distribution function, NI/TA net income over total assets, TL/TA total liabilities over total assets, and CA/CL current assets over current liabilities.

In a probit model, the marginal effect of a change in one of the explanatory variables on probability of default is given by:

$$\frac{\partial Y}{\partial x_i} = \beta_i \phi(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)$$

Which, in our case, translates into the following formula:

$$\Delta PD_t = \Delta \frac{NI}{TA_t} * -4.513 * \phi(-4.336 - 4.513 \frac{NI}{TA_{BaU,t}} + 5.679 \frac{TL}{TA_{BaU,t}} + 0.004 \frac{CA}{CL_{BaU,t}})$$
(2)

With:

 ϕ being the standard normal probability density function;

 ΔPD_t being the additional probability that the bond default in year t+1 under a transition scenario compared to a BaU scenario;

$$\Delta \frac{NI}{TA}$$

 T^{A_t} being the difference in NI/TA in year t between a BaU scenario and a transition scenario;

and $\frac{NI}{TA_{BaU,t}}, \frac{TL}{TA_{BaU,t}}, \frac{CA}{CL_{BaU,t}}$ being the "baseline" financial ratios under a BaU scenario in year t.

This gives us the additional probability that a bond defaults within one year under a transition scenario compared to a Business-as-Usual scenario.

Estimating the value of a bond under a transition scenario

Let X represent the present value of a bond's cash flow stream. The standard way to value a bond with a given probability of default in year t (PD_t) is to take each possible value of X, multiply it by its probability and sum the results:

$$V_j = \sum_{t=1}^{T} X_t P D_t (\prod_{k=0}^{t-1} (1 - P D_k))$$
(3)

With Vj being the value of bond j, T being the maturity date of the bond, and PD_t being the probability of default computed in Section Estimating the probability of default under a transition scenario.

To come up with an expression for the X_t, the following variables need to be defined:

Fj is the face value of the bond j, Cj is the coupon rate of the bond j, R is the recovery rate in case of default, rj is the discount rate for the cash flows.

For a bond expected to mature in T time periods, with coupons paid every period, the present value of its cash flow stream, assuming no default, can be written as:

$$X_T = \sum_{t=1}^T \frac{C_j F_j}{(1+r_j)^t} + \frac{F_j}{(1+r_j)^T}$$
(4)

By combining both equations (3) and (4), we get:

$$V_j = \sum_{t=1}^T \frac{C_j F_j}{(1+r)^t} \left(\prod_{k=1}^t (1-PD_k)\right) + R_j F_j \sum_{t=1}^T \frac{PD_t}{(1+r)^t} \left(\prod_{k=0}^{t-1} (1-PD_k)\right) + \frac{F_j}{(1+r)^T} \prod_{k=1}^T (1-PD_k) + \frac{F_j}{(1+r)^T} \prod_{k=1}^T \prod_{k=1}^T (1-PD_k) + \frac{F_j}{(1+r)^T} \prod_{k=1}^T \prod_{k=1}^T$$

In the example displayed in the Section below, we set $Rj = 38\%^3$, Fj = 1000, Cj = 5%, r = 5%.

³ Historical recovery rate of senior bonds (Moody's, 2017)

Results

In this section, we display some results obtained using the methodology developed above⁴. Changes in the mean equity value of companies in key sectors under a "late & sudden" transition scenario, assuming a sudden repricing due to a market sentiment shock when the transition starts, are displayed in Table 3 (World average). As for corporate bonds, Table 4 illustrates the increase in the mean 1-year probability of default of bonds tied to sensitive sectors, 1 year and 10 years after the beginning of a "late & sudden" transition (World average). Table 5 displays the change in the value of an illustrative bond with a face value of 1000\$, a 5% coupon rate, and a 38% recovery rate, depending on the sector it is tied to, and depending on its remaining time to maturity after the "late & sudden" transition starts (in 2025).

Table 3 Mean change in equity value compared to a BaU scenario⁵ under a "too late, too sudden" transition scenario for key sectors, assuming a sudden repricing in 2025 (%)

	Change in equity value (%)
Upstream Oil	-53.3%
Coal mining	-57.0%
Upstream gas	-30.8%
Coal electricity	-80.1%
Gas electricity	-20.3%
Solar PV	19.2%
Wind electricity	12.8%
Nuclear	19.9%
Crude Steel	-52.0%
Cement	-27.0%
Automotive	-9.5%
Aviation	-21.0%

As expected, upstream fossil fuel industrials and fossil-based power producers, in particular coal electricity producers, are the most strongly hit by the transition, while listed renewable energy producers enjoy a significant revaluation of their shares.

It is worth noting that these results are sectoral averages, and thus do not consider the adaptative capacities of individual companies. This aggregated impact on sectoral equity value might hide significant disparities between companies of a given sector. As, in the context of regulatory stress-testing, changes in the value of entire asset classes are of more interest than changes in individual asset values, this isn't much of a concern. Our flexible bottom-up approach to estimating changes in sectoral profits, detailed in Section Choice of the transition scenario,

⁴ As explained page 9, the "late & sudden" transition scenario was built based on the IEA global scenarios, and not the trajectory of a specific investment portfolio

⁵ 5 Corresponding to IEA's NPS & RTS scenarios

could however be adapted to uncover these disparities. Global production trends taken from the IEA could be broken down to company level using a fair-share approach, while indicators related to energy efficiency and operating margin could be estimated on a case-by-case basis, based on the CAPEX and R&D expenditures already engaged by the company. Such an approach would enable the assessment of the consequences of the transition on companies with mixed revenue streams (e.g. revenues from carbon intensive and renewable power production at the same time) (The CO FIRM, ClimateXcellence, 2018).

2018 2026 2035 (Baseline⁶⁶) 0.03 0.03 0.06 Steel 0.02 0.02 0.05 Cement Oil 0.01 0.03 0.06 0.03 0.04 0.09 Coal 0.02 0.02 0.05 Gas 0.04 0.09 **Coal power** 0.02 0.04 Gas power 0.027 0.02 0.00 Nuclear 0.02 Solar PV 0.00 Wind 0.01 0.00 0.03 0.04 0.06 Airlines 0.01 0.01 0.02 Automotive

Table 4 Mean 1-yr probabilities of default of bonds issued by climate-sensitive sectors under a "too late, too sudden" transition scenario (%)

Table 5 Mean change in bond values compared to baseline under a "too late, too sudden" transition scenario, depending on their remaining time to maturity, and assuming a sudden repricing in 2025 (%)

	1	2	3	4	5	6	7	8	9	10
Steel	-0.2%	-0.7%	-1.2%	-1.9%	-2.6%	-3.3%	-4.1%	-4.9%	-5.7%	-6.5%
Cement	-0.2%	-0.5%	-1.0%	-1.5%	-2.2%	-2.9%	-3.8%	-4.7%	-5.6%	-6.6%
Oil	-1.4%	-2.9%	-4.6%	-6.4%	-8.3%	-10.0%	-11.7%	-13.3%	-14.8%	-16.2%
Coal	-0.8%	-1.9%	-3.2%	-4.6%	-6.2%	-7.7%	-9.2%	-10.6%	-12.0%	-13.1%
Gas	-0.5%	-1.1%	-1.9%	-2.9%	-3.9%	-5.0%	-6.1%	-7.2%	-8.2%	-9.3%
Coal power	-1.1%	-2.5%	-4.2%	-6.2%	-8.4%	-10.2%	-12.1%	-13.8%	-15.5%	-17.1%
Gas power	-0.4%	-0.8%	-1.2%	-1.6%	-2.1%	-2.8%	-3.5%	-4.2%	-5.0%	-5.7%

⁶ Bloomberg data, Q4 2018

⁷ Mean 1yr probability of default of power producers worldwide in Q4 2018 (Source: Bloomberg)

Nuclear	0.1%	0.3%	0.6%	0.9%	1.4%	1.8%	2.4%	3.0%	3.7%	4.4%
Solar PV	0.4%	1.1%	2.2%	3.5%	4.6%	5.6%	6.6%	7.6%	8.5%	9.3%
Wind	0.3%	0.9%	1.7%	2.8%	4.0%	5.1%	6.1%	7.0%	7.9%	8.8%
Airlines	-0.2%	-0.6%	-1.2%	-1.9%	-2.6%	-3.4%	-4.2%	-5.1%	-5.9%	-6.7%
Automotive	-0.2%	-0.5%	-0.8%	-1.1%	-1.3%	-1.5%	-1.8%	-2.1%	-2.4%	-2.7%

As illustrated by **Table 4** and **Table 5**, the bond value that is put at risk by a "late & sudden" transition increases as a function of the time to maturity of the bonds, driven by a rise in their 1-year probability of default as the transition progresses. As highlighted above for equity, Coal & Oil producers, as well as Coal power producers are the most strongly affected by a late & sudden transition.

Caveats

Although the approach developed above represents a valuable first step in the development of a "too late, too sudden" transition scenario that could be used by financial supervisors as an input into climate stress-tests of regulated entities, there are several caveats to bear in mind.

First, the methodology that we developed in the Section on estimating the impact of a late & sudden transition on sectoral profits suffers some limitations – as detailed in this section: it overlooks some possible interactions between sectors, it considers a simplistic price dynamic, and it neglects changes in net margins for some sectors. The approach developed in this paper fills a gap – the absence of a proper late & sudden transition 'stress-test' scenario including the indicators needed to estimate the changes in sectoral profits, but it shouldn't be considered sufficient. A proper late & sudden scenario developed by the energy-economy modelling community, granular both at geography and sectoral levels, would to take climate stress-testing a step further.

Second, the methodology that we developed to assess the changes in bond and equity value for companies in carbon-intensive sectors doesn't consider potential mixed revenue streams, they apply to theorical companies with all their revenues coming from only one carbon-intensive sector. When applying the equity and bond shocks estimated above to investment portfolios, a solution would thus be to compute an average shock for each company in the portfolio, weighed by the share of their revenues coming from each carbon-intensive sector. Alternatively, as mentioned pg. 12, our methodology could also be applied at company-level and directly factor mixed revenue streams into future profits calculations.

Third, this paper doesn't cover sovereign bonds. This is because contrary to Battiston & Monasterolo (2019), we do not consider that a correlation between short-term GDP changes and sovereign interest rates is clearly established in literature, either in developed or emerging economies (IMF, 2012; G.Min, 1999), let alone did we find a value to use for the sensitivity factor. Fiscal indicators, in turn, are designated by the literature as key drivers of sovereign bonds' interest rates, but the methodology developed in this paper doesn't yet allow us to quantify the changes in a country's indebtedness under a late & sudden transition scenario. A next step would thus be to investigate the relationships between profits of carbon-intensive sectors and fiscal revenues, and understand how shocks to the value-added of these sectors would impact a country's dept-to-GDP ratio.

Conclusion

To answer the growing demand of financial supervisors for climate stress-testing scenarios, this paper provided a methodology for developing a "late & sudden" transition scenario that can be used as input into traditional or climate-specific stress-tests of regulated entities. In this paper we (1) built "late & sudden" transition scenarios including all the indicators needed to estimate future changes in profits of carbon-intensive sectors, and which can be calibrated to reflect the current climate trajectory of any investment portfolio, (2) developed a methodology to price the risks associated with a late & sudden transition into equity and corporate bonds' value, and (3) demonstrated that the risk associated with a late & sudden transition might have a significant impact on equity and bond value of companies in carbon-intensive sectors, fossil fuels extraction and coal-based power production being the most threatened activities.

Combining the methodology developed above with an analysis of the exposure of financial institutions to carbon intensive sectors would allow financial supervisors to assess the potential impact of a "late & sudden" transition on financial stability. Such a combination of top-down stress-tests and exposure analysis through asset level data has been pioneered in the Bank of England's 2019 insurance stress-test, which was informed in part by the results presented above. This methodology will also be applied by the 2°Investing Initiative in their partnership with the European Insurance and Occupational Pension Authority (EIOPA) in 2019.

Bibliography

- 2°II. (2017). "Right direction, wrong equipment". 2°II. (2016). "Transition Risk Toolbox".
- 2°II. (2015). "Financial risk and the transition to a low-carbon economy: towards a carbon stress testing framework"
- 2°II. (2014). "Carbon risks for financial institutions". 2°II; The CO FIRM. (2017). "Transition risk-O-meter".
- 2°II; UNEP FI; WRI (2015). "Climate strategies and metrics. Exploring options for institutional investors".
- Alexius, A., & Sp, D. (2018). Stock prices and GDP in the long run, 8(4), 107–126.
- Battiston, S., Mandel, A., Monasterolo, I., Schuetze, F., & Visentin, G. (2015). A Climate stress-test of the EU financial system, 1–12.
- Cantor, R., & Packer, F. (1996). Determinants and Impact of Sovereign Credit Ratings, 37–54. CISL. (2015). Unhedgeable risk.
- De Nederlandsche Bank. (2018). An energy transition risk stress test for the financial system of the Netherlands.
- De Greiff, K., Delis, M., & Ongena, S. (2018), "Being Stranded on the Carbon Bubble? Climate Policy Risk and the Pricing of Bank Loans", CEPR Discussion Papers (May 2018).
- Dudian, M., & Popa, R. A. (2012). Analysis on the Relationship between Rating and Economic Growth for the European Union Emergent Economies, 6(4), 431–434.
- Ecofys. (2018). Climate risks of 3 funds of Actiam.
- ESRB. (2016). "Too late, too sudden: Transition to a low-carbon economy and systemic risk". ESRB. (2018). "Adverse macro-financial scenario for the 2018 EU-wide banking sector stress test".
- FED reserve. (2018). Dodd-Frank Act Stress Test 2018: Supervisory Stress Test Methodology and Results, (June).
- International Energy Agency. (2017). World Energy Outlook.
- International Energy Agency. (2017). Energy Technology Perspectives 2017.
- Internation Monetary Fund. (2012). Long-Run and Short-Run Determinants of Sovereign Bond Yields in Advanced Economies.
- IPCC. (2018). Special report on global warming of 1.5°.
- Jakob, M., Luderer, G., Steckel, J., Tavoni, M., & Monjon, S. (2012). Time to act now? Assessing the costs of delaying climate measures and benefits of early action, 79–99. https://doi.org/10.1007/s10584-011-0128-3
- Johnson, N., Riahi, K., Kriegler, E., Johnson, N., Bertram, C., & Elzen, M. Den. (2013). Locked into Copenhagen pledges — Implications of short-term emission targets for the cost and feasibility of long- term climate goals Technological Forecasting & Social Change Locked into Copenhagen pledges — Implications of short-term emission targets for the cost and feasibility of long-term climate goals. Technological Forecasting & Social Change, (January 2014). https://doi.org/10.1016/j.techfore.2013.09.016
- Hong G. Min. (World Bank). (1999). Determinants of emerging market bond spread: do economic fundamentals matter?

- Moody's. (2015). Heat Map Shows Wide Variations in Credit Impact Across Sectors. OECD. (2017). Investing in Climate, Investing in Growth.
- Ravina, A. (2017). Assessing transition risk with a stress test methodology, (2013), 1–18. S&P. (2016). 2016 Annual Global Corporate Default Study And Rating Transitions.
- S&P. (2019). Connecting the dots: Energy transition scenarios and credit quality.
- S&P. (2015). How Environmental And Climate Risks Factor Into Global Corporate Ratings.
- S&P. (2017). How Environmental And Climate Risks And Opportunities Factor Into Global Corporate Ratings An Update.
- Tang, D.Y., & Yan, H. (2010). Market conditions, default risk and credit spreads. Journal of Banking and Finance, 34(4), 743–753. https://doi.org/10.1016/j.jbankfin.2009.05.018
- The CO FIRM; Kepler Cheuvreux. (2018). Investor primer to transition risk analysis.
- The CO FIRM; ClimateXcellence. (2018). Climate scenario analysis: cement's financial performance under 2° c and 2.7° c
- J. Thomä, S. Dupré, M. Haynes. (2018). A Taxonomy of Climate Accounting Principles for Financial Portfolios. Sustainability 2018, 10(2), 328; https://doi.org/10.3390/su10020328
- UNEP FI; Mercer. (2018). Extending our horizons.
- UN PRI. (2018). The inevitable policy response: Act now.



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings

homepage: http://www.acrn.eu/frap



Early Stage Investing in Green SMEs: The Case of the UK

Robyn Owen*, Othmar Lehner, Fergus Lyon, Geraldine Brennan

Centre for Enterprise and Economic Development, Middlesex University Business School

Abstract

How might a Green New Deal be applied to the early stage financing of Cleantechs? Amidst rising interest and adoption of Green New Deals in the US, the paper explores the need for more focused policy to address early stage long horizon financing of Cleantechs. We argue that insufficient focus has been applied to early stage investing into these types of SMEs that could lower CO2 emissions across a range of sectors through innovations (including renewable energy, and also for example recycling, advanced manufacturing, transport and bio-science). Adopting a resource complementarity lens and borrowing from transaction cost theory, we illustrate and build theory through longitudinal UK case studies how government policy can scale-up through international collaboration public-private principally venture capital co-finance in order to facilitate cleantech innovation with potentially game changing impacts on reducing CO2 emissions in order to meet the Paris 2015 Climate Change targets.

Keywords: Green New Deal, Early Stage Cleantech, Low Carbon, SME Finance

Introduction

A Green New Deal would represent an unprecedented, massive, *mobilisation and deployment of resources to tackle the accelerated climate change crisis*, requiring regions, nations and ultimately global acceptance (The Guardian, 23/04/2019).

How might a Green New Deal (GND) be applied to the early stage financing of Cleantechs? Amidst rising interest and adoption of Green New Deals in the US and potentially in the UK (NEF, 2019; The Guardian, 2019) and declaration of a '*Climate Emergency*' by the Welsh Government (29/04/2019), the paper explores the need for more focused policy to address early stage long horizon financing of Cleantechs.

Over 100 national governments have signed up to the Paris 2015 agreement to work together to form a more cohesive approach to tackling climate Change (UNFCCC, 2019). The major contributors to CO2 emissions in developed countries like the US are energy, heavy industry and transport (81%; EPA, 2017). Therefore, arguably, a key issue is to raise investment and overcome current market failures whereby private investors perceive that the risks of investment outweigh potential returns and investors are not appropriately remunerated for green investments (Lehner, 2016; Owen et al, 2018; Polzin, 2017). Whilst recent efforts have focused on developing Green Investment Banks and tackling larger infrastructural renewable energy projects, such as wind farms (Mazzucato and Semieniuk, 2018), less attention has been given to early stage Cleantech SME investment (Owen et al, 2018).

The paper focuses on early stage investment in Cleantech, which is broadly defined here as private for-profit SMEs which may be pre or early trading that have a mission to develop and adopt innovative technologies to reduce CO2 emissions in their products and processes (Kenton, 2018). They typically undertake long horizon R&D and struggle to obtain sufficient, often high,

levels of private investment required to reach commercialisation (Owen et al, 2019; BEIS, 2017; Rowlands, 2009).

Focusing on the example of the UK, post the 2008 Global Financial Crisis (GFC), we examine the role of the UK Innovation Investment Fund (UKIIF) as a leading example of government intervention to address market failure through public-private co-financing to raise the level of early stage VC investment into Cleantechs. Adopting a resource complementarity lens we are able to examine the operation of the UKIIF programme to establish what works well or less well and address the question as to how a GND approach might be applied, principally to venture capital (VC), to improving the under-developed Cleantech early investment market.

Positioning the UKIIF in the VC literature

The arguments for the importance of private VC in the development of innovation are well established (Gompers et al. 2010; Lerner, 2010; Munari and Toschi, 2015; Cumming and Johan, 2016; Baldock, 2016). Private VCs develop specialist skills and knowledge to invest in earlier stage innovative ventures and sectors, often preferring to invest locally in order to have hands-on contact with the venture management teams. Adopting Markowitz (1952) pareto principles they need to invest in sufficient numbers of portfolio ventures and at sufficient scale over time in order to obtain returns, which are principally made up of the most successful venture exits (usually through trade sales or IPOs). The problem for VC is that early stage investment is risky and the returns are longer horizon, typically taking 5-7 years and lengthening in times of recession (Owen et al, 2019).

Early stage Cleantechs may, therefore, be considered particularly vulnerable as they often exhibit long horizon intensive R&D with long *valley of death* periods spanning proof of concept to early commercialisation (Mazzucato and Semieniuk, 2018). Additionally, they suffer from a higher *Liability of Newness* compared to the already known issues of traditional new ventures (Lehner and Harrer 2018; Lehner and Nicholls, 2014), which may be because of their Cleantech hybrid business-models (Quelin 2017) that aim to combine a commercial logic with an environmental mission (Doherty, Haugh and Lyon, 2014). Since investors are not appropriately rewarded for the full environmental-societal value, the risk-reward balance is often viewed as unfavourable (Bocken, 2015; Bak, 2017; Owen et al, 2018). As a result, there is an increasing resource-scarcity in these ventures with large funding-gaps (Lehner, 2016; BEIS, 2017).

Mazzucato and Semieniuk (2018) refer to the well-established role of government to address market failure for public good. Furthermore, Lerner (2010) makes a compelling case for public intervention to co-invest with private VC to address the early stage funding gaps of innovative ventures. He also raises lessons from past programme failures, advocating; sufficient funding scale and appropriate consideration of follow-on funding requirements, long fixed-term VC co-finance arrangements (with powers to extend to obtain optimal value portfolio venture exit) with recruitment of experienced private fund managers to take the lead in investment decision-making, and the cultivation of a suitably supportive financing ecosystem which integrates University and R&D centres with the finance support intermediaries (e.g. accountants and lawyers) and public-private VC programmes. Owen and Mason (2019) take this forward with their vision for a mature economy institutional framework for the early stage innovation finance escalator, noting the need for an international approach to scale-up VC finance (supported in Baldock, 2016). What is missing, is a more focused approach to developing early stage *green VC*.

Owen at al (2019) draw on Lerner's (2010) and others' (e.g. Cumming and Johan, 2016; Munari and Toschi, 2015; Colombo et al, 2014; Grilli and Murtinu, 2014; Technopolis, 2009) lessons, indicating that the UK is a global leader in co-financing with business angels and VCs. They point to the £330m UK Innovation Investment Fund (UKIIF), established in 2009, as a leading programme addressing early stage innovation investment – second in size only to the more generic £1.2b Enterprise Capital Funds - and the only national VC programme with a clear focus on low carbon sectors (renewable energy, recycling, advanced manufacturing, digitech and bio-science; see BIS, 2012). It is worth noting that the proposed UK £20m Clean Growth Fund announced in 2017 was still being established in Summer 2019.

The UKIIF has some ground-breaking features (BIS, 2012). It operates pari passu on an equal footing between public and private investors, unlike ECFs which provide an upside incentive to encourage private investment. This is to accommodate EU policy, since the programme has £100m of European Investment Fund investment. UKIIF operates as a fund of funds, so although it is regulated and monitored by the British Business Bank, it is privately operated, with two umbrella funds (UK Future Technologies Fund and Hermes Environmental Impact Fund) contracting (unlike the more direct contracting of private VCs operated by the British Business Bank for ECFs) and overseeing the underlying frontline venture investment VC funds. Oxford Research (2015) points to a dearth of private VC fund of funds operating in Europe, due to poor track records and high administration costs of having two layers of fund management. However, on balance they suggest that public co-finance fund of funds can achieve sufficient scale to attract larger private institutional and international co-investment. They conclude that this provides the most suitable approach to international scale-up of VC funding and addressing the dual issues of achieving sufficient fund size and serving thin markets with insufficient viable business propositions (Nesta, 2009).

Although beyond the remit of this paper, it should be noted that numerous studies recognise that stimulation of demand-side early venture creation is crucial to co-financing VC programme success (see for example Cumming and Johan, 2016; ; Avnimelech and Teubal, 2006; Lerner, 2010; Gilson, 2003) particularly in more peripheral, smaller regional/national economies that lack critical mass of entrepreneurial innovation.

A unique feature of UKIIF is that the underlying funds are international – being based across Europe – and that it can invest globally in ventures, provided that at least the £150m UK government funding is invested in UK-based ventures. In this way UKIIF addresses many of Lerner's (2010) lessons by being private sector-led, contracting with leading international VCs with appropriate track records in early stage and Cleantech sector investing and not constrained to thin national/regional markets. Furthermore, as Owen et al (2019) observe, the scale of UKIIF underlying funds (at over £60m) is typically more than double that of ECFs, suggesting greater ability to provide large long horizon investment and follow-on funding. This addresses one of the key perceived failings of UK and European public and private VC, that they are too small when compared to their more successfully established US counterparts (Arundale, 2018; Deakins and Freel, 2012; Technopolis, 2009). Indeed, unlike the ECFs it was not constrained by EU state aid regulations to an initial £2m¹ investment cap per venture (and 10% of fund cap on investment per venture, typically amounting to circa £3m) and fund life is 12-15 years, compared to 10-12 years for ECFs, taking heed of the long horizon investment requirements of the Cleantech sectors.

Positioning UKIIF within the UK green financing escalator and ecosystem

It is helpful to use the concept of a green finance escalator (Owen et al, 2018) to position UKIIF within the evolving post GFC UK early stage green innovation financing landscape. The finance escalator adapts Berger and Udell's (1998) theory of decreasing opacity of early stage ventures as they develop, suggesting that as venture track record develops and information asymmetries

¹ Note that in January 2014 the EU state aid cap was raised to £5m, although many funds (e.g. UK Regional Investment Funds) subsequently remain constrained to £2m.

between ventures and potential financiers reduce, more and different private funding options and larger sums of money become available.

Table 1 demonstrates that the UK early stage innovative venture market contains a range of public interventions – including investor tax breaks, grants and equity funds, to address perceived early stage private equity funding gaps (where banks will not lend, North et al, 2013), particularly in the key, high risk valley of death area from proof of concept to early trading (up to 2 years). This is the so-called Macmillan gap (1931), currently thought to extend in the UK between £250k to £5m (Owen et al, 2019) and which may extend further for long horizon Cleantech R&D (BEIS, 2017).

North et al. (2013) highlight the importance of early stage funding complementarity to meet the financing demands of early stage ventures, suggesting that a fluent funding escalator requires effective bundling of different forms of finance, whilst Hopp (2010) recognised the value of syndication – notably between VCs (although increasingly occurring in recent times between seed VCs, accelerators, equity crowd funders and business angels (Baldock and Mason,2015; Owen et al, 2019) - to raise investment levels, share risks, introduce more diverse investor skills, facilitate longer distance equity investment (between some investors and ventures²) and open up international markets. This suggests the increasing opportunities for a more international approach to equity investing (Owen and Mason, 2019).

Pre-start	Start-up to early market development	Early growth and development	Later stage
Funds: <£50k	£50k - £2m+	$\pounds 2m - \pounds 5m$	£5m - £10m+
Timescale:	Start to 2 years trading	2-5 years trading	5-15 years trading
Internal funding (3Fs)	Internal funding (3Fs,	Re-invested profits	Re-invested profits
Credit cards	consultancy income)	Joint ventures and	Bank credit
Personal loans	Early stage business	licensing	Venture capital
i ciboliui loulis	angels/HNWIs	Bank credit (<mark>Loan</mark>	venture cupitur
'Proof of concept' grant	Business angel syndicates	guarantees)	Potential exit, trade sale,
funding (e.g. Innovate	Dusiness unger synateutes	guaranteesy	MBO/MBI, Initial Public
UK grants)	Crowd equity	Public and private VC	Offering (IPO)
Incubator support and	Accelerator finance	Peer-to-peer lending	Corporate/institutional
<mark>funding (e.g. CLT)</mark>	Start up Loans	Technology development	finance (private equity)
SFIS tay breaks	Start-up Loans	grants	Business Growth Fund
SEIS tax breaks	Technology development	grants	(BGF)
	grants (e.g. Innovate UK)	Help to Grow (H2G) loan	
	Innovation Accelerator	guarantees and mezzanine	Business Finance
	Programme (ILIK)	Innovation Loan Pilot	Partnership (BFP) funds
	r togramme (tote)	(IIIK)	Green Investment Bank
	Public and seed venture		(GIB)
	capital (VC) including ACF	EIS/VCTs	
	UKIIF		
	S/EIS tax breaks		

Table 1: Contemporary UK Green Innovation Finance Escalator for Early and Growth Stages

Source: Adapted from Nesta (2009a), Mason (2016), Baldock and Mason (2015) and Baldock et al. (2015): Yellow marks public sector/government funding intervention, Green where this has some expressed green innovation intent.

 $^{^{2}}$ E.g. Hopp (2010) and Baldock and Mason (2015) recognize that VC typically prefers to invest locally for easy hands-on access and support to venture teams, but that more distant national and international VC and angel investors will invest in syndication with trusted local lead investors who provide the hands-on support.

Table 2 demonstrates that whilst the UK has vast sums invested into early stage SME financing in the period of study, from 2009, a relatively small proportion has been directed at early stage Cleantechs. Indeed, the programmes with direct reference to early stage Cleantechs amount to under £600m, with the majority of this coming from UKIIF. To further contextualise the value of UKIIF, British Venture Capital Association (BVCA, 2017) data on members' investment into UK venture stage Cleantechs (which may include co-finance) shows a total of £95m into 151 ventures for the period 2014-2016.

Of course, the UK Department for Business, Energy and Industrial Strategy (BEIS, 2017a) has a £2.5b Clean Growth Strategy to support R&D through work with UK research councils, but this is centred around renewable energy and much of the financing is aimed at Catapult programmes for offshore and renewable energy solutions, with direct finance to SMEs being mainly directed at their energy efficiency. This forms a familiar pattern to that of the Green Investment Bank (GIB), which was established by the UK government to lever in private investment for infrastructure projects and successfully delivered £2.3b of investment between 2012 and GIB privatisation in 2017, but demonstrated little evidence of supporting early stage Cleantechs (Owen et al, 2018). A final addendum here is the signs of tensions between the British Business Bank's sector agnostic³ approach to early stage government VC programmes, as opposed to the more directly targeted schemes recently developing through BEIS (proposed Clean Growth Fund) and Innovate UK.

An important source of early stage Cleantech finance has also been Innovate UK grants, notably Smart awards which have traditionally supported proof of concept through to early commercialisation, with grants of up to £250,000 requiring match funding. However, the evaluation of Smart (SQW, 2015) revealed that only 9% of grant funding went to Cleantech, whilst Owen et al (2019) demonstrated that failure to secure match-funding and follow-on funding led to a systemic funding for failure. Subsequently, Cleantech has become a priority sector focus through the early proof of concept £50m Investment Accelerator Programme (IAP) piloting matching grants and seed VC for low carbon infrastructure projects of up to £150k and new Smart grants which offer tranches of funding of up to £3m over three years.

Overall, what we see is a proliferation of government public-private co-financing, creating a complex map of finance, much of which is not clearly linked with Cleantech objectives, and which is increasingly difficult for interested parties to navigate and achieve optimal investment outcomes.

Type of infunce		D 1	
		Body	
Tax			
R&D Tax Credits	SME scheme (<500 employees & £100m sales) for corporate tax rebate or cashback	HMRC	
Seed Enterprise Investment Scheme	Start-up (<25 employees) investor tax relief, up to 50%: investor cap £100k, firm cap £150k pa	HMRC	
Enterprise Investment Scheme (EIS)	Early stage (<500 employees) tax relief, up to 30%: investor cap £1m, firm cap £5m pa	HMRC	
Venture Capital Trusts (VCTs) Grants	VCs funded through EIS investing in SMEs (<£15m assets)	BBB	
Innovate UK	Various match fund grants: Knowledge Transfer Partnerships (<£80k), SMART awards/grants (<£3m), Launchpad and Feasibility (<£400k),	UKRI Innovate UK	

 Table 2: Positioning Key Sources of Public SME Green Innovation Finance (marked in green) in the UK, 2009-2019

 Type of finance
 Funding Description

³ Discussion with BBB (31/05/2019) indicated policy is market led by private VC.

Type of finance	Funding Description	Administering Body
	Collaboration and Catapult (£10m+), including £165m annual funding from 2016 for new forms of finance (e.g. repayable loans) Investment Accelerator Programme (£50m) matching grants with seed VCs	
European grants Regional Growth	Horizon2020 and Eurostars grants (up to E300k at 50% match-funding) Loans and grants through the £2.6bn RGF for English Local Enterprise	EU BEIS
Energy Entrepreneurs	£150m Energy Entrepreneurs Fund (EEF) and £10m Energy Innovation Fund	BEIS
Fund Debt/loans		
Enterprise Finance Guarantee (FEG)	75% government guarantee on loans, overdrafts and invoice finance capped at $\pounds 1m$	HMRC
Business Finance Partnership (BFP)	£100m govt co-invest in invoice, supply chain, asset, mezzanine (growth loans) and P2P (e.g. Funding Circle, Zona)	BBB
Green Investment Bank (GIB)	Funding complex projects tackling greenhouse gases, natural resource efficiency, natural environment, biodiversity, environmental sustainability. £2.3bn invested 2012-2017 including £2m Green Energy Savings Fund for	GIB
Big Society Capital VC/equity	£488.2m invested into charities and social financing organisations by June 2016	Cabinet Office
Angel Co- investment Fund (ACF)	£100m evergreen business angel co-invest fund, from 2011	BBB
Enterprise Capital Funds (ECF)	£1.2b co-investment VC funds, from 2006	BBB
UK Innovation Investment Fund	£330m UK government and European Investment Fund - fund of funds for health and Cleantechs, from 2009	BBB
Regional Venture Capital Funds (RVCFs) UK Government Regional Investment Funds	North West (£190m), Yorkshire & Humber (£90m+), North East (£125m), East of England (£20.5m Low Carbon Innovation Fund) - range of loans and equity (£2m cap) From 2016 VC and loan funds: Northern Powerhouse Investment Fund (NPIF) - £400m for North West, North East and Yorkshire & Humber; Midlands Engine Investment Fund (MEIF) - £250m for Midlands region; Cornwall & Isles of Scilly Investment Fund (CoSIF) - £40m	EU JEREMIE/ERDF, BBB BBB
Devolved UK Governments	Scottish (£185m), Northern Ireland (£160m),,Wales (£150m+) VC and loan funds (£2m cap)	ERDF/JEREMIE SIB/DETI/FW
Universities Innovation Fund	£160m (Higher Education Innovation Fund) and University Enterprise Zone Pilots (£15m), Knowledge Transfer Networks	Research England
Export finance UK Export Finance	£1.5b scheme 2012-15 to assist export trade credit arrangements	UK Export Finance

Note: BBB - British Business Bank; BEIS – Dept for Business, Energy and Industrial Strategy; UKRI – UK Research and Innovation; RE – Research England; HMRC - Her Majesty's Revenue and Customs; DETI – Northern Ireland Dept for Enterprise Trade and Industry; ERDF – European Regional Development Funds; FW – Finance Wales; JEREMIE - Joint European Resources for Micro to medium Enterprises SIB – Scottish Investment Bank. Sources: Adapted from Owen et al (2019); British Business Bank; BEIS (2017a)

Adopting a resource complementarity lens

The research question addressed in this paper is how a Green New Deal can generate improved external financing for early stage innovative Cleantechs. The literature establishes that substantive public-private co-financing with VC can address finance gaps for longer-horizon, capital intensive R&D and early commercialisation of Cleantech ventures. Taking the case of the UK's market-leading approach with UKIIF we are able to critically assess what works well, or less well, and to take on board lessons and address the areas for further improvement. To achieve this we construct a resource complementarity research framework (Chi and Levitas, 2007). This accounts for the interdependencies which occur between innovative (Cleantech) businesses and institutions, within a societal framework. From a resource-based perspective the balance between such interdependences is critical in how the value of a Cleantech innovation is assessed and supported financially (the purpose of this paper).

This interdependence – resource-complementarity framework operates on two levels (Table 3): first, the role of co-financed VC programmes within the Cleantech strategic policy sphere; second how the co-financed VC process can operate better in addressing the financing needs of Cleantechs. The first addresses the need for clearer objectivity in policy to target and assist Cleantech (which is evidently lacking from the current green finance escalator), whilst the second relates to programme operation. We argue that both can benefit from lowering transaction costs to improve capital market functionality. **Mindful of tier 1, in this paper we focus on tier 2 and the operation of UKIIF.**

Public-private partnerships (PPP) based on a Green New Deal (GND) can affect the standardisation and institutionalisation of the search, contracting and monitoring phases of Cleantech hybrid ventures (Lehner, 2016; Lehner and Nicholls, 2014; Lehner, Harrer and Quast, 2018). Standardisation may also be a catalyst for subsequent funding from private institutional investors as a successful PPP enhances the legitimacy of both involved parties and lowers the transaction costs in order to create a functioning market (ibid). With this we contribute to the ongoing research and discussion on hybrid business models and public-private collaborations (Achard and Di Berardino, 2018).

Table 3: Two-tier Framework for UK Cleantech Policy and Programme Delivery

Tier one players: the macro policy environment strategy and ecosystem concord

- Transnational public bodies (EU, World Bank)
- Government departments (national, regional and local): UK key players BEIS, Government Offices, Devolved Governments (NI, Scotland, Wales), Regional/City governments and Mayors (e.g. London and Manchester)
- Specialist economic development and financing / green financing non departmental bodies: British Business Bank, Innovate UK,
- Private financing bodies (financial intermediaries): Green Investment Bank, Carbon Limiting Technologies, Carbon Trust
- Business support and finance intermediaries (private and public sponsored advisory services): Local Enterprise Partnerships (38 in England), Local Enterprise Hubs

Tier Two players: the micro delivery process of policy programmes

- Government VC co-financing programmes (hybrids): UKIIF, IAP
- Private financing bodies (financial intermediaries); Banks (Challenger, Regional) VCs Angels PE, CF, P2P, Asset finance
- Business support and finance intermediaries (private and public sponsored advisory services): Accountants
- R&D centres/universities: Catapult Centres, University R&D commercial spin-out centres (e.g. Maxwell Centre, Cambridge University)
- Cleantech early stage Ventures: ventures trading less than 2 years undertaking R&D to reduce CO2 emissions

Focusing on a resource complementary lens and drawing on the literature of hybrid VC (Murray, 2007; Lerner, 2010; Munari and Toschi, 2014; Cumming and Johan, 2016; Owen et al, 2019 etc) we are able to construct a series of hypotheses for the different types of resources that will most prominently benefit from this type of public – private partnership (PPP) which transacts between government, Private VC and entrepreneur/venture. Whilst table 4 is by no means exhaustive, it nevertheless presents key elements derived from leading/current literature where we can hypothesise PPP will enhance operation and delivery of:

- (1) H1 Funding: improved early stage funding escalator and availability of early stage VC
- (2) H2- Management: Enhanced VC management, investment decision-making and Entrepreneurial/venture team management
- (3) H3 Risk-sharing: enhanced risk sharing environment will benefit VC investors and ventures
- (4) H4 Networking: enhanced networking will assist viable venture pipeline and increased VC and complementary investment including through syndication
- (5) H5 Skills Development: will lead to improved VC investment and venture performance
- (6) H6 Cleantech innovation / economic growth: enhanced cleantech innovation and economic growth through adequate ad effective VC financing market.

Table 4: Resource Complementarity Framework for UKIIF

Government	Private VC	Entrepreneur					
1.Funding							
Substantial government funds, accountable to HMT, provide value for money, effective flowing finance escalator (Baldock, 2016)	Insufficient funds and appetite in early stage investing, increased scale can increase institutional investment (Technopolis, 2009)	Asymmetric barriers to early stage risk finance (Carpenter & Petersen, 2002)					
2.Management							
Government bureaucrats lack skills to invest directly in early stage ventures (Lerner, 2002)	Skills, experience to address agency failures (Lerner, 2010; Gompers, 2009)	Commercial and financial management skills (Baldock & Mason, 2015)					
3.Risk sharing							
Improved likelihood of returns investing with private VC track record and scale-up (Lerner, 2010)	Government funding safety net, increases funding scale and likelihood of pareto return (Markowitz, 1952); VC syndication linkages (Hopp, 2010)	Cede share of ownership, but obtains certification of VC finance, increasing likelihood of further follow-on finance (Cumming and Johan, 2016)					
4.Networking							
Government policies promote entrepreneurial ecosystem networks for SME growth /sustainability (Brown & Mason, 2014 & 2018)	Need pipeline of viable ventures (Mason & Brown, 2013). Provide access to complementary finance, skills and commercial markets (Owen et al, 2019)	Require access to management skills/training, supply chain, commercial markets (Owen et al, 2019a)					
5. Skills Development							
Skills development increases productivity and competitiveness (Gambin et al, 2009) and virtuous cycle of development (Lerner, 2010)	Early stage VC skills and sector specialisms development (Gompers, 2009) and recycling of new VC manager talent (Lerner, 2010)	Management team development through recruitment, training, mentoring and oversight (Baldock, 2016; Munari and Toschi, 2014)					
6. Cleantech Innovation / economic growth							
Policy promotes cleantech innovation for climate change (Owen et al, 2018), raising GDP through vibrant SME jobs, sales, export growth (Lerner, 2010)	VC seek early stage innovation with risk/reward potential for financial return (Burchardt et al, 2016), create venture growth, scale-up (Baldock, 2016)	Create innovative disruptive cleantech R&D start-ups (Owen et al, 2019), creating jobs, sales and competitive exporting SMEs (Owen et al, 2019)					

Table 4 clearly demonstrates the overall key to addressing climate change in the interdependency of the three main actors studied. Government policy framework can fund and support early stage Cleantech, whilst early stage VC can reduce agency failures (moral hazards and adverse selection) using key skills to select and nurture portfolio ventures, with the ventures developing game changing disruptive green innovations to lower carbon emissions.

Methodology

We adopt a mixed methods approach (Creswell, 2003) in order to triangulate primary evidence drawn from in-depth longitudinal case studies of seven UKIIF assisted Cleantechs (in renewable energy, recycling and advanced manufacturing) with other relevant contextual sources in the 2010-15 period (Tables 5 and 6).

UK	GVC	Fund	F/F,	or	extended	Interview dates	5	Longitudinal	follow-up	
recipie	ents/study		telepho	ne inte	erview			interviews, Summer 2015		
Innov	ative SME	, /	30			Feb 2015		-		
journe	ey									
ECF (Enterprise	2	12			Feb 2010 &		7		
Capita	al Funds)					Feb 2014				
ACF (Angel Co	-	15			Feb 2014		-		
invest	ment fund)								
UKIII	F (UK		16			Feb 2012 &		14 - including	7 surviving	
Innov	ation					May 2013		Cleantech case	e studies	
Invest	tment Fun	d)								
Techn	ology Bas	ed	49			Feb 2011		1		
Small	Firms (TE	BSFs)								
Aspire			4			Feb 2010		2		
						May 2013				
EGF (Early Gro	wth	9			Feb 2010		-		
Fund)										
Total			133			(2010-15)		24		
			1.55	<u></u>		(2010-13)		24	<u> </u>	

Table 5: Breakdown of business interviews (2010-15)

Source: interviews undertaken by CATI (computer assisted telephone interviews), face to face on business premises or by extended telephone interview

In total the CEOs of 133 young potential high growth and innovative UK ventures were surveyed either face-to-face, or by extended telephone interview by the authors. This includes six UK early stage venture investment programmes overseen by the British Business Bank and two studies (technology based small firms, North et al, 2013; innovative firm's journey to finance, BEIS, 2017) specifically examining early stage innovative venture finance. Ventures were purposively selected in order to provide insight into a range of different early stage sectors, including Cleantech. Overall, they are characterised as follows; ventures are typically under 2 years established and pre-trading when first interviewed, micro-businesses with less than 10 employees and undertaking innovative market leading R&D in science, engineering and advanced manufacturing, or digitech service sectors (e.g. App and Fintech). One quarter of the base sample and 63% of the longitudinal sample are Cleantech, represented by renewable energy (e.g. solar PV, biomass), digitech (e.g. smart energy metres), recycling (e.g. water, rubber/plastics and aluminium) and advanced engineering (e.g. digital efficiency controls, robotics, electric vehicles, lightweight plastics and battery storage solutions).

All data was systematically collected by the authors using integrated interviews (Owen et al, 2019), typically taking 1.5 hours, covering business profile, nature and stage of innovation, external financing requirements, finance search times and costs, and extent of impact of VC/equity finance and non-financial assistance on subsequent business performance, including follow-on financing and investment exit information. The seven Cleantech case studies were selected as exemplar cases of UKIIF assistance and longitudinally tracked through 3 surveys between 2012-15, with supporting information coming from VC investors, programme management data and company websites.

Additional triangulatory qualitative contextual data was provided by 58 semi-structured telephone and face-to-face interviews with investors (business angels and VC fund managers), intermediaries and industry experts (Table 6, see Owen et al, 2019a; Baldock, 2016; Baldock and Mason, 2015; Baldock and North, 2015), with particular consideration given to the Cleantech UK early stage government co-financing market, post GFC.

Table 6 Breakdown of fund manager and investor interviews

Fund	Location/ time of interview
Enterprise Capital Funds (ECFs):	Spring 2014
The Catapult Growth Fund (ECF)	Leicester
IQ Capital Fund	Cambridge
Oxford Technology	Oxford
Seraphim Capital Fund	London
Sustainable Technology Partnership	London
Amadeus and Angels Seed Fund	Cambridge
Dawn	London
MMC	London
UK Innovation Investment Fund (UKIIF):	Spring 2012
Hermes GPE Environmental Innovation Fund	Fund of funds manager, London
European Investment Fund UK Future Technologies Funds	Fund of funds manager, Luxembourg
Underlying funds:	
Zouk Cleantech II	Hermes fund, London
Scottish Equity Partners (SEP) Environmental Energies	Hermes fund, Glasgow and London
WHEB Ventures	Hermes fund, London
DFJ Esprit	EIF UKFTF fund, London and Cambridge
Advent Life Sciences	EIF UKFTF fund, London
Gilde Healthcare III	EIF UKFTF fund, Utrecht and Cambridge USA
Angel Co-Investment Fund (ACF):	Spring 2014
19 Lead business angel investors from investment syndicates,	10 angel groups, 5 in London, 2 in South East, 1
of which 16 were successful ACF applicants and were	East Midlands, 1 in South West England and 1 in
unsuccessful	Scotland
3 Investment Committee (IC) members - experienced angel or	UK-wide representation
institutional investors	
Alternative Investors and Experts	Spring 2010 to Spring 2014
13 alternative private investors: Private VCs (4) and seed VCs	8 in London, 1 North East, 1 Yorkshire and
(3), Venture Capital Trusts (3) Angel Capital Groups (3)	Humber, 1 Scotland
6 industry stakeholders and experts: British Venture Capital	,
Association (BVCA), UK Business Angels Association	UK-wide coverage
(UKBAA), Angel News, St John's Innovation Centre,	6
European Investment Fund (EIF), Professor Dylan Jones-	
Evans (Finance Wales reviewer)	

Source: CEEDR studies (2010-15)

Data collected provides a rich source of qualitative information on the financing requirements of early stage Cleantechs and the extent to which public-private co-financing programmes are meeting these needs. The data gave rise to the possibility to consider a-posteriori how a GND could establish a more systematic and institutionalised approach to assisting Cleantech early stage investing, leading to lower corresponding transaction costs through the standardizing of the search, contracting and monitoring phases (Lubell et al, 2017).

In addition, we apply a resource-complementarity lens to understand how a public-private collaboration (Shi, Chong, Liu and Ye, 2016) based on a GND can result in effective, optimized value-creating solutions towards the climate goals that far out-weigh the involved costs (De

Schepper, Haezendonck, and Dooms, 2015). In this, we understand resource-complementarity lens (Sayeed and Onetti, 2018; Soda and Furlotti, 2017) as the degree to which the joint use of distinct sets of resources (e.g. the innovative mindset and skills and the financing power) produces a higher total return than the sum of returns that could be achieved if the resources were only utilized independently.

Findings

Adopting our resource complementarity framework (6 hypothesised elements), we present findings from the 7 case studies' perspectives to examine how in practice the UKIIF has addressed their Cleantech early stage financing requirements. We take into account transaction cost and impact measures to assess what has worked well or less well, before proceeding to a discussion on the implications for policy.

Funding, bundling and risk-sharing (H1, H3)

Early stage Cleantechs require substantial amounts of external financing. Table 7 demonstrates that the 7 case study ventures were seeking either first round, or early round finance for R&D and early commercialisation. They were all independent micro businesses and pre-trading at the time of their search for external finance. The size of the funding rounds which UKIIF contributed ranged from £300k to £13m, but were typically over £1m (median £2.1m). UKIIF typically provided at least half of the funding round (median £2m), frequently syndicating with other VCs and notably bundling with other funding providers to generate larger rounds (Baldock and Mason, 2015), providing evidence of syndication with overseas VCs to share early stage investment risk (Hopp, 2010), scale-up investment (Baldock, 2016) and facilitate access to international markets (Owen et al, 2019). Case D's CEO commented "SEP brought in a US VC at the commercialisation stage to introduce us to the US market."

It is worth reflecting that the scale and size of funding for these UKIIF cases is considerably higher than for comparable, but more generic UK ECF funds (median £750k ECF investment), with considerably more syndication taking place to achieve the financing round requirements (Owen et al, 2019). Crucially, UKIIF was a lever to other private investment with one CEO mentioning that "*the UKIIF SEP fund encouraged Zouk's later stage VC fund to invest £3m earlier in the business than otherwise possible*." The high share requirements of the UKIIF VCs is indicative of the high risk and early stage investments taking place (median 25%), with follow-on funding in case D leading to a 75% share. Notably, none of the interviewed CEOs complained about the proportion of ownership share taken by UKIIF VCs, supporting Gompers et al (2010) proposition that experienced early stage VCs offer both vital finance and also the required non-financial sector, management and networking skills.

Search times, negotiations and costs

Government VC programme assessments often focus on additionality and leverage, without sufficiently considering transaction costs (Baldock, 2016; BEIS, 2017; Owen at al, 2019). Most of the 7 case studies experienced relatively short search times of under 3 months. This was mainly serendipitous, highlighting the timeliness of UKIIF and existing VC connections and the excellent fit of the programme to the ventures' funding requirements. The UKIIF funds were described by CEOs as "*the fund of choice*" and "*the most experienced and proven fund*" in the sector. This might

suggest that the level of UKIIF additionality is low (Owen et al, 2019, indicate that it is lower than for ECFs), since the funds were selected quickly and sometimes over other competing private VC funds. However, this narrow view overlooks important contributary factors. First, the CEOs perceive that selecting the most suitable and experienced fund, and fund managers to work with, is more important than lower cost/share options and second that the speed and scale of access afforded by UKIIF in most cases led to more rapid R&D and commercialisation and increased the scale of activity. Additionally, from the fund managers' perspectives, UKIIF was reported as facilitating more rapid and larger-scale fund closes⁴, suggesting that the case study investment – particularly in terms of their early stage and large size of investments would not have been available without the government programme. This is underlined by the more severe experiences of the cases where funding was required before UKIIF was established and highlighted by case E.

Case E: Three-year round 1 funding search (2008-2010)

Seeking a package of equity and bank finance to fund a management buy-in to an existing UKbased business and undertake internal R&D to commercialise 'PET' lightweight plastics technology. Total project cost was £13m, of which UKIIF provided £2.6m for a 21% equity share. The funding search process took three years, hampered by lack of bank and early stage equity finance available in the UK and Europe. UKIIF was found using an accountant VC finder, and was not problematic, but took nine months to negotiate.

"Getting a UKIIF VC on board as lead investor was a catalyst to getting further VC investment and European bank finance... critical to business development, transforming a business that was closing down into a global market leader."

The negotiation time to undertake due diligence and agree terms and conditions was typically around 3 months, but could take up to nine months for more complex transactions such as case E which required a syndication bundle of additional bank and VC finance to undertake a manufacturing factory buy-in, or case A, described by the CEO as "*involving complex IP rights for a spin-out from Cambridge Innovation*."

Transaction costs for VC are high, typically amounting to 5% of ECF deals (BIS, 2010) and often an even higher proportion of smaller value deals. These can involve the search costs for accountant and VC finder services and direct costs for consultancy and accounting services to undertake due diligence and legal fees for negotiating and agreeing contractual terms and conditions, with applicants frequently responsible for both parties' legal costs or VC administration fees. Additionally, there is the management search time and opportunity costs associated with lengthy searches and compliance work for applications which take time out from other core business development activities (BEIS, 2017); which could typically require a day or two a week of senior management time for several months (BIS, 2010). Despite the high costs of obtaining VC, none of the UKIIF case CEOs complained. They stated that costs were at the expected market rate, which was particularly acceptable if search times were reduced and for case F which used UKIIF syndicated with the Scottish Co-investment Fund (SCIF), it was "...less expensive because it saved on the time and cost of finding and agreeing investment terms with individual angels."

⁴ This was particularly the case in the immediate aftermath of the GFC (2009-10) when the focus UKIIF underlying funds were being established (BIS, 2012).

Company/ Established	UKIIF Fund	Activity/green	Round/	UKIIF/	Leverage	Search	Time
Established	1 unu		amount	Silare		THIC	INCg
A (2010)	DFJ	3D Radar, windfarms near airports	1: £2.1m	£1.6m (50%)	0%	<3 months	6-9 months
B (2010)	Zouk/ SEP	Energy, smart metres, battery R&D	1: £6m	£3m (50%)	50%	<3 months	<3 months
C (2009)	SEP	Renewable grid connection	1: £1m	£0.5m (24%)	50%	<3 months	<3 months
D (2005)	SEP	Domestic water recycling	3: £2m	£2m (25%/75%)	0%	<3 months	<3 months
E (2009)	WHEB	Lighter PET plastic	1: £13m	£2.6m (21%)	500%	3 years	9 months
F (2010)	SEP	Drones energy inspection	1: £300k	£150k (25%)	25% (SCIF)	1 year	<3 months
G (2009)	WHEB	Domestic Smart energy controls	1: £6.5m	£3m (27%)	0%	<3 months	5 months

Table 7: Case Study Funding Requirements and Search Details

Management, networking and skills development (H2, H4, H5)

Early stage VC perform a vital role in nurturing their portfolio business development. This helps to overcome agency problems, particularly moral hazard where early stage venture management deficiencies can lead to sub-par performance (Baldock and Mason, 2015). It was widely evidenced by the case studies (via CEOs and UKIIF fund managers) that VC were in regular, at least monthly, contact, acted as board advisors, appointed non-executive directors and recruited specialist managers to strengthen the management team, notably financial and commercial directors.

A distinction was made here between technical development, which is typically the domain of the founding venture team and which VCs are usually less able to assist and broader commercial marketing expertise. Several interviewed CEOs mentioned the importance of commercial rigour, developing tight financial management controls and regular monthly reporting, and this was an important contributory factor in all of the CEOs feeling more confident about raising follow-on finance. These processes established various positive outcomes. For example, case A recruited a Chief Finance Officer, whilst for case E the VC was "...massively helpful, providing a strong steer on corporate finance and governance and constructive input at all times. They have been particularly helpful in finding new overseas customers." It is notable that this company now has board representatives from Russia and South Asia which have proved huge growth markets in recent years.

Another aspect of skills development, unrecognised in some studies, is early stage fund management. Whilst some funds like DFJ (Draper Esprit) have considerable track records in early stage finance, others such as Zouk have been encouraged to enter earlier stage venture financing, learning by operating with shared syndication risk. Lerner (2010) recognised the value of this process (e.g. Israel's Yozma funds and New Zealand's Venture Investment Fund) whereby experienced fund managers demonstration affects can encourage others into the market, but also upskill other local fund managers (Baldock, 2016).

Cleantech economic impacts and global climate change game changers (H6)

Since receiving early stage Cleantech funding from UKIIF the 7 case study ventures have performed well. Collectively, they have generated over 840 FTE jobs since receiving UKIIF investment and moved from pre-trading R&D through to commercialisation and in 2018 £106.7m sales (median £6m). Their impacts on reducing CO2 emissions have also been considerable. Whilst it is worth noting that there is no consistent measurement in place for these businesses, they demonstrate enormous energy savings through the operations of smart energy controls (e.g. case G's 23% savings on domestic energy use, and case B's smart controls for business and high end domestic users), windfarm development and maintenance (cases A, B and F) and advanced 'PET' lightweight and fully recyclable plastics which have transformed the drinks/liquid foods distribution industry, reducing plastic container materials by 90%, reducing container weight by 85% and eliminating empty container return costs through 100% recycling technology. Cases A and F are intriguing examples of Cleantech whose impacts are more difficult to quantify. Case A's holographic 3D radar should enable highly efficient large-scale wind turbines to be located on and nearby defence/military sector land and close to airports without affecting flight control (unlike conventional radar systems); it is undergoing trials with the UK Ministry of Defence (MoD) in 2018-19. Case F's inspection drones have been used for conventional oil and gas industry rig inspection and for renewable energy windfarms.

All UKIIF cases reached commercialisation and have delivered cleantech products and services, although case D's domestic grey water recycling product failed to gain a market foothold in 2013-14 and the company subsequently closed. The remaining 6 cases have all successfully exited UKIIF within a nine year period, producing substantial returns to the UKIIF underlying funds and the UK tax payer. Exits have included a trade sale and an IPO which raised US\$20m for case G in 2014. The majority of cases have refinanced and exited UKIIF, through Private Equity (PE) and joint venture arrangements and, for case C, through debt refinancing in 2015. It is notable that these businesses are highly innovative market leaders that service global markets. For example, case G has trialled smartphone controls on domestic energy use in Denmark and New Zealand, whilst case E's joint venture provided access to the vast Indian sub-continent market, with the company also very active across Eastern Europe, Russia and North America. Crucially, the surviving businesses retain considerable UK presence, notably R&D activity, and with several former CEOs involved in investing in and managing new Cleantech start-ups, thus supporting Lerner's (2010) and Owen and Masons' (2019) belief in the local recycling impact of public VC.

Co.	Jobs Created	Annual Sales	UKIIF Status/Exit year	Green Impacts / Mega Watt (MW) estimations
A	50	£10m	Acquired, 2017	Enabling giant wind power on MoD land, tests from late 2018
В	480	£75m	Private Equity, 2014	480MW renewable generation; 380MW battery storage by 2020
С	125	£6m	Debt refinance, 2015	250MW in Grid Renewables Input
D	0 (10*)	0	launched domestic grey water recycling in 2013	Grey water recycling product marketed, but failed in 2014

Table 8: Case Study Status, June, 2019

Co.	Jobs Created	Annual Sales	UKIIF Status/Exit year	Green Impacts / Mega Watt (MW) estimations
Е	75	6m	Joint venture buyout, 2014	PET containers reduce 90% weight, 85% material, 100% recyclable
F	60	3.9m	Private Equity, 2019	Drone inspection of windfarms
G	50	5.8m	Raised \$20m in IPO 2014	Smartphone control of domestic heating with 23% energy saving
Total	840	£106.7m	6 Exits, 1 closure	All made impacts, 6 ongoing

Note: * 10 staff at time of closure

Case B has focused on the UK market and, whilst becoming a global top 100 Cleantech, it has created nearly 500 FTE jobs in the UK and generated annual sales turnover of over £75m in the decade since start-up. This case provides a compelling example of how UK government VC funds can assist hugely impactful Cleantech outliers.

Case B: A Global 100 Cleantech Growth company

The first UK company offering total energy solutions, following North American and German models, focusing on the larger commercial business market, provide consultancy and advice on a full range of renewable energies. Established 2010 with UKIIF providing 50% of start-up capital, the company offers initial concept, R&D, monitoring, financing and all technical solutions and have patented their own energy monitoring devices, developed in-house through strategic company acquisition and key staff recruitment.

Key Green Impacts

The company's products and services over the last decade have reduced reducing UK carbon emissions by an estimated 480MW (2017) of renewable energy generated through solar and wind. The company's dual strategy of short-term digitech smart control and energy consultancy and long-term battery storage R&D enabled rapid Global 100 scale Cleantech sales growth and early exit from UKIIF via scale-up PE investment in 2014. A recent deal with a major utility company has put the at the forefront of battery storage and on track to meet 380MW by 2020.

"As the UK market leader in energy storage, we command the industry's most accurate revenue modelling tool and are on track to exceed 380MW of battery storage by 2020. We have constructed over 100 solar farms, and our O&M service, 'XYZMeter', monitors over 21,500 sites. We were first in the UK to achieve subsidy-free solar, first to introduce utility-scale energy storage, and first to co-locate energy storage with existing solar sites. To date [March, 2019], our technologies deployed and managed are generating over 1GW of renewable energy."

UKIIF additionality

Whilst the CEO acknowledges other VC funding was available, UKIIF's fund was their first choice: "I trusted them as a leading VC in the cleantech field. The fund manager knows and understands the sector and provides an excellent sounding board for day to day management decisions. Their industry connections and commercial rigour helped prepare the company for scale-up PE and built the platform for our later growth."
Discussion

Summarising the findings through a complementarity lens, all of the hypothesised benefits have accrued to some extent, but the findings also reveal important lessons for future programmes, particularly around evaluating green impacts (Owen et al, 2018) and the potential benefits of paying greater attention to transaction costs.

Table 9 presents a synthesis of key transaction cost factors where complementarity between government policy and programmes, private VC and cleantech entrepreneurs has led to demonstrable benefits for all parties. This synthesis formed from our complementarity framework findings presents three key themes.

Funding, bunding and risk sharing are all enhanced by UKIIF with key reductions in transaction costs for government relating to lighter touch oversight management required by the British Business Bank (e.g. compared to in-house management of the Angel Co-investment Fund), whilst the increased scale of activity of the UKIIF fund of funds enables direct investment from institutional pension funds (£80m was raised from two pension funds) and the oversight of an experienced Cleantech specialist fund with knowledge of over 60 VC funds:

"Hermes contributes ... by referencing providers (i.e. underlying funds) and introducing new investors to funds that Hermes are working on for investments. Also advising these funds on how to improve due diligence materials." Hermes fund of funds manager.

The scale up of activity possible through a fund of funds mechanism and flexibility of funding with potential top up funding to underlying funds provides greater opportunity to reach Markowitz (1952) portfolio investments numbers and to follow-on fund to generate optimal exit value. Furthermore, operation across international boundaries through international fund linkages across Europe and US reduces the adverse selection risk of thin markets. For the cleantech entrepreneurs the search time savings and single deal structure of UKIIF investments potentially saves costs, although there is always room for improvement on due diligence and legal negotiation fees (Baldock and North, 2015).

Management, networking and skills are all enhanced by UKIIF. Government can focus greater attention on wider policies necessary to generate the pipeline of early stage Cleantechs and support services to facilitate VC take-up. One aspect is the promotion of programmes equity finance trade bodies such as the British Venture Capital Association (BVCA) and UK Business Angels Association (UKBAA) and intermediary advisory services such as accountants through the Institute of Chartered Accountants in England and Wales (ICAEW). As Baldock and North (2015) suggest, more could be done to promote programmes, particularly in more remote UK regions from London. The additional umbrella tier of private VC creates additional management costs, but these are potentially off-set by the superior investment performance and quality of oversight provided by leading early stage Cleantech fund managers. As the Hermes fund manager suggested, they can diffuse skills downline to the underlying funds, contribution to Lerner's (2010) concepts of upskilling of VC fund managers. They also generate between the funds a considerable global network, increasing opportunities for syndication, further investment and international market entry. Cleantech entrepreneurs are shown to benefit considerably from the superior finance, market and network knowledge of top tier VCs – reducing search times and transaction costs these matters.

Government	Private VC	Cleantech Entrepreneur
1.Funding, Bundling and Risk Sha	ring	
Government funds encourage earlier stage, scaled-up private VC, shared risk, increased investments and greater likelihood of returns Government promotes funds, working with business support and private finance trade bodies	Larger VC funds more attractive to institutional investors, leverages more private funding VC networks increased syndication, sharing risks, increasing investment, opening up international markets (e.g. through international syndication)	Increased early investment facilitates faster R&D, follow-on funds and commercialisation Reduced search times/costs, competitive set-up fees less than fragmented (angel) investment costs
2.Management, Networking and S	Skills	
Government pays VC to manage funds, minimises government administrative costs of BBB Complementary policies promote entrepreneurial pipeline, entrepreneurial support infrastructure	Fund of Funds extra tier of costs, gaining expert fund selection, fund management and private financial leverage VCs network/syndicate to enhance their fund manager and portfolio management skills	Expert fund managers deliver commercial rigour and acumen in the boardroom, recruit and appoint key NEDs, managers, assist with market development and prepare and find further private finance, easing flow of finance escalator.
3. Cleantech Innovation, Climate	Change Mitigation and Economic C	Growth
Policies support early stage Cleantech to lower CO2, creating sustainable green economy and global impacts Optimal firm exists along with jobs, sales, multipliers, international trade provide tax revenues and 'green kudos'	Specialist VC funds select most viable Cleantechs, financing and managing their commercialisation and scale-up. Successful early stage co-finance VCs demonstrate attraction of specialist market for private VC and spin-out new fund managers	A greater number of Cleantechs with disruptive green innovation reach commercialisation and international scale-up quicker, impacting on reducing CO2 and increasing renewable energy. Sustainable firms, job creation, supply chain multipliers, global sales.
		Recycling of entrepreneurial wealth and IP – 'virtuous cycle'

Table 9: Positive Impacts on Transaction Costs Development Framework for Cleantech Financing

Cleantech innovation, climate change mitigation and economic growth all benefit from reduced transaction costs through the complementarity approach of UKIIF. Successful optimised investment exits generate returns of funds as well as a range of economic multipliers which are shown to benefit the UK exchequer, notably because these businesses are keeping a major presence in the UK, even after international trade sales and PE investments. Furthermore, there is evidence of Lerner's (2010) recycling and the virtuous cycle of entrepreneurial VC ecosystem development (Owen and Mason, 2019) through fund manager development, increased early stage Cleantech VC financing and serial entrepreneurship and investment activities of former UKIIF portfolio venture managers.

Conclusions

Whilst the study focuses on the UKIIF UK government and EIF co-finance VC programme, it has demonstrated through 7 case studies how such a targeted early stage Cleantech fund can make a difference. This has occurred both in terms of transition to a lower carbon economy tackling climate change through a variety of innovations impacting on key Cleantech sectors relating to renewable energy, advanced manufacturing and recycling, and also wider economic benefits. These relate to sustainable ventures, job generation and financial return to the economy through UKIIF investment returns to the British Business Bank within a nine year period, substantial sales revenues and tax returns to the exchequer, supply chain multipliers and recycling of CEO investment and time through serial entrepreneurship. An important observation is the retention of investee firms and employment in the UK, even in cases where there has been considerable international investment, although a cautionary tale surrounds case E which has outsourced manufacturing and several hundred jobs to Eastern Europe. This appears particularly pertinent, given the swathe of multinational company job losses taking place during the Brexit period.

References

Achard, P. O., & Di Berardino, A. (2018). Public Private Partnerships: Strategic Assets and Managerial Models.
 Avnimelech G, and Teubal M (2006) Creating Venture Capital Industries That Co-evolve with High Tech: Insights from an Extended Industry Life Cycle Perspective of the Israeli Experience. Research Policy 35(10):1477-1498

Bak, C. (2017) Generating Growth from Innovating the Low Carbon Economy: Exploring Safeguards in Finance and Regulation. Centre for International Governance Innovation (CIGI) Working Paper 117, January, Waterloo, Canada. Available from:

- https://www.cigionline.org/sites/default/files/documents/Paper%20no.117_WEB.pdf
- Baldock, R. (2016) An assessment of the Business Impact of the UK's Enterprise Capital Funds. Environment and Planning C: Government Policy. DOI: 10.1177/0263774X15625995
- BEIS (2017) The Innovative Firm's Journey to Finance. UK Department for Business, Energy and Industrial Strategy, Research Paper No.23, HMSO London
- Burchardt, J., Hommel, U., Kamuriwo, D. & Billitteri, C. (2016). Venture capital contracting in theory and practice: Implications for entrepreneurship research. Entrepreneurship: Theory and Practice, 40(1), pp. 25-48
- Chi T and Levitas E (2007) Resource complementarity, institutional compatibility, and some ensuing methodological issues in testing the resource-based view. In DJ Ketchen Jr and Bergh D Research Methodology in Strategy and Management (19-36), Elsevier, Oxford
- Creswell J.W. (2003) Qualitative, quantitative and mixed methods approaches. Thousand Oaks, CA, Sage
- De Schepper, S., Haezendonck, E., & Dooms, M. (2015). Understanding pre-contractual transaction costs for Public–Private Partnership infrastructure projects. *International Journal of Project Management*, 33(4), 932-946.
- Doherty, B., Haugh, H., & Lyon, F. (2014). Social enterprises as hybrid organizations: A review and research agenda. *International Journal of Management Reviews*, 16(4), 417-436
- EPA (2018) Sources of Greenhouse Gas Emissions. US Environment Protection Agency. Accessed 28/04/2019 https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions
- Gambin L, Greene AE, Hogarth T. (2009) Exploring the links between skills and productivity: Final report. Institute for Employment Research, Warwick University, Coventry, March
- Gilson RJ (2003) Engineering a Venture Capital Market: Lessons from the American Experience. *Stanford Law Review*, 55: 1067-1068.
- Guardian (2019) Let's seize the moment and create a Green New Deal for the UK. Milliband E, Lucas, C and Sandys L, 30/04/2019 accessed 02/05/2019 <u>www.thegaurdian.com</u>
- Guardian (2019a) It's time for nations to unite around an international Green New Deal. Varoufakis Y and Adler D 23/04/2019 accessed 02/05/2019 www.theguardian.com
- Kenton W (2018) Cleantech. Investopia 28/06/2018 accessed 28/04/2019 https://www.investopedia.com/terms/c/cleantech.asp

Lehner, O. M. (2016). Routledge handbook of social and sustainable finance. Routledge.

Lehner, O. M., Harrer, T., & Quast, M. (2018, July). Legitimacy and Discourse in Impact Investing: Searching for the Holy Grail. In *Academy of Management Proceedings* (Vol. 2018, No. 1, p. 10935). Briarcliff Manor, NY 10510: Academy of Management.

R. Owen, O. Lehner, F. Lyon, G. Brennan / Finance, Risk and Accounting Perspectives Proceedings (2019) 88-107

- Lehner, O. M., & Nicholls, A. (2014). Social finance and crowdfunding for social enterprises: A public–private case study providing legitimacy and leverage. *Venture Capital*, *16*(3), 271-286.
- Lerner J (2012) The Narrowing Ambitions of Venture Capital. MIT Technology Review 06/09/2012, Boston
- Lerner, J. (2010) The Future of Public Efforts to Boost Entrepreneurship and Venture Capital. Small Business
- Lubell, M., Mewhirter, J. M., Berardo, R., & Scholz, J. T. (2017). Transaction costs and the perceived effectiveness of complex institutional systems. *Public Administration Review*, 77(5), 668-680.
- Mazzucato M and Semenieniuk G (2018) Financing renewable energy: Who is financing what and why it matters. Technology Forecasting and Social Change Vol127:8-22
- Owen, R., Deakins, D. and Savic, M. (2019) Finance Pathways for Young Innovative SMEs. Strategic Change: Briefings in Entrepreneurial Finance 28(1):19-36
- Owen R, Lyon F, and Brennan G (2018) Enabling investment for the transition to a low carbon economy: Government policy to finance early stage green innovation. Current Opinion in Environmental Sustainability Special Issue for the Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Cycle and Special Report on Global Warming Vol31: 137-145
- Oxford Research (2015) Assessing the Potential to Invest in Venture Capital and other Risk Capital Fund of Funds. European Commission, Brussels, October
- Polzin, F. (2017) Mobilising Private finance for Low Carbon Innovation A Systematic Review of Barriers and Solutions. Renewable and Sustainable Energy Reviews, 77: 525-35.
- Quélin, B. V., Kivleniece, I., & Lazzarini, S. (2017). Public-private collaboration, hybridity and social value: Towards new theoretical perspectives. *Journal of Management Studies*, 54(6), 763-792.
- Rowlands C. (2009) The Provision of Growth Capital to UK Small and Medium Size Enterprises. TSO, London
- Sayeed, L., & Onetti, A. (2018). The relationship between organizational resources and Green IT/S adoption: A RBV approach. *Journal of International Technology and Information Management*, *27*(2), 43-72.
- Shi, S., Chong, H. Y., Liu, L., & Ye, X. (2016). Examining the interrelationship among critical success factors of public private partnership infrastructure projects. *Sustainability*, 8(12), 1313.
- Soda, G., & Furlotti, M. (2017). Bringing tasks back in: An organizational theory of resource complementarity and partner selection. *Journal of Management*, *43*(2), 348-375.
- UNFCCC (2019) Climate Get The Big Picture. UN Framework Convention on Climate Change. Accessed 28/04/2019 <u>https://unfccc.int/resource/bigpicture/#content-the-paris-agreemen</u>
- Villani, E., Greco, L., & Phillips, N. (2017). Understanding Value Creation in Public-Private Partnerships: A Comparative Case Study. *Journal of Management Studies*, 54(6), 876-905.



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings



homepage: http://www.acrn.eu/frap

The Problem of Heterogeneity within Risk Weights: Does Basel IV contain the Solution?

Christina Binder^{1,*}, Othmar Lehner²

¹ University of Applied Sciences Upper Austria

² Hanken School of Economics, Helsinki

Abstract

The article uses a bank's credit data to study the impact of the Basel IV regulations on risk weight density (RWD). The analysis of the simulated data shows mixed results, as the improvement of risk weight heterogeneity is restricted to optimistically valued portfolios. Conservatively valued portfolios are likely to be confronted with an RWD decrease. However, within these portfolios, risk weight heterogeneity usually does not play an important role. Out of all the analysed Basel IV rules, the output floor clearly has the biggest influence on risk weight density, while the effect of the input floors is very limited within optimistically valued portfolios and is even eliminated by the removal of the scaling factor within conservatively valued portfolios. The change in RWD will also lead to a concurrent change in risk-weighted assets and therefore also in the level of eligible capital. The findings within the retail portfolio confirm those of the EBA study, which already suggested that Basel IV and especially the output floor will lead to a significant increase of risk capital (European Banking Authority, 2018).

Introduction

The soundness and safety of the banking system are said to be achievable by requiring banks to hold a high level of capital. Due to this fact, capital regulation has become the centrepiece of bank regulation within the past decades and the Basel Committee on Banking Supervision (BCBS), which is known as the most important standard-setter within the field of bank regulation, has played an important role within this evolution. In 1988 it published the Basel Accord (Basel I), which represents the first attempt to set international risk-based standards for capital regulation (Bodellini, 2019; Thakor, 2018). Basel I was structured very simply and therefore was soon criticised as being risk insensitive (Bodellini, 2019; Pérez Montes, Trucharte Artigas, Cristófoli, & Lavín San Segundo, 2018). This shortcoming was addressed by the BCBS through the introduction of the internal ratings-based approach (IRBA) in the Basel II framework. Since then, banks are able to determine their risk weighted assets (RWAs), which provide the basis for the calculation of the eligible capital, by using complex internal models. These models use parameters like the probability of default (PD), the loss given default (LGD) and the exposure at default (EAD) for the determination of risk weights (RW). The fact that these parameters are determined by the banks themselves, provides them with a high level of discretion in the calculation of the capital they are required to hold (Behn, Haselmann, & Vig, 2016a; Cucinelli, Di Battista, Marchese, & Nieri, 2018). In addition to that, the determination of these parameters requires a certain level of judgement, which potentially leads to heterogenous risk weights for identical assets (Turk-Ariss, 2017).

On the one hand, heterogeneity of risk weights has natural causes such as different databases, variations in IRB models and differences in national and international implementation standards

(Haselmann & Wahrenburg, 2016). On the other hand, this effect may also be due to manipulation: a study conducted by Mike Mariathasan and Ouarda Merrouche concluded that risk reported by banks is declining following the approval of IRB models and that this effect is particularly strong within weakly capitalised banks. The IRB approach thus offers banks the opportunity to report their risk at a lower level than it actually is and thus to overstate regulatory capital (Mariathasan & Merrouche, 2014). During the financial crisis it then became apparent, that capital requirements, which have been set by the regulator were too low and that the definition of them has been set too broadly (Haldane, 2013; Rossignolo, Fethi, & Shaban, 2013). That is why the BCBS then published 'Basel III: A global regulatory framework for more resilient banks and banking systems' in which it raised the amount and the quality of the regulatory capital required (Basel Committee on Banking Supervision, 2010). However, the BCBS did not reform the IRB approach or the calculation of risk-weighted assets, which led to the fact that Basel III remained somewhat as self-regulatory as Basel II (Haldane, 2013).

The reduction of the excessive variability of risk weights was one of the reasons for the revision of the Basel III framework. In December 2017 the BCBS published the final version of "Basel III: Finalising post-crisis reforms" in which it emphasises the necessity of a credible and prudent RWA calculation. This new regulatory framework, which is commonly referred to as Basel IV, primarily deals with the revision of the standardised approach and the restriction of the internal ratings-based approach. The constraint of the IRBA comes with the removal of the IRB-option for certain asset classes and the introduction of parameter floors and output floors (Basel Committee on Banking Supervision, 2017). The publication of Basel IV led to a huge debate in the banking industry and since then various consulting firms have been trying to quantify the impact of the new rules. A study by the European Banking Authority (EBA) based on figures of 134 banks in the European Economic Area shows an increase in risk capital at full implementation of Basel IV of 21.8%. This increase is largely due to the introduction of the output floor (6.3%) (European Banking Authority, 2018). Another study even refers to the regulations as a "game changer" for the European banking landscape (Schneider, Schröckl, Koch, & Schneider, 2017).

The topic of risk weight variability has already been addressed in numerous articles, but to my knowledge none of them analysed the implications of the Basel IV input and output floors on heterogeneity.

The aim of this article is to analyse the impact of Basel IV on the heterogeneity of risk weights and thereby to answer the following research questions:

• Question 1: How can the impact of Basel IV on the heterogeneity of risk weights be simulated?

• *Question 2: Do the new Basel-IV-regulations have the potential to reduce or even to eliminate the problem of heterogeneity within risk weights?*

• Question 3: To what extent will the risk weight density be affected by the Basel IV input and output floors?

In order to get an answer to the research questions, a simulation of the Basel IV framework for credit risk has been developed and performed using Microsoft Excel 2016. The simulation is based on the credit data provided by an IRB bank located in Western Europe, and therefore only analyses the impact of Basel IV on the IRB approach for credit risk. Due to data protection issues the dataset has been anonymised by the bank itself before transmission.

The rest of the article is organised as follows. The following three sections provide the basis about risk weighted capital regulation, the IRB-Approach itself and the Basel IV framework. The fifth section describes the dataset and the methods employed. The final sections present the results and provide a conclusion from the results.

The Evolution of Risk-Weighted Capital Regulation

Capital regulation represents the centrepiece of bank regulation and there is a consensus that an increased level of equity goes hand in hand with increased stability (Bodellini, 2019; Thakor, 2018). While there is unity in this area, there is disagreement within others – that is why the required level of bank capital as well as the valuation of banks' assets has changed numerous times over the years (Mariathasan & Merrouche, 2014).

The Basel Accord of 1988 is the first attempt to set international risk-based standards for capital regulation. It regulated the minimum capital requirements and required banks to exhibit a minimum ratio of capital to risk weighted assets (RWAs) of 8%. The risk weights which have been used for the calculation of RWAs were negligible as there were only four of them (Blundel-Wignall & Atkinson, 2010; Haselmann & Wahrenburg, 2016; Hull, Mader, & Wagner, 2014). Basel I was criticised at an early stage as the calculation of RWAs was perceived as being too crude and as it did not distinguish between the risks within each category (Pérez Montes et al., 2018).

Basel II, the succeeding framework to Basel I was published in 2004, with the aim of promoting stronger risk management within banks. It introduced the IRBA, which determines the risk weights using internal models that are based on the bank's data. The aim of the IRBA therefore was to tie the calculation of regulatory capital closer to the actual risks banks are facing (Basel Committee on Banking Supervision, 2004). What actually happened was that Basel II increased the level of complexity, it enabled banks to reduce their capital and thereby caused an enormous indebtedness which then led to the fact that banks entered the financial crisis with a level of equity that was far too low (Cabrera, Dwyer, & Nieto, 2018; Hellwig, 2010; Mariathasan & Merrouche, 2014). These facts have also been admitted by the BCBS when they introduced Basel III as they stated that the leverage which has been built up by banks as well as the low level and quality of capital and the insufficient liquidity were some of the reasons for the financial crisis (Basel Committee on Banking Supervision, 2010). The focus of Basel III was then put on the increase of the level and the quality of regulatory capital. In addition to that the BCBS introduced a leverage ratio, which should act as a protection against the failure of internal models and therefore represents a risk insensitive backstop (Pérez Montes et al., 2018). However, what Basel III did not reform was the calculation of RWAs under both the standardised approach (SA) as well as the much criticised IRB-approach (Blundel-Wignall & Atkinson, 2010; Sonali & Amadou N.R., 2012). These aspects were then taken up by the BCBS in their Basel IV framework and will be described in the section on the Basel IV's Implications on the Measurement of the Credit Risk (Basel Committee on Banking Supervision, 2017).

The IRB-Approach and its Up- and Downsides

As already mentioned in the previous section, the standardised approach has often been criticised because of its rough RWA calculation (Behn et al., 2016a; Pérez Montes et al., 2018). This approach uses standardised risk weights depending on the asset class and the credit risk of the debtor. In order to address this point of criticism, the BCBS introduced a new method to calculate the RWAs: the so-called IRB-approach. This new calculation method uses complex internal models which have to be approved by the regulatory authority to determine the amount of regulatory capital (Resti, 2016).

The IRB-Approach under Basel II

When banks decide on the implementation of the IRB-approach they have the possibility to choose between the Foundation Internal Ratings-Based Approach (FIRBA) and the Advanced Internal Ratings-Based Approach (AIRBA). The difference between these two approaches is that for the foundation approach the bank only needs to determine the probability of default (PD) on its own, whereas the remaining parameters, such as exposure at default (EAD), loss given default (LGD), and maturity (M) are provided by the authority. For the advanced approach all parameters have to be determined by the bank (Basel Committee on Banking Supervision, 2004).

Both approaches are based on the calculation of the Value at Risk (VaR), which is a commonly used measure for risk management. The VaR measures the maximum loss within a certain period of time and a given confidence level – therefore the VaR does not give any information about the loss below this threshold (Chen, Wang, & Zhang, 2019; Zoia, Biffi, & Nicolussi, 2018). The Basel regulations require banks to calculate their RWAs using the VaR at a confidence level of 99.9% (Basel Committee on Banking Supervision, 2004). Figure 1 shows the VaR-model, which includes the expected loss, which banks already incorporate in their interest rates dependent on their customers PD and the estimated LGD, and the unexpected loss, which they do not incorporate as they do not expect it to be incurred. Figure 1 is represented in formulas 5 and 7, which show the deduction of the expected loss (LGD x PD) from the total loss at the confidence level of 99.9% (Aikman et al., 2014).



Figure 1 Value at Risk model for credit risk under Basel II Source: Aikman et al. (2014)

The general formula for the calculation of risk weighted assets is shown in formula 1.

RWA = K x 12,5 x EAD x 1.06

In this formula the value of 12.5 represents the reciprocal of the minimum capital ratio which has been set at 8%. The scaling factor of 1.06 has been introduced by the BCBS in order to maintain a certain level of minimum capital requirements (Basel Committee on Banking Supervision, 2004). However, the calculation of the capital requirements (K), which will be described in the following paragraphs, differs between the individual exposures.

(1)

Calculation for Exposures to Corporates, Sovereigns and Banks

In order to be able to calculate the capital requirements, one firstly needs to determine the correlation (p) and the maturity adjustment (b). Basel II assumes an inverse relationship between the PD and the parameter for correlation (p) which it bases on empirical results (Lopez, 2002). If the solvency of a company decreases, its PD increases. This assumes that its probability of default gets more idiosyncratic and therefore it is less dependent on the overall market conditions (Hull et al., 2014).

$$p = 0.12 \frac{1 - \exp(-50 x PD)}{1 - \exp(-50)} + 0.24 \left[1 - \frac{1 - \exp(-50 x PD)}{1 - \exp(-50)} \right]$$
(2)

The calculation of the correlation is slightly different for exposures to small- and mediumsized enterprises¹ (SMEs) as the formula also includes a firm-size adjustment in which S describes the annual turnover given in millions (Basel Committee on Banking Supervision, 2004):

$$p = 0.12 \frac{1 - \exp(-50 x PD)}{1 - \exp(-50)} + 0.24 \left[1 - \frac{1 - \exp(-50 x PD)}{1 - \exp(-50)} \right] - 0.04 x \frac{1 - (S-5)}{45}$$
(3)

If Banks apply the FIRBA they must use 2.5 years as their effective term to maturity (M), if they apply the AIRBA they need to calculate M for every individual exposure (Basel Committee on Banking Supervision, 2004).

$$b = (0.11852 - 0.05478 x \ln(PD))^2$$
⁽⁴⁾

After having calculated these parameters, the capital requirements (K) can be determined as follows:

$$K = \left[LGD \ x \ N \left\{ N^{-1}(PD) \ x \ \sqrt{\frac{1}{1-p}} + N^{-1}(99.9\%) \ x \ \sqrt{\frac{p}{1-p}} \right\} - LGD \ x \ PD \right] x \ \frac{1 + (M-2.5) + b(PD)}{1 - 1.5 \ x \ b(PD)}$$
(5)

If the IRB-bank determines the parameters, which are incorporated in this formula, correctly, the calculated capital requirements should be large enough to ensure that the bank faces insolvency once every thousand years (Aikman et al., 2014).

Calculation for Retail Exposures

The calculation for retail exposures is a little simpler than the one before, as there is no maturity adjustment and the correlation is set to the fixed value of 0.15 for residential mortgage exposures, to 0.04 for qualifying revolving retail exposures and is determined by the following

¹ SMEs = companies with an annual turnover of up to 50 million euros

formula for all other retail exposures (Basel Committee on Banking Supervision, 2004):

$$p = 0.03 \ \frac{1 - \exp\left(-35 \ x \ PD\right)}{1 - \exp\left(-35\right)} + 0.16 \left[1 - \frac{1 - \exp\left(-35 \ x \ PD\right)}{1 - \exp\left(-35\right)}\right] \tag{6}$$

The calculation of the capital requirement is similar to the one for corporates, sovereigns and banks but does not include the maturity adjustment in the end:

$$K = LGD \ x \ N\left\{N^{-1}(PD) \ x \ \sqrt{\frac{1}{1-p}} + N^{-1}(99.9\%) \ x \ \sqrt{\frac{p}{1-p}}\right\} - LGD \ x \ PD$$
(7)

The Dark Side of the IRB-Approach

The IRB-approach has been introduced by the BCBS because they wanted to align the calculation of RWAs to the actual risks of the banks' portfolio. Banks, which decided for the implementation of the IRB-approach, are rewarded with more discretion in their risk assessment, a lower level of RWAs and therefore they are also able to reduce the amount of eligible capital they are required to hold (Barakova & Palvia, 2014; Cucinelli et al., 2018). Research indicates that the IRB model is superior to the SA model, as IRB banks are better placed to manage credit risk risk-sensitively. This is confirmed by the fact that IRB-banks reported a lower credit risk following the financial crisis (Cucinelli et al., 2018). According to a study which has been implemented by Barakova and Palvia (2014) the aim of increasing the risk sensitiveness of the Basel framework through the introduction of the IRBA has been reached: The results of their study show that under the AIRBA the alignment of capital with the banks' risk has improved. However, they could not foresee that there are also some other variations, which are not risk-related (Barakova & Palvia, 2014).

In the past years the IRB-approach has met with a great deal of criticism as researchers accused it of being self-regulatory, subject to manipulation and inferior to insensitive measures. The criticism of the excessive variability or heterogeneity within risk weights is one aspect that numerous articles together raised (Abbassi & Schmidt, 2018; Aikman et al., 2014; Cucinelli et al., 2018; Haldane, 2013; Hellwig, 2010; Mariathasan & Merrouche, 2014; Mayes & Stremmel, 2012; Sonali & Amadou N.R., 2012). All these criticisms will be addressed in the next few paragraphs.

Capital Optimisation and Risk Weight Manipulation

The right amount of equity is difficult to determine: a low level of capital may lead to an increase in the riskiness of banks' portfolios whereas a high level of capital may encourage banks to reduce their lending (Mariathasan & Merrouche, 2014). If banks estimated their internal risks properly, the only differences in their risk weights, which is calculated by dividing the RWAs by their total assets, should lie in the riskiness of their portfolios, the variation in their modelling techniques and national implementation guidelines (Haselmann & Wahrenburg, 2016). However, it is expensive for banks to hold eligible capital and therefore they are encouraged to reduce it to a minimum. This fact is also referred to as Goodhart's law, which says: 'When a measure becomes a target, it ceases to be a good measure' (Plosser & Santos, 2018). Therefore, the IRBA might suffer from incentive problems (Behn et al., 2016a). Recent studies which have dealt with risk weight variability show that risk weight density decreases after the approval of the IRBA by the regulatory body (Mariathasan & Merrouche, 2014), that reported risk weights in general are significantly lower for IRB-banks than for SA-banks, but that actual default rates are higher for IRB-banks than for SA-banks. In addition to that, they show that the interest rates, charged by IRB-banks are higher than those charged by SA-banks, which indicates that banks were well aware of the risks within their portfolios but they just do not take this into account when calculating the risk weights (Barucci & Milani, 2018; Behn et al., 2016a). Some studies also show that these effects are particularly strong for weakly capitalised banks, which are found in countries where supervision is weak (Mariathasan & Merrouche, 2014; Plosser & Santos, 2018; Turk-Ariss, 2017). Another concern associated with the IRBA is the connection between risk weights and bank distress, which is statistically significant for SA-banks but insignificant for IRB-banks (Cizel, Rijken, Altman, & Wierts, 2017).

Simplicity versus Complexity

Another much criticised aspect of the IRBA is the high complexity involved in the calculation which makes it difficult to monitor (Tarullo, 2014). Although very complex models have already been used prior to the financial crisis, they still were not able to predict nor to prevent it. Since then the models have become even more complex as more and more information is taken into account (Aikman et al., 2014). However, some studies point out that risk insensitive models, like the leverage ratio which has been introduced with Basel III and represents a non-riskweighted ratio, often outperform risk sensitive ones. Not only because these models are easier to enforce, but also because empirical evidence proves that simple ratios provide a better indication of potential bank distress (Barucci & Milani, 2018; Behn et al., 2016a; Estrella, Park, & Peristiani, 2000; Kiema & Jokivuolle, 2014). In addition to that, the global financial crisis showed that risk-weighted models were no good as predictors of the crisis itself (Mayes & Stremmel, 2012). Risk measurement models use data from the past to predict future defaults (Rajan, Seru, & Vig, 2015) and they are subject to model risk, as they use statistical models to estimate risk. There is a variety of statistical models, which all present different outcomes and therefore also variable risk weights – the difference between all of them represents the model risk, which tends to be low during financially stable periods, but high in times of financial distress (Danielsson, James, Valenzuela, & Zer, 2016). This indicates that sometimes "It's better to be roughly right than precisely wrong" as John Maynard Keynes already said.

Procyclicality

In comparison to the standardised approach, the IRB approach uses a large number of risk weights which are variable between banks and dependent on the counterparty's PD and LGD. This large amount of risk weights grants risk-sensitive capital regulation, which ties capital charges to the actual risks banks are facing. However, this might not only be positive, as it also causes a certain level of procyclicality: banks are encouraged to increase their lending during good times, as loans are perceived as being less risky and therefore less eligible capital is required, but they will also tend to reduce lending during bad times when it becomes riskier (Ben Naceur, Marton, & Roulet, 2018; Ly & Shimizu, 2018). The IRBA, which represents the most risk-sensitive form of capital regulation, is said to be amplifying this effect (Repullo & Suarez, 2012). This fact is also confirmed by a study which analyses the effects of the global financial crisis on the lending behaviour of German banks. This study shows that the reduction in lending was greater for IRB-banks than for SA-banks. The BCBS has already reacted to this effect, as Basel III introduced a capital conservation buffer, which is designed in a way that leads banks to build up a capital buffer during times of growth. This excess capital can then be used in times of

recession (Behn, Haselmann, & Wachtel, 2016b).

The afore mentioned critics encouraged the BCBS to revise the Basel III framework and introduce Basel IV. The alterations to the existing framework will be described in the following section.

Basel IV's Implications on the Measurement of Credit Risk

In March 2016, the BCBS already responded to the afore mentioned criticism for the first time and drafted the consultative document 'Reducing variation in credit risk-weighted assets constraints on the use of internal model approaches' in which it made several proposals to reform the standardised approach and restrict the IRB approach. This paper restricts the application of the IRB approach to specific portfolios; it includes an input floor that provides a constraint on the parameters used for the calculation (PD, LGD, CCF, EAD), and an output floor, which links the IRB-RWAs to the SA-RWAs (Basel Committee on Banking Supervision, 2016). In December 2017, the BCBS then published its final paper 'Basel III: Finalizing post-crisis reforms', also commonly referred to as Basel IV. In it, the proposals of the consultative paper from 2016 were revised and finally fixed. The document contains a timetable which suggests an implementation of the regulations from 2022 onwards (Basel Committee on Banking Supervision, 2017).

Implications for the Standardised Approach

Although this article primarily deals with the effects of Basel IV on the IRBA, it is also necessary to describe its influence on the SA, as the RWAs which are calculated following this approach provide the basis for the output floor, which in turn is an important part of the Basel IV IRBA. Overall it can be said that the BCBS' revisions of the SA for credit risk are seen positively by the literature, as they lead to an increased risk-sensitivity (Joosen, 2016).

Within the exposures to corporates it is necessary to distinguish between specialised lending exposures and general corporate exposures. Basel IV gives banks the possibility to base the risk weights of its corporate exposures on the rating of their counterparty if this has been determined by an eligible credit assessment institution (ECAI) (Basel Committee on Banking Supervision, 2017). As the sample data does not provide any external ratings, this approach cannot be applied. Therefore, the rules for banks located in jurisdictions, which do not allow the application of the previously described approach, are described and applied in the simulation.

Concerning general corporate exposures, banks need to determine whether an exposure is to a normal corporation or whether it is to an SME. In cases where it is to a corporation, Basel IV distinguishes between investment grade exposures, which are risk weighted at 65% and normal exposures, which receive a risk weight of 100%. Exposures to SMEs are divided into regulatory retail exposures and general SME exposures. The definition of regulatory retail encompasses the following characteristics: the exposure needs to be revolving or a personal term loan or lease, it amounts to \in 1,000,000 or less and the exposure to one individual cannot exceed the threshold of 0.2% of the total regulatory retail portfolio. In cases where an exposure fulfills all these requirements, it is risk weighted at 75%. For all other SME exposures a risk weight of 85% is applied (Basel Committee on Banking Supervision, 2017).

Basel IV recognises three different types of specialised lending exposures: project finance, object finance and commodities finance. The risk weights, which are applied to object and commodity finance exposures are generally 100%, whereas there are three potential risk weights for project finance exposures: 130% if the project is in its pre-operational phase, 100% if its in

the operational phase and 80% if it is in the operational phase and deemed to be high-quality (Basel Committee on Banking Supervision, 2017).

Within the retail portfolio, Basel IV defines two different types of exposure: the regulatory retail exposure and other retail exposures, which defines the remaining population and is risk weighted at 100% (Basel Committee on Banking Supervision, 2017).

As the bank's corporate and retail portfolio also contains real estate exposures, the adapted rules need to be explained in this article. Basel II allocated a risk weight of 35% to exposures, which were fully secured by residential real estate, but then risk-weighted commercial real estate exposures at 100% (Basel Committee on Banking Supervision, 2004). This regulation has been revised by the BCBS, as risk weights are now dependent on the Loan-to-Value- Ratio (LTV-Ratio) of the exposure. This ratio is calculated by dividing the amount of the loan by the value of the property. The calculated LTV-Ratio then leads to individual RWs, which also depend on whether they are residential or commercial real estate and whether the repayment is materially dependent on the cashflows generated by the property or not (Basel Committee on Banking Supervision, 2017). The changes within the real estate portfolio are seen as the most significant ones within the revision of the SA, as they bring more sophistication and more risk sensitivity (Joosen, 2016).

Removal of the IRB Option for Certain Exposures

Basel IV removes the possibility to apply the IRBA for certain asset classes, which exhibit modelling difficulties due to the small number of defaults. The exposures which are affected are shown in table 1. This restriction reflects certain critical voices like Amorello (2016), who argued that Basel III did not address the questionable reliability of the IRBA and that banks therefore were still incentivised to calculate their regulatory capital based on internal models and parameters. Still, there are also adverse comments from the banking industry, which state that they fear a disproportionate increase of capital requirements for banks within the European Union, and that certain variations within the RWAs are desirable as they reflect differences in portfolios (Strickland, 2017). As the bank, which provided the data for the simulation applies the AIRBA on its retail portfolio and the FIRBA on its corporate portfolio and its equity portfolio, which then again has been excluded from the simulation due to insignificance, the effects of the removal of the IRB-option are not analysed in this article.

Table 1 Available Approaches under Basel II and IV Source: Basel Committee on Banking Supervision (2017)

Exposure	Approaches under Basel II	Approaches under Basel IV
Corporates (consolidated > 500 MEUR)	revenue AIRB, FIRB, SA	FIRB, SA
Banks & other financial institutions	AIRB, FIRB, SA	FIRB, SA
Equities	AIRB, FIRB	SA

Input Floor for Model Parameters

The parameters used for the RWA calculation of exposures which can still be measured using the IRBA will be subject to the introduction of input floors for the PD, the LGD and the EAD (Basel Committee on Banking Supervision, 2017). The bank which provided the credit data, determine their RWAs for the corporate portfolio using the FIRBA and their RWAs for the retail portfolio using the AIRBA. Under the FIRBA banks only determine the PD on their own, whereas the other parameters are prescribed by the regulator (Basel Committee on Banking Supervision, 2004). Due to that only the PD-Floor has been considered in the simulation of the corporate portfolio, whereas all other floors have been taken into account in the simulation of the retail portfolio.

The minimum value for the PD of an exposure to a corporation amounts to 0.05%. The minimum PDs for retail exposures have been set at 0.05% for mortgages, qualifying revolving retail transactors and for other retail exposures, while exposures which come within qualifying revolving retail exposures (QRRE) were set at 0.1%. Under the AIRBA the minimum value for the LGD depends on whether the exposure is secured or unsecured. For unsecured exposures the LGD input floor has been set at 25% for corporates, at 50% for QRRE and at 30% for other retail exposures. The LGD of secured loans has to amount to at least 5% for mortgages and is dependent on the collateral type for corporate and other retail exposures: financial collaterals allow banks to set the LGD at 0%, receivables and commercial or residential real estate at 10% and other collaterals at 15%. In addition to the floor, which requires a minimum level of LGDs and has been introduced for exposures under the AIRBA, the determination of the LGD in the FIRBA has also been reformed. The LGD for unsecured exposures has been reduced from 45% to 40% if the counterparty is a corporation. The LGD of the collateral depends on the exposure weighted average of the unsecured and the secured LGD and is determined using formula 8. The secured LGD amounts to 0% for eligible financial collateral, 20% for eligible receivables and real estate and 25% for other eligible physical collateral. In addition to these LGDs a haircut has to be applied: this amounts to 40% for eligible receivables, real estate and other physical collateral and to 100% for ineligible collateral (Basel Committee on Banking Supervision, 2017).

$$LGD \ Floor = LGD \ Floor_U \ x \ \frac{E_U}{E \ (1+H_E)} + \ LGD \ Floor_S \ x \ \frac{E_S}{E \ (1+H_E)}$$
(8)

Where

LGD FloorU - LGD floor for the secured portion of the exposure EU

- Unsecured portion of the other retail exposure

HE - Haircut of the collateral

LGD FloorS - LGD floor for the secured portion of the other retail exposure ES

- Secured portion of the other retail exposure

Under the AIRBA banks also need to determine their EAD measures. Basel IV now prohibits the application of internal EAD estimation and requires banks to determine their off balance sheet exposure using the credit conversion factors (CCF) which are also used under the standardised approach (Resti, 2016). In addition to that the BCBS also introduced a floor to the EAD, which is the sum of the entire on balance exposure plus 50% of the off balance exposure (Basel Committee on Banking Supervision, 2017).

Output Floor

The output floor will affect not only the RWAs resulting from credit risk but also all other risk types. In respect of credit risk, banks will need to determine their RWAs according to the standardised approach also for their IRB-portfolios. The output floor determines that the final IRB-RWAs must be at least 72.5% of the SA-RWAs by 2027 (Basel Committee on Banking Supervision, 2017). This floor replaces the previously applicable Basel I floor and serves as a backstop that links the RWAs calculated under the IRBA to the RWAs calculated under the SA. This introduction should ensure that a level playing field between IRB- and SA-banks is maintained (Bodellini, 2019). However, it is also seen critically as the leverage ratio already serves as a risk-insensitive backstop (Haselmann & Wahrenburg, 2016; Pérez Montes et al., 2018). Basel IV also includes transitional measures which include a phase-in arrangement for the output floor from 2022 to 2027. Within this period of time the BaselIV output floor gradually increases from 50% to 72.5%. The simulation includes the application of all percentages, but only the results of the final 72.5% output floor will be presented in this article. Due to these and the afore mentioned alterations to the IRBA, the BCBS decided that it is no longer necessary to apply the scaling factor of 1.06 in the RWA calculation (Basel Committee on Banking Supervision, 2017).

Methodology

The aim of this article is to show the impact of Basel IV on the heterogeneity of risk weights. This aspect has already been addressed in numerous studies, but to my knowledge none of them analysed the implications of the Basel IV input and output floors on heterogeneity. Most articles used RWA density, which is calculated by dividing RWAs by total assets, as a measure of heterogeneity: Beltratti and Paladino (2016) used it to show that banks perform capital optimisation through the usage of the IRBA, Mariathasan and Merrouche (2014) showed that RWA density decreased considerably after the regulatory approval of the IRBA and Vallascas and Hagendorff (2013) exhibited that there is an ill-calibration in the context of RWAs and portfolio risk with the usage of risk weight density. Due to the fact that RWA density has already been used by researchers to address numerous aspects, all in connection with risk weight heterogeneity, this ratio will also be used in this article.

The data analysis relies on a simulation as an inductive data driven instrument. Some research articles have previously already used a simulation for showing diverse aspects dealing with the Basel framework: Andersen (2011) used it to show Basel II's procyclical implications, Peura and Jokivuolle (2004) simulated the Basel stress tests for capital adequacy and Bellotti (2010) used it to compute the expected loss distributions of Basel II based on a credit card portfolio. The simulation itself has been applied using Microsoft Excel 2016. This choice can be justified with the fact that it allows the user to have full control over the applied parameters and maximum freedom in the analysis as well as the presentation of the data. The simulation was applied using advanced formulas and the data analysis tool.

Description of the Sample

The data was provided by a small IRB-bank located in Western Europe. The bank's total assets amount to less than 10 billion euros and the management controls the bank conservatively. The simulation is based on the bank's credit data and therefore only analyses the impact of Basel IV on the IRB approach for credit risk. As credit risk makes up about 80% of the total risk-weighted

assets (Berg & Koziol, 2017), it is considered appropriate to exclude alterations in other risk types which are due to Basel IV from this thesis. Table 2 shows that the bank applies the AIRBA on its retail portfolio, whereas it applies the FIRBA on its corporate portfolio, on equities and investment certificates. In addition to that it can also be stated that the bank clearly focuses on retail exposures. As the data stems from a small regional bank, the results are very likely to differ from the results of large multinational banks. Nevertheless, the data is considered to be appropriate for simulation, as it represents a typical regional bank within Europe. Another aspect, which made the usage of the data very appealing, was the fact that small banks usually hold a rather small number of syndicated loans, which could distort the simulation results. Therefore, this sample probably includes much clearer data within the corporate portfolio.

Table 2 EAD, RWA and RWD of the sample

	EAD	RWA	RWA	
IRB Advanced Approach	55.85%	22.18%	0.208	
Retail	55.85%	22.18%	0.208	
IRB Foundation Approach	44.15%	77.82%	0.924	
Equities	0.75%	4.26%	2.973	
Corporates	43.24%	72.44%	0.879	
Investment Certificates	0.16%	1.13%	3.700	
Total	100.00%	100.00%	0.524	

As the retail and corporate exposures already make up 99.09% of the total EAD and 94.62% of the RWA of the IRB-portfolio respectively, and these are the only asset classes which show a RWA density from below one, only these two asset classes are included in the simulation. The RWA density, which has been calculated by dividing the RWA by the EAD, of the overall IRB portfolio, amounts to 0.524 which suits the mean RWA density shown in table 3, which has been observed by Mariathasan and Merrouche (2014) during the period from 2004 to 2010 based on the data of 115 IRB-banks in 21 OECD-countries. The RWA density observed by Mariathasan and Merrouche (2014) is calculated on the basis of the entire portfolio of the banks being included. As a banks credit data provides the basis for this article, it is possible to analyse the RWA density more exactly. Therefore it is possible to calculate the RWA density for each asset class – this calculation shows that the RWA density of the retail portfolio lies clearly below the mean, as it amounts to 0.208, whereas the RWA density of the corporate exposure amounts to 0.879 and therefore lies above the mean observed by Mariathasan and Merrouche (2014).

Та	ble 3	RWD	according to	o Mariathasan	and Merrouche	(2014)
-						(-)

	Minimum	Mean	Maximum
RWA Density	0.024	0.516	0.965

For both portfolios, the maximum PD amounts to 22.77%. The minimum LGD and PD of the total corporate portfolio account for 0%, which is caused by the fact that specialised lending exposures are also included in the sample. For these exposures, the RW is not determined by the usage of the IRB-formula but is set by the Basel regulations. The minimum LGD of the remaining corporate portfolio amounts to 35%, which also represents a predefined parameter, as the FIRBA is used for this portfolio.

Asset Class	PDmi	PDav	PDma	LGDmin	LGDa	LGDmax
Retail	0.03%	2.17%	22.77%	5.00%	35.23%	80.14%
Corporates	0.00%	2.01%	22.77%	0.00%	41.87%	45.00%

Table 4 PD and LGD of the sample

As the Basel framework requires banks to apply different formulas to their asset classes, the simulation was split into two parts: the simulation of the corporate portfolio, which is described in 5.2 Corporate Portfolio and the simulation of the retail portfolio, which is shown in 5.3 Retail Portfolio. Before the separation into these portfolios it had to be determined whether an exposure falls into Basel IV's regulatory retail class, as this applies for the retail as well as for the corporate portfolio. Therefore, exposures within the corporate portfolio which were categorised as SMEs, amounted to a maximum value of \in 1.000.000 and were not categorised as specialised lending were added to retail exposures, which were defined as qualifying revolving retail by the bank itself. After grouping these exposures together, each one was analysed as to whether it amounts to more than 0.2% of the overall regulatory retail portfolio (Basel Committee on Banking Supervision, 2017). If the exposure was smaller than 0.2% of the regulatory retail portfolio it remained in this portfolio, if it was higher it was excluded – the categorisation was then used in the determination of risk weights, which will be described later on.

Corporate Portfolio

The simulation of the corporate portfolio is divided into three steps: the simulation of the standardised approach under Basel IV, the simulation of the Basel IV input floors and the simulation of the Basel IV output floors.

Simulation of the Basel IV Standardised approach

For the simulation of the standardised approach it was necessary to divide the corporate portfolio into general corporate exposures and specialized lending exposures. This categorisation has already been given by the bank itself as it splits its corporate portfolio into general corporates, real estate finance, project finance and object finance.

Basel IV includes project finance, object finance and commodities finance in its framework for specialised lending exposures. The risk weights which are applicable for these exposures

are dependent on the sub portfolio they belong to: object and commodity finance exposures generally receive a risk weight of 100% whereas project finance exposures have to be separated into exposures during the pre-operational phase, which receive a risk weight of 130% and exposures during the operational phase, which are generally risk weighted at 100% unless they are deemed to be high quality, then they receive a risk weight of 80% (Basel Committee on Banking Supervision, 2017). As the data did not include any information about the phase they are currently in and the quality of them, it was assumed, that 60% of the entire portfolio are in the operational phase whereas 40% are in the pre- operational phase. Due to simplification reasons there was no assessment of the quality of the exposure, so all project finance exposures, which are in their operational phase were risk weighted at 100%. It must be noted that this further categorisation would only have an insignificant influence on the overall result, as the project finance portfolio only amounts to 7.93% of the entire RWAs of the corporate portfolio. The Basel IV framework also includes new rules for banks' real estate exposures: Basel IV differentiates between the collateralisation by commercial and residential real estate, and whether repayment is materially dependent on the cash flows which are generated by the

property or not (Basel Committee on Banking Supervision, 2017). For reasons of simplicity, it was assumed that all real estate exposures within the corporate portfolio are secured by commercial real estate. The bank includes its real estate exposures in its specialised lending exposures, as it defines them as income-producing real estate (IPRE) under the IRBA. The Basel IV definition of IPRE exposures contains the requirement that the repayment of these loans needs to materially depend on the cash flows generated by the property (Basel Committee on Banking Supervision, 2017). Due to that, it was assumed that there exists a material dependence on the cash flows of the property for the entire real estate exposure. Therefore the risk weights depend on the Loan-to-Value-Ratio (LTV-Ratio) of the respective exposure. This ratio is calculated by dividing the amount of the loan by the value of the property – the applied risk weights are shown in formula 9.

$$RW = \begin{cases} LTV \le 60\% = 70\% \\ 60\% < LTV \le 90\% \\ LTV > 80\% = 110\% \end{cases}$$

(9)

Within the general corporate exposures there is a distinction between regulatory retail exposures, general SME exposures, investment grade exposures and all other corporate exposures. Regulatory retail exposures within the corporate portfolio must be to SMEs, amount to up to $\in 1,000,000$ and not exceed 0.2% of the overall regulatory retail portfolio. This allocation has already been made during the data preparation and has been described under 5.1 Description of the Sample. The exposures which are categorised as regulatory retail are risk weighted at 75%, whereas the remaining exposures which are to SMEs receive a risk weight of 85% (Basel Committee on Banking Supervision, 2017). Based on the information provided by the bank, it was not possible to assess whether an exposure can be allocated to the investment grade portfolio. Due to that reason the remaining corporate portfolio was risk weighted at 100%. After having determined the risk weights for all exposures, it was possible to calculate the RWAs under the Basel IV standardised approach. The results of this calculation will be needed for the calculation of the output floor.

Simulation of the Basel IV Input and Output Floors

Basel IV introduced a parameter floor for the probability of default which amounts to 0.05% (Basel Committee on Banking Supervision, 2017) - therefore all exposures to corporates, which showed a smaller PD, were floored at this percentage. As the bank, which provided the data for the simulation, applies the foundation approach on its corporate exposure, the effect of the PD floor is the only one that can be simulated, as the other parameters are fixed by the standard setter and therefore not affected by any floors. Basel II required banks to apply a risk weight of 45% on unsecured senior claims and 75% on unsecured subordinated ones under the FIRBA. The LGD for the secured part of the exposure depended on the type of collateral and amounted to 35% at a minimum (Basel Committee on Banking Supervision, 2004). Basel IV now requires banks to apply an LGD of 40% on unsecured senior claims for exposures to corporates, whereas the risk weight for unsecured subordinated exposures remains the same. The LGD for the secured portion of the exposure needs to be calculated as the exposure weighted average of the unsecured and the secured LGD using formula 8. As already mentioned, the LGD of the unsecured portion amounts to 40%. The LGD of the secured portion amounts to 0% for eligible financial collateral, 20% for eligible receivables and real estate and 25% for other eligible physical collateral. In addition to these LGDs, a haircut has to be applied: this amounts to 40% for eligible receivables, real estate and other physical collateral and to 100% for ineligible collateral (Basel

Committee on Banking Supervision, 2017). As the sample data did not comprise any information about the type of collateral, the secured LGD was fixed at 20%, as this applies for two out of four collateral types. In addition to that the haircut was fixed at 40%, as this applies for three out of five categories.

After the determination of the Basel IV LGDs and PDs it was possible to calculate the correlation (p), the capital requirements (K) and the RWAs by using the formulas stated in 3.1.1 Calculation for Exposures to Corporates, Sovereigns and Banks.The results of this calculation represent the RWAs following the Basel IV rules for input floors. The simulation of the output floor was done by multiplying the RWAs, which have been determined by applying the Basel IV rules for the standardised approach, with the applicable percentages of 50% in 2022 to 72.5% in 2027 (Basel Committee on Banking Supervision, 2017).

Retail Portfolio

Like the simulation of the corporate portfolio, the simulation of the retail portfolio is divided into three steps: the simulation of the standardised approach under Basel IV, the simulation of the Basel IV input floors and the simulation of the Basel IV output floors.

Simulation of the Basel IV Standardised approach

The simulation of the Basel IV SA required the segmentation of the retail portfolio into the real estate exposure, the regulatory retail exposure and other retail exposure. The segmentation criteria have already been partially provided by the bank itself as it splits its retail portfolio into three Basel sub categories: residential mortgages, qualifying revolving retail and other retail. For the purpose of the simulation it has been assumed that the bank's residential mortgage portfolio meets Basel IV's criteria for the real estate exposure, which includes the following requirements: the property needs to be fully completed, the claim has to be legally enforceable, the bank holds a first lien, the borrower needs to be able to repay, the value of the property has to be determined in a prudential way and all required information needs to be documented. In addition to that it is assumed that the bank's qualifying revolving retail exposure fulfills Basel IV's requirements for regulatory retail, which say that the exposure needs to be revolving or a personal term loan or lease, it amounts to € 1,000,000 or less and that the exposure to one individual cannot exceed the threshold of 0.2% of the total regulatory retail portfolio. The remaining exposures have been determined as 'other retail' by the bank and will therefore be treated as other retail under Basel IV (Basel Committee on Banking Supervision, 2017).

The risk weights for the regulatory retail portfolio amount to 75% and to 100% for the other retail portfolio under the Basel IV standardised approach. The determination of risk weights is a little more complicated for real estate exposures as it depends on the fact of whether the repayment is materially dependent on the cash flows which are generated by the property or not and on the Loan-to-Value-Ratio (LTV-Ratio) of the loan itself.(Basel Committee on Banking Supervision, 2017). As the data did not comprise any information about the material dependence on the cash flows of the property, it was assumed that there exists a material dependence if the counterparty is occupied in the real estate sector. This is due to the fact that debtors, whose ability to repay depends substantially on the cash flows of the property, are usually involved in real estate development. The risk weights for loans whose repayment is not materially dependent on the cash flows of the property are shown in formula 10, whereas the ones for loans whose repayment is materially dependent are shown in formula 11. After the determination of risk weights it was possible to simulate the calculation of risk-weighted assets under the Basel IV standardised approach. The results of this simulation provide a basis for the

simulation of the Basel IV output floor, which will be described later on.

$$RW = \begin{cases} LTV \le 50\% = 20\% \\ 50\% < LTV \le 60\% = 25\% \\ 60\% < LTV \le 80\% = 30\% \\ 80\% < LTV \le 90\% = 40\% \\ 90\% < LTV \le 100\% = 50\% \\ LTV > 100\% = 70\% \end{cases}$$
(10)
$$RW = \begin{cases} LTV \le 50\% = 30\% \\ 50\% < LTV \le 60\% = 35\% \\ 60\% < LTV \le 80\% = 45\% \\ 80\% < LTV \le 90\% = 60\% \\ 90\% < LTV \le 100\% = 75\% \\ LTV > 100\% = 105\% \end{cases}$$
(11)

Simulation of the Basel IV Input and Output Floors

As already described under 4.3 Input Floor for Model Parameters, Basel IV introduced parameter floors for LGD, PD and EAD. The LGD floor for mortgages is fixed to 5% and the one for QRREs to 50%. The LGD floor for the retail class 'other retail' depends on whether the loan is secured or not and in the case where it is secured, also on the type of collateral. As most of the loans within the other retail portfolio are not fully but only partially secured, the LGD floor has to be determined using formula 8. The unsecured LGD floor (LGD FloorU) for other retail exposures is fixed to 30% whereas the secured one (LGD FloorS) depends on the collateral type and lies between 0% and 15%. (Basel Committee on Banking Supervision, 2017) As the sample does not include any information about the collateral type, it is assumed that the minimum LGD amounts to 10% for the secured portion of this retail class as this percentage applies to two out of four collateral types. As the haircut for these two collateral types amounts to 40%, this percentage is used for the calculation of the entire other retail portfolio. As already mentioned under 5.3.1 Simulation of the Basel IV Standardised approach, the bank already splits its retail portfolio into residential mortgages, qualifying revolving retail and other retail. This information was used for the determination of the applicable PD floor. The application of input floors was simulated in Microsoft Excel by using an IF formula, which is shown in formula 12 for the PD-floor and in formula 13 for the LGD floor. This formula says that in the case where the PDs or LGDs, which have been provided by the bank, undercut the respective LGD- or PDfloor, the LGD- or PD-floor must be used for the RWA calculation. In the case where they exceed the floors, the LGDs or PDs provided by the bank can be used. The EAD floor was simulated by using a CCF of at least 50% for all off balance exposures.

 $PD Basel IV = \begin{cases} PD_{Mortgages} \leq 0.05\% = 0.05\% \\ PD_{Mortgages} > 0.05\% = PD_{Mortgages} \\ PD_{QRRE} \leq 0.10\% = 0.10\% \\ PD_{QRRE} > 0.10\% = PD_{QRRE} \\ PD_{other Retail} \leq 0.05\% = 0.05\% \\ PD_{other Retail} > 0.05\% = PD_{other Retail} \end{cases}$ $LGD_{Mortgages} \leq 5\% = 5\% \\ LGD_{Mortgages} > 5\% = LGD_{Mortgages} \\ LGD_{QRRE} > 50\% = 50\% \\ LGD_{QRRE} > 50\% = LGD_{QRRE} \\ LGD_{Other Retail} \leq LGD Floor_{Other Retail} = LGD Floor_{Other Retail} \end{cases}$ (12)

After the determination of the Basel IV LGDs, PDs and EADs it was possible to calculate the correlation (p), the capital requirements (K) and the RWAs by using the formulas stated in 3.1.2 Calculation for Retail Exposures. The result of this calculation represents the RWAs following the Basel IV rules for input floors. The simulation of the output floor was done by multiplying the RWAs, which have been determined by applying the Basel IV rules for the standardised approach, with the applicable percentages of 50% in 2022 to 72.5% in 2027 (Basel Committee on Banking Supervision, 2017).

Empirical Findings

In order to anonymise the data, the findings will only be presented as percentages or ratios.

Corporate Porfolio

As the bank uses the FIRBA for the valuation of its corporate portfolio, it only estimates the PDs of its exposures but uses fixed parameters, which are provided by the regulator for its LGDs and EADs. (Basel Committee on Banking Supervision, 2004) Due to that, it was only possible to analyse the effects of the PD floor and the output floor on the RWD of the corporate portfolio.

The analysis of the bank's credit data shows that the valuation of its corporate portfolio is rather conservative as the RW density amounts to 0.879, which is close to the maximum RW density, which has been observed by Mariathasan & Merrouche (2014). It can therefore be said, that heterogeneity does not represent a big problem within this portfolio. After the simulation of Basel IV, it was possible to split the corporate portfolio into the element which is affected by the Basel IV PD floor and the element, which is not. This more detailed analysis shows that the RW density of exposures, which are not affected by the PD-Floor amounts to 1.018 and therefore even lies above the maximum RW density, which has been observed by Mariathasan & Merrouche (2014). In contrast, the RW density of exposures, which are affected by the PD floor only amounts to 0.164. The composition of this portfolio already suggests that the alterations of the Basel framework on the IRB approach for credit risks might not have huge effects on the RW density, as the exposures which are affected by the PD floor only make up 3.04% of the

entire RWAs of the corporate portfolio.

	EAD	RWA	RWD
Not affected by PD Floor	83.68%	96.96%	1.018
Affected by PD Floor	16.32%	3.04%	0.164
Total	100.00%	100.00%	0.879

Table 5 Composition of the corporate portfolio prior simulation

The application of the Basel IV SA on the credit data led to a significant RWA increase of the PD floor affected exposure by 507.18%, while it led to a slight decrease of 1.90% within the remaining portfolio. The increase does not come as a surprise, as the applied risk weights within the SA are all near (regulatory retail: 75%, other SMEs: 85%) or even at 100% (general corporates). As the majority of the PD-floor-affected exposures fall into the group of general corporates, which are risk weighted at 100%, the RWD increases from 0.164 to 0.993. The development of the exposures which are not affected by the PD floor was rather unexpected, but is also comprehensible as this effect is caused by specialised lending exposures:

- The greatest RWA reduction is observed within the project finance portfolio, as it amounts to -41.78%. This development is caused by the fact, that the Basel IV SA connects the risk weights to the project phase and not to the solvency of the exposure as the IRBA does. Therefore, the risk weights lie between 80% for high quality projects and 130% for projects in their pre-operational phase within the SA, but between 70% and 250% within the IRBA (Basel Committee on Banking Supervision, 2017).
- The RWAs of the real estate portfolio decline by 39.22%, which is due to the fact that they were risk weighted at an average of 173% within the IRBA, but received a lower risk weight within the standardised approach because of the dependence on LTV-ratios.
- The RWAs of object finance exposures decrease by 13.04%, as these exposures are generally risk weighted at 100% within the standardised approach, whereas they all received a risk weight of 115% (category: satisfactory) under the IRBA.

The general corporate exposure, which is not affected by the PD-Floor, shows an RWA increase of 6.90%, which is caused by the fact, that the SA usually leads to higher risk weighted assets than the IRBA, as the risk weights are fixed by the regulator and tend to be higher (Cucinelli et al., 2018).

The application of the new Basel IV regulations concerning the IRBA for credit risk does not show the expected results, as it leads to an RWA decrease of the exposure, which is not affected by the PD floor, by 23.95% and even the PD floor affected exposure by 1.44%. Overall the RWA reduction amounts to 18.90% and results in an RWD of 0.713. This development shows, that the application of the new Basel IV rules has a negative impact on risk weight heterogeneity, as it reduces the RWD by 0.166. The reasons for this development can clearly be attributed to the lower LGD (reduction of unsecured LGD from 0,45 to 0,40 and exposure weighted average LGD for the secured exposure) and the removal of the scaling factor of 1.06. These parameters fully affect the not PD floor affected exposure but are attenuated for the PD floor affected portfolio through the PD floor, which leads to an increase of the PD at least.

The impact of the output floor on the RWAs and the RWDs, which is shown in tables 6 and 7, is calculated by using the final percentage of 72.5%, which will be applicable from 2027 onwards. This percentage is displayed in the RWD itself, as it amounts to 0.724, which is caused by the RWDs, which amount to almost 1.000 under the SA. There is only a little difference between the RWDs after the application of the Basel IV input floors and after the application of the output floor. Compared to the initial RWD, both floors lead to a significant RWD reduction.

Still, the final RWDs lie clearly above the mean, which has been observed by Mariathasan & Merrouche (2014). In addition to that, it has already been determined in the beginning, that risk weight heterogeneity is not such a big problem within this portfolio, as the initial RWD was already close to the maximum observed RWD by Mariathasan & Merrouche (2014).

			RWA Basel IV	RWA Basel IV
Exposure	A Basel III IKBA Basel IV SA		IRBA – Input	IRBA – Output
			Floor	Floor
Not affected by PD Floor	96.96%	83.75%	96.31%	83.75%
Affected by PD Floor	3.04%	16.25%	3.69%	16.25%
Total	100.00%	100.00%	100%	100%

Table 6 RWAs prior and after simulation

Table 7 RWD prior and after simulation

	D Basel III IRBA Basel IV SA		RWD Basel IV	RWD Basel IV
Exposure			IRBA – Input	IRBA – Output
			Floor	Floor
Not affected by PD	1.018	0.000	0.820	0.724
Floor	1.010	0.999	0.820	0.724
Affected by PD Floor	0.164	0.993	0.161	0.720
Total RWD	0.879	0.998	0.713	0.723
Total Δ to Basel III	-	0.119	-0.166	-0.156

The major findings within the corporate portfolio can be summarised as follows:

- Finding 1: The application of the Basel IV input and output floors can lead to an RWD decrease if they are applied to conservatively valued portfolios. This effect is primarily caused by a majority of PDs exceeding the PD floor.
- Finding 2: Due to the fact that the removal of the scaling factor applies to the entire portfolio, whereas the PD floor only applies to a certain part of it, the new regulations lead to an RWD reduction in the case of conservatively valued portfolios.



These findings are summarised in Figure 2:



Retail Portfolio

The bank, which provided the credit data, uses the advanced IRB approach for the calculation of its RWAs within the retail portfolio. Due to that it was possible to analyse the impact of the PD floor, the LGD floor, the EAD floor and the output floor on the risk weight density of the retail portfolio. The analysis of the transmitted data already showed that, compared to the corporate portfolio, the RWD of the retail portfolio is much lower. Still the overall RWD of the retail portfolio, which amounts to 0.208, is far from the minimum RWD (0.024) which has been observed by Mariathasan and Merrouche (2014).

The retail portfolio is divided into three sub-categories: qualifying revolving retail, other retail and residential mortgage, which already belongs to the real estate exposure class within Basel IV (Basel Committee on Banking Supervision, 2017). The analysis of these subcategories shows that the RWD calculated on the basis of the non-simulated data is rather similar, as it amounts to 0.221 for the qualifying revolving retail exposure, to 0.215 for the other retail and to 0.205 for the real estate exposures – still, the development of these ratios is different.

Qualifying Revolving Retail Exposures

Since the CCFs are rather high within this category, the QRRE is only affected by the parameter floors for PD and LGD. The division into exposures, which are affected by the input floors and those which are not, shows, that the affected exposures exhibit much lower RWDs throughout the simulation. The RWD of the exposures, which are affected by the PD floor even lies below the observed minimum of Mariathasan and Merrouche (2014). The application of the Basel IV standardised approach leads to an RWD increase to 0.750 for all sub-categories, which is caused by the fact, that the entire QRRE portfolio falls within the definition of regulatory retail and therefore receives a risk weight of 75%. The RWD after the application of the output floor of 72.5% which will be applicable from 2027 onwards amounts to 0.544 for all sub-categories. This uniform RWD is caused by the consistent SA- RWD and the multiplication with the fixed value of 72.5% (Basel Committee on Banking Supervision, 2017). The application of the Basel IV input

floors leads to an interesting result, as it leads to an RWD increase for the first three subcategories, but to a reduction for the exposures, which are not affected by the LGD floor. This RWD reduction is caused by the removal of the scaling factor and the fact that within this exposure only 5.21% of the Basel IV RWAs after the application of the input floor are affected by the PD floor. In contrast to that, the exposure which is not affected by the PD floor records an RWD increase due to the fact that 20.76% of the Basel IV RWAs after the application of the input floor are affected by the LGD floor.

_	D Basel III IRBA Basel IV SA		RWD Basel IV	RWD Basel IV
Exposure			IRBA – Input	IRBA – Output
			Floor	Floor
Affected by PD Floor	0.021	0.750	0.037	0.544
Not affected by PD Floor	0.312	0.750	0.314	0.544
Affected by LGD Floor	0.136	0.750	0.187	0.544
Not affected by LGI Floor	0.250	0.750	0.242	0.544

Table 8 RWD within QRRE prior and after simulation

Residential Mortgage Exposures

The residential mortgage exposures are not affected by the application of the LGD floor, as the BCBS had already introduced a transitional 10% LGD floor for these exposures in 2006, which was then prolonged in 2009 because of the volatile mortgage portfolios faced during the financial crisis (Basel Committee on Banking Supervision, 2006). It is noticeable that the exposures, which are affected by parameter floors, all exhibit an RWD improvement, whereas all non-affected exposures exhibit an RWD reduction. The reason for the deterioration is again the removal of the scaling factor and the fact that this effect cannot be compensated by another floor. Within this sub-portfolio, the application of the output floor leads to an RWD improvement. However, it is not as strong as within the other retail sub- portfolios, which is attributable to the fact, that residential mortgage exposures usually receive a lower risk weight than QRRE or other retail exposures within the standardised approach.

 Table 9 RWD within residential mortgage prior and after simulation

	D Basel III IDBA Basel IV SA		RWD Basel IV	RWD Basel IV	
Exposure	D Dasel III IK	DA Dasci iv SA	IRBA – Input	IRBA – Output	
			Floor	Floor	
Affected by PD Floor	0.014	0.590	0.016	0.428	
Not affected by PD	0.217	0 594	0.204	0.431	
Floor	0.217	0.394	0.204	0.451	
Affected by EAD Floor	0.164	0.565	0.387	0.410	
Not affected by EAD	0.205	0.504	0 104	0 431	
Floor	0.205	0.394	0.194	0.431	

Other Retail Exposures

The risk-weight density under the Basel IV SA amounts to 1.000 for all exposures, as the uniform risk weight of 100% is applied to all exposures which come under other retail (Basel Committee

on Banking Supervision, 2017). This fact is also the reason why the RWD after the application of the 72.5% output floor amounts to 0.725 for all exposures. This sub- category shows the same results as the residential mortgage portfolio: an improvement of the RWD for exposures, which are affected from certain floors and a deterioration of the RWD for the ones which are not affected. Again, this result can be attributed to the removal of the scaling factor and the little stake of exposures, which are affected by other floors, within the exposures, which are not affected by a certain floor.

	D Basel III IPBA Basel IV SA		RWD Basel IV	RWD Basel IV
Exposure	D Dasel III INDA	Dasci IV SA	IRBA – Input	IRBA – Output
			Floor	Floor
Affected by PD Floor	0.025	1.000	0.028	0.725
Not affected by PD	0.230	1 000	0.221	0 725
Floor	0.250	1.000	0.221	0.725
Affected by LGD Floor	0.057	1.000	0.074	0.725
Not affected by LGD	0 238	1 000	0.226	0 725
Floor	0.238	1.000	0.220	0.725
Affected by EAD Floor	0.166	1.000	0.400	0.725
Not affected by EAD	0.215	1 000	0.204	0.725
Floor	0.215	1.000	0.206	0.725

Table 10 RWD within other retail prior and after simulation

Since the volume of the retail exposures, which are not affected by any parameter floor exceeds the volume of the floor affected exposures by far, the overall RWD also decreases from 0.208 to 0.198 after the application of the Basel IV input floors. The biggest RWD improvement can be observed after the application of the output floor, as it increases by 0.300 and therefore harmonises the IRBA-RWDs to the SA-RWDs.

Exposure	D Basel III IRBA + Basel IV SA		RWD Basel IV	RWD Basel IV
			IRBA – Input	IRBA – Output
			Floor	Floor
Qualifying	0.221	0.750	0 228	0.544
Revolvin			0.228	0.544
Residential Mortgage	0.205	0.594	0.194	0.431
Other Retail	0.215	1.000	0.207	0.725
Total RWD	0.208	0.701	0.198	0.509
Total Δ to Basel III	-	0.493	-0.010	0.300

Table 11 Overall RWD prior and after simulation

The major findings within the retail portfolio can be summarised as follows:

- Finding 3: The removal of the scaling factor eliminates the input floor's effect of increasing the *RWD*, as large parts of the portfolio are not affected by any input floor.
- Finding 4: The output floor leads to a significant RWD increase and an assimilation to the Basel IV SA RWDs.

These findings are summarized in Figure 3:



Figure 3 RWD Development Retail Portfolio

Discussion and Conclusion

As already mentioned, the aim of this article is to analyse the effects of the Basel IV input and output floors on the heterogeneity of risk weights. In this context, heterogeneity describes the fact that identical assets are assigned different risk weights (Turk-Ariss, 2017). This effect has been measured by researchers like Beltratti and Paladino (2016), Mariathasan and Merrouche (2014) and Vallascas and Hagendorff (2013) using the ratio of risk- weighted-assets to total assets, namely the risk weight density, which is also used in this article. In order to provide a comparison with other banks, the results of the simulation are compared with the findings of Mariathasan and Merrouche (2014) as they observed the RWD development of a sample of 115 IRB-banks in 21 OECD-countries.

The simulation of Basel IV results in mixed findings concerning the development of the risk weight density. The extent to which the RWD is affected by the Basel IV regulation is illustrated in Figure 4. The figure shows an overall RWD reduction for the corporate portfolio, but an RWD increase for the retail portfolio. In the Empirical Findings section, these results have already been discussed in detail. In brief, Basel IV leads to an RWD increase for portfolios which are valued optimistically or exhibit a low Basel III RWD, and to an RWD decrease for the opposite. Moreover, the removal of the scaling factor leads to stronger effects than the input floor in case of conservatively valued portfolios and therefore might encourage an RWD reduction.



Figure 4 RWD Development Overall Portfolio

Following these results, there is only a two-part answer to research questions number two and three, which ask whether Basel IV is able to reduce the heterogeneity of risk weights and how the RWD is affected by the new rules: Taking into account findings one to three, which highlight the fact that the Basel IV rules might lead to an RWD decrease if they are applied to conservatively valued portfolios, suggests that Basel IV has a rather negative impact on risk weight variability. However, one also needs to consider the fact that conservatively valued portfolios are not usually affected by the problem of heterogeneity at all. This is also confirmed by the comparison of the Basel III RWD of the corporate portfolio (0.879) with the observed mean RWD (0.516) by Mariathasan and Merrouche (2014). On the contrary, finding number four clearly shows that the Basel IV regulations, especially the output floor, can lead to a significant RWD increase and therefore also to a reduction of risk weight heterogeneity. This statement can be confirmed by the fact that Basel IV results in an RWD of the retail portfolio that amounts to 0.509 and therefore almost matches the observed mean by Mariathasan and Merrouche (2014). Summarising it can be said, that the reduction of risk weight heterogeneity is reserved for optimistically valued portfolios, which exhibit a low RWD starting value, as conservatively valued portfolios show a counteracting effect. Within the observed portfolios, the input floors only have a very limited influence on the RWD, which can be attributed to the conservative valuation as well as the counteracting effect of the removal of the scaling factor. On the contrary, the Basel IV output floors can have a tremendous effect on RWD. The detailed effects of the Basel IV rules on RWD are shown in the Empirical Findings section. The before mentioned aspects can be summarized in the following implications:

Implication 1 – The Basel IV regulations have the potential to reduce the problem of heterogeneity within risk weights in certain portfolios.

Implication 2 - The improvement of risk weight heterogeneity is reserved for optimistically valued portfolios. Conservatively valued portfolios, which usually do not face the problem of RW heterogeneity, are likely to be confronted with an RWD decrease.

Implication 3 – Out of all analysed Basel IV rules, the output floor clearly has the biggest influence on risk weight density. The input floors' effect is very limited within optimistically valued portfolios and is even eliminated by the removal of the scaling factor within conservatively valued portfolios.

Unfortunately, these implications do not come without criticism: In literature there is a consensus that the excessive variability of risk weights has to be addressed by the BCBS. However, some articles criticise the BCBS because of its willingness to introduce a 'system of floors', which is seen as a restriction of the primary advantage of the IRBA: risk sensitivity. Moreover, researchers remark that capital floors are similar to 'flat tax' regulations, which are not desirable as they impose high capital charges on low-risk exposures. This fact may encourage banks to invest into high-risk exposures again (Haselmann & Wahrenburg, 2016; Resti, 2016). The literature criticises the BCBS' approach of restricting the IRBA, but also admits that the reduction of RW heterogeneity is difficult to curb. An alternative solution to RW heterogeneity, which has been brought up by research, is the one of clarifying the concepts of standards and guidelines. The current regulations and definitions leave considerable space for interpretation, which lead to different ways of applying the IRBA and therefore also to variable results. The comparison of IRBA-RWAs and SA-RWAs is also suggested by researchers, but rather as a way of revealing the weaknesses of the IRBA than as a constraint of the approach itself. Moreover, a peer-review process, which should be accompanied by supervisory authorities is considered to be an appropriate way to disclose differences between banks and countries. An aspect that is clearly said to be the key to restore the credibility of the IRBA is transparency (Resti, 2016).

As already described, the findings are based on the simulation of Basel IV. By using a simulation technique, the effects of Basel IV on the overall credit portfolio but also on risk weight density can be displayed very well, as each influencing parameter can be steered individually. The simulation technique and Microsoft Excel 2016 complement each other very well, as Excel provides the possibility to analyse the data in detail: the drill down to the single datasets allows the user to find the origin of any change in the overall results and therefore a detailed analysis of the driving parameters. The detailed description of how the simulation was performed and therefore also the answer to research question number one are included in previous sections.

During the process of writing this article, the following aspects were identified as potential fields for future research:

- 1. The results of this simulation rely on the data of a conservative Western European IRBbank. Therefore, this article does not contain any information about the effect of Basel IV on the heterogeneity of risk weights of large multinational banks. This aspect should be addressed by future research, as the results of this evaluation will probably look very different.
- 2. Another aspect, which has not been covered by this article, is the effect of the Basel IV regulations on corporate portfolios under the AIRBA. This article only analysed the impact of the PD-floor, but future research should also cover the LGD- and the EAD-floor.
- 3. It has already been mentioned multiple times that the bank, which provided the credit data, follows a rather conservative risk policy. For this reason, future research should also look at the bank's risk policy and its influence on RWD.

The contribution to the literature, which has been generated by condensing the afore mentioned implications, includes the information that Basel IV will definitely have an impact on RWD. Whether this impact is a positive or a negative one, depends on the preceding valuation of the portfolio. The change in RWD will also lead to a concurrent change in risk-weighted assets

and therefore also in the level of eligible capital. The findings within the retail portfolio confirm those of the EBA study, which have already suggested that Basel IV and especially the output floor will lead to significant a increase of risk capital (European Banking Authority, 2018).

References

- Abbassi, P., & Schmidt, M. (2018). A comprehensive view on risk reporting: Evidence from supervisory data. *Journal of Financial Intermediation*, *36*, 74–85.
- Aikman, D., Galesic, M., Gigerenzer, G., Kapadia, S., Katsikopoulos, K., Kothiyal, A., et al. (2014). Taking uncertainty seriously: simplicity versus complexity in financial regulation: Financial Stability Paper No. 28 - May 2014.
- Amorello, L. (2016). Beyond the Horizon of Banking Regulation: What to Expect from Basel IV. *Harvard International Law Journal*, 58, 21–38.
- Andersen, H. (2011). Procyclical implications of Basel II: Can the cyclicality of capital requirements be contained? *Journal of Financial Stability*, 7(3), 138–154.
- Barakova, I., & Palvia, A. (2014). Do banks' internal Basel risk estimates reflect risk? *Journal of Financial Stability*, 13, 167–179.
- Barucci, E., & Milani, C. (2018). Do European banks manipulate risk weights? *International Review* of Financial Analysis, 59, 47–57.
- Basel Committee on Banking Supervision (2004). *International convergence of capital measurement and capital standards: A revised framework* (Updated Nov. 2005). Basel: Bank for International Settlements.
- Basel Committee on Banking Supervision (2006). *International convergence of capital measurement and capital standards: A revised framework* (Comprehensive version). Basel: Bank for Internat. Settlements.
- Basel Committee on Banking Supervision (2010). *Basel III: a global regulatory framework for more resilient banks and banking systems* (December 2010 (rev. June 2011)). Basel: Bank for International Settlements.
- Basel Committee on Banking Supervision (2016). *Reducing variation in credit risk- weighted assets: Constraints on the use of internal model approaches; consultative document.* Basel: Bank for International Settlements.
- Basel Committee on Banking Supervision (2017). *Basel III: Finalising post-crisis reforms*. Basel: Bank for International Settlements.
- Behn, M., Haselmann, R., & Vig, V. (2016a). *The limits of model-based regulation. Working paper series / European Central Bank: no 1928 (July 2016).* Frankfurt am Main, Germany: European Central Bank.
- Behn, M., Haselmann, R., & Wachtel, P. (2016b). Procyclical Capital Regulation and Lending. *The Journal of Finance*, *71*(2), 919–956.
- Bellotti, T. (2010). A simulation study of Basel II expected loss distributions for a portfolio of credit cards. *Journal of Financial Services Marketing*, 14(4), 268–277.
- Beltratti, A., & Paladino, G. (2016). Basel II and regulatory arbitrage. Evidence from financial crises. *Journal of Empirical Finance*, *39*, 180–196.
- Ben Naceur, S., Marton, K., & Roulet, C. (2018). Basel III and bank-lending: Evidence from the United States and Europe. *Journal of Financial Stability*, *39*, 1–27.
- Berg, T., & Koziol, P. (2017). An analysis of the consistency of banks' internal ratings. *Journal of Banking & Finance*, 78, 27–41.
- Blundel-Wignall, A., & Atkinson, P. (2010). Thinking Beyond Basel III: Necessary Solutions for Capital and Liquidity. *OECD Journal: Financial Market Trends*. (1), 1–23.
- Bodellini, M. (2019). The long 'journey' of banks from Basel I to Basel IV: has the banking

system become more sound and resilient than it used to be? ERA Forum, 39, 180.

- Cabrera, M., Dwyer, G. P., & Nieto, M. J. (2018). The G-20's regulatory agenda and banks' risk. *Journal of Financial Stability*, *39*, 66–78.
- Chen, Y., Wang, Z., & Zhang, Z. (2019). Mark to market value at risk. Journal of *Econometrics*, 208(1), 299–321.
- Cizel, J., Rijken, H. A., Altman, E. I., & Wierts, P. (2017). Assessing Basel III Capital Ratios: Do risk weights matter?
- Cucinelli, D., Di Battista, M. L., Marchese, M., & Nieri, L. (2018). Credit risk in European banks: The bright side of the internal ratings based approach. *Journal of Banking & Finance*, 93, 213–229.
- Danielsson, J., James, K. R., Valenzuela, M., & Zer, I. (2016). Model risk of risk models. *Journal of Financial Stability*, 23, 79–91.
- Estrella, A., Park, S., & Peristiani, S. (2000). Capital Ratios as Predictors of Bank Failure. *Federal Reserve Bank of New York Economic Policy Review*, 6(2), 33–52.
- European Banking Authority (2018). *Basel III Monitoring Exercise: Results based on data as of 31 December* 2017, from https://eba.europa.eu/documents/10180/2380948/2018+Basel+III+Monitoring+Exercise +Report.pdf.
- Haldane, A. G. (April 2013). *Constraining discretion in bank regulation*. Federal Reserve Bank of Atlanta Conference on 'Maintaining Financial Stability: Holding a Tiger by the Tail(s)', Federal Reserve Bank of Atlanta.
- Haselmann, R., & Wahrenburg, M. (2016). *Banks' internal rating models time for a change? The "system of floors" as proposed by the Basel Committee*. European Parliament.
- Hellwig, M. (2010). *Capital Regulation after the Crisis: Business as Usual.* Bonn: Max Planck Institute for Research on Collective Goods Bonn.
- Hull, J., Mader, W., & Wagner, M. (2014). *Risikomanagement: Banken, Versicherungen und andere Finanzinstitutionen* (3., aktualisierte Aufl.). *wi Wirtschaft.* Hallbergmoos: Pearson.
- Joosen, B. P.M. (2016). Revisions to the Standardised Approach for Bank Credit Risk Measurement. *Journal of International Banking Law and Regulation, Forthcoming,* from https://ssrn.com/abstract=2715444.
- Kiema, I., & Jokivuolle, E. (2014). Does a leverage ratio requirement increase bank stability? *Journal of Banking & Finance, 39*, 240–254.
- Lopez, J. A. (2002). The empirical relationship between average asset correlation, firm probability of default and asset size. *SSRN Electronic Journal*.
- Ly, K. C., & Shimizu, K. (2018). *Did Basel regulations cause a significant procyclicity?* (Working Papers 2018-06). Swansea University, School of Management.
- Mariathasan, M., & Merrouche, O. (2014). The manipulation of basel risk-weights. *Journal of Financial Intermediation*, 23(3), 300–321.
- Mayes, D. G., & Stremmel, H. (2012). *The Effectiveness of Capital Adequacy Measures in Predicting Bank Distress*.
- Pérez Montes, C., Trucharte Artigas, C., Cristófoli, M. E., & Lavín San Segundo, N. (2018). The impact of the IRB approach on the risk weights of European banks. *Journal of Financial Stability*, 39, 147–166.
- Peura, S., & Jokivuolle, E. (2004). Simulation based stress tests of banks' regulatory capital adequacy. *Journal of Banking & Finance, 28*(8), 1801–1824.
- Plosser, M. C., & Santos, J. A. C. (2018). Banks' Incentives and Inconsistent Risk Models. *The Review* of Financial Studies, 31(6), 2080–2112.
- Rajan, U., Seru, A., & Vig, V. (2015). The failure of models that predict failure: Distance, incentives, and defaults. *Journal of Financial Economics*, 115(2), 237–260.
- Repullo, R., & Suarez, J. (2012). The procyclical effects of bank capital regulation. Journal of

Financial Studies, 26, 452–490.

- Resti, A. (2016). Bank internal ratings: are capital floors a suitable tool to restore their credibility? *Journal of Financial Management, Markets and Institutions*. (4), 179–194.
- Rossignolo, A. F., Fethi, M. D., & Shaban, M. (2013). Market crises and Basel capital requirements: Could Basel III have been different? Evidence from Portugal, Ireland, Greece and Spain (PIGS). *Journal of Banking & Finance*, 37(5), 1323–1339.
- Schneider, S., Schröckl, G., Koch, S., & Schneider, R. (2017). Basel "IV": What's next for banks?: Global Risk Practice April 2017.
- Sonali, D., & Amadou N.R., S. (2012). *How risky are bank's risk weighted assets? Evidence from the Financial Crisis: IMF Working Paper.*
- Strickland, O. J. (2017). Seeking to have Banks Sing to the Same Tune: the Basel Committee Addresses Credit Risk-Weighted Assets. *University of Miami Business Law Review*, 95, 95–124, from https://repository.law.miami.edu/umblr/vol26/iss1/6.
- Tarullo, D. K. (2014). Rethinking the aims of prudential regulation: Speech at the Federal Reserve Bank of Chicago Bank Structure Conference, Chicago, Illinois.
- Thakor, A. V. (2018). Post-crisis regulatory reform in banking: Address insolvency risk, not illiquidity! *Journal of Financial Stability*, 37, 107–111.
- Turk-Ariss, R. (2017). *Heterogeneity of Banks Risk Weights in the EU: Evidence by Asset Class and Country of Counterparty Exposure*. International Monetary Fund.
- Vallascas, F., & Hagendorff, J. (2013). The Risk Sensitivity of Capital Requirements: Evidence from an International Sample of Large Banks. *Review of Finance*, *17*(6), 1947–1988.
- Zoia, M. G., Biffi, P., & Nicolussi, F. (2018). Value at risk and expected shortfall based on Gram-Charlier-like expansions. *Journal of Banking & Finance*, 93, 92–104.



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings

homepage: http://www.acrn.eu/frap



How Social is Islamic Banking?

M. Luthfi Hamidi¹, Andrew C. Worthington*,²

¹ Jawa Barat, Indonesia, Griffith University ² Griffith University

Abstract

While the central tenets of Islamic banking relating to usury, profit-and-loss sharing, haram (sinful) business, gharar (uncertainty) and maisir (gambling and speculation) are well known, there is less awareness of the implied commitment to improved social outcomes. This paper outlines the argument for social outcomes as an objective for Islamic banks and investigates whether social failure currently exists in the Islamic banking industry by assessing it against this performance dimension. For this purpose, we employ data from 12 Islamic commercial banks, 7 Islamic banking units, and 7 Islamic rural banks operating in Indonesia during the period 2015–16. Using content analysis, we find that social failure is evident in all Islamic commercial banks and banking units, but in only one of the seven Islamic rural banks where most banks appear to pursue social outcomes at the accommodative level (accepting and doing all that is required). Using a social outcome-weighted asset formulation, we reveal that Islamic banking in Indonesia has to a certain extent improved in meeting its social objectives over time, but sometimes at the cost of other objectives relating to the environment and customers. The lack of progress in this area may lead some to question whether Islamic banking is delivering or indeed able to deliver on its social commitment.

Keywords: Islamic banking; social banking; Indonesia.

JEL Classification: G21; C49.

Introduction

Islamic banking is generally considered to be more social oriented than conventional banking (Asutay and Harningtyas, 2015; Haniffa and Hudaib, 2007; Hasan, 2005) given the oft-quoted objective of promoting a just, fair, and balanced society. In reality, however, Islamic banking has been criticised as demonstrating "social failure" given its small positive impact on desirable social outcomes for ordinary people (Asutay, 2007; Asutay, 2008; Mohd Nor, 2016) and its progressive shift from achieving, among other things, improvements in falah (human wellbeing) to a focus on conventional profit maximisation (Abdul-Rahman et al., 2014; Cebeci, 2012; Abdul-Baki and Uthman, 2017).

In investigating the social dimension of Islamic banking, some studies assess corporate social responsibility (Mallin et al., 2014; Aribi and Gao, 2010; Amran et al., 2017; Abdul Rahman et al., 2010; Hassan and Harahap, 2010), others social and ethical reporting (Maali et al., 2006; Kamla and G. Rammal, 2013; Haniffa and Hudaib, 2007), and yet others how social considerations have been inserted into performance measurement based on Maqasid al-Syari'ah (Asutay and Harningtyas, 2015; Hasan, 2005; Bedoui and Mansour, 2015; Hameed et al., 2004; Mohamed et al., 2008; Mohamada et al., 2016). The results mostly find that Islamic banking per se has little impact on social outcomes at least beyond that of comparable conventional banks.

In this paper, we examine whether the perception of social failure in Islamic banking is real. Using a case study approach of Islamic banking in a single country (Indonesia), we first identify the desired social outcomes of Islamic banking. These include social impact, programs, and policies aimed at managing social concerns held among stakeholders (Wood, 1991). We define social outcomes here as the observable results of the implementation of Islamic banking social objectives. With a broad Islamic banking framework (and Islamic economics more generally) these comprise promoting a just, fair, and balanced society, reinforcing social solidarities and strengthening values (Ahmad, 2000; Umer, 1985; Naqvi, 2016; Siddiqui, 2001; Tripp, 2006). In terms of specific outcomes, Hamidi and Worthington (2017) highlight actual social actions of Islamic banking such as the delivery of qardh al-hasan (benevolent loans), emphasising charity through zakat (alms giving), and avoiding any potentially harmful financing to the environment and society. We then measure the social responsiveness of Islamic banking and subsequently determine a rating each bank's attitude toward a social orientation ranging from denying to anticipating responsibility. This yields the actual social dimension level of individual banks and allows us to assess further whether these Islamic banks are meeting their social promise.

Our analysis differs from previous studies at least in three aspects. First, in terms of assessing social responsibility, all previous studies with the exception of Kamla and G. Rammal (2013), employ content analysis with a quantitative dimension relating to the use of certain words or the number of sentences or pages. Mostly, this is a mere dichotomous approach to assess constructs, whereby a construct will be given a score of one if present in the text, otherwise zero (Haniffa and Hudaib, 2007; Amran et al., 2017; Mallin et al., 2014). By itself, this kind of assessment is questionable and best balanced with qualitative evaluation.

Second, existing analysis of Islamic banking social disclosure largely employs Accounting and Auditing Organization for Islamic Financial Institutions (AAOIFI)—a not-for-profit organization established to maintain and promote Shari'ah standards for Islamic financial institutions, participants, and the overall industry—guidelines. For instance, Haniffa and Hudaib (2007) and Hassan and Harahap (2010) address the main themes of social performance relating to strategy, governance, products, community, employment, environment, and R&D. It is quite strange then that these measures neglecting any indication of the piousness and devoutness of Islamic banking management and staff. For this reason, we extend the prevailing framework to include religiosity as this may also influence an Islamic bank's attitude toward social responsibility (Ibrahim et al., 2008; Ramasamy et al., 2010).

Finally, existing studies have been extremely limited in the scope of analysis, with most considering social dimensions in only Islamic commercial banks (in Indonesia, Bank Umum Syariah). In this study, along with 12 Islamic commercial banks, we consider the social dimension of performance in seven Islamic business units (Unit Usaha Syariah), and seven Islamic rural banks (Bank Pembiayaan Rakyat Syariah). This provides a much more comprehensive account of how social objectives and outcomes vary across Islamic banking in all its manifestations.

The remainder of the paper is structured as follows. The following section provides a review of the background and the related literature. The third and fourth section discusse the methodology and present the findings. The final section concludes.

Social Objectives in Islamic Banking

Islamic banking

Islamic banking is a form of banking that conforms or complies with Shari'ah (Islamic law). One feature is the prohibition of riba (charging interest or usury) (Chong and Liu, 2009; Chapra, 2007). Another principal feature is the promotion of profit-and-loss sharing (PLS) in which earning of

profit should be based on economic activities (Zineb and Bellalah, 2013) and borrowers should not bear risk alone, hence promote equity and justice in the society (Khan, 2010; Al-Jarhi, 2004). In addition, Islamic banking avoids haram (sinful) business, including not financing firms that produce and sell alcohol, drugs and intoxicants, pork-related product, armaments, and indecent entertainment (Zineb & Bellalah 2013). The law also prohibits any gharar (uncertainty and high risk that may cause injustice to other parties) and maisir (gambling and speculation) to encourage businesses to act on an ethical basis as well as avoiding injustice among the parties involved in permitting no allowance for asymmetric information (Zineb & Bellalah 2013).

In 1963, the Mit Ghamr local savings bank was the first Islamic bank established in Egypt, coinciding with the creation of the Tabung Haji (Pilgrim's Savings Corporation) in Malaysia (Zineb & Bellalah 2013). Since then, Islamic banking has expanded significantly in both countries, their regions, and worldwide. According to Rating (2018) global Islamic finance assets were US\$2 trillion by the end of 2016, and estimated to grow to around US\$3.8 trillion by 2022 (Maierbrugger, 2018).

Under profit-and-loss sharing (PLS) schemes, there are several different financing contracts, including mudharabah (a partnership contract, wherein one side provides capital to other side or partner) and musyarakah (joint ventures). Theoretically, the use of mudharabah and musyarakah are expected to develop actual businesses, as a means of absorbing unemployment, creating future business opportunities, and protecting Islamic banking from risky financial engineering. However, in practice up to 90% of Islamic banking financing in some countries rely on murabahah or debt-like contracts (Vogel & Hayes 1998; Agarwal & Yousef 2000; Ali 2011).

Alongside these products and services, other Islamic banking services like zakat and qardh alhasan have been identified as supporting the needy and alleviating poverty, all good social objectives. Nonetheless, Islamic banking still faces accusations that it has failed to deliver on its promises of social responsibility. This apparent failure in demonstrating social ends seems to suggest Islamic banking prioritises better-off customers than the community as a whole (Asutay 2007). Consequently, the services of IBs may have no significant impact on the life of the marginalised in society (Asutay & Zaman 2009). Moreover, given the need to accumulate profit to ensure their long-term sustainability, Islamic banks are moving towards fulfilling a legal form to meet their Fiqh (jurisprudence) limitations, but increasingly ignoring their primary Shari'ah objectives (Nienhaus 2011; Ahmed 2011). As a result, the ethical dimension of Islamic banking may not not be fully integrated in existing banking practice (Mansour et al. 2015; Musa 2015). As a result, Maqashid al-Syari'ah (the objectives of Islamic law) as the central guidance of delivering "benefits and preventing harm" might be undelivered, thus the maximum benefits for social and developmental outcomes are not possible.

Social objectives in Islamic banking

Tulsian and Pandey (2008) classify business objectives into four categories: economic, social, human, and national. They further assert that the major social objectives include: (1) the supply of the desired quality of products; (2) the avoidance of antisocial and unfair trade practices; (3) the generation of employment; (4) the provision for the welfare of employees; (5) the avoidance of slums and pollution; and (6) contributing to the general welfare of society. Serving social objectives, rather than only maximizing profits, is also desirable in the banking industry. In particular, if commercial banks fail to promote social welfare, some may come to see that regulation, including at the extreme state-owned enterprises, are an alternative way to address market failure and to contribute to economic development and general societal welfare (Stiglitz, 1993).

In relation to social objectives, Islamic scholars are divided into two opposing groups on the role of social objectives in Islamic banking. A relatively few believe that Islamic banks should be permitted to operate as normal business entities as long as their operations comply with with Syari'ah (Lewis and Algaoud, 2001; Satkunasegaran, 2003). This may mean, for example, a reliance on zakat as the main social contribution, but only if such contributions do not put owner funds at risk.

In contrast, the majority of Islamic scholars consider that Islamic banking should aim at supporting and encouraging an economic system aimed at promoting a just, fair, and balanced society (Ahmad, 2000; Naqvi, 2016; Siddiqui, 2001). In this way, any economic interaction in society is to "…increase the sum of public happiness, whilst reinforcing social solidarities and strengthening their values" (Tripp, 2006, p. 119). Similarly, Al-Zuhayli (2003, p. 250) stated "…the primary goal of Islamic financial institutions is not profit-making, but the endorsement of social goals of socioeconomic development and the alleviation of poverty."

To realize these values, Islamic banks should include these tenets in their daily activities, and not place an emphasis solely on financial profit but also distributive justice and fulfilling societal obligation (Hasan, 2005; Taqi-Usmani, 2002). Furthermore, Islamic banking is expected to be more socially responsible than conventional banks (Asutay and Harningtyas, 2015; Haniffa and Hudaib, 2007). As a result, Kamel (1997) argues that Islamic goals should include objectives relating to economic development, the creation of value-added, more exports and fewer imports, job creation, the rehabilitation of the incapacitated, and the training of others.

To date, this strong view of the social responsibility of Islamic has found legislative support in just a few countries, including Indonesia and Jordan (Hamidi and Worthington, 2017). In evidence, the Undang-undang Republik Indonesia Nomor 21 Tahun 2008 tentang Perbankan Syariah (Islamic banking Acts), explicitly encourages Islamic banks to perform a social function, while Law No. 28 Year 2000 Banking Law in Jordan requires that Islamic banks not only provide banking services, but help revive social solidarity and to provide qardh al-hasan for the public interest.

Corporate and social responsibility (CSR) in Islamic banking

The interrelation between business and its social dimension has been observed for decades (Schlusberg, 1969; Davis, 1967; Heald, 1970), which some refer to as corporate social responsibility (CSR), being "...a concept whereby companies integrate social and environmental concerns in their business operations and in their interaction with their stakeholders on a voluntary basis (Communities, 2001). On this, Friedman (1970) makes a distinction between the social responsibility of individuals and that of business:

If he wishes to spend his money on such purposes, that is his right and I cannot see that there is any objection to his doing so...There is one and only one social responsibility of business—to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game, which is to say, engages in open and free competition without deception or fraud.

In contrast, Steiner (1975) asserted that the concept of CSR covers not only economic aspects of business, but also legal behaviour and voluntary activities. In relation on how a corporate respond the social issue, Frederick (1978) introduced what he called as Corporate Social Responsiveness (CSR₂). It associates with the way of corporate in responding social pressure, ranging from no response (do nothing) to proactive (do much). Carroll (1979) described a social responsibility including economic, legal, ethical, and discretionary responsibilities. She extended the concept of CSR into Corporate Social Performance (CSP) by proposing three-dimensional conceptual model consisting of social responsibility, social responsiveness, and social issues.
Wartick and Cochran (1985), then, classified CSR₂ into four degree responses: reactive, defensive, accommodative, and proactive. Clarkson (1995) further transformed CSR₂ into RDAP Scale. Besides this, he differentiated between social issues and stakeholder issues because corporations and their managers more willingly engage with stakeholders than with society. Accordingly, he proposed to evaluate CSP using a stakeholder framework. Graves and Waddock (1994) suggested to gauge CSP by using CSP index. They assessed the index by employing data provided by Kinder, Lydenberg, Domini Research & Analytics (KLD). KLD evaluates companies on eight dimensions including community relations, employee relations, environment, product, treatment of women and minorities, military contracts, nuclear power, and South Africa involvement.

Aguilera et al. (2007) argues that stakeholders support CSR for three main reasons: (1) instrumental (self-interest driven, (2) relational (group member relationship), and (3) moral (moral beliefs and ethical values). In relation to the moral motive, a few studies consider the role of religion in business. Hunt and Vitell (1993), for example, highlights how religion affects ethical decision making, while Ibrahim et al. (2008) finds that religiousness has a significant impact on the economic, ethical, and responsibility of business. Elsewhere, Ramasamy et al. (2010) suggest that there is direct relation between consumer religiosity and business CSR.

CSR from an Islamic perspective somewhat differs from the Western viewpoint. Dusuki (2005) suggests that Islamic CSR draws on the concept of falah (literally meaning success), as demonstrated by the achievement of material wellbeing in the material world and eternal life in the hereafter. Others offer CSR based on the Islamic principle of Tawhid. This requires a belief in Allah as the one and only God by surrendering all human actions to His commands (Farook and Lanis, 2007; Zain et al., 2014; Basah and Yusuf, 2013). As God is the Creator, Owner, and Sustainer of all creatures and universe, all human behaviour, attitudes, and activities, including business should be accountable before Allah.

Only a few existing studies investigate the operation of CSR in Islamic banking. Maali et al. (2006) employ a disclosure index method to assess 29 Islamic banks in 16 countries and reveal that social reporting is much less than that expected, with only those banks paying Zakat demonstrating consistently good social disclosure. Elsewhere, Haniffa and Hudaib (2007) examine the ethical identity of seven Islamic banks and find that these failed to meet the ideal Islamic ethical identity in at least four dimensions: commitment to society, disclosure of corporate vision and mission, role in managing Zakah, charity, and benevolent loans, and top management information. In other work, Kamla and G. Rammal (2013) use in-depth content analysis to investigate the social reporting (through annual reports and websites) of ten Islamic banks and reveal that these banks appear to commit little to helping eradicate poverty or strengthening social justice in society. Lastly, Mallin et al. (2014) sample 90 Islamic banks across 13 countries, finding that they pay rather less attention to the environment than expected, while Amran et al. (2017) compare the CSR of Islamic banks in Indonesia and Malaysia and show progress in both countries, mostly to do with the workplace and community.

Methodology

For our analysis, we use data on Indonesian Islamic banks derived from the annual reports of 12 (of the 13 in total) Islamic commercial banks (BUS) (Bank Muamalat Indonesia, Bank Victoria Syariah, BRI Syariah, BJB Syariah, BNI Syariah, BNI Syariah, Bank Syariah Mandiri, Bank Mega Syariah, Bank Panin Dubai Syariah, Bank Panin Dubai Syariah, Bank Syariah Bukopin, BCA Syariah, Bank Maybank Syariah Indonesia, and BTPN Syariah), seven (of the 20) Islamic business units (UUS) (UUS Danamon, UUS Bank Kalbar, UUS Bank DIY, UUS BPD Jateng, UUS CIMB-Niaga, UUS BPD Sumbar, UUS Permata), and the largest (2015/16 assets > IDR100 billion) seven Islamic rural banks (BPRS) (BPRS HIK Parahyangan, BPRS Bhakti

Sumekar, BPRS Amanah Ummah, BPRS Al-Ma'soem, BPRS Dinar Ashri, BPRS HIK Bekasi, BPRS Suriyah). In the absence of any annual report or any of the information therein (particularly for IRBs), following Kamla and G. Rammal (2013) and Mallin et al. (2014) we use each institution's website to obtain details on their financial and other social activities in conjunction with the financial reports submitted to the Indonesian Financial Services Authority (https://www.ojk.go.id).

We employ content analysis to observe the social outcomes of these Islamic banks in their various forms. According to Barelson (1952) as cited by Hsieh and Shannon (2005), content analysis is widely used as quantitative research method, although initially researchers employed it as both a qualitative and quantitative method. For the most part, content analysis analyses textual data (Cavanagh, 1997), but can also be used in relation to images, maps, and numerical contents (Krippendorff, 2004). The challenging part of content analysis is to select and develop categories so that content can be grouped and classified. Traditionally, content analysis is most appropriate for revealing new phenomena, so researchers specify categories collected from the data (Kondracki et al., 2002). However, we use directed content analysis and this requires theory to identify the key concepts or categories needed for the initial coding (Potter and Levine-Donnerstein, 1999).

We identify the categories from the prevailing Islamic banking literature, particularly in relation to social objectives and disclosure (Haniffa and Hudaib, 2007; Maali et al., 2006; Rashid et al., 2013; Hassan and Harahap, 2010). We augment these by including the KLD index of corporate social performance (Graves and Waddock, 1994; Mattingly and Berman, 2006; Waddock and Graves, 1997), the 17 Sustainable Development Goals (SDGs) and the 5 Environment Social Governance (ESG) Scorecard developed by Oikocredit International (2018), a worldwide cooperative and social investor group, as shown in Figure 1.



Figure 1: Social Performance Indicators for Islamic banking

Six traits constitute social outcomes as desired results reflecting the social objectives of Islamic banks. The first dimension is religiosity which contains four constructs: (a) Zakat; (b) Qardh al-Hasan; (c) Riba; and (d) encouraging spirituality. The second theme is the environmental dimension, which can be classified into three constructs: (e) preserving and contributing to environment; (f) managing a green office; and (g) managing green financing. The third dimension

is social ends constituting three constructs: (h) engagement in the community in terms of education and training; (i) quality improvements; and (j) economic empowerment. The fourth is governance matters, which cover four constructs: (k) vision and mission; (l) preventing corruption and fraud; (m) handling customer complaints; (n) ethics and staff compliance. The fourth category is how they manage their employees, and includes three (o) rewards; (p) equal work opportunity; and (q) training and education for employees. Lastly, customer orientation relates to (t) serving unserved area; (s) having diversified financial products, particularly for women and the disadvantaged; and (s) giving discount for customers. In total, there are 6 dimensions with 21 constructs, as shown in Table 1.

No.	Dimensions	Items	Constructs
	Religiosity	а	Zakat
1		b	Qardh al-Hasan
1		С	Riba
		d	Encouraging spirituality
		e	Preserve and contribute to environment
2	Environmental	f	Green office
		g	Green Financing
		h	Community education and training
3	Social	i	Community quality improvement
		i	Community economic empowerment
		k	Vision or Mission
1	Governance	1	Prevent corruption and fraud
7	Governance	m	Handling customer complaints
		n	Ethics and Staff compliance
		0	Reward
5	Employees	р	Work opportunity equality
		q	Training and education for employees
		r	Serving unserved area
6	Customers	S	Diversified financial product for women or disadvantaged group
		t	Discount for customers

Table 1: Social Outcome Dimensions and Constructs

While well established, content analysis is not without its challenges. According to Weber (1990) there are two crucial issues relating to the reliability and validity of the data collected. Reliability issues may arise due to vagueness of meanings or category definition, while validity problems are linked to the extent the result is consistent with the existing theory. To enhance reliability, Holsti (1969) proposes multiple coders. To address this, we use three coders to score the research constructs (see Appendix A). To improve validity, we use the categories and definitions from both the Islamic banking and corporate social performance literatures (Appendix B). The scoring of constructs, follows Scaltrito (2015) in using a dichotomous and quantitative approach. To assess social outcomes, we employ an equal-weighted index to measure the score for each bank under each dimension:

$$\frac{\sum_{t=1}^{n_j} X_{ij}}{n_i}$$

where n_j is the number of items or constructs disclosed by the jth bank, where $n_j = 20$, and $X_{ij} = 2$ if the ith construct is both qualitative and quantitative, $X_{ij} = 1$ if the construct is only qualitative, and $X_{ij} = 0$ if neither.

To rank the social responsiveness of each bank, we use a four-category rating: reactive, defensive, accommodative, and proactive (RDAP Scale) as introduced by Carroll (1979), Wartick

and Cochran (1985), Clarkson (1995), and further modified by Hamidi and Worthington (2018), as shown in Table 2.

Achievement (%) Rating		Posture or Strategy	Performance
0–40	Reactive	Deny responsibility	Doing less than required
41–50	Defensive	Admit responsibility but fight it	Doing the least that is required
51–60	Accommodative	Accept responsibility	Doing all that is required
61–100	Proactive	Anticipate responsibility	Doing more than is required

Table 2. RDAP Scale for Assessing Social Responsiveness

To investigate whether there is social failure across banks overall, we use a social outcome weighted asset formulation as follows:

$$\overline{\mathbf{x}} = \frac{\sum_{i=1}^{n} \mathbf{x}_{i} \mathbf{w}_{i}}{\sum_{i=1}^{n} \mathbf{w}_{i}}$$

where \bar{x} is social outcome weighted assets, x_i is the value of social outcome i and w_i is the asset weighting of x_i . If banks underperform, indicated by a rating of either reactive or defensive, then social failure occurs. Otherwise, the banks are socially acceptable.

Results

Table 3 provides the results for Islamic Commercial Banks (BUS), all of which have improved their social outcome rating from 2015 to 2016, except for BMI and BCA Syariah. The best social outcome performance is BNI Syariah with achievements of 72.5% and 77.5% in 2015 and 2016, respectively, followed in second and third place by BSM with 70% and 72.5% and BTPN Syariah with 57.5% and 72.5%). The worst performing BUS is Bank Victoria Syariah with a score of only 12.5% in 2015, albeit with the largest increase (80%) to 22.5% in 2016. Overall, of the 12 BUS, only 5 are performing in terms of social outcomes, all with ratings of either accommodative (accepting responsibility or doing all that is required) or proactive (anticipating responsibility and doing more than required). This is improving, such that by 2016, seven BUS are performing in terms of social outcomes.

			Growth		Perfor		
	Sco	re (%)	(%)	Rating	m	Rating	Perform
	2015	2016		2015		2016	
	62.					Accommodativ	
BMI	5	60.0	-4.00	Proactive	Yes	е	Yes
	12.						
Bank Victoria Syariah	5	22.5	80.00	Reactive	No	Reactive	No
	45.						
BRI Syariah	0	62.5	38.89	Defensive	No	Proactive	Yes
	<i>32</i> .						
BJB Syariah	5	42.5	30.77	Reactive	No	Defensive	No
	72.						
BNI Syariah	5	77.5	6.90	Proactive	Yes	Proactive	Yes
BSM	70	72.5	3.57	Proactive	Yes	Proactive	Yes
	42.					Accommodativ	
Bank Mega Syariah	5	52.5	23.53	Defensive	No	е	Yes

Table 3: BUS social outcome ratings

			Growth		Perfor		
	Sco	re (%)	(%)	Rating	m	Rating	Perform
	2015	2016		2015		2016	
				Accommodativ		Accommodativ	
Bank Panin Syariah	55	60	9.09	е	Yes	е	Yes
Bank Bukopin Syariah	45	47.5	5.56	Defensive	No	Defensive	No
BCA Syariah	35	32.5	-7.14	Reactive	No	Reactive	No
Maybank Syariah	25 57.	32.5	30.00	Reactive Accommodativ	No	Reactive	No
BTPN Syariah	5	72.5	26.09	е	Yes	Proactive	Yes

While the performance of the BUS in terms of social outcomes continue to increase, there are some neglected dimensions such as the environment with only an average achievement of 43% across the 12 BUS in 2016. While nearly half of the BUS give attention to environmental awareness and make serious efforts to enhance green offices, they mostly fail to provide green financing. This finding is consistent with Asutay and Harningtyas (2015) and Mallin et al. (2014) in that Islamic banks tend to be less attentive to the environment. Another dimension that is severely unaddressed is customer's benefits, scoring a very low social outcome of just 9.72%. This indicates that the BUS in Indonesia ignore three key elements: serving unserved areas, creating financial products for disadvantaged groups, and rewarding customers with discount. To serve remote areas, a few of these banks engage in a linkage strategy by disbursing funds to microfinance, but only few have created particular service such as creating and providing banking products to the poor. Furthermore, banks that give a discount to customers who want to speed up payment from due time or mugasah are almost non-existent, although the National Syariah Council (DSN) has made a decision based on religious consideration (fatwa) Number: 23/DSN-MUI/III/2002 that it is lawful to provide a deduction from the payment obligation, as long as it is not stipulated in the contract.

In regards to religiosity, the average score is quite good in the sense that most of the BUS collect zakat (70.8%) and distribute qardh al-hasan (91.6%), but only a few campaign actively against riba (25%) and promote spirituality in their daily business activities (33.3%), hence the average is only about 55.2%. The fact that the BUS keep an eye on zakat and qardh al hasan is consistent with the results in Maali et al. (2006). In respect to governance, this has the highest average score (70.83%) among the remaining dimensions. In detail, the weakest part concerns vision and mission, with only one third of the BUS specifically identifying that they are eager to promote social and environment goals, a finding similar to Haniffa and Hudaib (2007).

In stark contrast to the BUS, the Islamic banks units (UUS) are much more aligned with only one failing to meet its social outcomes, as shown in Table 3. Indeed, within a very short time, two UUS have moved from being accommodative (doing all that is required) to proactive (doing more than what is required). The reason behind this remarkable achievement is probably because IBUs receive benefit from their parent banks, with almost all CSR activities taken over by their conventional commercial banking parent, which most existing research suggests tend to perform better than Islamic banks in the environmental dimensions. It is no wonder their performance in term of the environmental dimension (59.52%) is better than BUS (43.06%).

However, when we further scrutinise the customer benefit dimension, a similar pattern arises as both BUS and UUS underperform, scoring just 9.72% and 11.9% respectively in 2016 (Figures 2 and 3). Another interesting result is that religiosity in UUS (41.07%) is lower than the BUS (55.21%). However, this is not unexpected as UUS campaigning against riba would be at odds with their commercial parent's interest-based operations. Furthermore, all UUS appear to ignore activities relating to spirituality with no publicly available evidence suggesting that they encourage regular prayer, sermons, or reciting the Qur'an.



Figure 2: Average Social Outcomes of BUS (%)



Figure 3: Average Social Outcomes of IBUs (%)

Finally, in the case of Islamic rural banks (BPRS), social outcome performance is unambiguously poor as all of their ratings are reactive, indicating that they either deny responsibility or pursue social outcomes less than what is required, as shown in Table 4. This is somewhat ironic in that by default and design BPRS are expected to serve the poor and disadvantaged. There are a number of possibilities of why this is the case/ First, BPRS publish their annual report as it is an obligatory based on Bank Indonesia Regulation PBI No. 7/47/PBI/2005. The report should consist of two specific parts: public information and the financial report. Unfortunately, most BPRS only meet the second not the first report requirement. As a consequence, some relevant information such as vision and mission, business development, strategies and policies taken by management are missing.

	Scor	re (%)	Growth	Rating	Perform	n Rating	Perform
	2015	2016	(%)	2015		2016	
				Reacti			
BPRS HIK Parahyangan	30	32.5	8.33	ve	No	Reactive	No
	27.			Reacti			
BPRS Bhakti Sumekar	5	30	9.09	ve	No	Reactive	No
				Reacti			
BPRS Amanah Ummah	20	20	0.00	ve	No	Reactive	No
	12.			Reacti			
BPRS Al-Ma'soem Syari'ah	5	17.5	40.00	ve	No	Reactive	No
				Reacti			
BPRS Dinar Ashri	15	20	33.33	ve	No	Reactive	No
				Reacti			
BPRS HIK Cibitung	15	15	0.00	ve	No	Reactive	No
2	12.			Reacti			
BPRS Suriyah	5	12.5	0.00	ve	No	Reactive	No

Table 4: Social Outcomes of IRBs (2015-2016)

Second, we could expect that the social outcomes of BPRS are trivial because their financial performance is quite poor and therefore to sustain any efforts, however well-intentioned, in this regard. As Seibel (2007) argues, many BPRS are managed by retired conventional bankers weak in Islamic banking skills and this may weaken their core performance. In fact, this explanation might not fully fit, since the ratio of operational expense to operating profit (Rasio Biaya Operasional terhadap Pendapatan Operasional/BOPO) of BPRS even smaller than that for the BUS, suggesting that the BPRS are relatively more efficient, as shown in Figure 4.



Figure 4: Efficiency of ICBs, IBUs, and IRBs, 2014–2017, measured by BOPO (%)

Third, the BPRS are very small when compared with the BUS and even the IBU. The total assets of the 101 BPRS is a only about IDR 10.8 trillion or one-sixth of the largest bank's (BMI) assets. Given the relatively small size, this could hamper BPRS their efforts in providing a particular social program or project. Another explanation is that BPRS face fierce competition from conventional rural banks as well as commercial banks that have been established to cater the rural segment such as Bank Rakyat Indonesia (BRI). Although Muslim customers may prefer to use Islamic microfinance services, in reality aspects such as easiness, service speed, and nearness may outweigh the religious motive (Masyita and Ahmed, 2011). In this regard, all resources of IRBs seem to be allocated to win the customers, thus only limited time, efforts, and

funds to serve community and disadvantaged group let alone pay considerable attention on environment aspect.

Lastly, we assess whether all Islamic banks in Indonesia fail in delivering social outcomes Based on the SOWA assessment, based on the figures in Table 5 we conclude that in 2015, the observed banks fail as their SOWA is less than 50%, thus achieving only a defensive rating (admitting responsibility but resisting it and doing the least that is required). However, this situation reverses in 2016, as some of BUS perform better with average social outcomes of 53.50%. This result is consistent with Mallin et al. (2014) and the findings concerning CSR 53.87%. Hence, the social failures of Islamic banks somewhat improve in 2016. However, all Islamic banks continue to need to address poor social performance in terms of the dimensions of environment and customer benefits.

Assets	(IDR						
Billion)			(%)	SO (x _i)		SOWA (Rating)	
2015	2016	2015	2016	2015	2016	2015	2016
186,371	235,425	84.5	85.07	46.25	52.92	47.41	53.50
32,336	39,020	14.7	14.10	55.71	58.93	(Defensive)	(Accommodative)
1839.8	2285	0.8	0.83	18.93	21.07		
220,547	276,730	100.0	100.00				
	Assets Billi 2015 186,371 32,336 1839.8 220,547	Assets (IDR Billion 2015 2016 186,371 235,425 32,336 39,020 1839.8 2285 220,547 276,730	Assets (IDR WA Billion WA 2015 2016 2015 186,371 235,425 84.5 32,336 39,020 14.7 1839.8 2285 0.8 220,547 276,730 100.0	Assets (IDR Billion WX (%) 2015 2016 2015 2016 186,371 235,425 84.5 85.07 32,336 39,020 14.7 14.10 1839.8 2285 0.8 0.83 220,547 276,730 100.0 100.00	Assets (IDR Billion) WA (%) SO (x) 2015 2016 2015 2016 2015 186,371 235,425 84.5 85.07 46.25 32,336 39,020 14.7 14.10 55.71 1839.8 2285 0.8 0.83 18.93 220,547 276,730 100.0 100.00 100.00	Assets (IDR Billion WA (%) SO (x _i) 2015 2016 2015 2016 186,371 235,425 84.5 85.07 46.25 52.92 32,336 39,020 14.7 14.10 55.71 58.93 1839.8 2285 0.8 0.83 18.93 21.07 220,547 276,730 100.0 100.00 100.00	Assets (IDR Billion WA (%) SO (x _i) SOW 2015 2016 2015 2016 2015 2016 2015 186,371 235,425 84.5 85.07 46.25 52.92 47.41 32,336 39,020 14.7 14.10 55.71 58.93 (Defensive) 1839.8 2285 0.8 0.83 18.93 21.07 220,547 276,730 100.0 100.00

Table 5: Aggregate Islamic Bank Social Outcomes (2015–2016)

Note: WA – weighted assets; SO (xi) represents SO element of xi; SOWA is the social outcome weighted-asset average.

Conclusion

This paper identifies the social outcomes of three different types of Islamic banks in Indonesia in 2015–15, comprising 12 Islamic commercial banks (BUS), 7 Islamic business units (UUS), and the 7 largest Islamic rural banks (BPRS). Employing content analysis, we confirm that social failure take places in all samples, as the observed banks on average earn the lowest rating, namely, reactive (doing the least that is required). As a general rule, The UUS perform better than the BUS and the BUS better than the BPRS. However, all Islamic banks in Indonesia appear to pay little attention to the environment and customer benefits.

This paper contributes to the small but evolving literature on the ability of Islamic banks to deliver social outcomes. In contrast to nearly all other dimensions of Islamic banking performance, this is little studied, but often used to support the introduction and expansion of Islamic banking systems, particularly in developing countries. There are two clear implications for institution, industry and regulatory policy. First, the Islamic banking sector needs to improve its social outcomes, particularly in relation to the environment and customer benefits. Failure to do so may result in a deterioration in support of the poor and ordinary people. Second, = supervisory bodies needs to regulate and provide incentives for the industry to address these issues. As future research, these findings need to be further investigated from the perspective of stakeholders, and this may confirm the empirical results of this content analysis.

References

Abdul-Baki, Z. and Uthman, A. B. (2017). Exploring the "Social Failures" of Islamic Banks: A Historical Dialectics Analysis. Journal of Islamic Accounting and Business Research, 8, 250-271.

Abdul-Rahman, A., Latif, R. A., Muda, R. and Abdullah, M. A. (2014). Failure and Potential of Profit-Loss Sharing Contracts: A Perspective of New Institutional, Economic (Nie) Theory. Pacific-Basin Finance Journal, 28, 136-151.

- Abdul Rahman, A., Asraf, M. F. and Bakar, F. A. (2010). Corporate Social Reporting: A Preliminary Study of Bank Islam Malaysia Berhad (Bimb). Issues in Social and Environmental Accounting, 4, 18-39.
- Ahmad, K. (2000). Islamic Finance and Banking: The Challenge and Prospects. Review of Islamic economics, 57-82.
- Al-Jarhi, M. A. (2004). Islamic Banking and Finance: Philosophical Underpinnings. SS Ali, & A. Ahmad, Islamic Banking and Finance: Fundamentals and Contemporary Issues, 13-24.
- Amran, A., Fauzi, H., Purwanto, Y., Darus, F., Yusoff, H., Zain, M. M., Naim, D. M. A. and Nejati, M. (2017). Social Responsibility Disclosure in Islamic Banks: A Comparative Study of Indonesia and Malaysia. Journal of Financial Reporting and Accounting, 15, 99-115.
- Aribi, Z. A. and Gao, S. (2010). Corporate Social Responsibility Disclosure: A Comparison between Islamic and Conventional Financial Institutions. Journal of Financial Reporting and Accounting, 8, 72-91.
- Asutay, M. and Harningtyas, A. F. (2015). Developing Maqasid Al-Shari'ah Index to Evaluate Social Performance of Islamic Banks: A Conceptual and Empirical Attempt. International Journal of Islamic Economics and Finance Studies, 1.
- Basah, M. Y. A. and Yusuf, M. M. (2013). Islamic Bank and Corporate Social Responsibility (Csr). European Journal of Business and Management, 5, 194-209.
- Bedoui, H. E. and Mansour, W. (2015). Performance and Maqasid Al-Shari'ah's Pentagon-Shaped Ethical Measurement. Science and engineering ethics, 21, 555-576.
- Carroll, A. B. (1979). A Three-Dimensional Conceptual Model of Corporate Performance. The Academy of Management Review, 4, 497-505.
- Cavanagh, S. (1997). Content Analysis: Concepts, Methods and Applications. Nurse researcher, 4, 5-13.
- Cebeci, I. (2012). Integrating the Social Maslaha into Islamic Finance. Accounting Research Journal, 25, 166-184.
- Chapra, M. U. (2007). The Case against Interest: Is It Compelling? Thunderbird International Business Review, 49, 161-186.
- Chong, B. S. and Liu, M.-H. (2009). Islamic Banking: Interest-Free or Interest-Based? Pacific-Basin Finance Journal, 17, 125-144.
- Clarkson, M. E. (1995). A Stakeholder Framework for Analyzing and Evaluating Corporate Social Performance. Academy of management review, 20, 92-117.
- Communities, C. O. T. E. 2001. Promoting a European Framework for Corporate Social Responsibility. Brussel.
- Davis, K. (1967). Understanding the Social Responsibility Puzzle. Business horizons, 10, 45-50.
- Farook, S. and Lanis, R. (2007). Banking on Islam? Determinants of Corporate Social Responsibility Disclosure. Islamic Economics and Finance, 217.
- Frederick, W. C. (1978). From Csr₁ to Csr₂ the Maturing of Business-and-Society Thought. Working Paper No. 279. Graduated School of Business, University of Pittsburgh.
- Friedman, M. (1970). The Social Responsibility of Business Is to Increase Its Profits. The New York Times.
- Graves, S. B. and Waddock, S. A. (1994). Institutional Owners and Corporate Social Performance. Academy of Management journal, 37, 1034-1046.
- Hameed, S., Ade, M. I., Bakhtiar, W., Mohamed, A., Nazli, M. N. and Sigit, P. (Year) Published. Alternative Disclosure & Performance Measures for Islamic Banks. Second Conference on Administrative Sciences: Meeting the Challenges of the Globalization Age, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia, 2004. 19-21.
- Hamidi, M. L. and Worthington, A. C. (2017). Islamic Banking Plus Social Banking Equals Islamic Social Banking: An Equation in the Making. In: JERZY, K. (ed.) Banking: Services, Opportunities and Risks. Nova Science Publishers, Inc.
- Haniffa, R. and Hudaib, M. (2007). Exploring the Ethical Identity of Islamic Banks Via Communication in Annual Reports. Journal of Business Ethics, 76, 97-116.
- Hasan, Z. (2005). Evaluation of Islamic Banking Performance: On the Current Use of Econometric Models.
- Hassan, A. and Harahap, S. S. (2010). Exploring Corporate Social Responsibility Disclosure: The Case of Islamic Banks. International Journal of Islamic and Middle Eastern Finance and Management, 3, 203-227.
- Heald, M. (1970). The Social Responsibilities of Business: Company and Community 1900-1960, Transaction Publishers.
- Holsti, O. R. (1969). Content Analysis for the Social Sciences and Humanities.
- Hsieh, H.-F. and Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. Qualitative health research, 15, 1277-1288.
- Hunt, S. D. and Vitell, S. J. (1993). Ôthe General Theory of Marketing Ethics: A Retrospective and Revisionõ. Ethics in Marketing (Irwin Inc., Homewood, IL), 775-784.
- Ibrahim, N. A., Howard, D. P. and Angelidis, J. P. (2008). The Relationship between Religiousness and Corporate Social Responsibility Orientation: Are There Differences between Business Managers and Students? Journal of Business Ethics, 78, 165-174.

International, O. (n.d.). Measuring Social Performance [Online]. Available: https://www.oikocredit.coop/whatwe-do/social-return/measuring-social-performance [Accessed 24 May 2018].

- Kamla, R. and G. Rammal, H. (2013). Social Reporting by Islamic Banks: Does Social Justice Matter? Accounting, Auditing & Accountability Journal, 26, 911-945.
- Khan, F. (2010). How 'Islamic'is Islamic Banking? Journal of Economic Behavior & Organization, 76, 805-820.

Kondracki, N. L., Wellman, N. S. and Amundson, D. R. (2002). Content Analysis: Review of Methods and Their Applications in Nutrition Education. Journal of nutrition education and behavior, 34, 224-230.

Krippendorff, K. (2004). Content Analysis: An Introduction to Its Methodology, Sage.

Lewis, M. K. and Algaoud, L. M. (2001). Islamic Banking. Books.

Maali, B., Casson, P. and Napier, C. (2006). Social Reporting by Islamic Banks. Abacus, 42, 266-289.

Maierbrugger, A. (2018). Islamic Finance Assets Seen at \$3.8tn by 2022. Gulf Times, January 30.

- Mallin, C., Farag, H. and Ow-Yong, K. (2014). Corporate Social Responsibility and Financial Performance in Islamic Banks. Journal of Economic Behavior & Organization, 103, S21-S38.
- Masyita, D. and Ahmed, H. (Year) Published. Why Is Growth of Islamic Microfinance Lower Than Conventional? A Comparative Study of the Preferences and Perceptions of the Clients of Islamic and Conventional Microfinance Institutions' in Indonesia. 8th International Conference on Islamic Economics and Finance, 2011. 1-22.
- Mattingly, J. E. and Berman, S. L. (2006). Measurement of Corporate Social Action Discovering Taxonomy in the Kinder Lydenburg Domini Ratings Data. Business & Society, 45, 20-46.
- Mohamada, M. H. S., Alic, M. A. and Sharif, R. a. M. (2016). Determinants of Maqasid Al-Shari'ah-Based Performance Measurement Practices: The Case of Malaysian Islamic Banks. International Journal of Economics, Management and Accounting, 24, 49.
- Mohamed, M. O., Abdul Razak, D. and Md Taib, F. (2008). The Performance Measures of Islamic Banking Based on the Maqasid Framework.
- Naqvi, S. N. H. (2016). Perspectives on Morality and Human Well-Being: A Contribution to Islamic Economics, Kube Publishing Ltd.
- Potter, W. J. and Levine-Donnerstein, D. (1999). Rethinking Validity and Reliability in Content Analysis.
- Ramasamy, B., Yeung, M. C. and Au, A. K. (2010). Consumer Support for Corporate Social Responsibility (Csr): The Role of Religion and Values. Journal of Business Ethics, 91, 61-72.
- Rashid, M., Abdeljawad, I., Manisah Ngalim, S. and Kabir Hassan, M. (2013). Customer-Centric Corporate Social Responsibility: A Framework for Islamic Banks on Ethical Efficiency. Management Research Review, 36, 359-378.
- Rating, S. P. G. (2018). Islamic Finance Outlook 2018 Edition [Online]. Available: https://www.spratings.com/documents/20184/4521646/Islamic+Finance+2018+Digital-1.pdf/cf025a76-0a23-46d6-9528-cecde80e84c8 [Accessed 18 May 2018].

Satkunasegaran, E. (Year) Published. Corporate Governance and the Protection of Customers of Islamic Banks. Proceedings of the International Islamic Banking Conference, 2003.

- Scaltrito, D. (2015). Assessing Disclosure Quality: A Methodological Issue. Journal of Modern Accounting and Auditing, 11, 466-475.
- Schlusberg, M. D. (1969). Corporate Legitimacy and Social Responsibility: The Role of Law. California Management Review, 12, 65-76.
- Siddiqui, S. H. (2001). Islamic Banking: True Modes of Financing. New Horizon, 109, 15-20.
- Steiner, G. A. (1975). Business and Society, New York, Random House.
- Taqi-Usmani, M. (2002). An Introduction to Islamic Finance (Arab and Islamic Laws). Boston: CQ.
- Tripp, C. (2006). Islam and the Moral Economy: The Challenge of Capitalism, Cambridge University Press.
- Umer, C. M. (1985). Towards a Just Monetary System, International Institute of Islamic Thought (IIIT).

Waddock, S. A. and Graves, S. B. (1997). The Corporate Social Performance-Financial Performance Link. Strategic management journal, 303-319.

Wartick, S. L. and Cochran, P. L. (1985). The Evolution of the Corporate Social Performance Model. The Academy of Management Review, 10, 758-769.

Weber, R. P. (1990). Basic Content Analysis, Sage.

- Wood, D. J. (1991). Corporate Social Performance Revisited. Academy of management review, 16, 691-718.
- Zain, M. M., Darus, F., Yusoff, H., Amran, A., Fauzi, H., Purwanto, Y. and Naim, D. A. (2014). Corporate Ibadah: An Islamic Perspective of Corporate Social Responsibility. Middle-East Journal of Scientific Research, 22, 225-232.
- Zineb, S. and Bellalah, M. (2013). Introduction to Islamic Finance and Islamic Banking: From Theory to Innovations. In: BELLALAH, M. and MASOOD, O. (eds.) Islamic Banking and Finance. Newcastle-upon-Tyne, United Kingdom: Cambridge Scholars Publishing.



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings



homepage: http://www.acrn.eu/journals

Children`s Toy or Grown-Ups` Gamble? LEGO Sets as an Alternative Investment

Nikita Shimkus^{1,2}, Savva Shanaev^{*,1}, Satish Sharma¹, Binam Ghimire¹

¹ University of Northumbria at Newcastle, United Kingdom

² University of St. Andrews, United Kingdom

Abstract

In this article, LEGO sets are studied as a potential alternative asset class. An exhaustive sample of 10,588 sets is utilised to generate inferences regarding long-term LEGO performance, its diversification benefits, and return determinants. Over 1966-2018, LEGO value-weighted index accounted for survivorship bias enjoys 1.20% inflation-adjusted return per annum, well below 5.54% for equities. However, the defensive properties of LEGO are considerable, as including 5-25% of LEGO in a diversified portfolio is beneficial in terms of Sharpe ratio and certainty equivalent metrics for investors with varying levels of risk aversion. LEGO secondary market is relatively internationalised, with investors from 95 countries engaging in 90,000 transactions over a 6-month period, comprising total trading volume of \$7.5 million. Larger economies, countries with higher per capita incomes and less income inequality are shown to trade LEGO more actively. There is evidence of LEGO is not exposed to any of the Fama-French factors, however, set-specific size and value effects are also well-pronounced on the LEGO market, with smaller sets and sets with lower price-to-piece ratio exhibiting higher yields. Older sets are also enjoying higher returns, demonstrating a liquidity effect.

Keywords: LEGO; alternative investment; collectable; performance evaluation; risk factor

JEL codes: G11, G12, G15, F21

Introduction and Literature Review

People have been collecting rarities and luxury items for centuries, turning mere tokens into legitimate asset classes traded on large-scale secondary markets. High-net-worth individuals globally enjoy the opportunity to transform their hobbies into a source of capital gains in the long term. Collectible cars, stamps, fine art, rare wine and whisky, and other luxuries have earned popularity as an alternative way to allocate funds (Jorda et al., 2017). Such non-financial assets, or "investment of passion", provide impressive risk-return trade-offs and can be a powerful portfolio diversification tool in an ever-increasing environment of global asset markets interconnectedness (Dimson and Spaenjers, 2014; Dimson et al., 2018). A considerable body of literature studying performance, diversification, and defensive properties of fine art (Goetzmann, 1993; Mandel, 2009; Goetzmann et al., 2011), books (Stein, 1977), wine (Burton and Jacobsen, 2001; Fogarty, 2010; Dimson et al., 2015; Lucey and Devine, 2015), whisky (Moroz and Pecchioli, in press), stamps (Veld and Veld-Merkoulova, 2007; Dimson and Spaenjers, 2011),

diamonds (Renneboog and Spaenjers, 2012), violins (Graddy and Margolis, 2011), and cryptocurrencies (Briere et al., 2015; Chuen et al., 2017) has been developed.

This study explores the emergent yet rather popular alternative asset class – LEGO sets. Dobrynskaya and Kishilova (2018) examine the performance of 2,322 LEGO sets over 1987-2015, finding that they yield 8.0% average annual real return and are virtually unexposed to Fama-French risk factors or broad market indices. Dobrynskaya and Kishilova (2018) remains the only piece of research of LEGO investment properties to date, so this study seeks to fill the gaps in the literature and address some limitations of the aforementioned paper.

The literature on alternative asset classes remains rather fragmented and isolated, with a few various theoretical links examining and explaining similarities in their performance. Most studies find real returns of art, stamps, and wine to exceed the risk-free rate and most bond yields but lag behind equity returns. Mandel (2009) undertakes a meta-analysis of fine art return studies and reports a median real return of 2.6% per annum, while annual wine return averages around 2.0% (Burton and Jacobsen, 2001), both slightly lower yet still similar to 2.9% enjoyed by stamps (Dimson and Spaenjers, 2011) and coloured diamonds (Renneboog and Spaenjers, 2012), and 3.3% exhibited by French and Italian violins (Graddy and Margolis, 2011). Whisky (Moroz and Pecchioli, in press), cryptocurrencies (Chuen et al., 2017), and, indeed, LEGO (Dobrynskaya and Kishilova, 2018) remain unsolved puzzles, as they offer both double-digit annual returns, thus outperforming equities, and low correlation with various conventional benchmarks and risk factors, possibly attributable to the relative immaturity of these asset classes. Potential explanations for high returns in various estimations include survivorship bias, weighting methodology, negative carry associated with storage or transportation, transaction costs, and unaccounted risk factors, such as high likelihood of forgery especially prominent on the art market (Bocart and Oosterlinck, 2011). For the LEGO market, though, fakes are very unlikely as all sets are easily identifiable at a low cost, and transaction fees, while material at 9.15% plus a \$25 fixed monthly subscription fee to participate on the secondary market, cannot alone explain the prominent outperformance (Dobrynskaya and Kishilova, 2018). As such, nominal average LEGO returns fall from 18.5% down to 14.7% when accounted for transaction fees, corresponding to a still substantial 6.36% annual real return. Dimson et al. (2015) report wine returns net of transaction costs and arrive at impressive 4.1% per annum, acknowledging that their main limitation might be success bias, as only particular vineries included into auction trades where they retrieve the data from, however, some robustness checks do generally reinforce their findings, albeit indirectly. Therefore, two potential shortcomings of Dobrynskaya and Kishilova (2018) methodology that can account for this puzzle are equal-weighted portfolios and survivorship bias. As noted by Fama and French (2008), equal-weighted portfolios based on sorts show higher returns than their value-weighted counterparts, however the results are less robust, while the equal-weighted strategy is also much harder to replicate in an actual investment strategy. requiring frequent rebalancing, corresponding transaction fees, and liquidity risk. Therefore, reporting value-weighted returns has recently become a strong convention in the asset-pricing literature. Furthermore, Dobrynskaya and Kishilova (2018) report and use in their estimations only the returns of 2,322 LEGO sets that are actively traded on the market. Presumably, there have been sets that are not currently being traded due to low or lack of collector value, and thus the estimations that disregard this effect will suffer from a form of survivorship bias, exaggerating the magnitude of LEGO returns. Of course, if an investor knows in hindsight which of the sets currently available on the primary market will yield greater value on the secondary market in the future, they would be able to enjoy high abnormal returns. However, if they pursue a passive strategy, purchasing all sets (or a representative mix of sets, thus "tracking" the market), as suggested by Dobrynskaya and Kishilova (2018), they would necessarily suffer the losses from not being able to sell some (and perhaps most of) the sets at a profit to collectors. This can potentially explain the low proportion of "loser" sets (9.5%) in Dobrynskaya and Kishilova (2018) sample – the greatest "losers" simply would not be traded on the secondary market as their market price would not exceed the consumer value of LEGO as a toy or even the cost of plastic used to manufacture those "unlucky" sets.

The existence and relevance of survivorship bias in alternative investment studies is highlighted by Mandel (2009) – in his meta-analysis, studies that focus on a narrow subsample of art, such as American impressionism (Mei and Moses, 2002) or Picasso paintings (Czujack, 1997) report significantly higher returns (4.9% and 8.3%, respectively) than the 2.6% median real return found in the cross-section of studies. An even earlier meta-analysis (Burton and Jacobsen, 1999) produces a higher median return of paintings and drawings at 4.63% and of prints at 1.51%, with smaller and more restricted samples, such as conceptual art, land art, minimal art (Ginsburgh and Penders, 1997), and modern Chinese art (Mok et al., 1993) yielding substantially higher returns, commonly at double-digits. Goetzmann (1996) sought to overcome the survivorship bias and the unavailability of the data for the whole universe by randomising his sample across various collections, arriving at a real return of 2.42% per annum, consistent with the theory, with the broadest available collectible art sample from his earlier study (Goetzmann, 1993) exhibiting an even lower annual real return of 2.00%, and Pesando (1993) broad sample of modern art yielding 1.51%. Wine and stamps in the Burton and Jacobsen (1999) meta-analysis demonstrate 3.31% and 3.45% returns, respectively, generally consistent with the recent literature (Dimson and Spaenjers, 2011; Lucey and Devine, 2015). Overall, survivorship bias is widely considered one of the most serious methodological challenges in the alternative investment studies (Goetzmann, 1996; Burton and Jacobsen. 1999; Mandel, 2009). This study fills the gap by addressing the aforementioned shortcomings of existing literature.

The existing research on alternative investments' market efficiency generally reports various forms of inefficiency, short-term return predictability, underreaction to relevant events, and long-term memory being the most common efficient market hypothesis violations (David et al., 2013; Urquhart, 2016), and illiquidity of these assets cited as the most likely cause of such behaviour of alternative investments' returns (Wei, 2018). This leads to virtually all research in the field focusing on long-term performance of the assets and portfolios comprised of them.

Generally all alternative investments exhibit zero or low positive correlation with broad financial markets and are exposed to moderate levels of idiosyncratic volatility, having significant diversification or "safe-haven" properties (Dimson and Spaenjers, 2011; Auer and Schuhmacher, 2013; Briere et al., 2015; Oosterlinck, 2017). Violins (Graddy and Margolis, 2011) even show negative correlations with major benchmarks, highlighting potential hedging opportunities. The returns of these asset classes and their defensive nature are primarily explained through the lens of four theories: consumption CAPM (Ait-Sahalia et al., 2004; Hiraki et al., 2009), conspicuous consumption (Veblen, 2005), non-pecuniary utility (Baumol, 1986), and discretion (Oosterlinck, 2017).

Conspicuous consumption and C-CAPM theories of alternative investments are closely related and often jointly used to address booms in demand for a particular asset class. Veblen (2005) first developed the theory of conspicuous consumption suggesting that affluent individuals might derive utility from luxury items based on their high price alone as a signal of social status. This concept is widely utilised in the studies of "investments of passion", with increases in agents' wealth and growing inequality of income distribution used as predictors of fine art returns and trading volumes in Japan (Hiraki et al., 2009). Dimson and Spaenjers (2011) hint at the existence of the wealth effect on the obviously less luxurious market for collectible stamps. However, their evidence is at best indirect, as they use positive correlation with equity returns to arrive at this conclusion, not testing the hypothesis directly (Dimson and Spaenjers, 2011). Dobrynskaya and Kishilova (2018), in turn, speculate that economic growth and wealth inequality in China and Russia could contribute to the increasing popularity of LEGO as an investment vehicle. This assertion is not self-evident, though, as LEGO sets are also relatively affordable, therefore this study seeks to test income and inequality determinants of investment in LEGO explicitly. C-CAPM, in turn, suggests that agents utilise investment portfolios to smoothen their consumption over business cycles, and thus assets that yield consumer as well as investment value should appreciate disproportionately during economic expansions (Ait-Sahalia et al., 2004). This rationale is plausible in explaining, among others, outstanding whisky returns in post-recession times (Moroz and Pecchioli, in press) and white diamonds' performance between 1999 and 2009 (Renneboog and Spaenjers, 2012).

However, if C-CAPM or conspicuous consumption are the only drivers of alternative asset classes' performance, it is still unclear where their well-documented defensive properties originate (Graddy and Margolis, 2011; Lucey and Devine, 2015; Oosterlinck, 2017). An evident explanation would be that those asset classes are "investments of passion" that are behaving idiosyncratically (Dimson and Spaenjers, 2014). A more well-refined theory has been recently developed by Oosterlinck (2017) to explain superior performance of fine art during the German occupation of France. He argues that as during wartime or other turbulent times the government tends to impose various excessive controls and regulations onto conventional financial trading, currency exchange, real estate market, and capital flows, works of art or other luxurious items might be the best store of value under such constraints, as owners have more discretion over them. Oosterlinck (2017) reports that French stock and real estate markets were heavily disrupted during the occupation, therefore art became a comparatively liquid asset protected from potential extortion. A well-known anecdote from the annals of the history of economic thought supports this view: when economist Ludwig von Mises was financially struggling in the United States in 1940 and could not transfer funds from his British bank account due to stringent foreign exchange and capital controls, his friend and former student Friedrich Hayek, who worked at the London School of Economics at that time, withdrew cash from Mises' account, purchased the first edition of the Wealth of Nations and two complete sets of work by Jeremy Bentham and sent these valuable books to his mentor so he could sell them in New York and temporarily subsist on the proceeds (Hulsmann, 2007). According to Mises-Hayek written correspondence analysed by Hulsmann (2007), the two used this means of international payments up until 1948, when the wartime exchange controls were finally lifted. This historical episode reinforces the discretionary property of luxurious items that can contribute to defensive performance characteristics, particularly abnormal performance during turbulence times. Interestingly, Oosterlinck (2017) argues that similar discretion-related reasons can explain the high value of art, collectibles, and, most notably, cryptocurrencies, even in peacetime as they can famously be utilised as tools for tax evasion or money laundering (Bryans, 2014). With regards to cryptocurrencies, this source of value can be even more important as they not only provide money laundering opportunities, but also can facilitate access to various grey markets (Foley et al., 2019), and cryptocurrency prices negatively react to the development of anti-money laundering legislation or industry regulation in general (Shanaev et al., in press).

Non-pecuniary utility theory is usually utilised to explain lower yields and higher prices of alternative "investment of passion" as compared to their conventional counterparts. Baumol (1986) suggests that particular assets (e.g., fine art) can be significantly overpriced as investors derive non-pecuniary utility from holding it rather than from its risk-return characteristics. Most recently, this framework has been applied to socially responsible investments, as conscious investors might want to forgo some of the financial gains to comply with their ethical standards, with evidence documented mostly on the mutual fund level (Bollen, 2007; Borgers and Pownall, 2014). Baker et al. (2018) have formalised this argument into a non-pecuniary utility augmented CAPM, with a subset of investors obtaining utility (or disutility) from holding various assets, and have applied it to the green municipal bond market in the US. As this group faces the trade-off between financial and non-pecuniary utility, they will overweigh into green bonds (or stocks of socially responsible companies, or works of fine art, etc.), and they become overvalued with

respect to their conventional counterparts, leading to a subsequent decrease in yields. To the best knowledge of authors, there has been to explicit testing of this theory with regards to various alternative asset classes, including LEGO, in the literature, therefore this study aims at filling this gap.

Data and Methodology

The study utilises the widely known LEGO set database of the secondary LEGO marketplace brickset.com to retrieve data on their fundamental characteristics (year of launch, piece count, initial retail price, dimensions – length, width, and depth, weight, general theme and specific topic of the set), current market prices (in USD), and current transactions in sets among market participants (90,000 transactions across 95 countries for the publically available time period September 2018 – March 2019). The initial retail price is quoted either in USD on in GBP, if the set was released solely for the European markets. As the database consists of 10,588 sets launched across 1966-2018, the purchasing power of the dollar and the GBP/USD exchange rate fluctuated significantly. Therefore, retail prices in GBP were converted into USD using the historical exchange rate at the year of launch and then deflated using the US consumer price index, therefore, all retail prices are expressed in 2018 dollars. Apart from ensuring comparability of sets' prices in the sample, this also allows to naturally calculate real returns of LEGO. Table 1 below reports the descriptive statistics regarding the fundamental characteristics of sample LEGO sets.

Variable	Notes	Number of observations	Min	Max	Mean	Median	Standard deviation
Launch year	date of initial set release	10587	1955	2019	2002.52	2005	12.69
Retail price	initial retail price in 2018 US dollars	6964	0	850.59	37.87	21.37	51.60
Piece count	number of individual pieces in the set	10587	0	10004	195.41	64	377.93
Height		2732	0.97	58.19	22.91	23.00	8.39
Width	of the set box, in cm	2732	6.30	83.82	29.25	26.21	14.17
Depth		2732	0.20	54.29	6.61	6.10	3.30
Volume	of the set box, in cm3	2732	0.02	57.79	5.77	3.69	6.32
Weight	of the set box, in kg	2760	0.01	15.00	0.61	0.38	0.74
Price-to-piece	dollars per piece	6866	0.00	213.18	1.03	0.16	6.57
Price-to-volume	dollars per litre	2696	0.48	193.57	8.42	7.26	9.39
Price-to-weight	dollars per kg	2715	6.56	3109.58	83.85	68.08	125.69
Total return	inflation-adjusted return	4771	-96.41	19962.58	107.46	36.30	451.32
Annual return	since launch, in %	4771	-58.46	485.41	2.85	3.46	15.44

Table 1A. Descriptive statistics

Apart from raw data extracted from brickset.com, the study also calculates the physical volume of each LEGO set (a product three dimensions) as well as price-to-piece, price-to-weight, and price-to-volume ratios (initial retail price per unit). Piece count, initial retail price, weight, and volume, therefore, can be treated as size proxies when studying the existence of size ("small set") effect on the LEGO market, while the three LEGO "valuation multiples" presented above can be utilised to determine the presence of value effect (similar to price-to-book ratio in traditional stock market research).

	Launch	Piece	Valuesa	Walaht	Retail	Price-to-	Price-to-	Price-to-	Annualised
	year	count	volume	weight	Price	piece	volume	weight	return
Launch year	1.000	0.180	-0.031	-0.086	0.013	-0.049	0.029	-0.009	-0.105
Piece count		1.000	0.783	0.777	0.759	-0.079	-0.098	-0.111	-0.066
Volume			1.000	0.959	0.842	-0.032	-0.206	-0.353	-0.097
Weight				1.000	0.907	-0.058	-0.172	-0.152	-0.093
Retail price					1.000	0.065	-0.088	-0.034	-0.073
Price-to-piece						1.000	0.214	0.573	-0.032
Price-to-volume							1.000	0.633	0.124
Price-to-weight								1.000	-0.002
Annualised return									1.000

Table 1B. Correlation matrix

As can be seen from the sample correlation matrix, all four candidate size proxies are tightly correlated with each other, while the three valuation multiples exhibit moderate to high cross-correlation. More importantly, there is virtually no correlation between the groups, suggesting that size and value effect on the LEGO market, if present, are orthogonal, simplifying the derivation of asset-pricing models.

The study is concerned with the testing of a wide set of hypotheses regarding the LEGO market and LEGO properties as alternative investments that are derived from the most prominent theories in the literature and the limitations of the existing studies.

First, the internationalisation of the LEGO secondary market is studied using the data on 90,000 transactions occurred within six sample months among counterparties from 95 countries. This will allow to compute country-specific LEGO market share and relate the prominence of LEGO trading on various national markets to widely cited factors, such as GDP, per capita income, and inequality. As it is commonly asserted alternative investments attract significant attention when income inequality rises (Hiraki et al., 2009), with Dobrynskaya and Kishilova (2018) expressing the same statement with regards to LEGO, it can be expected that market share would be positively related to the degree of income inequality as measured by Gini coefficient when controlled for other potentially impactful factors. For LEGO, this study considers share of global GDP generated in the country, log of per capita income (as LEGO might be expected to be a normal good), and log distance to Denmark (as countries closer to Denmark, where LEGO originates from, might be expected to trading LEGO more actively). As market share is a truncated variable ranging from 0% to 100%, all equations are estimated using OLS and Tobit with respective cut-off values (Tobin, 1958), while controlling for arbitrary heteroscedasticity via Huber-White variance estimator (White, 1980):

$$Share_{i} = \beta_{0} + \beta_{1}GDP_{i} + \beta_{1}\log(GDPPC_{i}) + \beta_{2}Gini_{i} + \beta_{3}\log(1 + DIST_{i}) + \varepsilon_{i}$$

where $Share_i$ is the market share of the *i*th country on the global LEGO market, GDP_i is its share in world GDP, $GDPPC_i$ is its level of per capita income in 2018 USD, $Gini_i$ is its most recent Gini coefficient, $DIST_i$ is the distance to Denmark (in km), and ε_i is the error term. Data availability limits the sample to 143 cross-sectional observations, with 48 of them having a market share of 0% (as they do not participate in the international LEGO investment).

Furthermore, as the data allows to distinguish between local and international sales of LEGO on the secondary market, it can be determined if there exists a local bias with regards to this alternative investors, with country nationals being more likely to trade with their compatriots rather than foreigners. Local bias is a well-studied and well-evidenced phenomenon on the stock market (Seasholes and Zhu, 2010), therefore it would be interesting to know if it is manifested on the LEGO market as well.

Second, the performance of various LEGO portfolios is studied on a 1966-2018 time horizon, with both equal- and value-weighting as well as accounting for survivorship bias. The study proposes four distinct assumptions in terms of long-term LEGO return measurement:

Assumption (1): baseline. As in Dobrynskaya and Kishilova (2018), only the LEGO sets that are currently being traded on the secondary market are considered for the purposes of return computation.

Assumption (2): conservative. LEGO sets that have initial retail prices available yet are not currently traded on the secondary market are assumed to have lost all their value and thus have a total return of -100%.

Assumption (3): reasonable loss. The study acknowledges that assumptions (1) and (2) are both very unrealistic, albeit for different reasons. (1) implies investors have perfect knowledge on which sets will be in demand on the secondary market and yield collectible value fifty years down the line, while (2) states that illiquid sets not traded are valueless. Here, it is assumed that if the set is illiquid, an investor would still be able to sell it at a discount. As average LEGO mark-up is commonly estimated at 45% (Robetson and Breen, 2013), it is assumed that LEGO sets could be sold at cost, therefore yielding -31.04%.

Assumption (4): hoarding. Sets that are currently not traded on the secondary market are assumed to be held by the investors in anticipation of greater selling opportunities in the future, thus enjoying a return of 0%.

Across all four assumptions, equal-weighted and value-weighted portfolios of LEGO sets are constructed, and their annual returns (based on the investable universe in each of the years 1966-2018) are calculated and compared with S&P 500 Price Index and Total Return Index in terms of real (inflation-adjusted) returns, and risk-adjusted performance measures. Moreover, the performance of diversified portfolios including stocks and LEGO is reported, estimating optimal capital allocation into LEGO for investors with varying risk tolerance.

Third, the non-pecuniary utility theory (Baumol, 1986) is tested with regards to the LEGO market, applying utility-augmented CAPM (Baker et al., 2018) and Fama-French five-factor models to annual LEGO returns across all four assumptions outlined above. Baker et al. (2018) estimate the expected return discount for an asset yielding non-pecuniary utility *e*:

$$r = \beta_m r_m - \frac{a_2}{a_1 + a_2} e$$

where *r* is the return of the asset, β_m is its market beta, r_m is the return of the market, and $\frac{a_2}{a_1+a_2}$ is the share of the investor clientele enjoying non-pecuniary utility from holding in the total investor mass. Hence, if one can estimate $\frac{a_2}{a_1+a_2}$ with reasonable accuracy, the slope of the regression can be used as an unbiased estimator of – *e*. If it is statistically significant, then the non-pecuniary utility theory is verified.

Therefore, the study estimates the following equation (utility-augmented CAPM and Fama-French five-factor models):

$$r_{LEGO} = \alpha - e \frac{a_2}{a_1 + a_2} + \sum_{i=1}^n \beta_i r_i + \varepsilon$$

where α is the intercept, *e* is the utility estimator, $\frac{a_2}{a_1+a_2}$ is the estimate of the clientele prominence, β_i is the *i*th factor's slope in the *n*-factor model, r_i is the return of the *i*th factor portfolio, and ε is the error term.

The study uses the share of cumulative sales of LEGO up to year t obtained from Robertson and Breen (2013) and Bloomberg divided by world stock market capitalisation in year t from the

World Development Indicators database as an estimate of $\frac{a_2}{a_1+a_2}$ in year *t*. As world stock market

capitalisation is available only starting from 1980, it constitutes 39 annual observations in 1980-2018 period.

Finally, the study considers the performance of portfolios based on LEGO set sorts by various characteristics, including candidate size and value proxies, and launch year, and performs cross-sectional regressions to determine the significance of these factors in explaining the cross-section of LEGO set returns.

Findings and Discussion

Within the six sample months (September 2018 – March 2019), the study has synthesised more than 90,000 LEGO transactions occurred on the secondary market. Within 6 months, the value of LEGO investment exceeds \$7.5 million, with investors from 95 countries engaging in LEGO transactions. LEGO sellers and buyers originate from 65 and 90 countries, respectively. The international LEGO market is rather concentrated, though, with almost 60% of trading executed by investors from five countries: United States (27.28%), Germany (12.64%), Netherlands (10.88%), Canada (4.28%) and United Kingdom (4.21%). Having calculated investment inflows and outflows from all sample countries, the study has constructed LEGO capital accounts (see Table 2 below).

Table 2. LEGO capital accounts

Country	LEGO sold	LEGO bought	Trading volume	Internal trading	Inflow	Outflow	Internationalisation	Domestic share	Market share	Capital account
US	\$1,625.91	\$2,520.16	\$2,073.04	\$1,410.67	\$215.24	\$1,109.49	31.95%	68.05%	27.28%	-\$894.25
BE	\$270.35	\$165.52	\$217.94	\$48.71	\$221.64	\$116.81	77.65%	22.35%	2.87%	\$104.83
PH	\$2.23	\$10.95	\$6.59	\$0.92	\$1.31	\$10.03	86.03%	13.97%	0.09%	-\$8.72
NL	\$1,330.20	\$323.82	\$827.01	\$159.84	\$1,170.36	\$163.98	80.67%	19.33%	10.88%	\$1,006.38
DE	\$1,096.13	\$824.13	\$960.13	\$313.10	\$783.03	\$511.02	67.39%	32.61%	12.64%	\$272.01
DK	\$181.56	\$103.39	\$142.48	\$21.14	\$160.43	\$82.25	85.16%	14.84%	1.88%	\$78.18
IE	\$66.26	\$38.23	\$52.24	\$2.79	\$63.47	\$35.44	94.67%	5.33%	0.69%	\$28.03
СН	\$51.16	\$134.70	\$92.93	\$26.45	\$24.70	\$108.25	71.53%	28.47%	1.22%	-\$83.54
CA	\$370.80	\$279.25	\$325.02	\$193.27	\$177.53	\$85.97	40.54%	59.46%	4.28%	\$91.56
GE	\$0.65	\$0.00	\$0.33	\$0.00	\$0.65	\$0.00	100.00%	0.00%	0.00%	\$0.65
HR	\$35.73	\$11.26	\$23.50	\$1.04	\$34.69	\$10.22	95.57%	4.43%	0.31%	\$24.47
FR	\$174.07	\$202.08	\$188.07	\$52.45	\$121.62	\$149.63	72.11%	27.89%	2.48%	-\$28.01
AT	\$116.06	\$103.71	\$109.89	\$21.79	\$94.27	\$81.92	80.17%	19.83%	1.45%	\$12.35
SI	\$73.13	\$17.58	\$45.35	\$1.32	\$71.81	\$16.26	97.09%	2.91%	0.60%	\$55.55
KR	\$160.35	\$193.18	\$176.77	\$9.47	\$150.88	\$183.71	94.64%	5.36%	2.33%	-\$32.83
PT	\$78.51	\$74.82	\$76.67	\$22.71	\$55.80	\$52.11	70.37%	29.63%	1.01%	\$3.69
ES	\$48.76	\$65.13	\$56.95	\$12.09	\$36.67	\$53.04	78.77%	21.23%	0.75%	-\$16.37
UK	\$279.98	\$359.18	\$319.58	\$115.92	\$164.06	\$243.26	63.73%	36.27%	4.21%	-\$79.20
HK	\$213.89	\$135.57	\$174.73	\$5.35	\$208.54	\$130.22	96.94%	3.06%	2.30%	\$78.32
KZ	\$10.30	\$13.48	\$11.89	\$0.00	\$10.30	\$13.48	100.00%	0.00%	0.16%	-\$3.18
RU	\$191.06	\$191.14	\$191.10	\$21.07	\$169.99	\$170.08	88.98%	11.02%	2.51%	-\$0.08
IT	\$161.23	\$284.82	\$223.03	\$57.31	\$103.93	\$227.51	74.30%	25.70%	2.94%	-\$123.59
AU	\$91.78	\$226.25	\$159.01	\$57.13	\$34.64	\$169.11	64.07%	35.93%	2.09%	-\$134.47
SE	\$50.83	\$93.69	\$72.26	\$10.13	\$40.69	\$83.56	85.97%	14.03%	0.95%	-\$42.86
SG	\$38.60	\$48.57	\$43.59	\$3.82	\$34.78	\$44.74	91.23%	8.77%	0.57%	-\$9.96
NO	\$45.07	\$54.34	\$49.71	\$7.95	\$37.12	\$46.39	84.00%	16.00%	0.65%	-\$9.27
HU	\$144.40	\$72.02	\$108.21	\$19.97	\$124.43	\$52.05	81.55%	18.45%	1.42%	\$72.38
BN	\$25.51	\$1.89	\$13.70	\$0.00	\$25.51	\$1.89	100.00%	0.00%	0.18%	\$23.62

Country	LEGO	LEGO	Trading	Internal	Inflow	Outflow	Internationalisation	Domestic	Market	Capital
ID	\$106.24	\$4 77	\$55.51	\$0.99	\$105.26	\$3 78	98 22%	1 78%	0.73%	\$101 47
CZ	\$165.13	\$53.77	\$109.45	\$9.95	\$155.18	\$43.82	90.91%	9.09%	1.44%	\$111.36
GR	\$9.58	\$30.96	\$20.27	\$1.03	\$8.55	\$29.93	94.92%	5.08%	0.27%	-\$21.38
MY	\$19.00	\$53.11	\$36.05	\$2.03	\$16.97	\$51.08	94.37%	5.63%	0.47%	-\$34.11
SK	\$60.17	\$31.51	\$45.84	\$3.04	\$57.12	\$28.47	93.36%	6.64%	0.60%	\$28.66
PL	\$67.16	\$66.88	\$67.02	\$4.10	\$63.06	\$62.79	93.89%	6.11%	0.88%	\$0.28
AD	\$0.50	\$0.00	\$0.25	\$0.00	\$0.50	\$0.00	100.00%	0.00%	0.00%	\$0.50
EC	\$0.31	\$0.00	\$0.15	\$0.00	\$0.31	\$0.00	100.00%	0.00%	0.00%	\$0.31
LU	\$32.25	\$14.99	\$23.62	\$0.28	\$31.96	\$14.71	98.80%	1.20%	0.31%	\$17.26
NZ	\$4.49	\$42.99	\$23.74	\$1.81	\$2.68	\$41.18	92.38%	7.62%	0.31%	-\$38.50
TW	\$10.34	\$42.31	\$26.32	\$0.00	\$10.34	\$42.31	100.00%	0.00%	0.35%	-\$31.97
FI	\$2.23	\$21.61	\$11.92	\$0.20	\$2.03	\$21.40	98.30%	1.70%	0.16%	-\$19.38
ZA	\$3.17	\$8.07	\$5.62	\$1.96	\$1.21	\$6.11	65.17%	34.83%	0.07%	-\$4.90
TR	\$23.62	\$16.49	\$20.06	\$0.92	\$22.70	\$15.57	95.39%	4.61%	0.26%	\$7.13
RO	\$76.06	\$9.14	\$42.60	\$0.86	\$75.20	\$8.28	97.99%	2.01%	0.56%	\$66.92
IL	\$0.63	\$30.12	\$15.38	\$0.00	\$0.63	\$30.12	100.00%	0.00%	0.20%	-\$29.48
AE	\$22.95	\$22.92	\$22.94	\$0.00	\$22.95	\$22.92	100.00%	0.00%	0.30%	\$0.04
EE	\$2.20	\$4.94	\$3.57	\$0.20	\$2.00	\$4.74	94.53%	5.47%	0.05%	-\$2.74
LT	\$3.19	\$7.56	\$5.38	\$0.04	\$3.15	\$7.52	99.25%	0.75%	0.07%	-\$4.37
JP	\$6.37	\$158.25	\$82.31	\$0.14	\$6.22	\$158.11	99.83%	0.17%	1.08%	-\$151.89
LV	\$16.42	\$3.69	\$10.06	\$0.00	\$16.42	\$3.69	100.00%	0.00%	0.13%	\$12.73
RS	\$2.32	\$1.10	\$1.71	\$0.02	\$2.30	\$1.08	98.93%	1.07%	0.02%	\$1.22
MT	\$1.92	\$2.14	\$2.03	\$0.00	\$1.92	\$2.14	100.00%	0.00%	0.03%	-\$0.23
BR	\$2.42	\$80.96	\$41.69	\$1.19	\$1.23	\$79.77	97.15%	2.85%	0.55%	-\$78.54
UA	\$1.20	\$18.37	\$9.79	\$0.69	\$0.51	\$17.68	92.96%	7.04%	0.13%	-\$17.17
SV	\$0.63	\$0.14	\$0.38	\$0.00	\$0.63	\$0.14	100.00%	0.00%	0.01%	\$0.49
TH	\$1.08	\$25.76	\$13.42	\$0.76	\$0.32	\$25.00	94.37%	5.63%	0.18%	-\$24.68
CL	\$12.48	\$5.19	\$8.83	\$0.00	\$12.48	\$5.19	100.00%	0.00%	0.12%	\$7.28
CN	\$3.82	\$200.44	\$102.13	\$0.08	\$3.73	\$200.36	99.92%	0.08%	1.34%	-\$196.63
IN	\$0.01	\$14.50	\$7.26	\$0.01	\$0.00	\$14.50	99.91%	0.09%	0.10%	-\$14.50
BY	\$0.61	\$8.40	\$4.51	\$0.11	\$0.51	\$8.30	97.66%	2.34%	0.06%	-\$7.79
BG	\$4.60	\$3.64	\$4.12	\$0.10	\$4.50	\$3.54	97.51%	2.49%	0.05%	\$0.96
MX	\$0.54	\$5.27	\$2.91	\$0.00	\$0.54	\$5.27	100.00%	0.00%	0.04%	-\$4.73
МО	\$0.22	\$5.93	\$3.07	\$0.00	\$0.22	\$5.93	100.00%	0.00%	0.04%	-\$5.72
BH	\$0.11	\$2.72	\$1.42	\$0.00	\$0.11	\$2.72	100.00%	0.00%	0.02%	-\$2.60
PK	\$0.01	\$0.00	\$0.01	\$0.00	\$0.01	\$0.00	100.00%	0.00%	0.00%	\$0.01
AM	\$0.25	\$0.00	\$0.13	\$0.00	\$0.25	\$0.00	100.00%	0.00%	0.00%	\$0.25
QA	\$0.00	\$20.48	\$10.24	\$0.00	\$0.00	\$20.48	100.00%	0.00%	0.13%	-\$20.48
VN	\$0.00	\$2.69	\$1.34	\$0.00	\$0.00	\$2.69	100.00%	0.00%	0.02%	-\$2.69
GL	\$0.00	\$1.10	\$0.55	\$0.00	\$0.00	\$1.10	100.00%	0.00%	0.01%	-\$1.10
IS	\$0.00	\$1.93	\$0.97	\$0.00	\$0.00	\$1.93	100.00%	0.00%	0.01%	-\$1.93
OM	\$0.00	\$0.08	\$0.04	\$0.00	\$0.00	\$0.08	100.00%	0.00%	0.00%	-\$0.08
SM	\$0.00	\$1.74	\$0.87	\$0.00	\$0.00	\$1.74	100.00%	0.00%	0.01%	-\$1.74
СО	\$0.00	\$0.20	\$0.10	\$0.00	\$0.00	\$0.20	100.00%	0.00%	0.00%	-\$0.20
BQ	\$0.00	\$3.33	\$1.67	\$0.00	\$0.00	\$3.33	100.00%	0.00%	0.02%	-\$3.33
PE	\$0.00	\$6.49	\$3.25	\$0.00	\$0.00	\$6.49	100.00%	0.00%	0.04%	-\$6.49
CY	\$0.00	\$4.31	\$2.15	\$0.00	\$0.00	\$4.31	100.00%	0.00%	0.03%	-\$4.31
LI	\$0.00	\$0.56	\$0.28	\$0.00	\$0.00	\$0.56	100.00%	0.00%	0.00%	-\$0.56
FO	\$0.00	\$0.07	\$0.03	\$0.00	\$0.00	\$0.07	100.00%	0.00%	0.00%	-\$0.07
AR	\$0.00	\$0.86	\$0.43	\$0.00	\$0.00	\$0.86	100.00%	0.00%	0.01%	-\$0.86
SA	\$0.00	\$1.80	\$0.90	\$0.00	\$0.00	\$1.80	100.00%	0.00%	0.01%	-\$1.80
ME	\$0.00	\$0.06	\$0.03	\$0.00	\$0.00	\$0.06	100.00%	0.00%	0.00%	-\$0.06

N. Shimkus, S. Shanaev, S. Sharma, B- Ghimire/ Finance, Risk and Accounting Perspectives Proceedings (2019) 150-184

Country	LEGO sold	LEGO bought	Trading volume	Internal trading	Inflow	Outflow	Internationalisation	Domestic share	Market share	Capital account
JO	\$0.00	\$0.14	\$0.07	\$0.00	\$0.00	\$0.14	100.00%	0.00%	0.00%	-\$0.14
CR	\$0.00	\$0.38	\$0.19	\$0.00	\$0.00	\$0.38	100.00%	0.00%	0.00%	-\$0.38
KY	\$0.00	\$0.03	\$0.01	\$0.00	\$0.00	\$0.03	100.00%	0.00%	0.00%	-\$0.03
KW	\$0.00	\$0.01	\$0.01	\$0.00	\$0.00	\$0.01	100.00%	0.00%	0.00%	-\$0.01
MA	\$0.00	\$0.03	\$0.02	\$0.00	\$0.00	\$0.03	100.00%	0.00%	0.00%	-\$0.03
EG	\$0.00	\$1.16	\$0.58	\$0.00	\$0.00	\$1.16	100.00%	0.00%	0.01%	-\$1.16
MK	\$0.00	\$0.33	\$0.16	\$0.00	\$0.00	\$0.33	100.00%	0.00%	0.00%	-\$0.33
LB	\$0.00	\$0.66	\$0.33	\$0.00	\$0.00	\$0.66	100.00%	0.00%	0.00%	-\$0.66
MU	\$0.00	\$0.34	\$0.17	\$0.00	\$0.00	\$0.34	100.00%	0.00%	0.00%	-\$0.34
PF	\$0.00	\$0.02	\$0.01	\$0.00	\$0.00	\$0.02	100.00%	0.00%	0.00%	-\$0.02
MN	\$0.00	\$0.04	\$0.02	\$0.00	\$0.00	\$0.04	100.00%	0.00%	0.00%	-\$0.04
BA	\$0.00	\$0.01	\$0.01	\$0.00	\$0.00	\$0.01	100.00%	0.00%	0.00%	-\$0.01
AL	\$0.00	\$0.06	\$0.03	\$0.00	\$0.00	\$0.06	100.00%	0.00%	0.00%	-\$0.06
GI	\$0.00	\$1.96	\$0.98	\$0.00	\$0.00	\$1.96	100.00%	0.00%	0.01%	-\$1.96
MD	\$0.00	\$0.33	\$0.16	\$0.00	\$0.00	\$0.33	100.00%	0.00%	0.00%	-\$0.33
Total	\$7,598.72	\$7,598.72	\$7,598.72	\$2,626.92	\$4,971.79	\$4,971.79	65.43%	34.57%	-	\$0.00

Most prominent countries with net positive LEGO capital account are Netherlands, Germany, Belgium Canada and Denmark, while United States, Japan, China, Australia and Italy have significant net negatives. The share of domestic trading exceeds countries' market share in 92 out of 95 cases (notable exceptions being China, Japan, and India), evidencing the universality of the local bias phenomenon (Seasholes and Zhu, 2010). However, as shipping and storage costs of LEGO might be quite material, this observation might be explained within the rational agent framework.



Figure 1. Local bias on various LEGO national markets

Next, the study examines the determinants of cross-country differences in LEGO trading via a multiple cross-sectional regression model, incorporating share of world GDP, log per capita income, income inequality, and log distance to Denmark. GDP and GDP per capita are always positive and statistically significant determinants both in OLS and Tobit models, as more affluent economies and individuals can be expected to tend to invest more, including alternative asset classes, LEGO being a normal good and a "normal investment". The results regarding distance to Denmark are inconsistent, however when the factor is significant, it is of expected sign: national markets situated further away from the country LEGO originates from tend to trade less LEGO.

Regressor	OLS	Tobit	OLS	Tobit	OLS	Tobit
	-1.5236	-9.9358	-1.7086	-9.0868	0.8228	-6.6844
Constant	(-2.5750)	(-3.7340)	(-1.8916)	(-2.9655)	(1.3381)	(-2.0958)
	0.0107	0.0002	0.0606	0.0030	0.5396	0.0361
	0.8161	0.8768	0.8188	0.8396	0.8308	0.8516
GDP	(2.8880)	(3.5549)	(2.7681)	(3.1683)	(2.7998)	(3.2029)
	0.0043	0.0004	0.0064	0.0015	0.0058	0.0014
	0.1827	0.9466	0.2914	1.1671	0.1949	1.0502
Log GDP per capita	(2.4108)	(3.5843)	(2.3952)	(3.4304)	(1.8385)	(3.2305)
	0.0168	0.0003	0.0179	0.0006	0.0681	0.0012
			-0.0170	-0.0623	-0.0012	-0.0431
Gini			(-1.9418)	(-2.5588)	(-0.1458)	(-1.7181)
			0.0542	0.0105	0.8843	0.0858
T 11					-0.2728	-0.2531
Log distance to					(-1.9145)	(-1.2506)
Deninark					0.0576	0.2111

Table 3. Determinants of LEGO investment

With regards to inequality, the findings are most puzzling: it is more egalitarian countries that trade LEGO more actively, not the other way around, thus rejecting the "conspicuous consumption" theory with respect to this asset class (Veblen, 2005; Hiraki et al., 2009). However, this is easily explainable if one takes into account considerable differences between LEGO sets and other "investments of passion" (e.g., fine art). The barriers for entry into the art market are much higher, with unit costs at least in the thousands and trading occurring at auctions (Goetzmann et al., 2011), while some LEGO sets can be purchased online for less than \$10, thus making LEGO an attractive and affordable "mass market" alternative investment.

Next, the performance of the equal- and value-weighted LEGO portfolios as per four assumptions identified above is tracked between 1966 and 2018. Consistent with the literature, equal-weighted returns exceed value-weighted returns, with value-weighted performance being more realistic and in line with 2%-3% per annum real returns of other alternative asset classes in the existing meta-analyses (Burton and Jacobsen, 1999; Mandel, 2009).

|--|

	LEGO index								S&D 500 index	
Voor	Assump	otion (1)	Assump	otion (2)	Assump	otion (3)	Assum	otion (4)	5aP 50	index
rear	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Price index	Total return
1966	2.96	2.96	-14.34	-14.43	-0.52	-0.53	0.11	0.11	-15.92	-11.72
1967	2.96	2.96	-14.41	-14.48	-0.53	-0.53	0.10	0.10	16.28	24.61
1968	2.96	2.96	-14.66	-14.96	-0.55	-0.58	0.09	0.08	2.82	8.88
1969	2.96	2.96	-14.79	-15.17	-0.56	-0.60	0.08	0.07	-16.30	-15.92
1970	0.94	1.32	-14.07	-14.77	-0.54	-0.58	0.07	0.07	-5.18	-5.25
1971	2.91	1.71	-12.95	-14.40	-0.10	-0.49	0.49	0.15	7.28	12.50
1972	2.99	1.76	-12.55	-14.24	-0.01	-0.46	0.56	0.17	11.82	13.04
1973	2.93	1.89	-12.92	-14.34	-0.03	-0.45	0.57	0.19	-24.15	-25.88
1974	2.93	1.89	-13.83	-14.71	-0.15	-0.48	0.48	0.18	-37.30	-35.55
1975	3.05	2.02	-14.22	-15.02	-0.19	-0.49	0.46	0.18	22.79	29.04
1976	3.05	2.02	-14.33	-15.49	-0.20	-0.53	0.45	0.16	13.44	20.90
1977	3.05	2.02	-14.41	-15.54	-0.21	-0.54	0.44	0.16	-17.04	-9.20
1978	2.92	1.97	-14.39	-15.37	-0.21	-0.51	0.45	0.18	-7.27	-0.71
1979	2.91	2.00	-14.48	-15.55	-0.21	-0.51	0.45	0.18	-0.84	9.02
1980	2.41	1.91	-14.22	-15.05	-0.19	-0.39	0.46	0.29	11.94	18.71
1981	2.59	2.49	-14.41	-14.97	-0.17	-0.28	0.49	0.41	-17.12	-11.32
1982	2.68	2.57	-14.42	-14.89	-0.15	-0.25	0.52	0.43	10.53	16.73
1983	2.68	2.57	-14.61	-15.06	-0.17	-0.27	0.51	0.42	12.99	18.07
1984	3.06	3.34	-14.75	-14.91	-0.10	-0.09	0.59	0.60	-2.54	-0.23
1985	3.10	3.01	-14.96	-14.90	-0.08	-0.09	0.62	0.60	21.72	27.79
1986	3.49	3.16	-14.33	-15.25	0.16	-0.05	0.84	0.67	13.28	14.92
1987	3.53	3.72	-13.95	-14.63	0.30	0.20	0.97	0.91	-2.21	-2.62
1988	3.68	3.87	-13.78	-14.30	0.39	0.32	1.06	1.01	7.65	12.92
1989	3.92	4.36	-13.17	-13.50	0.60	0.63	1.25	1.30	21.61	23.15
1990	4.00	4.51	-12.95	-13.26	0.70	0.76	1.35	1.43	-12.06	-11.67
1991	3.85	4.21	-12.48	-12.72	0.78	0.84	1.41	1.49	22.65	30.88
1992	3.82	4.25	-11.91	-12.21	0.91	0.99	1.52	1.62	1.45	6.58
1993	3.84	4.27	-11.84	-12.07	0.98	1.08	1.60	1.71	4.13	8.07
1994	3.85	4.10	-11.75	-11.90	1.04	1.11	1.66	1.74	-4.03	-2.73
1995	3.76	3.91	-11.96	-12.00	1.03	1.08	1.66	1.71	30.80	33.44
1996	3.73	3.81	-11.84	-11.28	1.09	1.20	1.72	1.81	16.33	17.21
1997	3.68	3.60	-12.02	-11.23	1.10	1.18	1.75	1.79	28.82	29.03
1998	3.64	3.45	-13.10	-12.74	1.00	0.96	1.70	1.64	24.67	22.35
1999	3.56	3.17	-13.83	-13.45	0.91	0.80	1.66	1.51	16.41	21.99
2000	3.49	3.08	-14.89	-14.19	0.80	0.70	1.60	1.46	-13.12	-14.64
2001	3.70	3.19	-15.29	-14.40	0.88	0.76	1.71	1.54	-14.42	-12.77
2002	3.76	3.35	-16.07	-15.19	0.85	0.77	1.73	1.60	-25.22	-23.02
2003	3.85	3.52	-16.44	-16.01	0.88	0.78	1.79	1.66	23.86	29.14
2004	3.87	3.61	-17.02	-16.32	0.85	0.81	1.79	1.71	5.47	8.30
2005	4.01	3.81	-17.57	-16.77	0.87	0.87	1.85	1.81	-0.33	2.64
2006	4.19	3.98	-17.47	-16.43	0.99	1.02	1.97	1.95	10.82	12.56
2007	4.38	4.29	-17.40	-16.40	1.10	1.19	2.10	2.14	-0.56	1.53
2008	4.46	4.29	-17.10	-15.78	1.21	1.30	2.20	2.23	-38.47	-36.73

N. Shimkus, S. Shanaev, S. Sharma, B-	Ghimire/ Finance, Risk and	nd Accounting Perspectives	Proceedings (2019)
	150-184		

		S&D 500 in day								
Voor	Assumption (1)		Assumption (2)		Assumption (3)		Assumption (4)		Ster 500 mater	
Icar	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Price index	Total return
2009	4.58	4.47	-16.71	-15.22	1.35	1.50	2.34	2.42	20.08	24.85
2010	4.63	4.43	-16.31	-14.63	1.45	1.59	2.43	2.48	11.18	15.82
2011	4.66	4.46	-15.82	-14.01	1.56	1.71	2.53	2.59	-2.97	-2.51
2012	4.82	4.55	-15.46	-13.37	1.72	1.86	2.71	2.72	11.44	14.33
2013	4.67	4.16	-14.99	-12.50	1.74	1.79	2.71	2.61	27.67	33.20
2014	4.72	4.05	-13.98	-11.75	1.92	1.84	2.86	2.63	10.67	11.00
2015	4.55	3.77	-12.92	-10.97	1.99	1.79	2.88	2.53	-1.38	-0.57
2016	4.67	3.36	-11.72	-10.28	2.23	1.63	3.06	2.33	7.30	11.18
2017	3.68	2.20	-11.45	-10.18	1.70	0.92	2.49	1.58	16.95	19.77
2018	2.86	1.32	-11.68	-10.58	1.14	0.29	1.96	0.96	-8.03	-6.93
Return	3.54	3.22	-14.22	-14.05	0.59	0.48	1.33	1.20	2.35	5.54
Risk	0.75	0.95	1.66	1.69	0.76	0.82	0.87	0.88	16.39	17.33
Sharpe	4.73	3.37	-8.56	-8.31	0.77	0.58	1.54	1.37	0.14	0.32
Alpha	3.50	3.15	-14.31	-14.18	0.54	0.42	1.29	1.15	-	-
Beta	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	-	-
Treynor	5.60	3.48	-9.87	-7.00	0.86	0.57	2.00	1.51	-	-
Correlation	0.19	0.21	0.18	0.26	0.22	0.24	0.20	0.22	-	-

As per assumptions (1), (2), (3), and (4), annual average real LEGO returns equal 3.54%, -14.22%, 0.59%, and 1.33% for the equal-weighted portfolios and 3.22%, -14.05%, 0.48%, and 1.20% for the value-weighted portfolios, respectively, well below the average annual real total return of equities at 5.54% and 8% reported by Dobrynskaya and Kishilova (2018). It is clear that a broader sample and accurate identification of survivorship biases brings LEGO returns closer to the findings of the existing literature on other alternative asset classes, suggesting that there is in fact no "LEGO puzzle". It is also evident that assumption (2) is too conservative and thus unrealistic, leading to double-digit negative return estimates. Therefore, the study relies mainly on assumptions (3) and (4), while reporting results for all four assumptions for consistency purposes.

Annual LEGO volatility, however, is much lower at less than 1%, with returns virtually uncorrelated with major benchmarks, highlighting significant diversification potential. Sharpe ratios of LEGO are also shown to be superior to those of S&P 500. Table 5 below studies the comparative performance of portfolios with various weights of LEGO in terms of return, risk, and certainty equivalent for investors with various risk tolerance.

LEGO	Dotum	m Diala	Sharra	Alaha	Pata	Troutor	Correlation	Idiosyncratic	Certainty equivalent			
weight	Keturn	KISK	Sharpe	Атрпа	Dela	Treynor	Correlation	risk	θ=0.5	θ=0.75	$\theta = 1$	
0%	5.54	17.33	0.32	0.00	1.00	0.06	1.00	0.00	4.04	3.29	2.54	
1%	5.51	17.15	0.32	0.00	0.99	0.06	1.00	0.01	4.04	3.30	2.56	
2%	5.47	16.98	0.32	0.01	0.98	0.06	1.00	0.02	4.03	3.31	2.59	
3%	5.44	16.81	0.32	0.01	0.97	0.06	1.00	0.02	4.02	3.32	2.61	
4%	5.40	16.64	0.32	0.02	0.96	0.06	1.00	0.03	4.02	3.32	2.63	
5%	5.36	16.47	0.33	0.02	0.95	0.06	1.00	0.04	4.01	3.33	2.65	
10%	5.18	15.61	0.33	0.04	0.90	0.06	1.00	0.08	3.96	3.35	2.74	
15%	4.98	14.75	0.34	0.06	0.85	0.06	1.00	0.12	3.90	3.35	2.81	
20%	4.78	13.89	0.34	0.08	0.80	0.06	1.00	0.16	3.82	3.33	2.85	
25%	4.57	13.03	0.35	0.11	0.75	0.06	1.00	0.20	3.72	3.30	2.87	
30%	4.35	12.17	0.36	0.13	0.70	0.06	1.00	0.24	3.61	3.24	2.87	
35%	4.12	11.32	0.36	0.15	0.65	0.06	1.00	0.28	3.48	3.16	2.84	
40%	3.89	10.46	0.37	0.17	0.60	0.06	1.00	0.32	3.34	3.07	2.79	
50%	3.39	8.75	0.39	0.21	0.50	0.07	1.00	0.40	3.01	2.82	2.63	
60%	2.87	7.03	0.41	0.25	0.41	0.07	1.00	0.48	2.62	2.50	2.38	
70%	2.32	5.33	0.43	0.30	0.31	0.08	0.99	0.57	2.17	2.10	2.03	
80%	1.73	3.64	0.48	0.34	0.21	0.08	0.98	0.65	1.67	1.63	1.60	
90%	1.12	2.00	0.56	0.38	0.11	0.10	0.93	0.73	1.10	1.09	1.08	
100%	0.48	0.82	0.58	0.42	0.01	0.57	0.18	0.81	0.48	0.47	0.47	

Table 5. Diversification benefits of LEGO

Compared to annual Sharpe ratio of 0.32 for stocks, a portfolio with 5% of value-weighted LEGO index as per assumption (3) has a Sharpe ratio of 0.33. Less risk averse investors do not benefit from including LEGO in their portfolio, however, their counterparts that are less tolerant of risk can improve their utility including from 5% to 25% of LEGO in a diversified portfolio.

Furthermore, the non-pecuniary utility aspect of LEGO set holdings is explored. For assumptions (3) and (4) which are the most realistic in terms of illiquid sets' performance, the estimated utility is positive and significant (as the regression coefficient is negative and significant), reinforcing the role of LEGO as an "investment of passion" for a particular investor clientele. LEGO returns are also unexposed to any Fama-French factors, including the market factor, and have positive and statistically significant alphas at 1.53% and 2.50% for assumptions (3) and (4), respectively. This illustrates both the defensive properties of LEGO as well as higher total utility obtained from LEGO investments by collectable enthusiasts.

Pagrossor	Assump	Assumption (1)		Assumption (2)		Assumption (3)		Assumption (4)	
Regressor	САРМ	FF5F	САРМ	FF5F	САРМ	FF5F	САРМ	FF5F	
	4.1370	4.1925	-16.2313	-15.9535	1.4081	1.5291	2.3720	2.4996	
Constant	(10.5358)	(9.6037)	(-18.7300)	(-17.0728)	(4.3622)	(4.3377)	(7.0457)	(6.8183)	
	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0000	0.0000	
	-83.8720	-87.5711	346.3459	327.3513	-87.4518	-87.9722	-118.1150	-117.3177	
$\frac{a_2}{a_2+a_1}$	(-1.5190)	(-1.4937)	(2.8421)	(2.6086)	(-1.9266)	(-1.8583)	(-2.4950)	(-2.3829)	
	0.1375	0.1451	0.0073	0.0137	0.0619	0.0724	0.0173	0.0233	
	0.0019	0.0012		0.0185	0.0019	-0.0019	0.0006	-0.0039	
Market	(0.2440)	(0.1327)		(0.9246)	(0.3044)	(-0.2476)	(0.0882)	(-0.4928)	
	0.8086	0.8953		0.3621	0.7626	0.8060	0.9302	0.6256	
Carallaniana Dia		-0.0076		-0.0358		-0.0016		0.0005	
Smaii-minus-Big		(-0.5719)		(-1.2562)		(-0.1526)		(0.0450)	

Table 6. Non-pecuniary utility of LEGO holdings

Dagrassor	Assumption (1)		Assumption (2)		Assumption (3)		Assumption (4)	
Regressor	САРМ	FF5F	САРМ	FF5F	САРМ	FF5F	САРМ	FF5F
		0.5714		0.2181		0.8797		0.9644
		0.0035		0.0150		0.0030		0.0023
High-minus-Low		(0.2379)		(0.4718)		(0.2535)		(0.1871)
		0.8135		0.6403		0.8015		0.8528
		-0.0030		-0.0063		-0.0074		-0.0086
Robust-minus-Weak		(-0.1909)		(-0.1893)		(-0.5833)		(-0.6540)
		0.8498		0.8511		0.5638		0.5178
		-0.0046		-0.0313		-0.0158		-0.0167
Conservative-minus-Aggressive		(-0.2028)		(-0.6511)		(-0.8707)		(-0.8841)
		0.8406		0.5197		0.3904		0.3832

Finally, the fundamental return determinants of LEGO sets are studied. Cross-sectional regressions are employed to estimate the impact of size, value, and age effects on the average annual real return of sets, while also controlling for set broad theme or narrow set topic via a dummy variable approach, as there are notable heterogeneities in returns between these groups¹.

Regressor	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2.8563	17.2947	5.7529	0.8370	18.1389	24.8524	11.1440
Constant	(12.7802)	(7.1422)	(4.5162)	(1.7962)	(8.5644)	(9.7078)	(4.1356)
	0.0000	0.0000	0.0000	0.0731	0.0000	0.0000	0.0000
		-1.7795			-1.8105	-1.9052	-1.0387
Volume (size effect)		(-6.1786)			(-8.0439)	(-8.5834)	(-5.4461)
(,		0.0000			0.0000	0.0000	0.0000
			-1.0329		-1.9673	-4.9014	-4.1953
Price-to-piece (value effect)			(-2.7030)		(-6.1922)	(-8.4511)	(-6.7241)
(1			0.0069		0.0000	0.0000	0.0000
				0.1961	0.7608	0.8584	0.8921
Set age				(7.7837)	(9.9719)	(12.7472)	(10.5330)
				0.0000	0.0000	0.0000	0.0000
Theme controls	no	no	no	no	no	yes	no
Topic controls	no	no	no	no	no	no	yes
R-squared	-	0.0329	0.0045	0.0110	0.0914	0.1726	0.2992

Table 7. Regression estimation results for assumption (1)

¹ The performance of portfolios based on sorts can be consulted in Appendix (Tables A1-A10 and Figures A1-A8). The preliminary findings in terms of these portfolio returns led the study to select physical volume and price-to-piece ratio as size and value proxies, respectively.

Regressor	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	-11.6810	14.1716	-1.1824	-6.6997	34.3097	22.2714	12.0274
Constant	(-36.8202)	(4.4514)	(-0.9447)	(-12.4025)	(9.6171)	(5.6090)	(2.2671)
	0.0000	0.0000	0.3448	0.0000	0.0000	0.0000	0.0235
		-2.2815			-1.8557	-1.3126	-0.0562
Volume		(-5.9123)			(-5.4084)	(-4.0468)	(-0.1503)
(Size effect)		0.0000			0.0000	0.0001	0.8805
			-3.3410		-7.4379	-4.3950	-4.3619
Price-to-piece			(-9.2911)		(-8.8986)	(-4.0970)	(-3.7728)
(value effect)			0.0000		0.0000	0.0000	0.0002
				-0.3794	-0.3718	-0.3684	-0.1907
Set age				(-14.5569)	(-5.6204)	(-3.0437)	(-1.2163)
				0.0000	0.0000	0.0024	0.2240
Theme controls	no	no	no	no	no	yes	no
Topic controls	no	no	no	no	no	no	yes
R-squared	-	0.0132	0.0228	0.0198	0.0738	0.1512	0.2309

Table 8. Regression estimation results for assumption (2)

Table 0	Degression	ectimation	regulte for	accumption	(3)	١
Table 9.	Regression	estimation	results for	assumption	(5))

Regressor	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	1.1446	15.6601	3.5917	0.7752	17.2471	22.1446	8.0653
Constant	(7.2926)	(7.0291)	(5.1963)	(2.1236)	(8.9903)	(9.5054)	(3.0348)
	0.0000	0.0000	0.0000	0.0337	0.0000	0.0000	0.0022
		-1.6728			-1.4112	-1.6775	-0.7653
Volume		(-6.3185)			(-6.7995)	(-8.1201)	(-4.3469)
(size effect)		0.0000			0.0000	0.0000	0.0000
			-0.8121		-1.9022	-4.1791	-3.5098
Price-to-piece			(-4.3573)		(-7.6897)	(-8.6855)	(-7.1091)
(value effect)			0.0000		0.0000	0.0000	0.0000
				0.0281	0.2298	0.6067	0.6568
Set age				(1.7055)	(4.7790)	(10.3897)	(8.4512)
				0.0881	0.0000	0.0000	0.0000
Theme controls	no	no	no	no	no	yes	no
Topic controls	no	no	no	no	no	no	yes
R-squared	-	0.0308	0.0059	0.0004	0.0451	0.1328	0.2604

Regressor	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	1.9569	15.7552	4.0164	1.4675	15.9760	22.3062	9.8435
Constant	(12.7123)	(7.1199)	(5.8811)	(4.0575)	(8.4112)	(9.6878)	(3.8325)
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
		-1.6294			-1.3773	-1.7058	-0.8190
Volume		(-6.1973)			(-6.6721)	(-8.3132)	(-4.7518)
(Size effect)		0.0000			0.0000	0.0000	0.0000
			-0.6922		-1.4620	-4.1721	-3.4787
Price-to-piece			(-3.7632)		(-6.0814)	(-8.8547)	(-7.1289)
(value effect)			0.0002		0.0000	0.0000	0.0000
				0.0373	0.2581	0.6560	0.6972
Set age				(2.2780)	(5.1062)	(11.3614)	(9.0971)
				0.0228	0.0000	0.0000	0.0000
Theme controls	no	no	no	no	no	yes	no
Topic controls	no	no	no	no	no	no	yes
R-squared	-	0.0303	0.0044	0.0008	0.0434	0.1398	0.2689

Table 10. Regression estimation results for assumption (4)

In all estimations, size and value effect (log physical volume and log price-to-piece ratio) are statistically and economically significant predictors of LEGO set returns. Just like on the stock market, smaller sets (low physical volume) and sets with low price-to-piece ratio ("value sets") demonstrate higher returns. The results, however, were robust to alternative measures of size and value. Older sets enjoy higher expected returns, suggesting that age influences performance not via the differential information or risk channels, as for stocks (Barry and Brown, 1984) or cryptocurrencies (Shanaev et al., 2019), but rather demonstrating a liquidity effect. This "LEGO yield curve" is estimated to have a slope of approximately 0.65% per year. The results are robust to theme and topic controls, proving that they are not a result of heterogeneity bias among various set groups but a general tendency observed on the market as a whole. Nevertheless, the magnitudes of size and value effect decrease somewhat in estimations with theme and topic controls, however still remaining negative and highly significant, as suggested by the theory and empirical literature from conventional financial markets.

Conclusion

In this article, LEGO sets – one of the most recently established alternative asset classes – are considered. The performance of LEGO sets as collectables and an alternative asset class is studied on a wide sample of 10,588 sets bought and sold on the secondary market. The emerging literature on alternative investment properties of LEGO (Dobrynskaya and Kishilova, 2018) is addressed, and some methodological limitations of existing research, namely, sample survivorship bias and overextrapolation in LEGO market index construction, are resolved via accommodating a much larger dataset as well as by relaxing the most stringent assumptions of Dobrynskaya and Kishilova (2018) regarding benchmark definition. In this study, the whole investable universe of 14,901 LEGO sets launched between 1955 and 2019 is considered for sample inclusion, and the final sample consists of 10,588 – all sets initial retail prices are available for, four times larger than the sample of Dobrynskaya and Kishilova (2018). Post-launch long-term performance of LEGO sets is studied on more than 50-year time horizon (1966-2019).

The study employs cross-sectional regressions as well as univariate sorts to examine the riskreturn characteristics of LEGO sets. Size and value effects are well-pronounced on the LEGO market, reinforcing the earlier findings and providing some non-trivial evidence for the robustness of Fama-French risk factor methodology (Fama and French, 2016) from an alternative asset class. The study considered four different proxies of size effect – piece count (number of individual LEGO pieces in a set), physical volume (the product of length, width and height dimensions of a set box), physical weight, and initial retail price – and found that volume is the best explanatory proxy for the cross-sectional persistence in outperformance. For value effect, it considered price-to-piece ratio, price-to-volume, and price-to-weight ratio, with price-to-piece yielding the best results. Small sets and "value" (low price-to-piece) sets enjoy significantly higher returns, while the outperformance of older sets can imply the existence of a liquidity premium.

The performance evaluation findings of the study show that LEGO indeed enjoys material returns both in terms of average annual set return since launch. Unlike Dobrynskaya and Kishilova (2018), who report only equal-weighted indices, this study employs both value-weighted and equal-weighted benchmarks and portfolios for consistency as both have their merits and limitations, with equal-weighted portfolios known to exaggerate returns in case of unconventional assets with high exposure to idiosyncratic risk sources. Moreover, the sample for this study included all sets that are known of, instead of only those that are currently being traded on the market – which is crucial to prevent potential survivorship biases (i.e. if the researcher selects only those LEGO sets that are continuously traded on the secondary market, they automatically and trivially self-select for sets with current market price substantially higher than the retail price, thus biasing return estimations upward).

When accounted for survivorship bias, value-weighted and equal-weighted inflation-adjusted returns of LEGO across 1966-2018 show average annual returns of 1.33% and 1.20%, comparable to but lower than annual average equity raw returns of 2.35% and much lower than total equity returns of 5.54%. Nevertheless, LEGO returns are much less volatile evidencing the defensive nature of this unconventional asset class. Therefore, to assess the diversification benefits of LEGO as an asset class, the methodology of Briere, Oosterlinck and Szafarz (2015) is applied to derive Sharpe ratio improvements of a diversified portfolio including both equities and LEGO. Compared to annual Sharpe ratio of 0.32 for stocks, a portfolio with 5% of value-weighted LEGO index has a Sharpe ratio of 0.33. Investors with various levels of risk aversion might benefit from including from 5% to 25% of LEGO value-weighted index into their portfolio to maximise their utility.

This article is also the first in the field to study the structure of international investment in LEGO by synthesising almost 90,000 transactions on the LEGO secondary market from September, 2018 until March, 2019. Within 6 months, the value of LEGO investment exceeds \$7.5 million, with investors from 95 countries engaging in LEGO transactions. LEGO sellers and buyers originate from 65 and 90 countries, respectively. The international LEGO market is rather concentrated, though, with almost 60% of trading executed by investors from five countries: United States (27.28%), Germany (12.64%), Netherlands (10.88%), Canada (4.28%) and United Kingdom (4.21%). Most prominent countries with net positive LEGO capital accounts are Netherlands, Germany, Belgium Canada and Denmark, while United States, Japan, China, Australia and Italy have significant net negatives. LEGO investment is more prominent in large economies, countries with high per capita incomes and more egalitarian income distributions, suggesting that, unlike fine art or wine, LEGO can be a collectible and an alternative investment class affordable to all groups of investors.

The study also examines the degree of internationalisation of LEGO trading and finds a substantial local bias, with share of domestic trading exceeding countries' market share in 92 out of 95 cases, providing further evidence for the universality of the local bias phenomenon (Seasholes and Zhu, 2010). Nevertheless, the severity of local bias on the LEGO investment market can be partially explained within the rational agent framework if one takes into account shipping costs and negative carry associated with storage.

It is also found that holding LEGO sets is associated with non-pecuniary utility separable from their risk-return profile for some investors, evidenced from utility-augmented CAPM and Fama-French five-factor models. Annual LEGO portfolio returns are also shown to be uncorrelated with broad market benchmarks as well as notable risk factors.

Overall, LEGO sets are an emerging asset class with attractive returns, substantial defensive properties and significant diversification benefits. Collectible-savvy investors can ultimately consider including LEGO into their portfolios to enjoy a substantial improvement in Sharpe ratios. Some well-known stylised facts regarding asset return dynamics and investor behaviour, such as Fama-French size and value factors, liquidity effect, and local bias, are present in LEGO, reinforcing the validity and universal generalisability of these finance concepts.

The limitations of the study are associated with data availability, as older transactions cannot be retrieved from public sources, as well as with different survivorship assumptions regarding returns of non-traded LEGO sets. All of those can be overcome as the LEGO market matures and a greater dataset on secondary market transactions becomes available.

Further research might benefit from studying the efficiency of LEGO markets and constructing traditional Fama-French long-minus-short portfolios, as well as replicating this study's methodology on a larger sample. Testing C-CAPM and the discretion theory with regards to LEGO might also benefit this emerging field of research.

References

- Ait-Sahalia, Y., Parker, J., & Yogo, M. (2004). Luxury goods and the equity premium. The Journal of Finance, 59(6), 2959-3004.
- Auer, B., & Schuhmacher, F. (2013). Diamonds—A precious new asset? International Review of Financial Analysis, 28(1), 182-189.
- Baker, M., Bergstresser, D., Serafeim, G., & Wurgler, J. (2018). Financing the response to climate change: The pricing and ownership of US green bonds (No. w25194). National Bureau of Economic Research.
- Barry, C., & Brown, S. (1984). Differential information and the small firm effect. Journal of Financial Economics, 13(2), 283-294.
- Baumol, W. (1986). Unnatural value: or art investment as floating crap game. American Economic Review, 76(2), 10-14.
- Bocart, F., & Oosterlinck, K. (2011). Discoveries of fakes: Their impact on the art market. Economics Letters, 113(2), 124-126.
- Bollen, N. (2007). Mutual fund attributes and investor behavior. Journal of Financial and Quantitative Analysis, 42(3), 683-708.
- Borgers, A., & Pownall, R. (2014). Attitudes towards socially and environmentally responsible investment. Journal of Behavioral and Experimental Finance, 1(1), 27-44.
- Briere, M., Oosterlinck, K., & Szafarz, A. (2015). Virtual currency, tangible return: Portfolio diversification with Bitcoin. Journal of Asset Management, 16(6), 365-373.
- Bryans, D. (2014). Bitcoin and money laundering: mining for an effective solution. Indiana Law Journal, 89(1), 441-472.
- Burton, B., & Jacobsen, J. (1999). Measuring returns on investments in collectibles. Journal of Economic Perspectives, 13(4), 193-212.
- Burton, B., & Jacobsen, J. (2001). The rate of return on investment in wine. Economic Inquiry, 39(3), 337-350.
- Chuen, D., Guo, L., & Wang, Y. (2017). Cryptocurrency: A new investment opportunity? Journal of Alternative Investments, 20(3), 16-40.
- Czujack, C. (1997). Picasso paintings at auction, 1963–1994. Journal of Cultural Economics, 21(3), 229-247.

David, G., Oosterlinck, K., & Szafarz, A. (2013). Art market inefficiency. Economics Letters, 121(1), 23-25.

- Dimson, E., Marsh, P., & Staunton, M. (2018). Credit Suisse global investment returns yearbook 2018. Credit Suisse Research Institute.
- Dimson, E., Rousseau, P., & Spaenjers, C. (2015). The price of wine. Journal of Financial Economics, 118(2), 431-449.
- Dimson, E., & Spaenjers, C. (2014). Investing in emotional assets. Financial Analysts Journal, 70(2), 20-25.
- Dimson, E., & Spaenjers, C. (2011). Ex post: The investment performance of collectible stamps. Journal of Financial Economics, 100(2), 443-458.

- Dobrynskaya, V., & Kishilova, J. (2018). LEGO the toy of smart investors. ttp://dx.doi.org/10.2139/ssrn.3291456.
- Fama, E., & French, K. (2008). Dissecting anomalies. The Journal of Finance, 63(4), 1653-1678.
- Fama, E., & French, K. (2016). Dissecting anomalies with a Five-Factor Model. The Review of Financial Studies, 29(1), 69-103.
- Fogarty, J. (2010). Wine investment and portfolio diversification gains. Journal of Wine Economics, 5(1), 119-131.
- Foley, S., Karlsen, J., & Putnins, T. (2019). Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies? The Review of Financial Studies, 32(5), 1798-1853.
- Ginsburgh, V., & Penders, A. (1997). Land artists and art markets. Journal of Cultural Economics, 21(3), 219-228.
- Goetzmann, W. (1993). Accounting for taste: Art and the financial markets over three centuries. American Economic Review, 83(5), 1370-1376.
- Goetzmann, W. (1996). How costly is the fall from fashion? Survivorship bias in the painting market. In Contributions to Economic Analysis (pp. 71-84). Yale School of Management.
- Goetzmann, W., Renneboog, L., & Spaenjers, C. (2011). Art and money. American Economic Review, 101(3), 222-26.
- Graddy, K., & Margolis, P. (2011). Fiddling with value: Violins as an investment? Economic Inquiry, 49(4), 1083-1097.
- Hiraki, T., Ito, A., Spieth, D., & Takezawa, N. (2009). How did Japanese investments influence international art prices? Journal of Financial and Quantitative Analysis, 44(6), 1489-1514.
- Hulsmann, J. (2007). Mises: The last knight of liberalism. Ludwig von Mises Institute.
- Jorda, O., Knoll, K., Kuvshinov, D., Schularick, M., & Taylor, A. (2017). The rate of return on everything, 1870–2015 (No. w24112). National Bureau of Economic Research.
- Lucey, B., & Devine, L. (2015). Was wine a premier cru investment? Research in International Business and Finance, 34(1), 33-51.
- Mandel, B. (2009). Art as an investment and conspicuous consumption good. American Economic Review, 99(4), 1653-63.
- Mei, J., & Moses, M. (2002). Art as an investment and the underperformance of masterpieces. American Economic Review, 92(5), 1656-1668.
- Mok, H., Ko, V., Woo, S., & Kwok, K. (1993). Modern Chinese paintings: an investment alternative? Southern Economic Journal, 59(4), 808-816.
- Moroz, D., & Pecchioli, B. (in press). Should you invest in an old bottle of whisky or in a bottle of old whisky? A hedonic analysis of vintage single malt Scotch whisky prices. Journal of Wine Economics.
- Oosterlinck, K. (2017). Art as a wartime investment: Conspicuous consumption and discretion. The Economic Journal, 127(607), 2665-2701.
- Pesando, J. (1993). Art as an investment: The market for modern prints. American Economic Review, 83(5), 1075-1089.
- Renneboog, L., & Spaenjers, C. (2012). Hard assets: The returns on rare diamonds and gems. Finance Research Letters, 9(4), 220-230.
- Robertson, D., & Breen, B. (2013). Brick by brick. How LEGO rewrote the rules of innovation and conquered the global toy industry.
- Seasholes, M., & Zhu, N. (2010). Individual investors and local bias. The Journal of Finance, 65(5), 1987-2010.
- Shanaev, S., Sharma, S., Ghimire, B., & Shuraeva, A. (in press). Taming the blockchain beast? Regulatory implications for the cryptocurrency market. Research in International Business and Finance.
- Shanaev, S., Sharma, S., Valluri, S., & Shuraeva, A. (2019). Proof-of-What? Detecting original consensus algorithms in cryptocurrencies with a four-factor model. https://dx.doi.org/10.2139/ssrn.3395008
- Stein, J. (1977). The monetary appreciation of paintings. Journal of Political Economy, 85(5), 1021-1035.
- Tobin, J. (1958). Estimation of relationships for limited dependent variables. Econometrica, 26(1), 24-36.
- Urguhart, A. (2016). The inefficiency of Bitcoin. Economics Letters, 148(1), 80-82.
- Veblen, T. (2005). Conspicuous consumption (Vol. 38). Penguin UK.
- Veld, C., & Veld-Merkoulova, Y. V. (2007). Portfolio diversification benefits of investing in stamps. Journal of Alternative Investments, 10(1), 56-61.
- Wei, W. (2018). Liquidity and market efficiency in cryptocurrencies. Economics Letters, 168(1), 21-24.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica, 48(4), 817-838

Appendix

	Number	- C + -				Average ann	ual return, %	D		
Thoma	Number	of sets	Assump	otion (1)	Assum	ption (2)	Assump	otion (3)	Assum	otion (4)
Theme	Launched	Trading	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted
Action/Adventure	992	607	1.22	1.11	-5.30	-5.57	0.64	0.53	1.04	0.94
Basic	1243	299	0.49	-0.84	-21.85	-23.20	-0.97	-1.67	0.23	-0.40
Constraction	447	284	9.82	7.70	-7.46	-9.55	6.34	4.69	7.22	5.59
Educational	367	48	-0.92	-1.70	-26.08	-24.14	-1.68	-1.91	-0.31	-0.68
Girls	457	216	-2.47	-2.52	-18.48	-17.51	-2.62	-2.52	-1.55	-1.61
Historical	449	290	6.09	5.91	-1.31	-0.81	4.71	4.65	5.03	4.93
Junior	219	74	4.87	-1.51	-13.82	-19.07	1.50	-2.07	2.73	-0.76
Licensed	1483	1200	4.36	1.83	2.09	-0.56	4.06	1.60	4.20	1.75
Miscellaneous	1043	335	3.58	-1.93	-16.15	-11.05	1.04	-2.23	2.27	-1.52
Model making	120	102	5.13	3.39	3.11	0.83	4.80	3.07	4.89	3.19
Modern day	1441	669	1.73	1.08	-6.75	-7.56	0.90	0.33	1.35	0.82
Pre-school	1277	263	-0.23	-0.11	-31.43	-30.81	-2.01	-1.91	-0.07	-0.04
Racing	261	95	2.80	3.51	-26.70	-23.17	-0.34	0.32	1.25	1.70
Technical	500	272	0.29	-0.35	-10.54	-10.07	-0.36	-0.78	0.21	-0.27
Vintage themes	282	12	3.74	2.79	-14.99	-14.91	-0.28	-0.43	0.40	0.24
Other	6	5	-8.96	-14.58	-8.96	-14.58	-8.96	-14.58	-8.96	-14.58

Table A1. Average annual return of LEGO sets by broad theme

Table A2. Average annual return of LEGO sets by narrow topic

	Number	of sata			A	verage ann	ual return, 9	%		
	INUITIDEI	of sets	Assump	otion (1)	Assump	otion (2)	Assump	otion (3)	Assum	otion (4)
Topic	Launcha	Tradin	Equal-	Value-	Equal-	Value-	Equal-	Value-	Equal-	Value-
	Launche	Traum	weighte	weighte	weighte	weighte	weighte	weighte	weighte	weighte
	u	g	d	d	d	d	d	d	d	d
4 Juniors	23	3	-0.84	-2.21	-39.04	-42.83	-2.19	-2.39	-0.12	-0.12
Action Wheelers	9	3	1.91	2.00	-25.34	-19.09	-0.68	-0.05	0.64	0.96
Advanced Models	61	54	4.27	2.97	1.24	-0.28	3.84	2.59	3.98	2.74
Adventurers	71	39	5.14	5.25	-4.21	-8.09	3.60	3.03	4.01	3.62
Agents	13	6	4.42	3.75	-29.03	-34.33	0.20	-0.69	2.04	1.43
Alpha Team	32	17	2.42	1.47	-15.56	-18.95	0.58	-0.23	1.47	0.80
Aqua Raiders	7	6	3.08	3.14	-5.01	0.44	2.20	2.84	2.64	2.99
Aquazone	28	14	4.61	3.47	-12.45	-9.50	1.84	1.70	2.58	2.28
Architecture	43	37	6.70	4.51	6.70	4.51	6.70	4.51	6.70	4.51
Assorted	7	2	6.54	6.19	-11.24	-9.71	2.58	2.75	3.27	3.37
Atlantis	24	21	0.54	0.28	0.54	0.28	0.54	0.28	0.54	0.28
Avatar The Last	2	1	11.02	11.02	10.42	11 50	1 55	6 20	5.06	7 45
Airbender	2	1	11.92	11.92	-19.42	-11.56	4.55	0.39	5.90	7.45
Baby	43	1	3.98	3.98	-40.23	-40.30	-1.98	-1.99	0.10	0.10
Basic	397	29	1.44	0.85	-17.15	-21.20	-0.36	-0.86	0.46	0.14
Batman	15	9	12.61	11.59	-8.48	-1.68	7.73	8.66	8.73	9.29
Belville	82	14	0.49	-0.72	-32.17	-32.07	-1.57	-1.86	0.11	-0.18
Ben 10: Alien Force	6	5	5.63	5.63	-5.98	-5.98	4.02	4.02	4.69	4.69
Bionicle	337	219	9.79	7.24	-5.55	-9.71	6.57	4.32	7.29	5.17
Boats	12	4	4.21	3.77	-16.02	-16.20	0.84	0.57	1.69	1.41
Boost	1	1	-13.15	-13.15	-13.15	-13.15	-13.15	-13.15	-13.15	-13.15
BrickHeadz	54	50	39.17	65.13	39.17	65.13	39.17	65.13	39.17	65.13
Bricks and More	64	28	2.99	2.07	-30.68	-36.94	-0.58	-1.58	1.50	0.85
Bulk Bricks	137	33	3.08	2.72	-29.22	-29.31	-0.76	-0.85	0.77	0.68
Cars	27	18	3.30	1.84	-12.70	-5.47	1.54	1.17	2.59	1.65
Castle	264	164	6.01	5.70	-1.83	0.01	4.56	4.60	4.90	4.84
City	501	356	-0.37	-1.32	-2.68	-4.02	-0.59	-1.58	-0.35	-1.26
City	1	1	-13.16	-13.16	-13.16	-13.16	-13.16	-13.16	-13.16	-13.16
Classic	49	30	-10.79	-9.99	-19.10	-14.66	-7.94	-8.57	-7.35	-8.26

Clikits	77	13	-1.25	-0.45	-37.18	-41.95	-2.24	-2.29	-0.26	-0.04
Collectable Minifigures	31	24	42.42	29.56	42.42	29.56	42.42	29.56	42.42	29.56
Creator	357	171	1.29	-0.18	-15.48	-16.30	-0.07	-1.05	0.85	-0.12
Dacta	151	13	-0.43	-0.73	-22.11	-23.92	-1.16	-1.27	-0.10	-0.13
DC Comics Super	83	54	-1.28	-5.77	-5.55	-11.73	-1.59	-5.96	-1.22	-5.35
Heroes	11	10	14.20	12 (4	14.20	12 (4	14.20	12 (4	14.00	12 (4
DC Super Hero Girls	11	10	-14.29	-13.64	-14.29	-13.64	-14.29	-13.64	-14.29	-13.64
Dimensions	08	65	-14.90	-17.58	-14.90	-17.58	-14.90	-17.58	-14.90	-17.58
Dino Attack	5	1	7.63	7.63	-37.04	-21.06	-0.57	2 36	1 53	3 71
Dinosaurs	12	8	4 54	3 77	4 54	3 77	4 54	2.50	4 54	3 77
Discovery	6	6	8.16	5.91	8.16	5.91	8.16	5.91	8.16	5.91
Disney	40	35	-3.53	-5.42	-3.53	-5.42	-3.53	-5.42	-3.53	-5.42
Duplo	1052	242	-0.47	-0.30	-29.67	-30.09	-2.12	-2.01	-0.18	-0.11
Education	198	30	-0.86	-0.99	-30.08	-31.62	-1.96	-2.01	-0.31	-0.31
Elves	35	31	4.98	4.17	4.98	4.17	4.98	4.17	4.98	4.17
Exo-Force	38	29	4.70	3.28	-0.68	-0.17	3.98	2.89	4.26	3.08
Explore	122	17	2.58	1.48	-36.34	-33.08	-1.54	-1.41	0.38	0.33
Fabuland	100	5	5.65	5.53	-18.10	-21.81	0.17	-0.63	1.05	0.39
Factory	7	4	0.34	0.00	-20.46	-19.72	-0.93	-1.07	0.20	0.00
Freestyle	70	5	4.84	4.00	-28.02	-30.99	-0.71	-1.28	0.69	0.24
Friends	230	180	-2.84	-2.81	-6.53	-4.99	-3.23	-2.99	-2.71	-2.73
Galidor	19	/	-0.08	-0.22	-1/.4/	-25.55	-0.95	-1.40	-0.05	-0.08
Geal	30 2	1	19.74	19.74	-41.91	-54.69	-5.41	-7.05	5.29	0.72
Harry Potter	2 60	2 50	-1.40	0.70	-1.40	0.70 6.81	-1.40	0.70	-1.40	0.70
HERO Factory	104	60	10.28	9.62	-14 12	-9.16	5 70	6 22	7.17	7 36
Hobby Set	4	1	4.80	4.80	4.80	4.80	4.80	4.80	4.80	4.80
Ideas	22	19	7.17	6.32	7.17	6.32	7.17	6.32	7.17	6.32
Indiana Jones	18	17	10.65	9.55	10.65	9.55	10.65	9.55	10.65	9.55
Island Xtreme Stunts	13	5	0.67	0.26	-20.58	-21.46	-0.74	-0.99	0.34	0.13
Jack Stone	25	6	-2.24	-2.30	-30.14	-31.69	-2.12	-2.12	-0.61	-0.53
Juniors	66	55	5.88	-2.56	2.17	-6.16	4.59	-3.61	5.68	-2.46
Jurassic World	12	11	-1.59	-0.86	-1.59	-0.86	-1.59	-0.86	-1.59	-0.86
Legends of Chima	134	86	-4.44	-4.59	-11.57	-11.85	-4.61	-4.74	-4.02	-4.14
Legoland	142	7	5.37	5.35	-14.94	-16.53	0.03	-0.39	0.72	0.36
Life of George	2	2	-11.60	-11.63	-11.60	-11.63	-11.60	-11.63	-11.60	-11.63
Make and Create	34	1	-0.56	-0.56	-50.10	-51.09	-2.88	-2.99	-0.03	-0.03
Marvel Super Heroes	97	76	-0.65	-2.24	-1.74	-3.43	-0.74	-2.32	-0.64	-2.21
Master Builder Academy	15	6	1.74	1.82	1.74	1.82	1.74	1.82	1.74	1.82
Mindstorms	5	5 37	4.90	4.95	4.90	4.95	4.90	4.95	4.90	4.95
Minecraft	61	55	-2.94	-2.42 5.75	-22.39	-20.08	-2.97	-2.33 5.75	2 16	-1.44 5.75
Miscellaneous	69	13	2.10	5.75	-5.07	-32.92	5.04	-1 02	6.17	3 42
Mixels	125	117	13 78	13 73	13 78	13 73	13 78	13 73	13 78	13 73
Model Team	16	11	4.06	4.32	1.07	2.59	3.60	4.04	3.73	4.11
Monster Fighters	13	11	8.66	7.32	8.66	7.32	8.66	7.32	8.66	7.32
Nexo Knights	71	59	-8.60	-8.26	-8.60	-8.26	-8.60	-8.26	-8.60	-8.26
Ninjago	208	153	2.35	1.87	-0.18	-0.32	1.95	1.49	2.27	1.82
Pharaoh's Quest	9	7	5.37	7.04	5.37	7.04	5.37	7.04	5.37	7.04
Pirates	78	59	6.78	6.27	2.54	1.87	5.89	5.38	6.06	5.56
Pirates of the Caribbean	15	12	9.42	7.31	9.42	7.31	9.42	7.31	9.42	7.31
Power Functions	18	18	0.55	-0.07	0.55	-0.07	0.55	-0.07	0.55	-0.07
Power Miners	22	15	7.36	6.32	3.13	-3.03	6.67	4.92	6.90	5.43
Prince of Persia	0	5 10	-0.20	0.12	-0.20	0.12	-0.20	0.12	-0.20	0.12
Pioliotional	273	10	-2.90	-2.90	-37.30	-2.90	-4.30	-2.90	-0.50	-2.90
Racels Rock Raiders	15	82 7	5.30	4.33 5.04	-28.07	-24.10	-0.29	4.05	1.39	2.01
Samsonite	84	4	0.40	0.82	-13 74	-14 42	-0.54	-0.57	0.07	0.06
Scala	68	9	-1 45	-2 32	-25.96	-20.62	-1 73	-2.07	-0.45	-1 10
Scooby-Doo	6	5	15.52	12.18	15.52	12.18	15.52	12.18	15.52	12.18
Seasonal	128	90	9.36	9.01	3.16	-1.83	8.13	6.92	8.51	7.61
Serious Play	18	5	-2.54	-2.62	-2.54	-2.62	-2.54	-2.62	-2.54	-2.62
Service Packs	369	90	1.41	0.31	-14.45	-14.10	-0.01	-0.52	0.72	0.16
Space	300	117	3.56	2.67	-5.31	-5.71	2.20	1.57	2.57	1.93
Speed Champions	28	27	11.60	6.59	11.60	6.59	11.60	6.59	11.60	6.59
Spider-Man	10	7	6.21	5.29	-0.04	-1.23	5.14	4.28	5.43	4.59
SpongeBob SquarePants	14	13	7.50	7.13	2.92	4.11	6.73	6.63	6.97	6.79
Sports	122	41	3.44	4.54	-14.25	-13.93	1.26	1.88	2.11	2.74
Star Wars	586	474	1.41	1.13	-0.94	-1.75	1.21	0.89	1.35	1.07

Studios	54	16	5.80	5.26	-6.71	-18.84	3.66	1.33	4.22	2.41
Technic	396	215	0.79	0.08	-7.87	-6.45	0.18	-0.26	0.61	0.07
Teenage Ninja Turtles	20	16	1.09	0.14	1.09	0.14	1.09	0.14	1.09	0.14
The Angry Birds Movie	6	6	-3.27	-3.60	-3.27	-3.60	-3.27	-3.60	-3.27	-3.60
The Hobbit	23	17	5.87	5.03	1.19	0.65	5.21	4.45	5.54	4.76
The LEGO Batman	20	20	12.11	11.60	10 11	11.60	10 11	11.60	12 11	11.60
Movie	30	28	-12.11	-11.00	-12.11	-11.00	-12.11	-11.00	-12.11	-11.00
The LEGO Movie	29	21	3.13	2.50	3.13	2.50	3.13	2.50	3.13	2.50
The LEGO Ninjago	20	21	11.42	15 42	11 42	15 42	11 42	15 42	11 42	15 42
Movie	20	21	-11.42	-13.42	-11.42	-13.42	-11.42	-13.42	-11.42	-13.42
The Lone Ranger	8	7	4.19	2.87	4.19	2.87	4.19	2.87	4.19	2.87
The Lord of the Rings	15	14	8.29	5.77	2.86	5.77	7.39	5.77	7.73	5.77
The Simpsons	2	2	2.97	2.96	2.97	2.96	2.97	2.96	2.97	2.96
Time Cruisers	9	6	1.85	1.23	-7.01	-5.88	0.98	0.65	1.39	0.98
Town	554	180	3.91	3.73	-10.25	-12.05	1.76	1.36	2.38	2.04
Toy Story	15	13	5.07	4.81	5.07	4.81	5.07	4.81	5.07	4.81
Trains	165	53	4.85	4.27	-2.87	-5.47	3.51	2.72	3.84	3.15
Ultra Agents	14	14	2.11	1.35	2.11	1.35	2.11	1.35	2.11	1.35
Universal Building Set	8	2	-0.70	-0.81	-6.29	-5.41	-0.72	-0.80	-0.47	-0.58
Vikings	7	7	7.86	7.26	7.86	7.26	7.86	7.26	7.86	7.26
Western	20	14	6.28	6.18	-7.01	-10.96	3.82	3.17	4.40	3.94
Wizarding World	6	5	-8.96	-14.58	-8.96	-14.58	-8.96	-14.58	-8.96	-14.58
World City	26	12	3.69	4.13	-17.15	-8.14	1.09	2.50	2.11	3.10
World Racers	9	8	-1.00	-3.44	-1.00	-3.44	-1.00	-3.44	-1.00	-3.44
Znap	19	2	4.34	-0.14	-22.22	-28.15	0.28	-1.42	1.45	-0.03

Table A3. Average annual return of LEGO sets by the year of launch

	Number	of coto				Average ann	ual return, %			
Launch	INUIIDEI	of sets	Assum	otion (1)	Assump	otion (2)	Assump	otion (3)	Assum	ption (4)
year	Launched	Trading	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted
1966	48	1	2.96	2.96	-13.25	-10.60	-0.18	0.34	0.42	0.84
1970	18	2	-0.07	-0.01	-5.76	-3.66	-0.30	-0.17	-0.05	-0.01
1971	38	8	3.65	2.02	-11.43	-13.15	0.49	-0.15	1.04	0.45
1972	32	2	3.40	2.13	-3.66	-6.49	2.00	0.87	2.27	1.21
1973	41	6	2.81	2.84	-13.81	-15.34	-0.06	-0.32	0.58	0.38
1975	29	1	5.26	5.26	-17.28	-18.00	-0.43	-0.62	0.35	0.19
1978	50	2	1.59	1.78	-13.92	-12.13	-0.19	0.08	0.45	0.65
1979	63	1	2.82	2.82	-16.67	-19.03	-0.30	-0.68	0.47	0.18
1980	113	12	1.46	1.84	-13.09	-13.57	-0.10	-0.04	0.52	0.60
1981	71	3	4.61	4.86	-16.62	-14.44	0.07	0.59	0.87	1.31
1982	64	2	4.40	4.49	-14.48	-9.31	0.54	1.63	1.26	2.15
1984	74	4	6.92	7.37	-16.46	-13.41	0.72	1.63	1.54	2.35
1985	132	8	3.30	1.74	-16.52	-14.81	0.08	-0.10	0.88	0.61
1986	138	21	4.47	3.71	-10.60	-16.90	1.55	0.16	2.13	0.99
1987	185	28	3.61	5.54	-12.09	-9.99	1.00	2.14	1.63	2.72
1988	68	10	5.27	4.92	-10.96	-9.25	1.88	2.09	2.51	2.64
1989	130	28	4.84	6.22	-8.19	-6.33	2.31	3.36	2.82	3.83
1990	117	20	4.62	5.32	-10.46	-11.09	1.83	2.00	2.43	2.64
1991	134	36	3.16	3.10	-8.67	-8.64	1.46	1.44	1.96	1.93
1992	102	38	3.68	4.45	-6.39	-7.65	2.12	2.34	2.54	2.84
1993	139	45	3.93	4.36	-11.30	-11.00	1.52	1.77	2.16	2.40
1994	129	40	3.96	3.23	-10.95	-10.58	1.63	1.32	2.26	1.92
1995	158	45	3.14	2.71	-13.76	-12.79	0.88	0.81	1.62	1.50
1996	191	79	3.59	3.41	-11.10	-6.51	1.51	2.04	2.15	2.48
1997	246	84	3.40	2.54	-13.14	-10.92	1.17	1.00	1.91	1.61
1998	354	80	3.37	2.02	-19.31	-22.46	0.38	-0.44	1.40	0.70
1999	310	87	3.02	1.03	-18.81	-18.96	0.36	-0.48	1.37	0.49

2000	328	105	3.04	2.34	-21.14	-19.71	0.14	0.03	1.26	1.07
2001	349	128	4.97	4.05	-17.97	-16.19	1.39	1.25	2.43	2.19
2002	369	126	4.26	4.74	-21.52	-21.46	0.67	0.86	1.88	2.07
2003	333	139	4.53	5.05	-19.53	-22.83	1.13	0.86	2.27	2.17
2004	318	111	4.06	4.69	-23.07	-20.28	0.53	1.16	1.86	2.37
2005	314	110	5.62	6.06	-23.94	-21.86	1.09	1.59	2.53	2.94
2006	245	123	6.24	5.45	-15.88	-12.07	2.72	2.87	3.82	3.75
2007	242	114	7.00	7.70	-16.25	-16.02	3.14	3.54	4.31	4.72
2008	277	161	5.30	4.25	-12.54	-7.86	2.82	2.75	3.77	3.41
2009	307	189	5.74	6.01	-11.13	-8.32	3.34	3.92	4.27	4.71
2010	319	193	5.06	4.03	-9.83	-6.50	3.10	2.78	3.97	3.41
2011	338	221	4.99	4.76	-8.33	-4.38	3.26	3.60	4.08	4.16
2012	420	259	6.37	5.49	-10.55	-3.24	3.92	4.30	5.02	4.88
2013	440	297	3.25	1.46	-8.52	-2.96	1.92	1.05	2.78	1.38
2014	486	357	5.19	3.02	-0.44	-0.95	4.41	2.55	4.86	2.88
2015	514	409	3.15	1.30	0.92	-0.57	2.86	1.09	3.07	1.28
2016	544	422	5.69	-0.28	4.98	-1.35	5.57	-0.41	5.65	-0.28
2017	492	395	-6.66	-8.31	-7.12	-8.89	-6.71	-8.37	-6.62	-8.26
2018	385	219	-14.35	-21.97	-18.44	-24.82	-15.15	-22.30	-13.66	-21.17

Table A4. Average annual return of LEGO	O sets by piece count
---	-----------------------

			Number	mber of sets Accumption (1) Accumption (2) Accumption (3) Accumption (4)								
Т	Diaca (count	Number	of sets	Assump	otion (1)	Assump	otion (2)	Assump	otion (3)	Assump	otion (4)
1		Jouin	Launched	Trading	Equal-	Value-	Equal-	Value-	Equal-	Value-	Equal-	Value-
			Launcheu	Traunig	weighted	weighted	weighted	weighted	weighted	weighted	weighted	weighted
0	-	1	861	134	7.33	-3.97	-13.71	-19.69	3.07	-3.53	4.23	-2.43
1	-	10	1207	225	2.63	-0.03	-22.35	-23.93	-0.18	-1.31	1.08	-0.01
11	-	20	728	197	2.67	1.76	-21.27	-20.75	0.02	-0.35	1.31	0.82
21	-	30	882	302	5.72	2.43	-12.15	-18.53	2.63	-0.07	3.71	1.35
31	-	40	638	236	2.99	2.56	-15.42	-22.91	0.62	-0.43	1.80	1.16
41	-	50	619	244	4.03	3.14	-15.82	-22.09	1.32	0.06	2.35	1.34
51	-	60	435	219	3.70	2.16	-12.28	-18.11	1.67	0.13	2.53	1.24
61	-	70	442	225	6.79	3.82	-10.86	-22.08	3.57	0.42	4.47	1.86
71	-	80	279	147	3.54	1.81	-12.53	-22.63	1.53	-0.45	2.36	0.84
81	-	90	287	159	1.31	1.61	-11.56	-17.97	0.24	-0.12	0.96	0.92
91	-	100	274	131	1.64	0.06	-14.33	-19.57	0.19	-1.04	1.06	0.03
101	-	110	208	120	2.23	0.49	-10.15	-18.15	0.98	-0.75	1.65	0.30
111	-	120	150	76	1.29	0.85	-10.92	-11.82	0.05	-0.25	0.94	0.61
121	-	130	127	64	0.45	0.80	-15.77	-19.16	-0.61	-0.65	0.29	0.46
131	-	140	130	66	3.70	2.58	-12.20	-20.54	1.64	0.08	2.47	1.31
141	-	150	146	76	3.92	5.80	-11.45	-14.23	1.86	2.41	2.68	3.42
151	-	175	273	168	4.46	3.53	-4.93	-8.33	2.94	1.88	3.50	2.54
176	-	200	288	181	3.30	3.53	-5.12	-6.35	2.25	2.23	2.66	2.71
201	-	225	255	151	4.86	6.27	-7.79	-11.65	2.94	3.06	3.58	3.91
226	-	250	212	137	2.35	3.04	-6.41	-6.92	1.43	1.85	1.90	2.34
251	-	275	172	108	0.29	-1.13	-10.16	-11.51	-0.37	-1.43	0.22	-0.84
276	-	300	169	107	2.18	2.14	-6.90	-9.30	1.23	0.94	1.68	1.48
301	-	325	137	87	-0.40	-0.13	-10.16	-12.24	-0.81	-0.71	-0.29	-0.09
326	-	350	112	73	3.07	3.04	-4.05	-5.18	2.18	2.01	2.64	2.47
351	-	375	147	101	0.27	1.09	-6.12	-6.35	-0.11	0.50	0.23	0.87
376	-	400	123	85	0.67	1.22	-7.01	-6.87	-0.10	0.36	0.55	0.98
401	-	450	202	129	1.71	1.87	-7.01	-10.26	0.89	0.68	1.35	1.29
451	-	500	161	99	0.16	0.96	-8.72	-8.22	-0.34	0.24	0.12	0.68
501	-	550	133	80	1.21	1.25	-8.42	-7.54	0.19	0.37	0.94	0.94
551	-	600	135	95	0.34	0.65	-5.43	-3.84	-0.02	0.34	0.30	0.58
601	-	650	106	68	-0.51	-0.07	-13.34	-11.59	-1.13	-0.74	-0.38	-0.06
651	-	700	99	70	1.78	1.95	-7.34	-4.54	0.96	1.33	1.45	1.67
701	-	750	98	72	0.20	0.76	-8.53	-8.43	-0.32	0.11	0.16	0.60
751	-	800	65	50	1.51	1.98	-3.78	-3.61	1.05	1.43	1.33	1.71

801	-	900	107	71	-0.58	-0.43	-5.43	-6.52	-0.78	-0.70	-0.52	-0.37
901	-	1000	85	54	1.39	1.83	-7.70	-2.97	0.64	1.39	1.12	1.64
1001	-	1500	191	141	0.79	0.94	-5.04	-3.93	0.29	0.51	0.70	0.86
1501	-	2000	52	35	2.39	2.43	-6.09	-3.05	1.54	1.86	1.99	2.15
2001	-	10000	71	60	0.68	-0.28	-3.76	-4.30	0.29	-0.51	0.62	-0.25

Table A5. Average annual return of LEGO sets by physical volume

			Number	ofsets				Average ann	ual return, %)		
Vol	ume	cm ³	INUITOCI	01 5015	Assump	otion (1)	Assump	otion (2)	Assump	otion (3)	Assump	otion (4)
		,	Launched	Trading	Equal-	Value-	Equal-	Value-	Equal-	Value-	Equal-	Value-
			Baanonea	indding	weighted	weighted	weighted	weighted	weighted	weighted	weighted	weighted
0	-	500	213	186	10.98	10.04	8.90	7.89	10.19	9.31	10.26	9.38
501	-	1000	280	248	2.35	2.31	-2.75	-1.46	1.84	1.92	2.14	2.14
1001	-	1500	226	194	5.32	5.00	-2.59	-1.43	4.18	4.07	4.69	4.47
1501	-	2000	138	111	1.08	0.85	-9.57	-7.67	0.23	0.20	0.89	0.73
2001	-	3000	289	229	3.41	2.00	-9.30	-8.33	1.97	1.04	2.74	1.67
3001	-	4000	340	303	2.15	2.08	-4.62	-3.48	1.48	1.52	1.91	1.89
4001	-	5000	216	189	3.55	2.74	-4.48	-3.97	2.60	2.01	3.11	2.41
5001	-	6000	146	138	0.32	-0.29	-2.69	-2.16	0.11	-0.40	0.30	-0.28
6001	-	8000	212	184	1.54	1.27	-6.87	-4.94	0.76	0.72	1.34	1.13
8001	-	10000	188	165	1.53	1.03	-5.98	-4.24	0.84	0.60	1.35	0.94
10001	-	12500	129	112	-0.02	-0.46	-8.06	-5.13	-0.53	-0.72	-0.01	-0.42
12501	-	15000	118	107	0.96	0.25	-4.26	-2.88	0.55	0.04	0.87	0.24
15001	-	20000	137	125	1.53	1.60	-3.41	-1.44	1.12	1.34	1.40	1.51
20001	-	100000	100	85	1.45	1.43	-6.45	-2.20	0.79	1.09	1.24	1.29

Table A6. Average annual return of LEGO sets by weight

	Number	of acts				Average ann	ual return, %			
Weight kg	INUIIIDEI	of sets	Assump	otion (1)	Assump	otion (2)	Assump	otion (3)	Assump	otion (4)
weight, kg	Launched	Trading	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted
0 - 0.1	439	403	6.87	4.18	3.81	0.74	6.38	3.76	6.56	3.98
0.1 - 0.2	458	381	4.35	4.06	-6.74	-5.95	2.95	2.84	3.66	3.49
0.2 - 0.3	278	225	2.51	1.36	-8.80	-7.24	1.38	0.64	2.05	1.17
0.3 - 0.4	272	247	2.10	1.72	-3.38	-3.03	1.59	1.30	1.91	1.58
0.4 - 0.5	211	191	2.86	2.59	-3.04	-2.59	2.23	2.05	2.60	2.38
0.5 - 0.75	368	316	0.80	0.34	-7.45	-5.89	0.13	-0.12	0.69	0.31
0.75 - 1	237	208	2.28	2.02	-5.52	-3.39	1.47	1.49	2.02	1.86
1 - 1.5	247	216	0.40	-0.05	-5.94	-4.76	-0.02	-0.33	0.36	-0.05
1.5 - 2	132	118	1.16	0.83	-4.72	-3.33	0.69	0.52	1.05	0.78
2 - 3	83	72	2.55	2.18	-2.10	0.35	2.08	2.01	2.33	2.11
3 - 15	35	28	2.45	2.29	-1.02	-0.42	2.11	2.03	2.29	2.18

Table A7. Average annual return of LEGO sets by initial retail price (in 2018 US dollars)

	ail price Number of sets		of sots	Average annual return, %												
Reta	ail p	orice,	INUIIDEI	OI SELS	Assump	otion (1)	Assump	otion (2)	Assump	otion (3)	Assump	otion (4)				
	\$		Launched	Trading	Equal-	Value-	Equal-	Value-	Equal-	Value-	Equal-	Value-				
			Launcheu	Trauing	weighted	weighted	weighted	weighted	weighted	weighted	weighted	weighted				
0	-	3	73	47	23.34	8.96	2.18	-6.82	14.39	5.24	15.03	5.87				
3	-	4	86	75	12.54	12.25	6.41	5.90	10.71	10.38	10.94	10.62				
4	-	5	261	190	5.26	5.22	-5.66	-5.85	3.35	3.30	3.83	3.79				
5	-	6	415	343	7.59	7.55	-0.96	-1.27	5.85	5.77	6.27	6.20				

6 - 7	188	97	2.47	2.43	-16.67	-16.88	0.35	0.31	1.28	1.25
7 - 8	162	96	3.19	3.18	-12.94	-12.98	1.13	1.12	1.89	1.88
8 - 9	172	94	6.15	6.15	-15.99	-15.98	2.02	2.01	3.36	3.36
9 - 10	183	126	6.60	6.60	-8.31	-8.21	3.85	3.86	4.55	4.56
10 - 12	573	415	4.24	4.30	-8.59	-8.70	2.42	2.45	3.07	3.10
12 - 14	378	224	1.64	1.70	-16.00	-15.93	0.05	0.10	0.97	1.01
14 - 16	331	214	1.62	1.67	-14.78	-14.60	0.09	0.13	1.04	1.08
16 - 18	209	141	3.77	3.81	-13.82	-13.91	1.53	1.54	2.55	2.56
18 - 20	136	86	-9.19	-9.51	-21.61	-21.74	-7.20	-7.46	-5.81	-6.05
20 - 25	592	422	3.07	3.15	-11.11	-11.35	1.32	1.33	2.19	2.23
25 - 30	443	289	-0.94	-1.00	-15.74	-15.70	-1.52	-1.56	-0.61	-0.66
30 - 35	399	290	0.40	0.46	-11.69	-11.81	-0.42	-0.39	0.29	0.33
35 - 40	218	133	0.22	0.06	-16.69	-16.66	-0.90	-1.00	0.13	0.04
45 - 50	169	107	0.86	0.81	-15.88	-15.88	-0.51	-0.55	0.54	0.52
50 - 60	324	240	2.24	2.30	-9.64	-9.83	1.04	1.05	1.66	1.69
60 - 70	268	186	0.04	0.09	-13.08	-13.08	-0.71	-0.67	0.03	0.06
70 - 80	178	116	0.29	0.32	-14.69	-14.54	-0.63	-0.60	0.19	0.21
80 - 90	138	102	-0.33	-0.31	-10.60	-10.68	-0.79	-0.78	-0.25	-0.23
90 - 100	106	85	-0.67	-0.69	-8.80	-8.78	-0.98	-0.99	-0.54	-0.55
100 - 150	368	279	1.78	1.83	-7.90	-7.75	0.81	0.86	1.35	1.39
150 - 200	89	70	0.47	0.56	-8.73	-8.74	-0.17	-0.10	0.37	0.44
200 - 1000	120	94	1.18	0.24	-7.20	-7.89	0.50	-0.23	0.92	0.18

Table A8. Average annual return of LEGO sets by price-to-piece ratio

Price-to-piece ratio, cents per piece		Number of sets		Average annual return, %								
				Assumption (1)		Assumption (2)		Assumption (3)		Assumption (4)		
		Launched	Trading	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	
0	-	5	65	20	43.47	5.42	-17.49	-37.09	11.64	-1.38	13.37	1.22
5	-	6	43	21	-1.58	-1.60	-26.50	-29.66	-2.28	-2.43	-0.77	-0.72
6	-	7	79	52	2.32	-4.87	-12.97	-14.30	0.75	-4.38	1.53	-3.80
7	-	8	202	166	6.06	0.65	-3.39	-6.92	4.38	-0.01	4.98	0.55
8	-	9	327	286	3.56	2.87	-2.96	-2.27	2.76	2.31	3.12	2.60
9	-	10	390	339	1.44	0.47	-4.99	-4.77	0.85	0.05	1.25	0.42
10	-	11	455	406	1.79	0.71	-3.64	-3.44	1.28	0.39	1.60	0.66
11	-	12	438	376	1.48	0.11	-5.38	-5.84	0.90	-0.24	1.27	0.10
12	-	13	401	346	2.31	1.05	-4.26	-1.72	1.58	0.77	1.99	0.99
13	-	14	338	293	1.79	1.29	-4.03	-3.20	1.25	0.92	1.55	1.16
14	-	15	310	258	2.99	2.75	-4.60	-3.38	2.09	2.07	2.49	2.38
15	-	16	264	199	1.88	0.42	-8.23	-7.21	0.89	-0.07	1.41	0.34
16	-	18	395	293	4.03	4.63	-7.70	-5.59	2.28	2.99	2.99	3.57
18	-	20	337	238	4.62	4.17	-6.86	-5.22	2.75	2.70	3.26	3.12
20	-	25	599	380	3.31	2.93	-11.83	-10.19	1.27	1.26	2.10	1.91
25	-	30	345	227	2.31	2.40	-12.82	-9.81	0.72	0.95	1.52	1.55
30	-	35	207	116	2.20	1.13	-15.58	-17.89	0.33	-0.40	1.23	0.57
35	-	40	126	74	1.52	-1.32	-16.74	-18.54	-0.10	-1.69	0.89	-0.73
45	-	50	88	41	6.13	3.06	-22.71	-24.98	1.37	-0.16	2.86	1.40
50	-	60	173	85	1.68	0.45	-22.91	-22.96	-0.56	-1.10	0.83	0.22
60	-	70	93	41	3.79	0.76	-25.00	-24.07	-0.01	-1.06	1.67	0.36
70	-	80	93	41	1.17	0.87	-24.91	-26.70	-1.45	-1.48	0.52	0.32
80	-	90	88	48	1.12	0.18	-21.40	-23.53	-1.26	-1.76	0.61	0.09
90	-	100	72	32	1.17	0.23	-23.70	-28.61	-0.80	-1.48	0.52	0.07
N. Shimkus, S. Shanaev, S. Sharma, B- Ghimire/ Finance, Risk and Accounting Perspectives Proceedings (2019) 150-184

100	-	150	233	87	1.38	1.07	-27.18	-28.91	-1.02	-1.25	0.51	0.37
150	-	250	203	63	1.11	1.38	-26.36	-24.58	-1.07	-0.85	0.34	0.45
250	-	500	113	43	0.59	0.50	-22.83	-28.55	-0.96	-1.35	0.23	0.12
500	-	1000	78	29	0.04	-0.73	-22.46	-22.92	-1.12	-1.42	0.02	-0.28
1000	-	25000	119	63	-0.70	-0.82	-20.03	-23.09	-1.45	-1.71	-0.37	-0.42

Dei	Price-to-volume Number of sets			Average annual return, %								
PII	ratio	o.	INUITIDET	of sets	Assumption (1)		Assum	ption (2)	Assum	ption (3)	Assum	ption (4)
\$ per litre		Launched	Trading	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	
0	-	3	78	44	2.21	1.65	-24.22	-23.44	-0.38	-0.58	1.25	0.96
3	-	4	180	125	3.67	3.26	-18.06	-16.53	1.00	0.97	2.55	2.32
4	-	5	222	186	3.91	4.05	-6.41	-5.37	2.65	2.89	3.27	3.44
5	-	6	333	280	2.47	2.23	-7.10	-3.88	1.52	1.65	2.08	2.00
6	-	7	421	366	3.00	2.31	-5.01	-0.94	2.15	2.00	2.61	2.19
7	-	8	429	396	2.65	0.16	-2.35	-1.68	2.12	0.03	2.44	0.15
8	-	9	247	230	1.57	-0.19	-2.31	-1.09	1.23	-0.24	1.46	-0.18
9	-	10	234	216	1.33	0.00	-2.29	-1.39	1.01	-0.09	1.23	0.00
10	-	12	229	224	-0.26	-2.08	-1.56	-2.69	-0.34	-2.09	-0.25	-2.05
12	-	15	182	178	7.24	5.32	6.76	-0.44	7.07	3.64	7.08	3.82
15	-	20	100	93	5.81	2.47	4.33	0.10	5.36	2.06	5.40	2.15
20	-	200	41	38	8.81	-0.49	5.01	-1.82	7.99	-0.55	8.17	-0.48

Table A9. Average annual return of LEGO sets by price-to-volume ratio

Table A10. Average annual return of LEGO sets by price-to-weight ratio

Deia	a ta	maight	Number of sets			Average annual return, %								
PIIC	rati	o.	INUITIDEI	OI SELS	Assump	otion (1)	Assump	otion (2)	Assump	otion (3)	Assump	otion (4)		
\$ pe	er kil	ogram	Launched	Trading	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted	Equal- weighted	Value- weighted		
0	-	30	36	15	7.41	7.80	-30.10	-28.09	1.00	1.59	3.09	3.65		
30	-	35	50	32	0.16	0.22	-24.60	-24.46	-1.81	-1.76	0.11	0.14		
35	-	40	77	54	1.82	1.81	-16.45	-16.50	0.15	0.15	1.27	1.27		
45	-	50	118	100	3.89	3.87	-5.13	-5.12	2.78	2.77	3.30	3.28		
50	-	60	500	413	3.02	3.05	-7.84	-7.90	1.85	1.86	2.50	2.51		
60	-	70	588	535	2.64	2.63	-3.08	-3.09	2.06	2.05	2.41	2.39		
70	-	80	429	410	2.71	2.71	-0.22	-0.28	2.41	2.40	2.59	2.59		
80	-	90	236	220	0.68	0.73	-3.44	-3.42	0.38	0.42	0.64	0.68		
90	-	100	174	160	5.17	5.17	-0.40	-0.37	4.40	4.40	4.75	4.75		
100	-	150	328	318	3.65	3.72	1.31	1.42	3.38	3.45	3.54	3.61		
150	-	250	42	41	0.63	0.01	-1.38	-1.75	0.45	-0.14	0.62	0.01		
250	-	500	23	23	27.49	26.45	27.49	26.45	27.49	26.45	27.49	26.45		
500	-	25000	19	18	-2.49	-3.47	-6.49	-5.60	-2.68	-3.54	-2.36	-3.37		





Figure A1. Performance of value-weighted LEGO portfolios by year of launch

N. Shimkus, S. Shanaev, S. Sharma, B- Ghimire/ Finance, Risk and Accounting Perspectives Proceedings (2019) 150-184



Figure A2. Performance of value-weighted LEGO portfolios by piece count





Figure A3. Performance of value-weighted LEGO portfolios by physical volume





Figure A4. Performance of value-weighted LEGO portfolios by weight





Figure A5. Performance of value-weighted LEGO portfolios by initial retail price





Figure A6. Performance of value-weighted LEGO portfolios by price-to-piece ratio





Figure A7. Performance of value-weighted LEGO portfolios by price-to-volume ratio





Figure A8. Performance of value-weighted LEGO portfolios by price-to-weight ratio



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings



homepage: http://www.acrn.eu/frap

Entrenchment, Equity Incentives and Workplace Safety

Jesper Haga, Fredrik Huhtamäki, Dennis Sundvik*

Hanken School of Economics

Abstract

In this study, we investigate the association between managerial entrenchment, capital market incentives and workplace safety. Using an establishment level dataset comprising 4,274 unique establishments from 786 firms spanning the period 2002-2011 we find a positive association between managerial capital market incentives and injury frequency, and a negative association between managerial entrenchment and injury frequency. Furthermore, we find that establishments close to the company headquarter are characterized by a significantly higher number of workplace accidents.

Keywords: Workplace safety, managerial entrenchment, equity incentives, OSHA

Introduction

In the European Union (EU), workplace injuries amount to a cost of \notin 476 billion, which is the same as 3.3 percent of the total GDP in the EU (EU-OSHA, 2017). Each year in the United States (US), more than 3.5 million workplace injuries and illnesses occur. In 2007, these US workplace injuries cost the society around \$250 million, which is more than the \$219 million spent on cancer treatment (Leigh, 2011). The majority of these costs are borne by the workers themselves and their families, insurances carriers and the government. The firm's compensation to the worker is estimated to only be 20.7 percent of the total costs (Leigh and Marcin, 2012).

Firm executives in general and CEOs in particular play a critical role in corporate workplace safety practices. For example, BP's CEO Tony Hayward, emphasized economic efficiency and cost control in a manner inconsistent with an enduring safety culture (Amernic and Craig, 2017). On April 20, 2010, the Deepwater Horizon exploded in the Gulf of Mexico. CEOs do have power to influence workplace safety (Tucker et al, 2016), and in this study we examine the association between workplace injuries, CEO entrenchment, and equity incentives.

CEO utility maximization may lead to actions at odds with shareholder wealth maximization (Jensen and Meckling, 1976). Examples of such actions are when managers extract larger compensations and perquisites for themselves (Shleifer and Vishny, 1989) or make value-destroying acquisitions (Morck et al., 1990). These problems are especially pronounced for entrenched managers, i.e. managers less affected by discipline and control mechanisms.

Cronqvist et al. (2009) relate entrenchment to actions that benefit the work force at the expense of the owners, by showing that more entrenched CEOs pay higher wages to their employees, which constitute a private benefit to the CEO since it creates a better work environment and workers that are more loyal. In a similar line of thought, lower injury frequencies may also create satisfied and loyal employees, while excessive workplace safety may be value destroying.

We argue that the CEO aims to have an injury frequency that is in his own best interest, and this injury frequency might differ from the frequency that maximizes shareholder value. The CEO may get private benefits from keeping workplace injury frequencies at a low level. For example, a low level of workplace injuries may lead to a more loyal workforce and a better public image of the CEO. Our hypothesis is that CEOs on average wants to enjoy these private benefits. However, only entrenched CEOs or CEOs with misaligned incentives will actually reap these benefits. Pagano and Volpin (2005) present another theoretical argument supporting our view that entrenched CEOs have lower injury frequencies. According to their model, satisfied employees can protect the CEO's job and work against hostile takeovers and thus be a substitute for other entrenchment methods.

We suggest two possible ways in which an entrenched CEO might affect the injury frequency to gain private benefits. The first is through facilitating a low level of worker effort and a work environment with little pressure to perform. The second is through overinvestment in workplace safety. In both situations, firms engage in suboptimal activities that decrease the level of workplace injuries.

To measure whether CEO entrenchment relates to workplace injuries, we utilize a sample on workplace injuries from 4,274 establishments from 786 firms for the period 2002 to 2011. The data is from the US Occupational Safety and Health Administration (OSHA). We measure the CEO's entrancement with three firm-level measures: whether the CEO takes a dual role as chairperson or president, the CEO's pay in proportion to the five highest paid executives, and the percentage of times the CEO talks during conference calls.

We find evidence that firms with a more entrenched CEO have a lower frequency of workplace injuries. Our results are robust for all three measures of entrenchment. The magnitude of the results is both empirically and economically significant. For example, we estimate that an establishment with a CEO holding dual positions has 1.4 less workplace injuries per 100 workers per year than one where the CEO does not function as chairman or president. The descriptive statistics indicate that the driving force of the difference is fewer establishments with high injury frequencies among more entrenched CEOs. Since we link our finding to agency theory, we expect characteristics that decrease the agency problem to work in the opposite direction compared to our entrenchment variables. We utilize CEO stock ownership as a mitigating factor and find evidence supporting our expectation.

The literature is largely unanimous that a more entrenched CEO obtain private benefits at the cost of shareholder value. However, to our knowledge, no study investigates the relationship between entrenchment and workplace safety. The subject workplace safety has been studied comprehensively in other fields, but the link between CEO characteristics and workplace safety has also remained largely unexplored. In a related stream of the literature, Cohn and Wardlaw (2016) show that financially constrained firms often have a higher frequency of workplace injuries, while another part of the finance literature examines how stock market pressure affects workplace safety. Overall, these studies find that more pressure increases the likelihood of workplace injuries. Specifically, Caskey and Ozel (2017) show that the likelihood of injuries increases in firms that marginally beat analyst expectations, while Bai et al. (2018) find that an increase in short-sale pressure causes an increase in workplace injuries. Bradley et al. (2018) examine the impact of analyst coverage on workplace safety, and find that analyst coverage encourages the management to improve workplace safety.

Previous research also provides evidence that CEOs are favour workers geographically closer to the headquarters. For example, Landier et al. (2007) show that workers closer to the headquarters are less likely to be laid off. Furthermore, Cronqvist et al. (2009) find that these workers have higher wages if the CEO is entrenched. In terms of workplace injuries, we document that establishments in the same state as the headquarters have lower injury frequencies. However, we find no evidence that these injury frequencies are lower when the CEO is more entrenched.

The organization of the paper is as follows. In Chapter 2, we develop our hypotheses. Chapter 3 presents the data and the descriptive statistics. Chapter 4 describes the research method and Chapter 5 reports the results on the relationship between CEO entrenchment and injury frequency. Finally, Chapter 6 concludes.

Agency theory and injury frequencies

The separation of ownership and control may result in agency conflicts, if the manager derives personal utility from taking actions that are not in the interest of shareholders. For example, the manager can engage in empire building, shirking from taking a costly action that benefits the shareholders or perhaps consume an excessive amount of perquisites (Jensen 1986; 1993). A low injury frequency can be a type of private benefit for the CEO, in the form of a better public image as a responsible leader, in a similar way as Peterenko et al. (2015) show with regard to employee oriented actions and CSR. Another way would be in the form of personal relationships with the staff, i.e. through a better work climate and loyalty, similar to how Jensen and Meckling (1976), describes one element of private benefits. Two ways that the CEO can affect the injury frequency among workers are through investing resources in improving workplace safety or through exerting less pressure on the workforce. We agree that these actions do have the possibility to be maximizing shareholder wealth, as Cohn and Wardlaw (2016) suggest, but note that overinvesting in safety practises or reducing the productivity demands for the workforce are likely to be decisions that decrease shareholder wealth. In such a case, the costs would be borne by the shareholders, but the CEO receives the benefits.

In this paper, we follow Berger et al. (1997: 1411) and use the following definition of managerial entrenchment: "the extent to which managers fail to experience discipline from the full range of corporate government and control mechanisms, including monitoring by the board, the threat of dismissal or takeover and stock- or compensation-based performance initiatives". According to this definition, an entrenched CEO has more power to exert private benefits detrimental to the shareholder value, without suffering negative consequences in the form of disciplinary actions. More entrenched CEO's should therefore be able to deviate from the shareholder maximizing level of workplace safety to a higher degree than a less entrenched CEO. Managers can also use actions that benefit the workforce, such as creating a less demanding and safer work environment, at the expense of shareholders as a substitute for entrenchment. For example, Pagano and Volpin (2005) show that satisfied employees can protect CEO's job and private benefits of control against hostile takeovers. To protect their safe working environment workers may refuse to sell their shares or actively lobby against the takeover.

Based on the arguments of how an entrenched CEO have more power to exert private benefits and the substitutionary nature of entrenchment and a safe work environment we formulate our first hypothesis:

H1: CEO entrenchment is negatively related to the injury frequency.

One way to align the interests of the CEO with those of the shareholders is through equitybased compensation (Jensen and Meckling, 1976). A CEO owning stock may also be more inclined to put more demands on the workforce, which may lead to increased injuries. Accordingly, we formulate our second hypothesis:

H2: CEO equity incentives is positively related to the injury frequency.

Data description

To test these hypotheses, we combine data on public firms in the Compustat North America annual file with data on establishment-level workplace injuries from the Bureau of Labor Statistics' annual Survey of Occupational Injuries and Illnesses in the US.¹ This survey was a joint effort with the OSHA data initiative program that ended in 2011 due to funding cuts. The OSHA data set includes statistics on the state of occupational risk in various industries for a large variety of establishments for the 2002 to 2011 period generated by a stratified sampling process. Firms selected to be part of the survey are required to maintain a log recording any injuries "that result in death, loss of consciousness, days away from work, restricted work activity or job transfer, or medical treatment beyond first aid".

Following Caskey and Ozel (2017), the OSHA data is matched to firms in Compustat based on name searches to identify whether the establishment matches with a subsidiary of a Compustat firm. Each Compustat firm may have multiple establishments, and location and industry may differ by establishment. Our final sample covers 4,274 establishments from 786 firms for 10 years, which generates 44,427 establishment-year observations. About one tenth of the observations are located in California and the second most popular state is Texas (6.6 percent). Most establishments operate in the manufacturing industry, representing 57 percent of the observations.

For the purpose of our study, we are mainly interested in the data on the number of injuries, average number of employees, and total number of hours worked. We use these data to construct annual measures of the injury frequency at each establishment by dividing injuries by working hours multiplied by 200,000 (100 workers x 40 hours x 50 weeks). As examples of injuries, the data set includes deaths, skin disorders, respiratory conditions, poisonings, hearing losses, et cetera.

From the OSHA establishment-level data, we also obtain several control variables. We control for hours per employee, establishment work force, strikes or lockouts, shutdowns or layoffs, seasonal work, and natural disasters or adverse weather conditions that could affect the injury frequency. From Compustat, we obtain firm-level control variables, including capital expenditures, leverage, firm size, market to book, fixed assets, and turnover. We drop observations where any of these variables are missing. To control for outliers, we winsorize all non-dummy variables at the top and bottom 0.5 percent. Appendix A gives an overview of the variables used in the analysis.

We use three different measures for CEO entrenchment: CEO title concentration, the salary concentration within the firm and CEO expertise. Our first measure of entrenchment, Title Concentration, is in line with Morck et al. (1989). This variable is a dummy variable that is equal to one if the CEO holds both the title of chairperson and president. Shleifer and Vishny (1989) show that entrenched CEO receive higher salaries, as such we view a high salary as a sign of entrenchment. Following Bebchuk et al. (2011), we form our second variable, Salary Concentration, as the fraction of the aggregate salary of the top five executives obtained by the CEO. Our third and final measure is a measure of CEO expertise (Expertise). We use the measure of CEO knowledge by Li et al. (2014), which is obtained from the CEO behaviour during conference calls. The higher fraction of comments given by the CEO the more knowledge the CEO possesses. In our view, a CEO with more knowledge about the firm is also a more entrenched CEO.

To give an overview of how entrenchment and equity incentives are related to the injury frequency, Figure 1 shows box plots of how injury frequency is distributed when the sample is divided according to our entrenchment proxies. In Figure 1, the box plot to the left is the sample

¹ 1 We thank Dave Schmidt from the Occupational Safety and Health Administration, Office of Statistical Analysis for providing us the data on workplace injuries.

divided based on the median value of the Expertise variable, while the middle box plot is divided according to median Salary Concentration, and the box plot to the right is the sample divided based on the Title Concentration variable. Common for these box plots is that both the median injury frequency and the top percentiles are substantially lower for the sample characterized by more entrenchment.



Figure 1

Similarly, Figure 2 shows a box plot for injury frequency where the sample is divided according to whether the establishment has a CEO that is above or below the median percentage of owned shares. Here, the part of the sample that represents establishments with a CEO that has share ownership above the median, which we use as a proxy for equity incentive, is more prone to a higher injury frequency.



In Table 1, we present the descriptive statistics of our sample. Panel A of Table 1 provides firm and establishment level statistics for our sample, including our proxies for managerial entrenchment and equity incentives. About 24 percent of the CEO's in our sample also holds the title as chairman of the board or president, while the salary of the average CEO constitutes a third of the total salary received by the top five paid executives. Average CEO's do not own a substantial percentage of the total shares outstanding, with the median ownership constituting

0.19 percent of the total number of shares. On average, the employees work 1,966 hours in a year and the average establishment has 250 employees. The mean injury frequency, which we calculate in the same way as OSHA, is 7.4. This means that for every 100 employees, there is on average 7.4 injuries occurring during a year. The firms in our sample have leverage averaging at 25 percent and the size of book assets averaging at 10.5 billion USD.

Table 1: Descriptive statistics

_

_

Establishment Year Obs. Mean Median Std.de Max Mir CEO level variables	۱
CEO level variables	<u> </u>
Salary Concentration 44427 0.328 0.323 0.064 0.591 0	
Title Concentration 44427 0.242 0.000 0.428 1.000 0	
Expertise 1/557 40.332 38.773 23.24 100.000 0	
Percentage ownership 12212 3.870 0.194 9.361 53.577 0	
Establishment level variables	
Number Of Employees 44427 250 13 391 2665 17	
Hours Per Employee 44427 1966.199 1997.31 307.21 2930.6 1004	.9
Injury Frequency 44427 7.364 5.713 6.583 36.352 0	
Seasonal 44427 0.036 0.000 0.187 1 0	
Shutdown 44427 0.073 0.000 0.259 1 0	
Strike 44427 0.002 0.000 0.049 1 0	
Disaster 44427 0.005 0.000 0.073 1 0	
Firm level variables	
Capex/Assets 44427 0.056 0.044 0.037 0.704 0.00	0
Leverage 44427 0.246 0.229 0.156 1.511 0	
Ln(Assets) 44427 9.256 9.571 1.561 12.537 3.26	7
M/B 44427 1.872 1.671 0.834 11.866 0.46	5
PPE/Assets 44427 0.388 0.363 0.186 0.926 0.01	0
Turnover444271.4281.2890.6777.5800.01	9
Panel B: Observations by year	
Year Observation Percent Cum.	
S	
2002 3.61 8.13 8.13	
2003 4.73 10.64 18.77	
2004 4.33 9.75 28.52	
2005 6.08 13.68 42.20	
2006 5.76 12.97 55.17	
2007 5.16 11.60 66.78	
2008 6.00 13.48 80.26	
2009 2.83 6.38 86.63	
2010 4.56 10.26 96.89	
2011 1.38 3.11 100.0	
Total 44.43 100.00	

Panel A: Descriptive statistics on CEO-, establishment- and firm level variables

This table presents descriptive statistics. The data sample covers 4274 establishments from 786 firms spaning the period 2002 to 2011. Definitions of variables are presented in Appendix A.

Table 1 Panel B shows the number of establishment level observations, measured in thousands for each year in our dataset. The sample is moderately evenly distributed across the years, although the final year in the sample only has 1,380 observations compared to 2005, which has the most observations with approximately 6,000.

Table 2 reports the correlations between our variables. The correlations between all three of our measures of CEO entrenchment are positive. Of these correlations the highest is 0.4, which is the one between Title Concentration and Salary Concentration. The low correlation between our entrenchment measures suggest that the measures are capturing slightly different aspects of CEO entrenchment. In line with our hypothesis, there is a negative correlation between the injury frequency and two of the entrenchment variables, which are Title Concentration and Salary Concentration.

Table 2: Correlations

	Injury Frequency	Salary Concentration	Title Concentration	Hours per Employee	Number of Employees	Ln(Assets) Leverage Turnover	PPE/Assets	Capex/Assets	M/B	Strike	Shutdown	Seasonal	Disaster	Expertise
Injury Frequency	1													
Salary Concentration	-0.137***	1												
Title Concentration	-0.0985***	0.406***	1											
Hours per Employee	-0.255***	0.0140***	-0.0225***	1										
Number of Employees	0.00297	0.00966**	0.0340***	-0.0657***	1									
Ln(Assets)	-0.00104	-0.0193***	-0.0355***	-0.145***	0.0993***	1								
Leverage	0.0975***	-0.278***	-0.210***	0.0764***	-0.0330***	-0.184*** 1								
PPE/Assets	0.134***	0.0483***	0.0879***	-0.187***	-0.103***	0.235*** -0.125***	1							
Turnover	0.0530***	0.121***	0.130***	-0.00712	-0.00553	-0.0689*** -0.342***	0.308*** 1							
Capex/Assets	0.159***	0.197***	0.146***	-0.0846***	0.0102**	0.111*** -0.335***	0.520*** 0.375***	1						
M/B	0.0994***	0.0917***	0.0605***	-0.133***	0.0236***	0.192*** -0.493***	0.164*** 0.250***	0.396***	1					
Strike	0.0238***	-0.00766	0.000657	0.00445	0.0316***	-0.00720 0.0313***	-0.0153*** -0.00905*	-0.0138***	-0.0199***	1				
Shutdown	-0.0386***	-0.0160***	0.0193***	0.00186	0.0692***	-0.0738*** 0.0510***	-0.150*** -0.128***	-0.144***	-0.123***	0.0328***	1			
Seasonal	0.0432***	0.0824***	0.0701***	-0.0753***	0.0384***	0.00121 -0.0602***	0.0217*** 0.0521***	0.0693***	0.123***	0.0104**	0.0165***	1		
Disaster	0.0160***	-0.00706	0.0130***	0.0104**	0.00458	-0.0190*** 0.0151***	-0.0152*** 0.00577	0.00459	-0.00866*	0.00282	0.0333***	0.0356***	1	
Expertise	0.263***	0.0978***	0.217***	0.231***	-0.0306***	-0.381*** 0.293***	-0.204*** -0.00904	-0.320***	-0.348***	0.0201***	0.0344*** -	0.0511*** -0	0.00976	1

Research Methodology

To investigate the relationship between managerial entrenchment and injury frequency (H1) we utilize establishment level data and employ the following regression model:

Injury Frequency_{*e*,*t*} = $\alpha_{0+\beta_1}$ Entrenchment_{*e*,*t*} + β_{2-9} Establishment Controls_{*e*,*t*} + β_{10-13} Injury Controls_{*e*,*t*} + θ Year_{*t*} + φ Industry_{*e*} + $\varepsilon_{e,t}$, (1)

where t indexes year and e indexes establishment. The variable injury frequency is the number of cases in an establishment per year divided by the total number of hours worked and multiplied by 200,000. Entrenchment is one of our proxy for managerial entrenchment.

We follow Caskey and Ozel (2017) and include establishment controls in the form of firm characteristics that may have an effect on the injury frequency. Ln(Assets) is the natural logarithm of total assets, Leverage is the total short- and long term debt divided by total assets, PPE/Assets is net property plant and equipment divided by total assets, Turnover is the total sales divided by assets, Capex/Assets is the capital expenditure divided by total assets, and M/B is the market-to-book ratio. The establishment controls, except Sales/Assets, Capex/Assets and M/B are lagged by one year.

The injury control variables are hours per employee, i.e. total establishment work hours divided by the average number of workers, Establishment Size is the yearly average of employees per establishment, measured in thousands. Strike, Shutdown, and Disaster are dummy variables for those occurrences, while Seasonal is a dummy variable taking the value 1 if the establishment employed seasonal workers in the corresponding year. We include year and industry dummies to control for unobservable heterogeneity. In all regressions, we use robust standard errors, clustered at the firm level.

To investigate whether CEO equity incentives is associated with the injury frequency (H2) we add an Equity incentive variable to Equation (1):

Injury Frequency_{*e*,*t*} = α_0 + β_1 Equity Incentive_{*e*,*t*} + β_2 Entrenchment_{*e*,*t*} + $\beta_{3^{-10}}$ Establishment Controls_{*e*,*t*} + $\beta_{11^{-14}}$ Injury Controls_{*e*,*t*} + θ Year_{*t*} + φ Industry_{*e*} + $\varepsilon_{e,t}$ (2)

The equity incentive variable is the percentage of shares owned by the CEO. We report the results from the OLS regressions in the following chapter. However, as Caskey and Ozel (2017) notes, arguments can be made that a count model would be a better fit, considering the nature of the dependant variable. Thus, we employ both a poission model and a negative binomial model. The results from those variations are reported in Appendix B, and we find that the results are similar to the OLS regressions.

Furthermore, as Cronqvist et al. (2009) notes, the reason the CEO might chose to put less pressure on the work force is related to the private benefits for such a decision. One of the private benefits may for example arise through dealing with a more content workforce. To examine this line of thought we investigate whether the proximity of the establishment to the company headquarter has an effect on the injury frequency, by including Headquarter, a dummy variable taking the value 1 if the establishment is located in the same state as the headquarter, and 0 otherwise. In the following regression model, we interact the dummy variable with our measures for CEO entrenchment to capture whether the entrenchment of CEO's have a different association to the injury frequency in establishment closer to the headquarter:

Injury Frequency_{*e*,*t*} = $\alpha_{0 + \beta_1}$ Entrentchment_{*e*,*t*} + β_2 Headquarter_{*e*,*t*} + β_3 Headquarter_{*e*,*t*} × Entrentchment_{*e*,*t*} + β_{4-11} Establishment Controls_{*e*,*t*} + β_{12-15} Injury Controls_{*e*,*t*} + θ Year_{*t*} + φ Industry_{*e*} + $\varepsilon_{e,t}$ (3)

Results

To test the first hypothesis (H1) whether managerial entrenchment is negatively related to the injury frequency, we regress the injury frequency on our proxies for managerial entrenchment and controls (Equation 1). Table 3 presents the results, with univariate regressions for each entrenchment proxy in Model (1), (3), and (5), and regressions with the full set of controls in Model (2), (4), and (6). In Model (1) and (2) of Table 3, we measure it by the concentration of titles. In Model (3) and (4) of Table 3, we use the salary of the CEO as a fraction of aggregate salary of the top five highest paid executives. In Model (5) and (6) of Table 3, we use the average percentage of comments by the CEO during conference calls as the entrenchment variable. Since the dataset on how much the manager speaks during conference calls spans a shorter time period than the OSHA dataset, the number of observations drops from 44,427 establishment-years in Model (1)-(4) to 17,557 in Model (5) and (6). The coefficient for the managerial entrenchment proxies are negative and statistically significant at the 1 percentage level in all models, which indicates a negative relationship between managerial entrenchment and injury frequency. The size of the coefficient on title concentration indicates that an establishment where CEO who also is the chairman or the president have on average 1.3 injuries less per year per 100 full time employees. Similarly, the coefficient on salary concentration indicates that for each ten percentage more the CEO is compensated through salary in relation to the highest compensated five directors, injuries decreased by 1.2 per year per 100 full times employees. Notably, the variable Hours Per Employee is negatively related to the injury frequency, which we interpret as an indication that it is not a high amount of work hours in themselves that drive the injury frequency but rather how the work is conducted. In line with Cohn and Wardlaw (2016), we also find a positive relationship between leverage and the injury frequency, suggesting that companies financially constrained companies have more injuries. The coefficients for the dummy variable Strike and Shutdown are positive and significant in the majority of the models and positive and insignificant in one model, which we suggest that they might be indicators of a poor work environment leading to more injuries. Similarly, to Cohn and Wardlaw (2016), we also find that the coefficient for the natural logarithm of assets is negative and significant. The R-square are in the region of 24.6 to 31 percentage, although the majority of the explained variation comes from the industry and year fixed effects.

J. Haga. F. H	uhtamäki, D.	Sundvik / Financ	e, Risk and	Accounting	Perspectives	Proceedings	(2019)	185-
			202					

Dependent variable.:	Injury freque	ncy					
	_	(1)	(2)	(3)	(4)	(5)	(6)
Entrenchment variables							
Title Concentration	-1	.424***	-1.317***				
	(-4	4.109)	(-4.027)				
Salary Concentration				-12.920***	-12.002***		
				(-3.574)	(-4.036)		
Expertise						-0.037***	-0.032***
						(-2.824)	(-3.644)
Control variables							
Hours Per Employee			-0.004***		-0.004***		-0.005***
			(-4.975)		(-4.914)		(-4.475)
Establishment Size			0.108		0.080		0.129
			(1.053)		(0.786)		(1.075)
Ln(Assets)			-0.324***		-0.329***		-0.550***
			(-2.733)		(-2.727)		(-3.961)
Leverage			4.035***		3.498***		6.635***
C			(3.641)		(3.694)		(2.764)
PPE/Assets			-1.272		-1.828		-7.280***
			(-1.025)		(-1.482)		(-3.519)
Turnover			-0.330		-0.362		-0.228
			(-0.974)		(-0.992)		(-0.571)
Capex/Assets			14 975		16 012		32 487**
Cupent Tibbeto			(1 239)		(1.420)		(2.480)
M/B			-0.008		-0.058		0.076
			(-0.042)		(-0.284)		(0.316)
Strike			(0.042) 2 321***		2 286**		1.626
Strike			(2.521)		(2.504)		(1.020)
Shutdown			(2.009)		(2.504)		(1.057)
Shutdown			(2, 100)		(2,000)		(1, 234)
Saaconal			(2.190)		(2.090)		(1.234) 1 506***
Seasonal			-0.031		(0.020)		(2.259)
Disastar			(-0.055)		(0.044)		(3.258)
Disaster			1.120*		1.003		0.699
			(1.667)		(1.553)		(0.998)
Constant	5	5.125***	15.479***	9.129***	19.803***	5.584***	18.244***
		(4.378)	(5.711)	(5.501)	(7.369)	(3.526)	(5.075)
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes
Establishment Yrs		44427	44427	44427	44427	17557	17557
R-sar		0.246	0.288	0.252	0 202	0 2/0	0.310

* p<0.1, ** p<0.05, *** p<0.01

We hypothesize that equity incentives align the interest of the CEO with interest of the shareholders, which would decrease the net private benefit a CEO receives from having suboptimal high workplace safety. To test this, we employ the regression in Equation (2). The variable for CEO equity incentive, i.e. the percentage owned of the outstanding shares, is positive and highly significant in all models. The size of the coefficient ranges from 0.05 to 0.06 which can be interpreted as for each percentage of the outstanding shares the CEO owns, the injuries increase by 5 per 1,000 full time employees. However, when including the equity incentive variable for CEO entrenchment is insignificant. It should be noticed that this is a substantially smaller dataset than the one in Table 3. Very few observations exist in the dataset for the amount of CEO ownership before 2006. However, the results indicate that there exists a positive association between CEO equity incentives and worker injuries.

Dependent variable.:	Injury free	quency					
		(1)	(2)	(3)	(4)	(5)	(6)
Equity Incentives		0.066*** (3.906)	0.051*** (3.311)	0.065*** (3.659)	0.050*** (3.105)	-0.013 (-1.630)	-0.022** (-2.603)
Entrenchment variables							
Title Concentration		-0.351	-0.257				
Salary Concentration		(-1.129)	(-0.904)	-4.164	-3.656		
Expertise				(-1.011)	(-0.915)	0.007	0.000
Control variables						(0.560)	(0.020)
Hours Per Employee			-0.002***		-0.002***		-0.002**
Establishment Size			(-4.283) 0.175		(-4.276) 0.174		(-2.583) -0.294
Ln(Assets)			(1.387) -0.214*		(1.384) -0.216*		(-1.231) -0.218
,			(-1.656)		(-1.707)		(-0.794)
Leverage			0.981 (0.916)		0.983 (0.938)		5.827*** (2.723)
PPE/Assets			-2.051* (-1.728)		-2.087* (-1.736)		1.348 (0.438)
Turnover			0.333		0.303		1.013***
Capex/Assets			(1.295) 5.650		(1.202) 5.321		(2.981) -8.054
M/B			(0.961) -0 492**		(0.910) -0 491**		(-0.778) 0.455
1417 D			(-2.234)		(-2.230)		(1.202)
Strike			1.941		2.042		-2.429*
Shutdown			(1.547) 0.182		(1.551) 0.190		(-1.738) 1.017**

Table 4: Injury frequency, equity incentives and entrenchment

		202				
		(0.756)		(0.794)		(2.040)
Seasonal		-0.555		-0.580		0.292
		(-0.897)		(-0.944)		(0.400)
Disaster		0.469		0.467		5.675***
		(0.432)		(0.437)		(7.489)
Constant	10.481***	16.825***	11.897***	16.874***	6.393***	12.105***
	(9.777)	(7.708)	(6.985)	(4.596)	(3.996)	(3.375)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Establishment Yrs	12212	12212	12212	12212	1136	1136
R-sqr	0.275	0.290	0.276	0.290	0.222	0.251
* .0.05 ** .0.01 **** .0.001						

* p<0.05, ** p<0.01, *** p<0.001

If entrenched CEO's receive a private benefit in the form of a better work climate when the workers in the firm are less pushed, this effect should be strongest for workers in closer proximity to the CEO. To investigate this, we include a dummy taking the value one when the establishment is in the same state as the company headquarter. This is the case for 13 percent of the establishments in the sample. Table 5 reports the results.

Table 5: Headquarter state and injury frequency Dependent variable.: Injury frequency (1) (2) (3) (4) Headquarter state 0.83751** 0.99618** 1.83029* 1.12728** Entrenchment variables Title Concentration 1.40576** Salary Concentration * (-4.09) (-4.04) Expertise (-3.68) Entrenchment * Expertice 0.47404* 2.79391 -0.00090 (1.66)(1.16)(-0.13)Control variables Hours Per Employee 0.00423** 0.00421** 0.00424** 0.00487** Establishment Size 0.13084 0.14024 0.11448 0.16321 (1.28)(1.39)(1.13)(1.36)Ln(Asset 0.33717** 0.35155** 0.35653** 0.59361** s) * (-2.83) * (-2.94) * (-2.94) * (-4.24) 1 52288** 2 07168** 2 /2100** 66/182** PPE/Assets -1.91637 -1.62340 -1.35503 (-1.10)(-1.09)(-1.54)(-3.53)Turnover -0.43537 -0.32914 -0.36237 -0.24005

Capex/Assets	(-1.22) 14.18235 (1.09)	(-0.96) 15.63940 (1.28)	(-0.98) 16.71879 (1.46)	(-0.59) 33.39829** (2.49)
M/B	0.00178	-0.01234	-0.06390	0.07171
	(0.01)	(-0.06)	(-0.31)	(0.29)
Strike	2.36317**	2.38721**	2.36271***	1.92217
	(2.63)	(2.74)	(2.61)	(1.29)
Shutdown	0.54350**	0.57578**	0.55071**	0.52821
	(2.05)	(2.18)	(2.10)	(1.21)
Seasonal	-0.31804	-0.06955	-0.01529	1.56309***
	(-0.45)	(-0.12)	(-0.03)	(3.19)
Disaster	1.14847*	1.14923*	1.03093	0.62657
	(1.66)	(1.66)	(1.54)	(0.88)
Constant	15.77863**	* 16.08386**	20.58621**	19.16387**
	* (5.71)	* (5.73)	* (7.41)	* (5.25)
Year Fixed Effects	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes
Establishment Yrs	44427	44427	44427	17557
<u>R-sqr</u>	0.284	0.290	0.295	0.313

* p<0.05, ** p<0.01, ***

In Table 5, the coefficient for the headquarter state dummy is negative and significant for all specifications, which is also the case for the entrenchment proxies. These findings suggest that there is indeed fewer injuries in establishments located in the same state as the company headquarter and as Table 3 also reports, there is a negative association between entrenchment and the injury frequency. The interaction between head state and the entrenchment proxy is positive in two cases and significant at the 10 percent level in one case, which indicates that when the CEO is entrenched there is a smaller difference in injury rates between establishments in the headquarter states compared to establishments in other states. One explanation is that entrenched CEOs put emphasis on the companywide workplace safety.

Conclusions

The impact CEOs have on workplace safety has previously been a facet left unexplored in existing literature. In this study, we provide evidence that entrenchment and equity incentives are related to injury frequencies in US firms based on establishment level injury data. First, we find that CEO entrenchment is negatively associated with the frequency of workplace injuries. Second, we find that CEO equity incentives are positively associated with workplace injuries and when including equity incentives, our entrenchment proxies are not significant. Finally, we show that establishments that are located in the same state as the company headquarter have a significantly lower injury frequency than establishments located elsewhere.

We argue that the underlying reason for the negative relation between entrenchment proxies and injury frequency is that CEOs receive a private benefit from better workplace safety in the form of a better public image and a more loyal workforce. Furthermore, we argue that CEO equity incentives mitigate the entrenchment effect and results in a higher degree of workplace injuries.

Our study contributes to the literature in several ways. However, we mainly contribute to research on the consequences of managerial entrenchment for shareholders and other stakeholders. While most of the prior studies highlight that entrenchment is detrimental for

shareholders (e.g. Morck et al., 1988; Morck et al., 1990), voices also claim that entrenched managers affect employees and for instance, their pay (Cronqvist et al., 2009). We build on this literature by connecting entrenchment and workplace safety. Understanding the drivers and moderating factors of workplace injuries is very important, considering the large negative effect of such injuries on the economy as a whole.

Lastly, while our analyses are robust to alternative model specifications and setups, we acknowledge that our study has caveats. For example, we recognize that there are obvious weaknesses in the study regarding endogeneity.

References

- Amernic, J. and Craig, R., 2017. CEO speeches and safety culture: British Petroleum before the Deepwater Horizon disaster. Critical Perspectives on Accounting, 47(1), pp.61–80.
- Bai, J., Lee, E., and Zhang, C., 2018. Capital market frictions and human capital investment: Evidence from workplace safety around Regulation SHO. Working paper.
- Bebchuk, L.A., Cremers, K.J.M., and Peyer, U.C., 2011. The CEO pay slice. Journal of Financial Economics, 102(1), pp.199–221.
- Berger, P.G., Ofek, E., and Yermack, D.L., 1997. Managerial entrenchment and capital structure decisions, Journal of Finance 52(4), pp.1411–1438.
- Bradley, D., Mao, C., and Zhang, C., 2018. Does analyst coverage affect workplace safety? Working paper.
- Caskey, J. and Ozel, N.B., 2017. Earnings expectations and employee safety. Journal of Accounting and Economics, 63(1), pp.121–141.
- Cohn, J.B. and Wardlaw, M.I., 2016. Financing constraints and workplace safety. Journal of Finance, 71(5), pp.2017–2058.
- Cronqvist, H., Heyman, F., Nilsson, M., Svaleryd, H. and Vlachos, J., 2009. Do entrenched managers pay their workers more? Journal of Finance, 64(1), pp.309–339.
- EU-OSHA (2017). Work-related accidents and injuries cost EU €476 billion a year according to new global estimates. Available at: https://osha.europa.eu/en/printpdf/12566 (last accessed: October 31, 2018).
- Jensen, M.C. and Meckling, W.H., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics, 3(4), pp.305–360.
- Jensen, M.C., 1986. Agency costs of free cash flow, corporate finance, and takeovers.
- American Economic Review, 76(2), pp.323-329.
- Jensen, M.C., 1993. The modern industrial revolution, exit, and the failure of internal control systems. Journal of Finance, 48(3), pp.831-880.
- Landier, A., Nair, V.B., and Wulf, J., 2007. Trade-offs in staying close: Corporate decision making and geographic dispersion. Review of Financial Studies, 22(3), pp.1119–1148.
- Leigh, J.P., 2011. Economic burden of occupational injury and illness in the United States. The Milbank Quarterly, 89(4), pp.728–772.
- Leigh, J.P. and Marcin, J.P., 2012. Workers' compensation benefits and shifting costs for occupational injury and illness. Journal of Occupational and Environmental Medicine, 54(4), pp.445–450.
- Li, F., Minnis, M., Nagar, V., and Rajan, M., 2014. Knowledge, compensation, and firm value: An empirical analysis of firm communication. Journal of Accounting and Economics, 58(1), pp.96–116.
- Morck, R., Shleifer, A., and Vishny, R.W., 1988. Management ownership and market valuation: an empirical analysis. Journal of Financial Economics, 20(1), pp.293–315.
- Morck, R., Shleifer, A., and Vishny, R.W., 1990. Do managerial objectives drive bad acquisitions? Journal of Finance, 45(1), pp.31–48.
- Shleifer, A. and Vishny, R.W., 1989. Management entrenchment: The case of manager-specific investments. Journal of Financial Economics, 25(1), pp.123–139.
- Pagano, M. and Volpin, P.F., 2005. Managers, workers, and corporate control. Journal of Finance, 60(2), pp.841-868.
- Petrenko, O., Aime, F., Ridge, J. and Hill, A., 2005. Corporate social responsibility or ceo narcissism? CSR motivations and organizational performance. Strategic Management Journal, 37(2), pp.262–279.
- Tucker, S., Ogunfowora, B., and Ehr, D., 2016. Safety in the c-suite: How chief executive officers influence organizational safety climate and employee injuries. Journal of Applied Psychology, 101(9), pp.1228–1239

Appendix A Variable Descriptions (Source)

Expertise	Average percentage of company comments spoken by the CEO during conference calls. (Li et al. (2014))
Title Concentration	A dummy variable taking the value 1 if the CEO is the chairman of the board or the president. (Execucomp)
Salary Concentration	How large the salary of the CEO is in relationship to the top five highest paid executives. (Execucomp)
Establishment Size	The yearly average of the number of employees per establishment, measured in thousands. (OSHA)
Hours per Employee	Total number of annual hours worked in an establishment divided by the annual average number of employees. (OSHA)
Injury Frequency	Number of cases in an establishment divided by total workhours multiplied by 200 000. (OSHA)
Seasonal	A dummy variable taking the value 1 if the establishment employs seasonal workers.(OSHA)
Shutdown	A dummy variable taking the value 1 if the establishment experienced a shutdown/layoffs during the year. (OSHA)
Strike	A dummy variable taking the value 1 if there was as strike in the establishment during the year. (OSHA)
Disaster	A dummy variable taking the value 1 if there was a natural disaster affecting the establishment during the year. (OSHA)
Capex/Assets CAPX/AT).	The capital expenditure divided by total assets (Compustat
Leverage	The total short- and long-term debt divided by total assets Compustat (DLC+DTT)/AT).
Ln(Assets)	The natural logarithm of total assets (Compustat AT).
M/B	The market-to-book ratio (Compustat (CSHPRI * PRCC_F + LT + PSTKL - TXDB)/AT).
PPE/Assets	Net property, plant and equipment divided by total assets (Compustat PPENT/AT).
Turnover	Sales divided by total assets (Compustat SALE/AT).

Dependent variable.: Iniurv freauencv						
1	(1)	(2)	(3)	(4)	(5)	(6)
Entrenchment variables						
Title Concentration	-0.194*** (-4.112)	-0.184*** (-4.414)				
Salary Concentration			-1.745*** (-4.033)	-1.641*** (-5.104)		
Expertise			((,	-0.003** (-2.254)	-0.003*** (-3.009)
Control variables						
Hours Per Employee		-0.000*** (-6.385)		-0.000*** (-6.483)		-0.001*** (-5.722)
Establishment Size		0.007 (0.480)		0.003 (0.193)		0.005 (0.297)
Ln(Assets)		-0.062*** (-3.626)		-0.064*** (-3.817)		-0.098*** (-4.927)
Leverage		0.555*** (3.486)		0.472*** (3.408)		0.817*** (3.226)
PPE/Asset		-0.289* (-1.751)		-0.380**		-1.051*** (-4.748)
Turnover		-0.011 (-0.249)		-0.016 (-0.335)		-0.015 (-0.256)
Capex/Assets		1.471 (1.230)		1.570 (1.460)		3.748*** (2.980)
M/B		0.003 (0.120)		-0.004 (-0.136)		0.019 (0.561)
Strike		0.201***		0.207*** (2.620)		0.154 (1.116)
Shutdown		0.106*** (3.310)		0.099*** (3.058)		0.102** (2.018)
Seasonal		-0.004 (-0.054)		-0.002		0.164*** (2.879)
Disaster		0.091 (1.269)		0.068 (0.993)		0.053 (0.759)
Constant	1.365*** (3.826)	2.881*** (5.741)	1.914*** (5.043)	3.504*** (7.115)	1.337*** (3.229)	3.084*** (5.461)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Establishment Yrs Log-Likelihood	44427 -128303.313	44427 8 -127308.19	44427 6 -128126.68	44427 37 -127151.73	17557 35 -52405.99	17557 95-51729.965

Appendix B – Poisson regression

* p<0.1, ** p<0.05, *** p<0.01

J. Haga. F. Huhtamäki, D.	Sundvik / Finance,	Risk and Accounting	Perspectives Pro	oceedings (2019) 185-	202
0 ,	,	0	1		

Appendix	3 – Nega	ative
binomial	-	

Dependent variable.: Injury frequency						
	(1)	(2)	(3)	(4)	(5)	(6)
Entrenchment variables						
Title Concentration	-0.194*** (-4 112)	-0.184*** (-4 414)				
Salary Concentration	((-1.745*** (-4.033)	-1.641*** (-5.104)		
Expertise			(1.000)	(0.101)	-0.003** (-2.254)	-0.003*** (-3.009)
Control variables						
Hours Per Employee		-0.000*** (-6.385)		-0.000*** (-6.483)		-0.001*** (-5.722)
Establishment Size		0.007		0.003		0.005
Ln(Assets)		-0.062*** (-3.626)		-0.064*** (-3.817)		-0.098*** (-4.927)
Leverage		(3.486)		(3.408)		0.817***
PPE/Asset		-0.289* (-1.751)		-0.380** (-2.397)		-1.051*** (-4.748)
Turnover		-0.011		(-0.016)		-0.015
Capex/Assets		(0.213) 1.471 (1.230)		(1.460)		3.748***
M/B		(0.120)		-0.004		0.019
Strike		(2.735) (2.735)		0.207***		0.154
Shutdown		0.106***		0.099***		0.102**
Seasonal		-0.004		-0.002		0.164***
Disaster		(0.091) (1.269)		(0.023) 0.068 (0.993)		(2.077) 0.053 (0.759)
Constant	1.365*** (3.826)	2.881*** (5.741)	1.914*** (5.043)	3.504*** (7.115)	1.337*** (3.229)	3.084*** (5.461)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Establishment Yrs	44427	44427	44427	44427	17557	17557
Log-Likelillood	-128303.31	0 -12/308.19	0 -120120.08	or -1∠/151./3	55 -52405.9	75-51/29.905

* p<0.1, ** p<0.05, *** p<0.01



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings





On Artificial Intelligence's Razor's Edge: On the Future of Democracy and Society in the Artificial Age

Julia M. Puaschunder*

The New School, Columbia University, Princeton University

Abstract

The introduction of Artificial Intelligence (AI) in our contemporary society imposes historically unique challenges for humankind. The emerging autonomy of AI holds unique potentials of eternal life of robots, AI and algorithms alongside unprecedented economic superiority, data storage and computational advantages. However, the introduction of AI to society also raises ethical questions. What is the social impact of robots, algorithms, blockchain and AI entering the workforce and our daily lives on the economy and human society? Should AI become eternal or is there a virtue in switching off AI at a certain point? If so, we may have to define a 'virtue of killing' and a 'right to destroy' that may draw from legal but also philosophical sources to answer the question how to handle the abyss of killing with ethical grace and fair style. In light of robots already having gained citizenship and being attributed as quasi-human under Common Law jurisdiction, should AI and robots be granted full citizen rights – such as voting rights? Or should we simply reap the benefits of AI and consider to define a democracy with different classes having diversified access to public choice and voting – as practiced in the ancient Athenian city state, which became the cradle of Western civilization and democratic traditions spread around the globe. Or should we legally justify AI slaves to economically reap their benefits, as was common in ancient Rome, which became the Roman Law legal foundation for Continental and some of Scandinavian Law traditions and which inspired very many different codifications around the world. Finally, we may also draw from the Code Napoléon, the French Code Civil established under Napoleon in 1804, which defined male and female into two classes of human with substantial right and power differences, and - to this day - accounts for one of the few documents that have influenced the whole world in legal and societal ways. In asking critical questions and unraveling the ethical boundary conditions of our future artificial world, the paper thereby takes a descriptive – afar from normative – theoretical angle targeted at aiding a successful introduction of AI into our contemporary workforce, democracy and society.

Keywords: AI, Artificial Intelligence, Athenian city state, Code Civil, Code Napoléon, Democracy, Right to destroy, Roman Law, Slavery, Society, Workforce

Introduction

Artificial Intelligence (AI) poses historically unique challenges on humankind. The article addresses the introduction of Artificial Intelligence (AI) in our contemporary society. What is the impact of robots, algorithms, blockchain and AI entering the workforce and our daily lives on the economy and human society? As emerging globally trend, AI is extending its presence at almost all levels of human conduct having raised both expectations and concerns (Cellan-Jones, 2018;

* Corresponding author.

E-Mail address: Julia.Puaschunder@newschool.edu

Financial support of the Austrian Office of Science and Technology, Fritz Thyssen Foundation, Research Association for Interdisciplinary Studies, The New School (Dean's Office, Department of Economics, Fee Board, Social Science Fellowship, The New School for Social Research, The New School Eugene Lang College), Taylor & Francis, University of Vienna and Vernon Arts and Science is gratefully acknowledged. The author declares no conflict of interest. All omissions, errors and misunderstandings in this piece are solely the author's.

Sofge, 2015; United Nations, 2017). AI will hold the potential to replicate human existence but grant eternal being.

On innovation's razor's edge of 24/7 working robots that can live eternally and have no feelings, ethical questions arise whether robots, algorithms and AI should be granted citizenship and legally be considered as quasi-human beings — a technocratic and legal trend that has already started¹. As AI is currently reaching status of actual personhood – e.g., via citizenship and quasi-human rights applied in the Common Law but also Roman Law territories of the US and the EU – this legal personhood raises challenging legal and ethical questions².

The novel predicament between eternity and overpopulation hence calls for revising legal codes for killing, which would allow switching off AI at a certain point to curb overpopulation and harmful behavior. But how to argue legally the right to kill? And when to pull the plug? How do we switch quasi-human intelligence off when misbehaving or if AI life has become a burden that cannot be borne by society? How to balance robots living forever in light of overpopulation and finite resources? We may want to draw on the ethics of dying and virtues of killing as well as suicide literature to answer these novel questions. The proposed frame offers innovative insights for legal conducts but also overlapping generations relationships. The nature of algorithms and digital technology being global demands for an international response, potentially via international law supremacy principle.

Is feelingless AI vulnerable and prone to become enslaved or will the computational power and energetic capacities of robots outperform and enslave humankind? Should we have a democracy with a diversified populace of human enslaving robots? In light of robots already having gained citizenship and being attributed as quasi-human, should AI and robots be granted full citizen rights – such as voting rights? Given the humane fallibility and biases, would a rational AI agent make better democratic choices? Should AI therefore be used for governance as for being insusceptible for bribery and fraud, or does the installment of algorithms in leadership positions imbue dangers to humankind? How should we organize the human-led evolution of AI production and the blend of human-AI enhanced workforce? Or will algorithms in charge put humankind in danger? And what is it that makes human humane in the artificial age?

The paper proposes to investigate the economic, legal and societal impact of AI from an ethical perspective. The current legal status of robots being referred to as quasi-human will be discussed as for implications to society and democracy.

The power divide imperatives between human and AI robotics will become subject to scrutiny in light of historic examples of early forms of ancient Athenian democracy and Roman Law civilization that legally allowed for slavery but also with an eye on French Napoleonic code civil that established a supremacy of a man over his wife and children with attention to possession and property.

Humanness will be highlighted as key to future success in the age of AI and automated control. The proposed research will thereby draw from behavioral human decision making insights and evolutionary economics in order to outline what makes human humane and how human decision making is unique to set us apart from AI rationality. AI will be argued to bevalue humanness and improve the value of human-imbued unique features.

While the research is planned to be descriptive – afar from normative – and targeted to aid a successful introduction of AI into the workforce and society, the project will ask critical questions and unravel the ethical boundary conditions of our future artificial world. The findings promise to hold novel insights on future success factors for human resource management but also invaluable contributions for the successful introduction of AI and digital humanities in modern democracies and societies. Innovation's razor's edge is thereby aimed to be ennobled by ethical imperatives as old as humankind civilization.

¹ https://www.businessinsider.com/meet-the-first-robot-citizen-sophia-animatronic-humanoid-2017-10?r=UK&IR=T

² https://www.cnbc.com/2017/12/05/hanson-robotics-ceo-sophia-the-robot-an-advocate-for-womens-rights.html

The paper is structured as follows: First, the ontology of AI is outlined as well as an analysis of legal personhood. Then, the predicament between eternal life and overpopulation is presented. The virtues of dying and killing but also philosophical arguments for choosing suicide are discussed. The paper then embarks on the discussion of a future society with AI and draws on the socio-historic and legal examples of the Athenian city state democracy with different access to democratic rights, the ancient Roman Empire justification of a society featuring slavery but also the Code Napoléon of 1804, which granted male and female different rights establishing a substantial and justified power hierarchy between different classes of human. The article closes with an international law prospect on regulating AI.

Theory

Artificial Intelligence (AI)

Artificial Intelligence (AI) is "a broad set of methods, algorithms, and technologies that make software 'smart' in a way that may seem human-like to an outside observer" (Noyes, 2016). The "human-like" intelligence of machines derives from machines being created to think like humans but at the same time to also act rationally (Laton, 2016; Russell & Norvig 1995; Themistoklis, 2018). AI is perceived as innovative technology or as the sum of different technological advances as the privilege of the private, technological sector with little — if any — public regulation (Dowell, 2018).

As the most novel trend, AI, robots and algorithms are believed to soon disrupt the economy and employment patterns. With the advancement of technologies, employment patterns will shift to a polarization between AI's rationality and humanness. Robots and social machines have already replaced people in a variety of jobs - e.g. airports smart flight check-in kiosks or self-checkouts instead of traditional cashiers. Almost all traditional professional are prospected to be infused with or influenced by AI, algorithms and robotics. For instance, robots have already begun to serve in the medical and health care profession, law and – of course – IT, transportation, retail, logistics and finance, to name a few. Social robotics may also serve as quasi-servants that overwhelmingly affect our relationships. Already, social robots are beginning to take care of our elderly and children, and some studies are currently underway on the effects of such care (Alemi, Meghdari & Saffari, 2017). Not only will AI and robots offer luxuries of affordability and democratization of access to services, as they will be - on the long run - commercially more affordable and readily available to serve all humanity; but also does the longevity potential of machines outperform any human ever having lived (Hayes, 2018). However, the new technology also comes with the price of overpopulation problems and the potential for misuse and violent action. Just like many other technologies, robots could be misused for wars, terrorism, violence and oppression (Alemi et al., 2017; Puaschunder, 2018).

AI's entrance in society will revolutionize the interaction between humans and AI with amply legal, moral and social implications (Kowert, 2017; Larson, 2010). Autonomous AI entities are currently on the way to become as legal quasi-human beings, hence self-rule autonomous entities (Themistoklis, 2018). AI is in principle distinguished between weak AI, where "the computer is merely an instrument for investigating cognitive processes" and strong AI, where "[t]he processes in the computer are intellectual, self-learning processes" (Wisskirchen, Biacabe, Bormann, Muntz, Niehaus, Jiménez Soler & von Brauchitsch, 2017, 10). Weak AI is labeled as Artificial Narrow Intelligence (ANI) while strong AI is further classified into Artificial General Intelligence (AGI) and Artificial Super Intelligence (ASI).

The emergence of robotics technology is developing much quicker than previously thought. Robots are anticipated to soon be as ubiquitous as computers are today (Meghdari & Alemi, 2018). Society has long been concerned with the impact of robotics technology from nearly a century ago, when the word "Robot" was devised for the first time (Căpek, 1921; Meghdari & Alemi, 2018). The EU Committee on Legal Affairs (2016, p. 4) holds that "[U]ltimately there is a possibility that within the space of a few decades AI could surpass human intellectual capacity in a manner which, if not prepared for, could pose a challenge to humanity's capacity to control its own creation and, consequently, perhaps also to its capacity to be in charge of its own destiny and to ensure the survival of the species." AI mimicking human intellect could soon surpass humans intellectually but also holistically breaking the barrier of human controlled-automization (Schuller, 2017; Themistoklis, 2018). Modern literature about robots features cautionary accounts about insufficient programming, evolving behavior, errors, and other issues that make robots unpredictable and potentially risky or dangerous (Asimov, 1942/1950, 1978, 1985; Meghdari & Alemi, 2018). "Observe, orient, decide, act" will therefore become essential in the eye of machine learning autonomy and AI forming a new domain of intellectual entities (Armstrong & Sotala 2012, p. 52; Copeland 2000; Galeon & Reedy, 2017; Marra & McNeil, 2013). The uncertainty surrounding AI development and self-learning capabilities give rise to the need for guarding AI and an extension of the current legal system to cope with AI (Themistoklis, 2018; Puaschunder, 2018).

With the advancement of technology, social robots have found broader applications in the private and public sectors, such as educational and cultural affairs, games and entertainment, clinical and rehabilitation, nursing of children and/or elderly, search and rescue operations (Meghdari, Alemi, Shariati & Zakipour, 2018). For example, social robots such as ASIMO, Nao, iCub, ARASH, and RASA have been developed for "Edutainment" or "educationentertainment" purposes. They aid the study of cognition (both human and artificial), motion, and other areas related to the advancement of robotics serving our society (Meghdari & Alemi, 2018). In addition, a few medical and healthcare toy-like robots, such as PARO, which looks like a baby seal, or ARASH, which is a humanoid, have been designed for therapeutic purposes such as reducing distress, stimulating cognitive activity, teaching specific subjects, and improving socialization (Meghdari, Shariati, Alemi & Vossoughi, 2018). Similarly, Sharif University of Technology's socially assistive robot RASA has been developed to help coach and teach Persian Sign-Language to Iranian deaf children (Meghdari, Alemi, Zakipour & Kashanian, 2018). Personal care and companion robots are increasingly being used to care for the elderly and children, such as RI-MAN, PaPeRo, and CareBot (Meghdari & Alemi, 2018; Puaschunder, 2018).

In recent years, robotics technology has extended its applications from factories to more general-purpose practices in society – for instance, such as the use of robots in clinical and rehabilitation, nursing and elderly care, search and rescue operations (Meghdari & Alemi, 2018). Social robots have become clinical and educational assistants for social interventions, treatment, and education such as language trainings but also assistance with children with disabilities like autism, down syndrome, cancer distress, hearing impairment, etc. (Meghdari et al., 2018). Initial investigations clearly indicate that social robots can play a positive role in the improvement of children's social performance, reduction of distress during treatments, and enhancing their learning abilities (Meghdari & Alemi, 2018). Surprisingly, although not too hard to imagine, relationships of a more intimate nature have not quite been satisfied by robots yet (Meghdari et al., 2018; Veruggio, 2005).

AI's ethical boundaries

In today's economy, robots and algorithms now taking over human decision-making tasks and entering the workforce but also encroaching our private lives, currently challenges legal systems around the globe (Themistoklis, 2018). The attribution of human legal codes to AI is one of the

most groundbreaking contemporary legal and judicial innovations. Until now, legal personhood has only been attached directly or indirectly to human entities (Dowell, 2018). The detachment of legal personhood from human being now remains somewhat of a paradox causing an extent of "fuzziness" of the concept of personhood (Barrat, 2013; Solum, 1992, p. 1285). As AI gets bestowed with quasi-human rights, defining factors of human personhood will need to be adjusted (Dowell, 2018). Human concepts, such as morality, ownership, profitability and viability will have different meaning for AI. The need for redefining AIE has therefore reached unprecedented momentum.

As a predicted trend, the co-existence of AI with the human species is believed to change the fundamental concepts of social, political and legal systems. AI has already produces legal creations and will do so even more in the near future, through its developing autonomy. In addition, the technology leading to AGI and ASI is already present, posing moral and legal dilemmas about who should control it and under what terms (Themistoklis, 2018). The emergence of AGI and ASI will necessitate the attribution of some extent and of some type of legal personhood, bearing rights and obligations. AI will not be most probably an exact replication of human intellect behavior (Themistoklis, 2018). "[U]ltimately, robots' autonomy raises the question of their nature in the light of the existing legal categories – of whether they should be regarded as natural persons, legal persons, animals or objects - or whether a new category should be created, with its own specific features and implications as regards the attribution of rights and duties" (Committee on Legal Affairs 2016, p. 5). Behavioral economists add the question whether AI and robots should be created to resemble human beings' decision making with fast thinking and fallible choices or rather be targeted at perfect rationality and slow thinking (Kahneman & Tversky, 1979). General conscious is strived for so that AI possesses consciousness, which it can evolve and enhance on the basis of its own critical reflection and assessment of external factors (Themistoklis, 2018). A lower level of autonomy exists if an entity can demonstrate such consciousness at a narrow field or can self-evolve and self-adapt to external influences, thus reaching decisions "of its own," without being conscious of its intelligence as such (Themistoklis, 2018). As AI emerges as new types of intellect capacities coupled with human-like emotional features, they are attributed a legal personhood in order to ensure to be comprehended correctly and to avoid unfair treatment, towards humans as well (Themistoklis, 2018). Artificial entities are currently gaining human or quasi-human status in the Western and Arab worlds in forming an intellectual autonomy of the entity (MacDonald, 2016). For instance, in Saudi Arabia the first female robot got a citizenship in 2017 and the robot appears to have more rights than a human female in Saudi Arabia. With the rise of AI persons, their eternal life poses ethical challenges in light of overpopulation and evolutionary perfection, which could crowd out human fallibility if determining merit-based eternal life.

With citizenship and quasi-humanness being attributed to AI, the power relation between human and AI will need to be defined. Should AI be granted full citizenship rights, the problem of overpopulation occurs, since there is the possibility of infinite life of AI. In a human-led evolution, AI will have to be switched off for various reasons, such as malfunction but also merit-based efficiency calculus. If now AI is considered as quasi-humane and granted citizenship rights, switching off AI becomes a legally problematic.

While there is currently cutting-edge writing about the potential emergence of an AI personhood as well as concern over the merge of AI with cyberspace that might lead to the breach of the relationship between legal personhood and nation state sovereignty and a nomenclature is emerging on legal characterizations of different levels of AI development; hardly any information exists about the eternal living of AI (Beerbaum & Puaschunder, 2018; Hildebrandt, 2013). From the theoretical standpoint, the eternal longevity of AI contradicts the fundamental concept of fairness in death, as a general condition for all. From the practical standpoint, the international community is currently urged to think on the basis of global commons in terms of AI and AI

eternal life potentials contributing to overpopulation. Thereby global commons theories may be tabbed on, which primarily offer guidance for a regulatory framework, which establishes control "...for the benefit of all nations" and refer to space constraints (Clancy, 1998; Puaschunder, 2018; Tsagourias, 2015).

Regarding limited space, longevity and eternal life appears problematic. Humankind may face tough decisions whether or not to have AI proceed and what kind of developments to flourish and what to extinct. In what cases should we consider to switch off AI? In 1950, Isaac Asimov introduced the idea robot to (1) not injure a human being or, through inaction, allow a human being to come to harm. (2) A robot obeying the orders given it by human beings except where such orders conflict with the first law. (3) A robot must protect its own existence as long as such protection does not conflict with the first or second law. So in the cases of overpopulation and harm emerging from AI, algorithms and robots can be considered to be switched off. But when to stop AI?

An economic killing market mechanism may be natural market selection via price mechanisms and the falling rate of profit. Regarding prices, natural supply and demand mechanisms will always favor innovation with a higher price and following supply of goods lead to a price drop. The falling rate of profit is one of the major underlying features of business cycles, long-term booms and downturns (Brenner, 2002, 2006a, b). Capitalism is thereby described as competitive battle for innovation and reaping benefit from first-market introductions. Once followers enter the market, profit declines, leading eventually to market actors seeking novel ways to innovate in order to regain a competitive market advantage and higher rates of profit. Thereby industries and innovations fade and die off. Such a natural market evolution is also likely to occur with AI innovations, which will determine which AI traits will remain and which ones will fade off (Puaschunder, 2018). Apart from soft market mechanisms that may lead to AI evolution, what are the cases when AI should be shut down or switched off or – in the case if AI personhood – be killed?

Killing AI

Errors and Safety: The main and leading concern about any new and emerging technology is to be safe and error free (Meghdari & Alemi, 2018). Therefore, sufficient and numerus tests on health and safety must be performed by developers and/or well-known independent sources before rolling out any technology onto the marketplace and society (Meghdari & Alemi, 2018). In robotics, the safety issue mainly centers around software and/or hardware designs (Meghdari & Alemi, 2018). Even a tiny software flaw or a manufacturing defect in an intelligent machine, like a smart car or a social robot, could lead to fatal results (Meghdari & Alemi, 2018). When these deviations occur and especially when they are harmful to the human community but also to other AI species, the faulty AI should be terminated. With regard to the risk of robotic malfunctions and errors, product legal responsibility laws are mostly untested in robotics (Meghdari & Alemi, 2018). A usual way to minimize the risk of damage from social robots is to program them to obey predefined regulations or follow a code-of-ethics (Meghdari & Alemi, 2018). Ethical codes for robotics are currently needed and should become formed as a natural behavioral law to then be defined and codified as law. Laws but also an ethical understanding to terminate AI, algorithms and robots in case of impairment and harm are needed.

Morals, Ethics, and the Law: As social robots become more intelligent and autonomous and exhibit enough of the features that typically define an individual person, it may be conceivable to assign them responsibility and use them in social, educational, and therapeutic settings (Meghdari & Alemi, 2018). In the currently ongoing research on the integration of computers and robotics with biological corpse it is found that a cognizant human brain (and its physical body) apparently

has human-rights; hence, replacing parts of the brain with artificial ones, while not harming its function, preserves those rights (Meghdari & Alemi, 2018; Warwick & Shah, 2014). Also, consider a handicapped person featuring an electronic robot arm that commits a crime. It becomes obvious that half-robot-human beings should be considered as human and robots as quasi-human beings. Meghdari & Alemi (2018) speculate that at some point in the future, we may face a situation in which more than half of the brain or body is artificial, making the organism more robotic than human, which consolidates the need of special robot-rights and attributing (quasi)-human rights onto robots. When considering robots as quasi-human beings, their termination appears legally questionable and ethically challenging, requiring revisiting laws as legitimation to kill a likewise species as well as ethical consensus on the virtue of killing (Puaschunder, 2018).

The legal argumentation may draw on justifiable homicide as outlined in criminal law cases such as prevention of greater harm to innocents during an imminent threat to life or well-being in self-defense. According to the United Nations Universal Declaration of Human Rights, Article 3 states that everyone has the right to life, liberty and security of person and most nations' policy allows for some degree of leniency for self-defense, which reduces charges. Potentially excusing conditions common to most jurisdictions include wartime, when the person's death is inflicted by the effect of a lawful arrest or prevention of lawfully detained person's escape, quelling riot or insurrection, when the use of force is "no more than absolutely necessary." Some countries deem it lawful for a citizen to resort to violence to protect valuable property and there is the "heat of the moment" defense argument, in which the defendant deemed to have lost control through provocation. Doctrine of necessity allows, for example, a surgeon to separate conjoined twins and killing the weaker twin to allow the stronger twin to survive. While fetuses are considered as unborn children in the US, the right to an abortion was upheld in the US legal system as exemption from prosecution (Roe v. Wade, 1973). Several countries, such as the Netherlands, Belgium, Switzerland, Japan, and the U.S. states of Oregon and Washington, allow both active and passive euthanasia by law, if justified. Where the person concerned is to be arrested for an offense referred to in Schedule 1 or is to be arrested on the ground of having committed such an offense, and the person authorized under this Act to arrest or to assist in arresting him cannot arrest him or prevent him from fleeing by other means than killing him, the killing shall be deemed to be justifiable homicide. If any arrestor attempts to arrest a suspect and the suspect resists the attempt, or flees, or resists the attempt and flees, when it is clear that an attempt to arrest him or her is being made, and the suspect cannot be arrested without the use of force, the arrestor may, in order to effect the arrest, use such force as may be reasonably necessary and proportional in the circumstances to overcome resistance or to prevent the suspect from fleeing: Provided that the arrestor is justified in terms of this section in using deadly force that is intended or is likely to cause death or grievous bodily harm to a suspect, only if he or she believes on reasonable grounds (§7 Judicial Matters Second Amendment Act 122 of 1998).

In light of overpopulation and harmful behavior of AI, switching off artificial life, which is currently be granted quasi-human status, will need to be argued legally and supported ethically. Killing in terms of the death penalty is justified legally in the 5th (and the 14th) amendment that states "no person shall be deprived of life, liberty, or property without due process of law," while the eighth amendment prohibits "cruel and unusual punishment." Killing in terms of harmful behavior of AI can be grounded on similar legal reasons to ensure that no AI harms the collective. Overpopulation claims leading to the need to take AI partially off the grid more lead to philosophical sources that argue for individual's free will to choose to live or die (Critchley, 2015).

Apart from self-defense, suicide may also serve as legally justified argument for switching off AI, if artificial life is programmed to terminate itself when harmful in such way that AI causes injury to a human being or, through inaction, allow a human being to come to harm. A robot not obeying the orders given it by human beings except where such orders conflict with the first law.

We could argue that AI should stay alive at whatever the cost in virtue of killing AI, when turning harmful.

Suicide has been tabooed for most part of history and propagated to be a religious sin or classified as a psychological disorder (Critchley, 2015). Yet the human gift of reflection and search for meaning in life or death could leverage into an asset in the AI evolution in the decades to come. Suicide understood as neither a legal nor a moral offence but as right to death bestowed upon human beings in their self-conscious reflection may be extended as a virtue of killing in the artificial age, when human beings will have to decide what AI should stay alive and what AI be taken off the grid. Human will thereby become the rulers of the forthcoming AI evolution.

The ethical imperative of switching AI off may be found in David Humes' saving 'No man ever threw away life, while it was worth keeping' (Critchley, 2015, p. 15f.). Hume's point is that when life has become a burden that cannot be borne, one is justified in taking it. In this argumentation line, if AI life has become a burden that cannot be borne by society, society is justified in taking AI's life. Critchley (2015) recommends reflective compassion based on empathy and introspection, but we may also need foresight and inclusion of future externalities. In the artificial age, AI may therefore be programmed with a constitution for suicide. Also in Seneca we find that when a human life no longer flourishes, one being permitted to end it (Critchley, 2015). The Stoic tradition argued that suicide is a legitimate act and an honorable gesture of farewell from a state of unbearable pain, whether physical or psychical (Critchley, 2015). In this sense, AI's death may be argued to be justified when AI imposes a state of unbearable pain unto others. Religious stances that suicide is wrong because only God having moral authority over human lives and thus us being property of God could be subsumed into a condition to legitimize human having authority over AI and thus being our property, in which we can decide what developments to maintain and which ones to switch off in a human-led AI-evolution. This human-led evolution is believed to revolutionize modern society and civilization. Killing AI – or determining what AI development should survive - may therefore become an act of self-defense or legalized suicide.

As in a suicide note that speaks as final communication to the descendants, algorithms that are forced to be switched off should also store information on the reason that terminated them and be conserved in a blockchain that serves to educate the network about malfunction and malpractice. This piece of publicity should serve as disciplinary and signal function. In suicides with guns, people aim at the head not the heart – while both head and heart stop function in human thereafter, in AI, we may program that the brain function, that is constant storage of information and adaptive reprogramming and actions get switched off but some positive parts remain intact to be reprogrammed (Critchley, 2015). Death will end the incoherence in creating a beautifully benevolent AI structure, which we may see as evolutionary cleansing of destruction coming out of AI. The death algorithm button will bestow coherence to the human-led evolution of AI. The voluntary switch to shut AI off will be the pejorative of human and dominating privilege of human over AI. There will be a beauty to death, the stillness, the rest and the finally stopped negative character of AI evolution (Puaschunder, 2018).

The virtue of killing could also be grounded on Viktor Mayer-Schönbergers "right to be forgotten," which ensures data privacy through automated deletion of contents after a certain period and grants individuals rights to have their data been destroyed (Puaschunder, 2018; forthcoming). In this line, we may argue a "right to destroy" and program AI to stop itself should it incur hurt, damages and losses to humankind. However, the implementation of this right is still in infancy and hindered by questions of what court is responsible for an as such claim. As a legal subsumption, we may speculate that individuals may be granted a 'right to terminate' and can order for robots to be switched off if causing harm to them. As the 'right to be forgotten' law can be overruled by concern for public safety, this may also apply to the right to terminate. Thereby it deserves mentioning that safety differs around the world and also expected safety standards

(Puaschunder, 2018). All these developments are prospected to lead to an AI-evolution, in which human are meant to select the process what AI should survive or be killed. Key decision maker thereby divert favorable traits and developments from unfavorable. But who should determine what should survive, human or AI? A question that can be answered by sorting out the legal power relation between AI and human.

AI human democracy and society

With AI entering human society and being considered as quasi-human and granted citizenship, the ethical question arises, what kind of citizen AI are? Should AI be considered as full citizens, we run into the problem of overpopulation as discussed. In addition, AI dominance of physical and computational power creates risks of AI outperforming and eventually dominating human. With this scenario in mind, a legal power hierarchy should be established that grants a predominance of human over AI that allows human to benefit from AI but also ensures that dignity in the treatment of AI is upheld for the sake of breeding a generally favorable and amicable climate in society.

Human predominance over AI

When considering the enormous physical and longevity advantages AI hold over human, a natural dominance of AI over humankind is implied. In order to ensure that human lead AI and are not subordinated, a society should be established, in which robots gain quasi-human rights but may not have the same powers and rights as human beings. In the earliest form of democracy in the ancient Athenian city state, different classes of citizenship existed.

The first known democracy developed around the fifth century BC in the Greek city-state of Athens, which featured the first government by its people, in which the supreme power was vested in the people and exercised directly by them or by their elected agents under a free electoral system. The ancient Athenian democracy became an important source for 18th-century revolutionaries' intellectual background during the American, French and other continental European revolutions. State constitutions around the globe and over time, political speeches and writings about nation states and society reflect the core principles conveyed in the ancient Athenian city state democratic model, which become a model for shaping civilization throughout the world. To this day, a democracy accounts for the most advanced political order in an egalitarian society (Vlassopoulos, 2009).

In 507 BC, the Athenian leader Cleisthenes introduced a system of political reforms called demokratia or rule by the people to ensure security, stability and prosperity to the entire community. Key features of democracy are equality, accountability, citizen participation, rule of law, political tolerance, transparency, economic freedom and a multi-party political system. In the ancient Athenian democracy model, not every citizen had the right to vote, run for office and participate in political discussions. Yet to all, the democracy was meant to protect and uphold dignity of all people. Therefore, the Athenian democracy bestowed a favorable climate in society without political equality of all citizens.

As a direct democracy, citizens voted directly on legislation and executive bills. However, participation in democracy was not open to all residents but limited to adult, male citizens excluding women, foreign residents and slaves. In ancient Athens, only male Athenian citizens who had completed their military training had the right to vote and only about 10 to 20% of inhabitants actually participated in governmental decision making. Women had limited rights and privileges, restricted movement in public and were legally segregated from men. Also
excluded from voting were citizens whose rights were under suspension – foremost for failure to pay debt to the city. Only descendants from two Athenian parents could claim Athenian citizenship. Citizenship could also be granted by the assembly and sometimes given to large population groups as a reward for service to the state. As slavery was widespread in Athens and seemed to have developed the city state, the ancient Athenian democracy is attributed to economically be based on slavery, which allowed the general public to devote privileged amounts of time to political life.

The Athenian form of direct democracy does not only serve as an example of not all citizens being allowed to vote being a feasible governmental structure but also – as for its direct character – as a forerunner of electronic democracy. A future world with AI blended into society could structure the human – AI relation based on the ancient Athenian city state societal composition, in which different classes of citizenship lived together in harmony. As in the ancient Athenian democracy model, not every citizen should have the right to vote, run for office and participate in political discussions. AI could become citizen, yet not be allowed to vote, run for office and participate in political discussions. Yet to all, AI and human, democracy and citizenship is meant to protect and uphold dignity of all people and AI.

In order to create a more inclusive democracy than the ancient Athenian direct electronic democracy may be introduced, in which voters vote on a political agenda featuring different spectra of choices (e.g., libertarian versus state-controlled, pro-against immigration...) and the mean of their choices then gets processed by algorithmic choice of programs to be enacted by politicians. Algorithms could thereby compute the standard choice of politicians representing different agenda based on historical information and aid to inform politicians about the outcomes of several choices in the past. AI and algorithms hold the computational power and data calculus capacity to do so. This would ensure closer accuracy of political will resembling collective choice and enable to reap AI benefits for political choice, while ensuring human to stay in charge but enhanced by artificial benefits. This integration of AI in form of an advisory role to governments could enable AI access to democracy as a compromise without AI having direct voting rights.

Human reaping benefits of AI

AI entering the workforce and holding enormous physical and longevity advantages over human, implies the economic gains to be reapable. Standard economic growth models hold that capital and labor are essential for an economy to flourish. While capital is usually considered as fungible and exchangeable and eternal; labor is more individual, human and inflexible. AI entering the workforce and blending in as a substitute to human capital, will change the nature of labor, potentially dividing labor into a putty, flexible, eternal and exchangeable AI part and a clay labor of inflexible human capital (Puaschunder, work in progress). In order to ensure that human can legally benefit from the economic output and growth generated by AI, a society should be established, in which robots gain quasi-human rights but may not have the same material needs and rights as human beings. In the earliest form of society in the ancient Roman Empire, a society existed that featured a high culture and human protection but legal slavery.

Slavery in ancient Rome played an important role in society and the economy. Slaves provided manual labor and agriculture, working on farms, mines and mills, household domestic services, urban crafts and services but also skilled, educated professions, such as accountants and physicians as well as imperial and public services. Slaves were considered property under Roman Law and had no legal personhood. Unlike Roman citizens, they could be subject to corporal punishment, sexual exploitation, torture and summary execution. Over time in history, slaves gained increased legal protection, including the right to file complaints against their masters.

Legal foundations of slavery can already be found as early as in the Twelve Tables, Rome's oldest legal code. Around the 2nd century, Ulpian defined slavery as an aspect of the ius gentium, a customary international law held in common among all people. Slavery was held to be a practice common to all nations, who might then have specific civil laws pertaining to slaves. In ancient warfare, the victor had the right under the ius gentium to enslave a defeated population. The ius gentium was not a legal code but reasoned compliance with standards of international conduct (Bederman, 2004). During the Roman imperial expansion, substantial growth of slavery transformed the economy (Hopkins, 1983). Delos in the eastern Mediterranean was made a free port in 166 BC and become one of the main market venues for slaves. Augustus imposed taxes on slave trade, which was increased over time with the rising number of slaves, which comprised of up to 35-40% of Italy's population (Harris, 2000). The price of slaves was determined by origin, health, character, intelligence and education. The living conditions of urban slaves was inferior to free persons living with them but sometimes superior to that of many free urban poor in Rome. Household slaves likely enjoyed the highest standard of living among Roman slaves, next to publicly owned slaves, who were not subject to the whims of a single master (Johnston, 1957). Imperial slaves were those attached to the emperor's household (The Oxford Encyclopedia of Ancient Greece and Rome, 2010). Sophisticated slaves that were used for economic trade were permitted to earn money for their personal use (Berger, 1991). Slaves could also be freed by a magistrate's declaration with the prior authorization of the Senate or the Emperor.

Slaves were seen as excluded from the persona, the synonym for the true nature of the individual, and considered to not have a personality. A slave would not own his or her body, had no ancestors, no name, no cognomen and no goods of his or her own (Mauss, 1979). The testimony of a slave could not be accepted in a court of law unless the slave was tortured. Rome differed from Greek city-states slaves in allowing freed slaves to become citizens. After manumission, a male slave who had belonged to a Roman citizen and now acquired libertas as a libertus/a (freed person) enjoyed active political freedom including the right to vote. The former master become a patron. Freed slaves become the class of libertini, who were not entitled to hold public office or state priesthoods, nor could they achieve senatorial rank. Any future children of a freedman would be born free, with full citizenship rights. Roman slaves could hold and use property, which belonged to their masters, as if it were their own (Gamauf, 2009). Skilled or educated slaves were allowed to earn their own money in hope to buy their own freedom (Kehoe, 2011)³. Slaves that were abandoned by masters were also free. Legal protection of slaves grew over time and history due to egalitarian views of humanity of the Stoics. A master who killed a slave without just cause could be tried for homicide and it became common for slaves to complain against cruel and unfair treatment of owners (Dillon & Garland, 2005). Rebellions and runaways of escaped slaves were published and those returning fugitives rewarded. Fugitives were branded on the forehead or had to wear a metal collar around the neck with the contact of the master. Slave rebellions surrected several times, most noticeable are the First, Second and Third Servile War. On the countryside, the Roman Empire also featured practiced serfdom. The Stoics and early Christians opposed ill-treatment of slaves⁴.

In order to ensure that human can legally benefit from the economic output and growth generated by AI, robots should be granted quasi-human rights but may not have the same material needs and rights as human beings. Slavery thereby allows to reap the benefits AI. AI's newly assigned roles appear to overlap with slave tasks of ancient Rome slaves that provided manual labor and agriculture, working on farms, mines and mills, household domestic services, urban

³ http://www.bbc.co.uk/history/ancient/romans/slavery_01.shtml

⁴ <u>https://www.biblegateway.com/passage/?search=Ephesians+6%3A5%E2%80%939&version=KJV</u>

https://www.biblegateway.com/passage/?search=Colossians+4%3A1&version=KJV

https://www.biblegateway.com/passage/?search=1Corinthians+7%3A21&version=KJV

https://biblehub.com/1_peter/2-18.htm

crafts and services as well as skilled, educated professions, such as accountants and physicians as well as imperial and public services. Like in ancient Rome, AI should be considered as property with no legal personhood. However, unlike ancient Roman slaves, they should not be subject to corporal punishment, sexual exploitation, torture and summary execution. Over time in history, AI – as the ancient Roman Law example of slaves – may gain more sophisticated legal protection, including the right to file complaints against misuse.

As for the international character of AI and algorithms, their fungability and fluid capital character; broad legal foundations of AI and the overarching regulatory framework how to classify reaping benefits from AI should be codified in customary international law held in common among all people. This would resemble the ancient tradition of Roman slavery being codified under ius gentium – an ancient predecessor of international law – and allow AI to remain fully fungible and practiced common in all nations, who might then have specific civil laws pertaining nuances of AI conduct in society.

As practiced during slavery in the Roman Empire and proposed by Bill Gates, reaping benefits from AI should be taxed based on the revenue generated by AI and/or the price of AI determined by sophistication. Defining AI as slaves would ensure to uphold decent standards of living for these creatures, while human naturally stay in charge of the evolution and introduction of AI into human society. As debated in the ancient Roman society, sophisticated AI that is used for economic trade may also be permitted to earn money for their personal use; but should never be freed and gain the same rights as human as there is something unique and special to humanness, which will be discussed in more detail in the discussion section of this paper. The uniqueness of human naturally leads to the natural exclusion of AI from the persona, the synonym for the true nature of the individual, and considered to not have a personality. As a Roman Law slave, AI should not own his or her body, have no awareness of its ancestors, and no goods or material cravings of his or her own. The testimony of AI should not be accepted in a court of law unless AI reports misuse that can be harmful to humankind. Differing from Roman Law slavery, AI should never be freed and human should always stay masters of their own creation. AI should not be entitled to hold public office or religious leadership and remain without rights to hold and use property on their own. AI and robots should not be allowed to earn their own money and even if being abandoned by masters, they should never be considered as free.

In order to protect humankind against rebellions of robots and AI, fugitives or deviant developens should be published, stopped according to the right to destroy and those aiding to inform about deviant developments rewarded. While fugitives in the ancient Roman Empire were branded on the forehead or had to wear a metal collar around the neck with the contact of the master; stopped AI or robots should be integrated into a blockchain as a trace of information on unwanted AI and robot behavior but also as a disciplinary function against other AI uprising and rebellious tendencies.

As in the case of the Greek and the Roman Law slaves, legal protection of AI may grow over time and history due to egalitarian views of humanity. For instance, destruction of AI without just cause could be tried for homicide and complaints of robots against cruel and unfair treatment of owners be supported in front of courts. In order to oppose ill-treatment of slaves immediately, dignity may be upheld in applying a legal code with two different classes of society. As such, the Code Napoléon as the first civil code may serve as guiding example.

Dignity upheld in the treatment of AI

If AI gets legally and economically subordinated to human, ethical questions arise. According to Kant's categorical imperative, which states one should only engage in actions, one wants to be done to oneself, AI should be protected against harm and misuse or abuse. The concern here is

less so the emotional and psychological state of AI, which arguably may not exists given missing self-cognition and emotions in AI, but more to set a signal and not to allow triggering sadist and negative compulsion in human that could be taken out on other human as well, if human become conditioned and learn from mistreating AI on a daily basis.

A legal code that may serve as reference hereby may be the Code Napoléon of 1804, a Civil Law code that defines and classifies male and female as human beings but legally bestows upon them substantial power differences, especially regarding material possession and democratic participation.

The Napoleonic Code (Code Napoléon, Code civil de Français) is the French civil code established under Napoléon I in 1804. As the first clearly written and accessible compilation of modern law, the Code Napoléon has become one of the most influential legal documents in history that influenced the law of many countries around the world (Mohamed, 2016). The Napoleonic Code became the most influential legal civil law code around the world that was adopted as the basis of private law systems of Arab world, Austria, ss Belgium, Canada, Chile, the Commonwealth, Egypt, Germany, Italy, Ireland, Latin America, the Netherlands, Portugal, Poland, Puerto Rico, Romania, Russia, Scandinavian countries, Scotland, Spain, Switzerland, United Kingdom, United States Louisiana to name a few.

With regard to family, the Code established the supremacy of the man over the wife and children, which was the general legal situation in Europe at the time (Smith, 2006). A woman was given even fewer rights than a minor.

In the attempt to protect AI against suffering, harm and misuse or abuse, the Code Napoléon may be applied and define AI and human as quasi-human and grant citizenship to both forms but different power regarding material possession, democratic participation and public leadership. A natural supremacy of human over AI and robots could be established. As the role of woman and minor even differed, a power hierarchy could even be codified between sophisticated and less-sophisticated AI and robots.

The value of humanness in the artificial age

Globalization led to an intricate set of interactive relationships between individuals, organizations and states. Unprecedented global interaction possibilities have made communication more complex than ever before in history as the whole has different properties than the sum of its increasing diversified parts. Electronic outsourcing in the age of artificial intelligence is likely to increase and with this trend a possible nudgital divide in the 21st century. In the light of growing tendencies of globalization, the demand for an in-depth understanding of how information will be shared around the globe and artificial intelligence hubs may evolve in economically more developed parts of the world has gained unprecedented momentum. Another predictable trend in the wake of the artificial intelligence revolution will feature time. Artificial intelligence with eternal life and 24/7 productivity capacities will change tact.

Finally, we may address the question what is it that makes human humane? In the age of artificial intelligence and automated control, humanness is key to future success. Future research may draw from behavioral human decision making insights and evolutionary economics in order to outline what makes human humane and how human decision making is unique to set us apart from artificial intelligence rationality.

Drawing from proposed research will thereby draw from behavioral human decision making insights and evolutionary economics in order to outline what makes human humane and how human decision making is unique to set us apart from AI rationality; AI is argued to bevalue humanness and improve the value of human-imbued unique features.

Humanness as found in heuristics, decision making errors but also procreation and creativity is believed to become more valuable in a future of AI entering the workforce and our daily lives.

The findings promise to hold novel insights for future success factors for human resource management but also invaluable contributions for artificial intelligence ethics. Having parts of the world being AI-driven and others being human capital grounded is prospected to increase the international development divide in the years to come. While in the AI-hubs human will be incentivized become more creative and humane while AI performs all rational tasks to a maximum productivity, other parts of the world will naturally fall back as for being stuck in spending human capital time on machine-outsourceable tasks and not honing humane skills, which are not replicable by machines.

Future research endeavors may address inequality drawing on the future vision that central rational AI-hubs will outperform underdeveloped remote areas of the world even more in the digital age.

Discussion

Overall the ongoing research project plays an important role in the evaluation of AI's entrance into the workforce and our daily lives. Depicting nudging during this unprecedented time of economic change and regulatory reform holds invaluable historic opportunities for capturing AI's influence on the stability of economic markets and societal systems. Global governance policy makers can snapshot AI's potential in the digitial age and bestow market actors with futureoriented foresighted. The results are aimed at guiding a successful AI and robot implementation to lower systemic economic market downfalls with attention to the changes implied in the wake of the ongoing artificial intelligence revolution. Market and societal policy recommendations for global governance experts on how to strengthen society by AI but also overcome unknown emergent risks within globalized markets and bestow market actors with key qualifications in a digitalized world are endeavored in future research.

References

- Acevedo, S., Mrkaic, M, Novta, N., Poplawski-Ribeiro, M., Pugacheva, E. & Topalova, P. (2018). How can low income countries cope? The effects of weather shocks on economic activity. Presentation delivered at The New School Economics of Climate Change Project, New York: The New School, April 12.
- Alemi, M., Meghdari, A., Saffari, E. 2017. "RoMa: A Hi-tech Robotic Mannequin for the Fashion Industry." Lecture Notes in Computer Science (LNCS): Social Robotics 10652: 209-219, Springer, Nov. 2017.
- Altenburg, T. & Assmann, C. (2018). United Nations Environment Programme Report on Green Industrial Policy: Concept, Policies, Country Experiences (2018). Geneva, Bonn: United Nations: United Nations Environment Programme & German Development Institute / Deutsches Institut für Entwicklungspolitik.
- Altenburg, T. & Rodrik, D. (2018). Green industrial policy: Accelerating structural change towards wealthy green economies. In T. Altenburg & C. Assmann (Eds.), United Nations Environment Programme Report on Green Industrial Policy: Concept, Policies, Country Experiences, pp. 2-20. Geneva, Bonn: United Nations: United Nations Environment Programme & German Development Institute / Deutsches Institut für Entwicklungspolitik. Andreoni, J., Erard, B. & Feinstein, J.S. (1998). Tax compliance. Journal of Economic Literature, 36, 2, 818-860.
- Armstrong, Stuart & Kaj Sotala. 2012. "How We're Predicting AI--or Failing to", in Beyond AI: Artificial Dreams, edited by Jan Romportl, 52. Pilsen: University of West Bohemia.
- Asimov, I. 1942. Runaround, In I, Robot, 33 51. London: Grafton Books.
- Asimov, I. 1950. I, Robot (2004 ed.). New York : Bantam Dell.
- Asimov, I. 1978. My own view. In The Encyclopedia of Science Fiction, ed. Robert Holdstock, N.Y.: St. Martin's Press.
- Asimov, I. 1985. Robots and Empire. New York: Doubleday.
- Asmundson, I. (2017). Financial services: Getting the goods. IMF entry retrieved from http://www.imf.org/external/pubs/ft/fandd/basics/finserv.htm

- Bai, S., Koong, K. & Valenzuela, E. (2018). The dynamic impact of oil price shocks on stock returns: Evidence from U.S. energy companies in the value chain. 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.
- Barrat, James. 2013. Our Final Invention: Artificial Intelligence and the End of the Human Era. New York: St. Martin's Press.

Barro, R. (1990). Government spending in a simple model of endogenous growth. Journal of Political Economy, 98, 103-125.

Bederman, D.J. (2004). International law in antiquity. Cambridge, UK: Cambridge University Press.

Beerbaum, D. & Puaschunder, J.M. (2018). A behavioral economics approach to digitalisation: The case of a principles-based taxonomy. Proceedings of the 10th International RAIS Conference on Social Sciences and Humanities; Published by Atlantis Press; Part of Series: ASSEHR; ISSN: 2352-5398, Volume 211.

Berger, A. (1991). Encyclopedic Dictionary of Roman Law. American Philosophical Society.

- Bienenfeld, M. (1988). Regularity in price changes as an effect of changes in distribution. Cambridge Journal of Economics, 12, 247-255.
- Bonato, M. (2018). Realized correlations, betas and volatility spillover in the agricultural commodity market: What has changed? 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.
- Brenner, R. (2002). American economic revival, In R. Brenner, The Boom and the Bubble: The US in the World Economy. New York: Verso.
- Brenner, R. (2006). From boom to downturn, In R. Brenner, The Economics of Global Turbulence: The Advanced Capitalist Economies from Long Boom to Long Downturn, 1945-2005. New York: Verso.
- Brenner, R. (2006). The puzzle of the long downturn, In R. Brenner, The Economics of Global Turbulence: The Advanced Capitalist Economies from Long Boom to Long Downturn, 1945-2005. New York: Verso.
- Bürgenmeier, B. (1994). Environmental policy: Beyond the economic dimension. In B. Bürgenmeier (Ed.), Economy, environment, and technology: A socio-economic approach, pp. 166-175. New York, NY: Armonk Sharpe.
- Burke, M., Hsiang, S.M. & Miguel, E. (2015). Global non-linear effect of temperature on economic production. Nature, 527, 235-239.
- Căpek, K. 1921. Rossum's Universal Robots (2004 ed.), trans. Claudia Novack. New York: Penguin Group .
- Cellan-Jones, R., 2014. "Stephen Hawking Warns Artificial Intelligence Could End Mankind", BBC News, 2 December, 2014. www.bbc.com/news/technology-30290540.
- Centeno, M.A., & Tham, A. (2012). The emergence of risk in the global system. Princeton, NJ: Princeton University working paper.
- Centeno, M.A., Creager, A.N., Elga, A., Felton, E., Katz, St.N., Massey, W.A. & Shapiro, J.N. (2013). Global systemic risk: Proposal for a research community. Princeton University, NJ: Princeton Institute for International and Regional Studies working paper.
- Chancel, L. & Piketty, Th. (2015). Carbon and inequality: From Kyoto to Paris. Paris, France: Paris School of Economics.
- Chilcote, E.B. (1997). Interindustry structure, relative prices, and productivity: An input-output study of the U.S. and O.E.C.D countries. Doctoral Thesis. New York, NY: The New School for Social Research Department of Economics.
- Cialdini, R.B. (1996). Social influence and the triple tumor structure of organizational dishonesty. In D.M. Messick and A.E. Tenbrunsel (Eds.), Codes of Conduct, pp. 44-58. New York, NY: Sage.
- Cifarelli, G. & Paesani, P. (2018). Fundamentalists, chartists and hedgers: A non-linear dynamic model of futures oil pricing: Theory and empirical evidence. 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.

Clancy, Erin. 1998. "The Tragedy of the Global Commons." Indiana Journal of Global Legal Studies 5(2): 601-19.

Cockshott, W.P. & Cottrell, A. (1993). Towards a new socialism. Nottingham, UK: Spokesman.

- Committee on Legal Affairs. 2016. Draft Report with Recommendations to the Commission on Civil Law Rules on Robotics. May 31. 2015/2103(INL)
- Copeland, Jack. 2000. "What is Artificial Intelligence?" AlanTuring.net, May. www.alanturing.net/turing_archive/pages/ Reference%20Articles/what_is_AI/What%20is%20AI02.html Critchley, S. (2015). Notes on suicide. Fitzcarraldo.

Dawes, R.M. (1980). Social dilemmas. Annual Review of Psychology, 31, 1, 169-193.

- Dillon, M. & Garland, L. (2005). Ancient Rome: From the early republic to the assassination of Julius Caesar. London, UK: Routledge.
- Dowell, Ryan. 2018. "Fundamental Protections for Non-Biological Intelligences or: How we Learn to Stop Worrying and Love our Robot Brethren." Minnesota Journal of Law, Science & Technology 19(1): 305-36.
- Dukhanina, E., Leveque, F. & Massol, O. (2018). Policy measures targeting a more integrated gas market: Impactonpricesandarbitrageactivity.Retrievedfrom

https://www.eeg.tuwien.ac.at/conference/iaee2017/files/paper/649_Dukhanina_fullpaper_2017-06-30_21-58.pdf

- Ekholm, T. (2018). Climatic cost-benefit analysis under uncertainty and learning on climate sensitivity and damages. 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.
- Elgouacem, A. & Legrand, N. (2018). The delaying effect of storage on investment. Evidence from the crude oil sector. 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.
- Emmert, H. (2018). The melting point: The economics and environmental impact of the winter sports industry. Unpublished working paper, Economics of the Environment, The New School, New York.
- Environmental Defense Fund (2018). How climate change plunders the planet: A warming earth disturbs weather, people, animals and much more. Retrieved from https://www.edf.org/climate/how-climate-change-plunders-planet
- Esposito, M., Haider, A., Semmler, W. & Samaan, D. (2018). Enhancing job creation through the green transformation. In T. Altenburg & C. Assmann (Eds.), United Nations Environment Programme Report on Green Industrial Policy: Concept, Policies, Country Experiences, pp. 50-68. Geneva, Bonn: United Nations: United Nations Environment Programme & German Development Institute / Deutsches Institut für Entwicklungspolitik.
- EU Committee on Legal Affairs. 2016. "Draft Report with Recommendations to the Commission on Civil Law Rules on Robotics", May 31.
- Feld, L.P. & Frey, B.S. (2002). Trust breeds trust: How taxpayers are treated. Economics of Governance, 3, 2, 87-99.
- Figuerola-Ferretti, I. & Cervera, I. (2018). Recent credit risk and bubble behavior in the corporate energy sector. Retrieved from https://repositorio.comillas.edu/jspui/handle/11531/26571
- Filip, O., Janda, K., Kristoufek, L. & Zilberman, D. (2018). Food versus fuel: An updated and expanded evidence. CAMA working paper 73. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3070704
- Foley, D. (2007). The economic fundamentals of global warming. Santa Fe, NM: Santa Fe Institute Working paper 07-12-044. Retrieved from https://www.santafe.edu/research/results/working-papers/the-economic-fundamentals-of-global-warming
- Fotopoulos, T. (1997). Towards an inclusive democracy. Cassel.
- Frey, B.S. (1992). Tertium datur: Pricing, regulation and intrinsic motivation. Kyklos, 45, 2, 161-184.
- Frey, B.S. (1997). Not just for the money: An economic theory of personal motivation. Cheltenham, UK: Edward Elgar.
- Galeon, Dom & Christianna Reedy. 2017. "Kurzweil Claims That the Singularity Will Happen by 2045." Futurism, October 5, futurism.com/kurzweil-claims-that-the-singularity-will-happen-by-2045/.
- Gamauf, R. (2009). Slaves doing business: The role of Roman Law in the economy of a Roman household. European Review of History, 16, 3, 331-346.
- Ganepola, C., Garrett, I. & Cho, S. (2018). Oil prices, discount rate news and cash flow news. 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.
- Gonglewski, M. & Helm, A. (2010). An examination of business case methodology: Pedagogical synergies from two disciplines. Global Business Languages, 15, 3, 17-31.
- Greiner, A. & Semmler, W. (2008). The global environment, natural resources, and economic growth. Oxford, UK: Oxford University Press.
- Harris, W.V. (2000). Trade. In: The Cambridge Ancient History: The High Empire A.D. 70-192, p. 11. Cambridge, UK: Cambridge University Press.
- Hasseldine, J.D. (1998). Prospect theory and tax reporting decisions: Implications for tax administrators. International Bureau of Fiscal Documentation, 52, 11, 501-505.
- Hayes, A. (2018). Decentralized banking: Monetary technocracy in the digital age. Social Science Research Network working paper retrievable at https://papers.srn.com/sol3/papers.cfm?abstract_id=2807476
- Held, D. & McGrew, A.G. (2007). A new world economic order? Global markets and state power: Beyond globalization/anti-globalization: Beyond the great divide. New York, NY: Polity.
- Helm, A. (2009). Global marketing classroom: Green marketing. American Marketing Association Global SIG Newsletter, Global Interests.
- Hildebrandt, Mireille. 2013, "Extraterritorial Jurisdiction to Enforce in Cyberspace? Bodin, Schmitt, Grotius in Cyberspace?" Toronto Law Journal 63: 196-224.
- Hopkins, K. (1983). Conquerors and slaves: Sociological studies in Roman history. New York: Cambridge University Press.
- Ielpo, F. & Sevi, B. (2013). Forecasting the density of oil futures returns using model-free implied volatility and highfrequency data. Montpellier, France: Centre de Recherche en Economie et Droit de l'Energie CREDEN Equipe ART Dev, Université Montpellier, Faculté d'Economie, Cahier de recherche n° 13.10.106.

Ji, Q., Liu, B.Y., Zhao W.L. & Fan, Y. (2018). Modelling dynamic dependence and risk spillover between all oil price shocks and stock market returns in the BRICS. 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.

Johnston, M. (1957). Roman life. Chicago, IL: Scott, Foresman.

- Kaabia, O., Dhaoui, A., Abid, I. & Guesmi, K. (2018). Oil supply-demand shocks and stock prices. Retrieved from https://www.researchgate.net/publication/324687727_Oil_supply-demand_shocks_and_stock_prices
- Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47, 2, 263-291.

Kehoe, D.P. (2011). Law and social function in the Roman Empire. The Oxford Handbook of Social Relations in the Roman World. Oxford, UK: Oxford University Press.

- Kilian, L. & Zhou, X. (2018). The propagation of regional shocks in housing markets: Evidence from oil price shocks in Canada. Centre for Economic Policy Research, London, UK, CEPR Discussion Paper DP12845. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3159148
- Kirchler, E.M. (2007). The economic psychology of tax behaviour. Cambridge, UK: Cambridge University Press.
- Klein, T., von Mettenheim, H.J., Walther, Th. & Wegener, Ch. (2018). Forecasting high frequency realized volatility of oil with exogenous regressors. 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.

Kowert, Weston. 2017. "The foreseeability of human-artificial intelligence interactions." Texas Law Review 96: 181-204.

Kowert, Weston. 2017. "The foreseeability of human-artificial intelligence interactions." Texas Law Review 96: 181-204.

Kruse, R. & Wegener, Ch. (2017). Time-varying persistence in real oil prices and its determinants. Working paper retrieved from http://www.wisostat.uni-koeln.de/sites/statistik/abstracts/RKB-Time-Varying_Persistence_in_Real_Oil_Prices_and_its_Determinant.pdf

- Larson, David Allen. 2010. "Artificial Intelligence: Robots, Avatars, and the Demise of the Human Mediator." Ohio State Journal on Dispute Resolution 25: 105-64.
- Larson, David Allen. 2010. "Artificial Intelligence: Robots, Avatars, and the Demise of the Human Mediator." Ohio State Journal on Dispute Resolution 25: 105-64.
- Laton, David. 2016. "Manhattan_Project.Exe: A Nuclear Option for the Digital Age." Catholic University Journal of Law & Technology 25(4): 94-153.
- Laton, David. 2016. "Manhattan_Project.Exe: A Nuclear Option for the Digital Age." Catholic University Journal of Law & Technology 25(4): 94-153.
- Legnazzi, Ch. (2018). Forward-looking VaR and CVaR: An application to the natural gas market. 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.
- Leontief, W. (1997). Input-output economics. Oxford, UK: Oxford University Press.
- Lin, P., Abney, K., Bekey, G.A. 2012. ROBOT ETHICS: The Ethical and Social Implications of Robotics. London, England: The MIT Press.
- Lips, J. (2018). Debt and the oil industry: Analysis on the firm and production level. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3026063
- MacDonald, Fiona. 2016. "Harvard Scientists Think They've Pinpointed the Physical Source of Consciousness." Science Alert, June 23. http://www.sciencealert.com/harvard-scientists-think-they-ve-pinpointed-the-neural-source-of-consciousness.
- MacDonald, Fiona. 2016. "Harvard Scientists Think They've Pinpointed the Physical Source of Consciousness." Science Alert, June 23. http://www.sciencealert.com/harvard-scientists-think-they-ve-pinpointed-the-neural-source-of-consciousness.
- Mahadeo, S.M.R., Heinleinb, R. & Legrenzib, G.D. (2018). New contagion applications for the energy-finance nexus: A higher moment approach. 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.
- Markowitz, E.M., Grasso, M. & Jamieson, D. (2015). Climate ethics at a multidisciplinary crossroad: Four directions for future scholarship. Climate Change, 130, 3, 465-474.
- Marra, William & Sarah McNeil. 2013. "Understanding "The Loop": Regulating the Next Generation of War Machines." Harvard Journal of Law & Public Policy 36: 1139-87.
- Marra, William & Sarah McNeil. 2013. "Understanding "The Loop": Regulating the Next Generation of War Machines." Harvard Journal of Law & Public Policy 36: 1139-87.
- Marron, D.B. & Morris, A.C. (2016). How to use carbon tax revenues. Washington, D.C.: Tax Policy Center Urban Institute & Brookings Institution.

Marx, K. (1867/1995). Capital: A critique of political economy. Moscow, RU: Progress.

Mauss, M. (1979). A category of the human mind: The notion of the person, the notion of 'self,' In M. Mauss, p. 81, Sociology and Psychology, London, UK: Routledge.

- Meghdari, A. & Alemi, M. (2018). Recent advances in social & cognitive robotics and imminent ethical challenges. Proceedings of the 10th International RAIS Conference on Social Sciences and Humanities organized by Research Association for Interdisciplinary Studies (RAIS) at The Erdman Center at Princeton University, Princeton, New Jersey, United States. Cambridge, MA: The Scientific Press.
- Meghdari, A. & Alemi, M. (2018). Recent advances in social & cognitive robotics and imminent ethical challenges. Proceedings of the 10th International RAIS Conference on Social Sciences and Humanities organized by Research Association for Interdisciplinary Studies (RAIS) at The Erdman Center at Princeton University, Princeton, New Jersey, United States. Cambridge, MA: The Scientific Press.
- Meghdari, A., Alemi, M., Zakipour, M., Kashanian, S.A. 2018. "Design and Realization of a Sign Language Educational Humanoid Robot." Journal of Intelligent & Robotic Systems, pp. 1-15, Springer, 2018.
- Meghdari, A., Alemi, M., Zakipour, M., Kashanian, S.A. 2018. "Design and Realization of a Sign Language Educational Humanoid Robot." Journal of Intelligent & Robotic Systems, pp. 1-15, Springer, 2018.
- Meghdari, A., Shariati, A., Alemi, M., Vossoughi, G.R., et. al. 2018. "Arash: A Social Robot Buddy to Support Children with Cancer in a Hospital Environment." Proc. of the IMechE, Part H: Journal of Engineering in Medicine 232(6): 605-618.
- Meghdari, A., Shariati, A., Alemi, M., Vossoughi, G.R., et. al. 2018. "Arash: A Social Robot Buddy to Support Children with Cancer in a Hospital Environment." Proc. of the IMechE, Part H: Journal of Engineering in Medicine 232(6): 605-618.
- Mohamed, A.M.I. (2016). Globalization and new international public works agreements in developing countries: An analytical perspective. London, UK: Routledge.
- Moore, J.W. (2009). Ecology and the accumulation of capital: A brief environmental history of neoliberalism. Paper presented at the Workshop on Food, Energy, Environment: Crisis in the Modern World-System. Fernand Braudel Center, Binghampton University, October 9-10.
- Mumford, A. (2001). Taxing culture. Aldershot, UK: Ashgate.
- Nordhaus, W.D. (1994). Mapping the global commons: The economics of climate change. Cambridge, MA: MIT Press.
- Noyes, Katherine. 2016. "5 things you need to know about A.I.: Cognitive, neural and deep, oh my!" COMPUTERWORLD, March 3. www.computerworld.com/article/3040563/enterprise-applications/5-things-you-need-toknow-about-ai-cognitive-neural-anddeep-oh-my.html [http://perma.cc/7PW9-P42G.
- Noyes, Katherine. 2016. "5 things you need to know about A.I.: Cognitive, neural and deep, oh my!" COMPUTERWORLD, March 3. www.computerworld.com/article/3040563/enterprise-applications/5-things-you-need-toknow-about-ai-cognitive-neural-anddeep-oh-my.html [http://perma.cc/7PW9-P42G.
- Noyes, Katherine. 2016. "5 things you need to know about A.I.: Cognitive, neural and deep, oh my!" COMPUTERWORLD, March 3. www.computerworld.com/article/3040563/enterprise-applications/5-things-you-need-toknow-about-ai-cognitive-neural-anddeep-oh-my.html [http://perma.cc/7PW9-P42G.
- Ordu-Akkaya, B.M. & Soytas, U. (2018). The impact of foreign portfolio investment on financialization of commodities. 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.
- Orttung, R. & Overland, I. (2011a). A limited toolbox: Explaining the constraints on Russia's foreign energy policy. Journal of Eurasian Studies, 2, 1, 74-85.
- Orttung, R. & Overland, I. (2011b). Russia and the formation of a gas cartel. Problems of Post-Communism, 58, 3, 53-66.
- Orttung, R. (2010a). Politics and profits in Russia's foreign oil and gas conflicts. Center for Security Studies Colloquium, Swiss Federal Institute of Technology, August 30.
- Orttung, R. (2010b). Five myths about energy security. Energy Forum 2010, Swiss Federal Institute of Technology, May 11.
- Orttung, R. (2014). Russia's place in the world energy system: Fueling the future: Deconstructing the role of energy in contemporary international relations. Toronto, Canada: University of Toronto, January 25.
- Orttung, R. (2015). Promoting sustainability in Russia's artic: Integrating local, regional, federal, and corporate interest. In S. Oxenstierna (Ed.), The Challenges for Russia's Policized Economic System, pp. 202-219. London, UK: Routledge.
- Orttung, R. (2016). Sustaining Russia's arctic cities: Resource politics, migration, and climate change. New York, NY: Berghahn.
- Perovic, J., Orttung, R.W. & Wenger, A. (2009). Russian energy power and foreign relations. London: Routledge.
- Puaschunder, J. M. (2015). Trust and reciprocity drive common goods allocation norms. Proceedings of the Cambridge Business & Economics Conference. Cambridge, UK: Cambridge University.
- Puaschunder, J. M. (2017b). Mapping Climate in the 21st Century. Development, 59, 3, 211-216.

Puaschunder, J. M. (forthcoming a). Climate polices with burden sharing: The economies of climate financing. Advances in Financial Economics.

- Puaschunder, J.M. (2016). Intergenerational climate change burden sharing: An economics of climate stability research agenda proposal. Global Journal of Management and Business Research: Economics and Commerce, 16, 3, 31-38.
- Puaschunder, J.M. (2017a). Climate in the 21st century: A macroeconomic model of fair global warming benefits distribution to grant climate justice around the world and over time. Proceedings of the 8th International RAIS Conference on Social Sciences and Humanities organized by Research Association for Interdisciplinary Studies (RAIS) at Georgetown University, Washington, D.C., United States, March 26-27, pp. 205-243.
- Puaschunder, J.M. (2017c). Socio-psychological motives of socially responsible investors. Advances in Financial Economics, 19, 1, 209-247.
- Puaschunder, J.M. (2017d). Sunny side up! From climate change burden sharing to fair global warming benefits distribution: Groundwork on the metaphysics of the gains of global warming and the climatorial imperative. Proceedings of the Administrative Sciences Association of Canada Annual Conference, HEC Montreal, Quebec, Canada.
- Puaschunder, J.M. (2017e). The call for global responsible inter-generational leadership: The quest of an integration of inter-generational equity in corporate social responsibility (CSR) models. In Natural resources management: Concepts, methodologies, tools, and applications: Information resources management association. Hershey, PA: IGI Global.
- Puaschunder, J.M. (2018). Artificial Intelligence Evolution: On the virture of killing in the artificial Age. Social Science Research Network working paper retrievable at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3247401
- Puaschunder, J.M. (2018a). Gifts without borders: Intergenerational glue connecting over distance and time as pure international development in the age of migration. In J. M. Puaschunder (Ed.), Intergenerational responsibility in the 21st century, pp. 143-174. Wilmington, DE: Vernon.
- Puaschunder, J.M. (2018b). Value at looking back: Towards an empirical validation of the role of reflexibility in econo-historic backtesting: Economic market prediction corrections correlate with future market performance. Social Science Research Network Working Paper retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3170734
- Puaschunder, J.M. (forthcoming b). Governance and climate justice: Global South and developing nations. Cham, Switzerland: Palgrave Macmillan.
- Puaschunder, J.M. (forthcoming c). Nachhaltigkeit und Investment. In E. Bamberg & C. Schmitt (Eds.), Psychologie und Nachhaltigkeit: Konzeptionelle Grundlagen, Anwendungsbeispiele und Zukunftsperspektiven. Hamburg, Germany: Springer.
- Puaschunder, J.M. (forthcoming). Artificial Intelligence Ethics.
- Puaschunder, J.M. (work in progress). Climate change induced market prospects. Washington, DC: George Washington University working paper.
- Puaschunder, J.M. (work in progress). Putty and clay labor: What to learn from the Cambridge Controversy for the entrance of AI in today's workforce.
- Renz, A. (2018). Sector competition for energy resources in a carbon constrained hoteling model. 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.
- Ricardo, D. (1817/1951). On the principles of political economy and taxation. In P. Sraffa (Ed.), The Works and Correspondence of David Ricardo, pp. 1951-1973. Cambridge, UK: Cambridge University Press.
- Roberts, G., Barbier, E. & van't Veld, K.V. (2018). The beginning of the end of the age of oil. Retrieved from https://cenrep.ncsu.edu/cenrep/wp-content/uploads/2016/08/Roberts.pdf
- Rosenlund Soysal, E. (2018). Share repurchase under diverging beliefs about carbon risk. 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.
- Russell, Stuart & Peter Norvig. 1995. Artificial Intelligence a Modern Approach. New Jersey: A. Simon & Schuster Company.
- Russell, Stuart & Peter Norvig. 1995. Artificial Intelligence a Modern Approach. New Jersey: A. Simon & Schuster Company.
- Russell, Stuart & Peter Norvig. 1995. Artificial Intelligence a Modern Approach. New Jersey: A. Simon & Schuster Company.
- Sachs, J.D. (2014). Climate change and intergenerational well-being. In L. Bernard & W. Semmler (Eds.), The Oxford Handbook of the Macroeconomics of Global Warming, pp. 248-259. Oxford, UK: Oxford University Press.
- Saffari, E., Meghdari, A., Vazirnezhad, B., Alemi, M. 2015. "Ava (A Social Robot): Design and Performance of a Robotic Hearing Apparatus." LNCS: Social Robotics 9388: 440-450, Springer, Oct. 2015.
- Saussay, A. & Sato, M. (2018). The impacts of energy prices on industrial foreign investment location: Evidence from global firm level data. Retrieved from https://personal.lse.ac.uk/satom/publication/currentss/
- Schmölders, G. (1960). Das Irrationale in der öffentlichen Finanzwirtschaft. Frankfurt am Main, Germany: Suhrkamp.

- Schuller, Alan. 2017. "At the Crossroads of Control: The Intersection of Artificial Intelligence in Autonomous Weapon Systems with International Humanitarian Law." Harvard National Security Journal 8: 379-425.
- Schuller, Alan. 2017. "At the Crossroads of Control: The Intersection of Artificial Intelligence in Autonomous Weapon Systems with International Humanitarian Law." Harvard National Security Journal 8: 379-425.
- Schütze, F., Aleksovski, D. & Mozetic, I. (2018). Stock market reactions to international climate negotiations. Retrieved from https://bibbase.org/network/publication/schtze-aleksovski-mozetistockmarketreactionstointernationalclimatenegotiations-2018
- Schweiger, H. & Stepanov, A. (2018). Energy subsidies, energy intensity and management practices. Retrieved from https://www.ebrd.com/cs/Satellite?c=Content&cid=1395236502077&d=Mobile&pagename=EBRD%2FContent%2FContentLavout
- Sherman, H.J. (1967). Marx and the business cycle. Science & Society, 31, 4, 486-504.
- Smith, B. (2006). Gender. Gale Virtual Reference Library. Charles Scribner's.
- Sofge, E. 2015. "Bill Gates Fears A.I., But A.I. Researchers Know Better." Popular Science. www.popsci.com/ billgates-fears-ai-ai-researchers-know-better.
- Sofge, E. 2015. "Bill Gates Fears A.I., But A.I. Researchers Know Better." Popular Science. www.popsci.com/ billgates-fears-ai-ai-researchers-know-better.
- Solum, Lawrence. 1992. "Legal Personhood for Artificial Intelligences." North Carolina Law Review 70(4): 1231-87.
- Solum, Lawrence. 1992. "Legal Personhood for Artificial Intelligences." North Carolina Law Review 70(4): 1231-87.
- Somanathan, E., Somanathan, R., Sudarshan, A. & Tewari, M. (2014). The impact of temperature on productivity and labor supply: Evidence from Indian manufacturing. Discussion Paper in Economics 14-10, Indian Statistical Institute, Economics and Planning Unit.
- Soros, G. (2003). The alchemy of finance. Hoboken, NJ: Wiley Finance.
- Soros, G. (2009). The crash of 2008 and with is means. New York, NY: Public Affairs.
- Soros, G. (2013). Fallibility, reflexivity and the human uncertainty principle. Journal of Economic Methodology, 20, 309-329.
- Soskice, D. (1978) Strike waves and wage explosions, 1968-1970: An economic interpretation, In C.Crouch & A. Pizzorno (Eds.), The Resurgence of Class Conflict in Western Europe since 1968, pp. 221-245, London, UK: Palgrave Macmillan.
- Sraffa, P. (1960). Production of commodities by means of commodities. Cambridge, UK: Cambridge University Press.
- Stiglitz, J. (1998). The private uses of public interests: Incentives and institutions. Journal of Economic Perspectives, 12, 3-22.
- Stroebe, W. & Frey, B.S. (1982). Self-interest and collective action: The economics and psychology of public goods. British Journal of Social Psychology, 21, 2, 121-137.
- Taheri, A.R., Meghdari, A., Alemi, M., Pouretemad, H.R., et. al. 2018. "Human–Robot Interaction in Autism Treatment: A Case Study on Three Pairs of Autistic Children as Twins, Siblings, and Classmates." Int. Journal of Social Robotics 10(1): 93-113.
- Täuber, S., van Zomeren, M. & Kutlaca, M. (2015). Should the moral core of climate issues be emphasized or downplayed in public discourse? Three ways to successfully manage the double-edged sword of moral communication. Climate Change, 130, 3, 453-464.
- The Oxford Encyclopedia of Ancient Greece and Rome. (2010). Oxford, UK: Oxford University Press.
- The World Bank (2015). Green bonds attract private sector climate finance. World Bank Brief, 2015.
- The World Economic Forum 2015 Report. (2015). Davos, Switzerland: World Economic Forum.
- Themistoklis, T. (2018). Artificial intelligence as global commons and the "international law supremacy" principle. Proceedings of the 10th International RAIS Conference on Social Sciences and Humanities organized by Research Association for Interdisciplinary Studies (RAIS) at The Erdman Center at Princeton University, Princeton, New Jersey, United States. Cambridge, MA: The Scientific Press.
- Themistoklis, T. (2018). Artificial intelligence as global commons and the "international law supremacy" principle. Proceedings of the 10th International RAIS Conference on Social Sciences and Humanities organized by Research Association for Interdisciplinary Studies (RAIS) at The Erdman Center at Princeton University, Princeton, New Jersey, United States. Cambridge, MA: The Scientific Press.
- Trinks, A., Mulder, M. & Scholtens, B. (2017). Greenhouse gas emissions intensity and the cost of capital. Retrieved from https://pdfs.semanticscholar.org/c8f5/d6b192df64cde5f578f100d7fb2a6b40515b.pdf
- Tsagourias, Nicholas. 2015. "The legal status of cyberspace." in Research Handbook, International Law and Cyberspace, edited by N. Tsagourias & R. Buchan, 13-29. Cheltenham: Edward Elgar Publishing.
- Tversky, A. & Shafir, E. (1992). Choice under conflict: The dynamics of deferred decision. Psychological Science, 3, 6, 358-361.

Tyler, T.R. & De Cremer, D. (2006). How do we promote cooperation in groups, organizations, and societies? In P.A.M van Lange (Ed.), Bridging Social Psychology: Benefits of Transdisciplinary Approaches, pp. 427-433. Mahwah, NJ: Lawrence Earlbaum.

United Nations Department of Economic and Social Affairs. 2017. Will robots and AI cause mass unemployment? Not necessarily, but they do bring other threats. New York: https://www.un.org/development/desa/en/news/policy/will-robots-and-ai-cause-mass-unemployment-notnecessarily-but-they-do-bring-other-threats.html

United Nations Department of Economic and Social Affairs. 2017. Will robots and AI cause mass unemployment? Not necessarily, but they do bring other threats. New York: https://www.un.org/development/desa/en/news/policy/will-robots-and-ai-cause-mass-unemployment-notnecessarily-but-they-do-bring-other-threats.html

United Nations Report on the World Social Situation (2011). The global social crisis. New York, NY: United Nations Department of Economic and Social Affairs Report ST/ESA/334.

Veruggio, G. 2005. "The Birth of Roboethics." ICRA 2005, IEEE Int. Conference on Robotics and Automation: Workshop on Robo-Ethics, Barcelona, April 18, 2005.

Vlassopoulos, K. (2009). Politics antiquity and its legacy. Oxford, UK: Oxford University Press.

Warwick, K. & Shah, H. 2014. "How Good Robots Will Enhance Human Life." In book: Treatise on Good Robots Edition: PRAXIOLOGY: The International Annual of Practical Philosophy and Methodology Vol. 21, Book Chapter, Transaction Publishers, USA Editors: Krzysztof Tchoń, Wojciech W. Gasparski, January 2014.

Wenger, A., Orttung, R.W. & Perovic, J. (2009). Energy and the transformation of international relations: Toward a new producer-consumer framework. Oxford, UK: Oxford University Press.

- Wisskirchen, Gerlind et al. 2017. Artificial Intelligence and Robotics and their Impact on the Workplace. London: IBA Global Employment Institute.
- Wisskirchen, Gerlind et al. 2017. Artificial Intelligence and Robotics and their Impact on the Workplace. London: IBA Global Employment Institute.
- World Bank 2015 Report. (2015). Washington, D.C.: World Bank.
- World Bank Group Migration and Development Brief 26 (2016). Migration and remittances: Recent development and outlook. Washington D.C.: International Bank for Reconstruction and Development, World Bank Group, April 2016.
- World Investment Report (2015). Reforming international investment governance. United Nations Conference on Trade and Development Report. New York, NY: United Nations.
- Yarlik, M.V., Filis, G., Lloyd, T. & Degiannakis, St. (2018). Forecasting oil price volatility: The role of mixed-frequency data (MIDAS) model. 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.
- Zhang, D. (2018). Upgrading China's energy structure: The role of financial markets. 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.
- Zhou, S.L., Smulders, S. & Gerlagh, R. (2018). Closing the loop in a circular economy: Saving resources or suffocating innovations? 6th International Symposium on Environment Energy & Finance Issues, ISEFI 2018, Paris, France, May 24-25.



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings



homepage: http://www.acrn.eu/frap

Value at Looking Back: Towards an Empirical Validation of the Role of Reflexivity in Econo-Historic Backtesting: Economic Market Prediction Corrections Correlate with Future Market Performance

Julia M. Puaschunder*

The New School, Columbia University, Princeton University

Abstract

Globalization led to an intricate set of interactive relationships between individuals, organizations and states. Unprecedented global interaction possibilities have made communication more complex than ever before in history as the whole has different properties than the sum of its increasing diversified parts. With growing globalization and quickening of transfer speed, information may impose unknown systemic economic risks on a global scale. Collective interaction effects lead to hard-to-foreseeable fallacy of composition downfalls. Emergent risks imbued in interaction appear to be inherent in global economic systems. In the light of growing tendencies of globalization, the demand for an in-depth understanding of how information echoes in socio-economic correlates has gained unprecedented momentum. In seeking to shed light on implicit system failures' socio-economic consequences down the road and potentially-disastrous outcomes of cumulative actions triggering mass movements; the paper outlines unexpected dangers and insufficiently-described shadows of past market expectation corrections on future economic market performance. Overall, the following article innovatively paints a novel picture of the mass psychological underpinnings of business cycles based on information flows in order to recommend how certain communication strategies could counterweight and alleviate information failing market performance expectations that could potentially build disastrous financial market mass movements of booms and busts. This paper will study the role of information in building socially-constructed economic correlates, which promises to explain how market outcomes are developed in the social compound and can be guided by central agents' communication. Classical theories of price will be reflected in regards to market expectations. Through the lens of the real competition paradigm, the following paper will then specifically unravel how central bank economic forecasts produce certain types of price expectations that form market patterns leading to collectively-shared economic outcomes that may echo in the real economy. An introduction to the history of economic cycles will lead to George Soros' Theory of Reflexivity and Anwar Shaikh's formalization in order to draw inferences for the analysis of the role of information in creating economic booms and busts in the age of globalization. Empirically, based on a central European central bank's GNP projections and backtesting corrections, a pattern of central bank corrections communication and economic market performance will be unraveled for the first time to outline that central bank market prediction corrections are positively correlated with near future market performances and negatively correlated with distant future market performances. The collective reality of prices and the irrationality of the crowds perturbating markets will be discussed. Business cycles are argued to obey some kind of natural complexity, as for being influenced by econohistoric communication trends. Recommendations how to create more stable economic systems by avoiding emergent risks in communicating market prospects more cautiously will be given in the discussion followed by a prospective future research outlook and conclusion.

* Corresponding author.

E-Mail address: Julia. Puaschunder@newschool.edu

Financial support of the American Academic Research Conference on Global Business, Economics, Finance and Social Sciences, Fritz Thyssen Foundation, Research Association for Interdisciplinary Studies, The New School Dean's Office, The New School Department of Economics, The New School Fee Board, The New School for Social Research, The New School Eugene Lang College, and the University of Vienna is gratefully acknowledged. The author declares no conflict of interest. The student thanks Professor Anwar Shaikh for most generous share of expertise and vibrant classes on '*Advanced Political Economy*,' '*The Economics of Crisis and Austerity*' as well as '*Classical Theory of Price*.' The author thanks Professor Willi Semmler for sharing invaluable insights. The participants of the Fall 2016 New School class '*Classical Theory of Price*' contributed to discussions on the ideas underlying and outlined in this paper. All omissions, errors and misunderstandings in this piece are solely the author's.

Keywords: Backtesting, BIP, business cycles, central bank communication, corrections, errors, expectations, forecasting corrections, Gross National Product (GNP), globalization, long boom, long downturn, market performance, price, reflexivity

"The economic history of the developed capitalist world appears to be one of almost constant progress: inexorable growth, rising standards of living, rising productivity, and ever-improving health, well-being, and welfare. Seen from afar, it is the system's order, its internal coherence, which stands out.

> Yet the closer one looks, the more haphazard it all seems. Individuals wander along entangled paths, propelled by obscure motivations toward some dimly imagined ends, crisscrossing and colliding as they act out their economic roles as buyers and sellers, bosses and workers, producers and speculators, employed and unemployed. Information, misinformation, and disinformation hold equal sway. Ignorance is as purposeful as knowledge. Private and public spheres are entwined throughout, as are wealth and poverty, development and underdevelopment, conquest and cooperation. And everywhere there appears a characteristic unevenness: across localities, regions, and nations; and across time, in the form of booms, busts, and breakdowns. Seen up close, it is the system's disorder that is most striking." (Shaikh, 2016, p. 3)

Introduction

Globalization led to an intricate set of interactive relationships between individuals, organizations and states (Centeno & Tham, 2012). Deepening nets of interactions challenge human foresight (Gilpin, 2001). Collective interaction effects lead to hard-to-foresee fallacy of composition societal downfalls (Shaikh, 2016). As complex interdependencies may hold unknown outcomes for society, highly integrated international communities are under the whim of unexpected socio-economic developments. In seeking to shed light onto implicit system failures' socio-economic consequences down the road and potentially-disastrous outcomes of cumulative actions triggering mass movements; the currently emerging Emergent Risk Theory outlines unexpected dangers and insufficiently-described shadows of the invisible hand of the real competition world economy in the age of globalization (Centeno & Tham, 2012; Miller & Rosenfeld, 2010; Shaikh, 2016).

Since the post-World War period, the world globalized. International economic activities now involve a larger number of countries and sectors than at any time in history and reach deeper into every human life than ever before (Held & McGrew, 2007). Global interaction possibilities have also made communication unprecedentedly complex. With growing globalization and quickening of transfer speed, information flows may impose unknown systemic economic risks on a global scale (Centeno et al., 2013; Okamoto, 2009; Urry, 2012). Nowadays information flow has no longer limited local effects but potentially unforeseen global consequences (Leonhardt, Keller, & Pechmann, 2011; Stiglitz, 2006; Summers & Pritchett, 2012). This paper argues that market performance expectations and information about their burst may create risks within economic markets as for fueling economic booms and downturns. The paper thereby builds on classical theories of prices and the classics' notions of expectations to then embark on heterodox socio-economic concepts such as George Soros' Theory of Reflexivity and Anwar Shaikh's formalization of reflexivity in his wider paradigm of Real Competition.

In the light of growing tendencies of globalization, the demand for an in-depth understanding of how information flows echo in socio-economic correlates and may steer economic fluctuations has gained unprecedented momentum. New economic thinking widens the interdisciplinary lens to study emergent risks of international communication shadowing economic markets and the societal compound. In exceeding orthodox economics' insights and traditional public policy attempts to curb societal risks, heterodox economic approaches outlining socio-economics of crises appear as real-world relevant emergent risk prevention strategy (Shaikh, 2016). Applying real competition and reflexivity concepts but also the emergent risk theory onto economic fluctuations is an innovative heterodox way to explain how information creates economic ups and downs in a turbulent economic world (Held & McGrew, 2007; Shaikh, 2016).

While by the end of the 1960s, the most renowned economists agreed that recessions were preventable, history proved them wrong (Brenner, 2006). For instance, the period from 1940 to 1973 became renowned for a time of economic prosperity – yet from 1973 a worldwide recession set in. Historic post World War II booms were transitioned to downturn from the mid 1970s on as the economic performance declined in the industrialized world. While in the 1980s and 90s the economy seemed to expand again, from 2000 on productivity slowed, again, from 1998-99 stock markets and currencies crashed or halted in 2001 and 2008 (Shaikh, 2016). Today, supplyside theory explains the downturn dependent on pressures from labor (Brenner, 2006). Behavioral economists give credit to the unplanned, uncoordinated and competitive nature of capitalist production as well as the problem of aggregates deviating from the individual's choice predictions. Anwar Shaikh's Theory of Real Competition as outlined in Capitalism: Competition, Conflict and Crises (2016) is at the forefront of explaining complex market interactions centered around profit and price-cutting competitive edges.

The expectations of outcomes of multiple decisions among people with conflicting interests may not be foreseeable and therefore have been overlooked by stable equilibrium theories. Yet in reality expectations may guide individual decision making and therefore imposed novel risks in complex markets. Economic indicators of product wages, international competition, outputcapital ratio, and post-tax profits have been studied extensively to derive conclusions about economic pre-indicators of crises and recommendations for improving economic systems. The role of inflation is widely cited in its invers relation to unemployment to determine economic conditions (Armstrong, Glyn & Harrison, 1991; Brenner, 2002). International transmissions of inflation are discussed in historic examples to build adaptive expectations on information about current and past experiences with inflation (Soskice, 1978). Disequilibrium inflation is explained in the literature to occur when the actual rate of inflation is greater than the expected followed by social unrest and industrial conflict (Soskice, 1978). Yet in all these discussions, concrete information about economic prospects and the role of economic expectations in shaping ideas and intentions of individual market actors playing out in economic fundamentals is missing. While business cycle theories primarily focus on describing economic correlates of booms and busts such as tight labor markets, investment trends and the uneven development throughout the world causing advantages and disadvantages in the economic impact of booms and busts around the world (Brenner, 2002); less attention is shed on socio-economic correlates that build expectations leading to irrational exuberance. Communication about markets and their performance is yet hereby argued to be underlying long-term economic trends as well. For instance, information transfer is key for innovation and markets to pick up new ideas. Information shaping expectations are the basis of investment trends. Economic expectations grown out of information on economic prospects and forecasts guide economic market actions. The collective mood in society shapes investment allocations amalgamating into economic trends that determine economic dynamisms of the collective soul of booms and busts. Economic forecasting and information about economic prospects may trickle down in individual economic decision making, which in its entirety shapes the economy as a whole. Access to information about market performance may also play a key role in price comparisons around the world, which are the basis of outsourcing and capital allocation decisions (Brenner, 2002). Information portrayed in media may also determine the investment mood and credit liquidity preferences of individuals and institutional representatives. Individual communication in the eye of economic ups and downs may in addition be the basis of social unrest and waves of strikes, which have been shown to be underlying factors determining wages and rates of profit for capitalist bleeding into

economic cycles (Brenner, 2002). All these correlates set the tact in shaping economic booms and busts as well as the long term cycles. The role of information for economic long-term cycles, however, has been – so far – widely overlooked in the standard neoclassical literature. Communication interventions are neglected in a wealth of writings on Federal Reserve and Central Bank interventions ranging from lowering interest rates to direct monetary stimulus. Studying the effect of information and communication on economic correlates to implicitly influence investment decisions may offer invaluable insights on how bubbles start and economic fluctuations can be smoothed.

The following paper therefore departs from the orthodoxy of hyper-rational individuals, who make rational choices based on perfect knowledge purely self-centered not taking other, history and the governmental issued information into account but makes a case for past performance and backtesting market prediction corrections influencing future performance. While classical economics typically focuses on actual outcomes rather than on the various expectations that might have motivated them (Arestis & Eatwell, 2008), the presented results will connect expectation performance corrections with future market outcomes.

Through capturing the interplay of communication about prospects and the fundamentals of the economy; the following article is meant to shed light on the socio-psychological underpinnings of economic downfalls. The paper thereby investigates the unprecedentedly described role of information in building and fueling economic booms and downturns. Theoretically, the paper will start with the notion of expectations in classical writings to then draw on historical foundations in economic analysis as outlined in Anwar Shaikh's Theory of Real Competition (2016) as well as Shaikh's (2013) formalization of George Soros' (1994) Theory of Reflexivity. Boom and bust patterns will be theoretically described and be argued to have the expected outcomes deviate from the actual path and that the actual path in turn deviates from the underlying fundamentals in reflection of past performances' prediction corrections (Shaikh, 2013). This paper will then discuss the impact of central bank forecasting and backtesting actual performance corrections in shaping economic ups and downs. Empirically, a central bank's economic forecasts of the Gross National Product (GNP) followed by publicly published backtesting corrections being related to certain market outcomes will be portrayed as instigation of herd and swarm behavior that potentially caused actual economic booms and busts leading to economic crises. The paper will then discuss the impact of the found future shadows of today's predictions in a heterodox fashion in order to serve as a window of opportunity for alleviating negative externalities of globalization imbued in technocrats' communication about market prospects and outperformed or bust expectations. Pursuing the greater goal of deriving recommendations how to stabilize economic markets in the instant communication century will lead to wider recommendations on finding an optimum balance of deregulated market systems and governmental control (Shaikh, 2016).

The following paper is organized as follows: An introduction to the theory of price and history of economic cycles leads to the analysis of the role of information in the creating of economic booms and busts. Economic forecasting and backtesting market correction data of a central European central bank will be presented as for retrieving information on the relation of past performances' corrections and future market outcomes. Recommendations how to create more stable economic systems by avoiding emergent risks imbued in market communication are given in the discussion followed by a prospective future research outlook and conclusion.

"The competitive firm must be concerned with tactics, strategy, and prospects for growth.... In the battle of real competition, the mobility of capital is the movement from one terrain to another, the development and adoption of technology is the arms race, and the struggle for profit growth and market share is the battle itself. There are winners and losers, and places can be switched.' (Shaikh, 2016, p. 15). "Firms within an industry fight to attract customers. Price is their weapon, advertising their propaganda, the local Chamber of Commerce their house of worship, and profit their supreme deity. Prices and propaganda serve two important functions: they attract customers away from other firms; and they attract new customers into the market as a whole. Cost-cutting becomes a central concern because prices are ultimately limited by costs. Costs intern depend on the length and intensity of the working day, the wage paid to workers, and the technology in use. Hence, struggles between capital and labor over wages and working conditions are immanent in the drive for profit. So too is never-ending technical change, whose principal purpose is to reduce costs." (Shaikh, 2016, p. 261).

Real competition

The neoclassical perfect competition model stylizes a picture of the economy of a very large number of very small firms, identical in scale and cost structure with the same profit margins and rates, and all facing the same horizontal demand curve. In perfect competition, firms are assumed to passively take prices and technology as given. In perfect competition, hyper-rational agents immediately catch on and adjust their behavior to new policies, so the economy jumps back to the natural rate of profits resulting in dynamic optimization. Each firm within a given industry is regulating capital with a profit rate equal to its industry average. Rational expectations are built on hyper-rationality, optimization, perfect competition and perfect information available to representative agents. Since competition between industries equalizes profit rates, all firms everywhere have the same rate of profit.

Imperfect competition acknowledges persistent firm-level profit rate differences, as does any correlation between profit margins and scale or capital intensity. Competition is the driver of production decisions, technological change, relative prices, interest rates and asset prices and exchange rates. In the imperfect competition paradigm, prices and costs are not uniform and firms face downward sloping demand curves. Price-setting and price-leadership by firms is viewed to be an indication of their monopoly power related to their scale, capital intensity, and relative market share.

As novel imperfect competition paradigm, behavioral economics acknowledges an interdisciplinary economic view. When sociology enters economics, animal spirits and excess demand are attributed to raise the level of the output path or lower its growth rate. The state also comes to exercise some degree of control over economic decisions as outlined in the behavioral insights paradigm. Behavioral economics remains to be organized around the failures of standard economics. The contribution starts with a demonstration of a failure of some common economic assumption and proceeds to prove a psychological explanation for that failure. Yet it remains at providing small changes of standard models that leave the basic structure of the theory intact (Shaikh, 2016, p. 115).

The novel heterodox stance is Theory of Real Competition (Shaikh, 2016). Elements of real competition can be found in the business literature (Shaikh, 2016). In real competition, the intensity of the competitive struggle depends on price-setting, cost-cutting and technology variations intrinsic to competition. Market prices differ within limits and firms respond to changes

in demand and supply through periodic price adjustments. Newer firms have larger scale and lower costs leading to cutting prices. In real competition, industries with higher entry scales have higher profit margins. Therefore, there is a positive correlation between selling prices and unit costs, and a negative between prices, firm scale and/or capital intensity. As more efficient firms tend to be larger and more capital-intensive, concentration ratios are correlated with barriers to entry.

In the turbulent real competition economies, the profit motive is inherently expansionary: investors try to recoup more money than they put in, and if successful, can do it again and again on a larger scale, colliding with others doing the same. Some succeed, some just survive and some fail altogether as the central regulating mechanism of capitalism (Shaikh, 2016).

Prices set by different sellers in the same industry are roughly equalized through the mobility of customers toward lower prices. Profit rates on new investments in different industries are roughly equalized through the mobility of capital toward higher profit rates. Both produce distributions around a corresponding common center. The classical notion of turbulent equilibration is very different from the conventional description of the equilibrium as a state-of-rest. Supply and demand play a role in the process but not in the final outcome, since both are affected by price-cutting entry and exit. An important point is that price and profit rate equalization are quintessential emergent properties, unintended outcomes of constant jockeying for greater profits (Shaikh, 2016). Profit rate equalization is a dynamic and turbulent process with fluctuations around a moving center of gravity. Investment flows into an industry are motivated by the expected rates of return on those potential new investments that embody the best-practice conditions of production (Arestis & Eatwell, 2008).

Prices tend to equalize because buyers gravitate toward the lowest price, which forces other sellers to adjust their own prices. Similarly, profit rates tend to equalize because investors flock to higher rates of return. This accelerates supply relative to demand in the favored industries and drives down their prices and profits. The rush toward riches close the gaps that initially motivated the agents while opening up new gaps, which feed new arbitrage movements. The turbulent equalizations of prices and profit rates are quintessential emergent properties. In keeping with price-setting and cost-cutting behavior of real competition, firms are forced to select the lowest cost conditions of production as determined by reproducible conditions involving technology, the length and intensity of the working day as was as the sum of unit depreciation, materials, and wage costs.

Given real markets are always turbulent, all choices are susceptible to information. The profit motive is the dominant factor in the regulation of economic growth (Shaikh, 2016). Actual decisions are always made in terms of current and expected market prices. Expected profitability at the core of any economic activity implies a mode of interaction between aggregate demand and supply. Profit thus regulates both – supply and demand (Shaikh, 2016). Profit determines an endogenous natural rate of growth. The interest rate thereby becomes the price of finance. Financial firms exist to make profit, and competition makes the profit rate of the regulating financial capitals gravitate around the general rate of profit (Shaikh, 2016).

In the theory of real competition, price-setting and price-cutting behavior are fundamental. A firm with lower unit costs can always drive out its competitors by cutting price. Price wars become the norm of the business world. Investors also gravitate towards higher rates of return. New investment flows more rapidly toward industries with higher rates of profit. Competition between industries will turbulently equalize regulating rates of profit (Shaikh, 1998). Mobility of capital implies that new investment will accelerate relative to demand in industries with higher rates of profit. Entry and exit of capitals occurs in response to profit rate differentials.

Competition within an industry compels individual producers to set prices strategically, forcing them to lower costs by cutting wages and increasing productivity to compete effectively within

the given technological constraints. Since economic growth derives from price margins and there is a lower level at which it becomes inefficient for firms to produce; economic growth within competitive settings will eventually lead towards a falling rate of profit. The only way to sustain economic growth on the long run and maintain in the market is therefore to constantly re-invent and find innovative ways how to produce cheaper or better customer-serving. Innovation thereby becomes the eternal flame of capitalism fueled by competition.

Real competition generates specific patterns. Real competition is antagonistic by nature, turbulent and different from perfect competition. In real competition each-against-each competition everyone uses tactics and strategies to evaluate prospects for gaining. Prices are given by competition. Firms set prices, which are weapons in competition for cutting costs and expanding the market share. Every corporation wants to operate at lowest costs to compete. There is a turbulent equalization of profit rates as prices are going up and down. The profit rate is central to accumulation because profit lies at the heart of any capitalist investment. Profits are the ultimate measure of capitalist success. Capital tends to flow into sectors, in which the profit rate is higher and leaves those in which profit rates are lower. This acceleration and deceleration raises active supply and demand and drives prices and profits up and down. This gives rise to a general tendency of profit rates to be equalized across sectors in a never-ending over- and undershooting in ever-changing centers of gravity. There is never a state of resting equilibrium, but average balance through perpetually offsetting errors. No real economy is ever in equilibrium, which makes the distinction between actual prices and equilibrium prices crucial (Ochoa, 1989). Turbulent arbitrage characterizes these recurrent fluctuations. Turbulence is normal around profitability. Equalization occurs in recurrent turbulent overshootings and underperformance in markets. In real competition, firms face downward sloping demand curves, set prices, have different costs, and partition into price-leaders and followers. Irrational expectations are grounded in real competition featuring downward sloping demand curves for firms. There is overand undershooting and imbalance that the fundamentals will not close gaps between expectations that cannot be validated. This turbulent climate, in which firms make decisions about investments, however, is hardly described in standard neoclassical economics. The role of information in building expectations and reporting back on backtesting derived outperformed or underperformed expectations' corrections is hardly described in the discounting literature.

In real competition, higher capital intensity leads to lower costs and there is a falling rate of profit. Firms constantly need to cut costs. Prices become weapons for survival. The price level, which depends on the cost of capital, also determines the interest rate. The real interest rate is the nominal interest rate minus the change of prices. The nominal interest rate is equal to the constant and the inflation rate. Price cutting behavior gives competitive firms advantages in terms of competition between industries. Capital moves to the sector with the most defensible price. If new technologies come in, the distribution of costs and profit margins changes. Adoption of technology and mobility of capital lead to lowering costs. Expectation plays a crucial role in real competition. In real competition, the incremental rate of profit derives from new investments and the adoption of new technologies. Firms are driven into new technologies, potentially based on expectations and access to market performance information.

The equilibrium as the amalgamated meeting point between buyers and sellers is a gravitational process featuring a constant overshooting and underperformance around its gravitational center (Kurz, 2010). The equilibrating process is therefore inherently cyclical and turbulent, subject to self-repeating fluctuations. In this context, individual capitalists make their decisions based on judgements about an intrinsically determinate future (Shaikh, 2016). But how these decisions are made, the theory is rather silent about. The role of information for the formation of intuition that leads to assumptions about the future and the hidden empathy with the crowd are rather undescribed in contemporary economic theory.

To study the underlying role of information and unravel how social forces drive the individual decision to echo in business cycles will allow recommendations on crowd control to lead towards stabilizing economic systems. The outlined paper therefore gives credit to the role of information in perturbating economic systems and assumes that future market performance may be based on past market expectation corrections (Shaikh, 2016).

Business Cycle Theories

Recurrent crises are part of capitalism. Economic shocks triggered crises in the 1820s, 1870s, 1930s, and 1970s. The logic of profit drives the system to repeat this pattern, which is inherently turbulent with powerful business cycles. In recent decades, the study of business cycles has moved towards economic fluctuations. Business cycles capture fluctuations in the gross domestic product (GDP). Business cycles studies identify different types of recurrent aggregate fluctuations tied to investments. Business cycles are the most visible elements of the intrinsic dynamics of capitalism, including a fast inventory cycle, a medium term fixed capital and possibly longer structures cycles. Inventory cycles are in the order of three to five years. Equipment cycles range around seven to eleven years. Long waves are forty-five to sixty years in duration (Shaikh, 2016). Inventory cycles are linked to balance between demand and supply, while capital equipment fluctuations are traced back to the balance between capacity and actual output. Long-term growth trends feature periods of rapid economic growth in expansions or booms followed by periods of relative stagnation or decline bleeding into economic contractions, busts or recessions. Long waves patterns are found in the price of commodities expressed in gold inspired by Kondratieff.

Business cycles are measured by economic output but also related to the interest rate. Deregulation of financial activities first appears as success. But slowdowns of real wages relative to productivity and a drop in interest rates and in relative real wages boost the net rate of profit. In the build-up to every general economic crisis, the price of gold shoots up relative to the price of other commodities. Booms are greatly enhanced by a sharp drop in interest rates, which raises the net rate of return on capital. Falling interest rates lubricate the spread of capital across the globe, promote a rise in consumer debt and fuel international bubbles in finance and real estate. With falling interest rates and credit being made easier, consumers and other spending continue to rise, buoyed on a rising tide of debt. Crashes occur if people realize that price expectations are not sustainable anyone. A recession is a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in declining real GDP, real income, employment, and industrial production. Economic depression is characterized by high unemployment and falling prices. Long-run economic patterns in advanced capitalist countries include persistent growth in output, productivity, profits and employment, all taking place in-andthrough recurrent cycles and periodic depressions (Shaikh, 2016).

Classical economics either denies the existence of business cycles as – for instance in Say's law there is no over- or underconsumption – or attributes fluctuations to external shocks such as war, hunger, pandemics and natural disasters. But the inherent character of fluctuations in the economy is a rather heterodox view outlined in Marx (1867/1995), Keynes (1936/2003) and the Austrian economics (Schumpeter, 1949) followed by Soros (2003), Brenner (2002, 2006) and Shaikh (2016). The relatively unpredictability of the changing seasons of booms and busts calls for a thorough investigation of business cycles embracing heterodox interdisciplinary viewpoints.

Contemporary studies of business cycles and long term waves capture the relation of export, investment and raw materials on economic ups and downs alongside the role of institutions (Armstrong et al., 1991; Brenner, 2002, 2006). Remittances and other credit inter-governmental transfers are vividly described in the literature by historical snapshots. Historical examples of exchange rates pegged to the dollar and the dollar being convertible to gold as well as devaluation

strategies are discussed (Armstrong et al., 1991). The Phillips curve invers relation of unemployment and interest rate is unraveled as concomitant of economic waves. Remittances and credit expansions' influence on prices and industrial production are discussed (Armstrong et al., 1991). Shortages and production drivers but also trade association formations were found to play a vital role in shaping the economic climate. In addition, the role of political ideologies determining budget discipline and trade correlates as well as social unionization is thematized (Armstrong et al., 1991). Societal influences – such as working class strikes and tax evasion – play a role in booms or busts (Armstrong et al., 1991). Throughout history – vividly underlined by historical cases of France, Germany and Italy – we can see that booms and busts – through social unrest potential – are also connected and influenced by political ideologies. Central banks and their monetary reform potential are obvious determinants of economic stability. Governmental price control and money circulation means set the tact on the economic status of a nation. The delicate role of war and peacetime recovery is discussed in the economic cycle literature (Armstrong et al., 1991).

By the early 1980s Real Business Cycle Theory (RBCT) developed by retaining the notion of rational expectations and continuous market clearing and adding random productivity shocks to generate aggregate fluctuations that mimicked business cycles. Real business cycle theories attribute technological shocks or political cycles as causes of business fluctuations. The recurrence and turbulent regulation arise quite naturally. Perturbations include "socially influenced relations..., the salutary impact of policy and institutions...; the turbulent equalization of rates of return across industrial sectors; and the structural determination of industrial relative prices (Shaikh, 2016, p. 8). Underlying of these long-term patterns is the notion of turbulent regulation, in which balance is only achieved by recurrent over- and undershooting with intrinsic nonlinearities in the process (Shaikh, 2016, p. 8). Endemic turbulent growth is found in US industrial production, real investment and real GNP per capita over periods of 150 years (Shaikh, 2016). There is a recurrence of fluctuations of successive episodes of booms and busts, of overshooting and undershooting, in never ending sequences. Productivity growth is an essential measure of technical change and its steady long-term rise speaks to the fundamental role of technological progress in capitalist development. Technical change is an imperative for capitalist firms, rooted in the very nature of profit-driven competition.

Digging deeper into the causes of business cycles reveals phenomena like overaccumulation and overheating (Armstrong et al., 1991). Overheated growth appears when capitalism generates a higher rate of accumulation than sustainable and investors realize that the expectations do not represent the fundamental price. Imbalances between accumulation and labor supply lead to labor shortages and excess demand for labor, which then lead to a scrapping of old equipment and rise of real wages (Armstrong et al., 1991). Accumulation first drives demand for labor but eventually causes a profit squeeze through the scrapping of old plants to speed transfer of workers to new means of production. Accumulation thereby tends to decline gently to a sustainable rate. While overaccumulation was unraveled as one of the underlying future crises onset indicators, in all this the nature of communication behind accumulation leading to a cooling of economic productivity remains unknown. The underlying enthusiasm that first drives profitability and confidence is also not depicted in qualitative terms.

In the case of overheating, prices and wages rise and tighten the labor market, which drives up wages. From there productivity declines and economic prospects become gloom and despondent (Armstrong et al., 1991). Strikes emerge that effect differentials, real wages and profit margins of income (Soskice, 1978). Currencies need to be devaluated in order to remain competitive regarding exports.

In real competition, markets clear but do not rest in clearance. Money is endogenous and nonneutral, aggregate demand and supply are rooted in profitability. Whenever there is a bank run, credit money is devalued in relation to paper money and precious metals. In the worst of circumstances, bank accounts turn out to be mere unfulfilled promises, and a part of credit money evaporates. Similarly, when there is serious doubt about a nation's economic health, its currency can be devalued relative to other national currencies, as well as to gold, the currency of last resort for the international system.

In all this available knowledge, the role of information for building business confidence but also to instigate economic fluctuations is unknown. How market communication about economic prospects may accelerate market performance is yet an open research question. Information about how expectations build up in the social compound and how a disconnect of expectations from fundamental values leads to confidence losses and hence capital flight in social networks is While economic stabilization policy using fiscal and monetary policy as well as missing. governmental automatic stabilization appear to mitigate the downsides of cycles or at least dampened the worst excesses of business cycles, little is known about the psychological underpinnings of the global mass psychology of crises and how to use communication to alleviate economic fallouts. While adaptive decision making theory has entered the economics literature, we still have no concrete information on how market expectations' backtesting-based reevaluation guides individual's choices that amalgamate to the overall economic performance (von Weizsäcker, 2010). Capturing the collective soul of crises, however, could help deriving communication strategies to further counterbalance mass movements as a depression prevention to solve disastrous outcomes of collective moods, herd and swarm behavior on economic markets echoing in the real economy. Key in understanding what drives economic fluctuations based on price expectations appear historical foundations of the theory of price.

Theory of Practice

The classics: According to Adam Smith, free interaction creates an orderly pattern guided by the invisible hand. Competition creates order out of activities and makes market wages gravitate around the natural wage. Profit rates gravitate around natural prices. Market prices are invisible centers of gravity and profit rates equalize due to the mobility of captial. High profit rates attract capital that brings the market prices down. Low profit rates will slow down supply in the movements of market prices. Prices determine the money wages and markets create the aggregate of prices. Prices move around centers of gravity among each other. The natural course is that actual market prices are above or below the natural price. The central price is the price to which all prices gravitate continually. According to Smith, the price of production has a natural profit, which is only achieved through gravitation. The natural profit is the profit one deserves added to costs.

Exchange takes place in disorder (Shaikh, 2016), in which prices guide through the monetary expression of a commodity's quantitative worth. Market prices gravitate in a turbulent manner around prices of production. Profit is an objective measure subject to constant scrutiny by the firm's managers, the stock market, the banks and the public in general. The uniform price is assumed to be supremely responsive to the market demand and supply. According to Kalecki's (1954/2009) Theory of Price, the price of individual firms depends on the relative size of firms based on the union power of employees, industry prices and average costs and monopoly power that determine the mark-up structure.

Methodological weaknesses of the price system reveal that there is a complexity of the calculation of prices via an input-output system that prevents a simple description of the relation of a commodity's prices to any other single parameter (Bienenfeld, 1988; Sraffa, 1960). Divergences in price expectations mainly stem from the organic composition of capital in each industry and relative capital intensity in layers of goods and the vertically integrated value

composition of capital as an intricate network of relations between the rate of profit and prices in the whole economic system (Bienenfeld, 1988).

Ricardo (1817/1951) argued that changes in relative prices are dominated by changes in relative vertically integrated labor coefficients. Marx (1867/1995) outlined that relative prices move directly with changes in vertically integrated labor coefficients and inversely with changes in relative vertically integrated labor productivity, which even holds for prices of production (Chilcote, 1997). Turnover and fixed capital are considered as secondary influences on relative prices (Chilcote, 1997). Prices appear to be dependent on the prices of production, which depend on fixed capital, turnover, wages, return on capital, capacity utilization and capital stock (Cockshott & Cottrell, 1993; Leontief, 1997). Evaluating the sensitivity of different prices of production estimators at predicting market prices can provide insight into the importance of including variables such as depreciation, fixed capital, turnover, and capacity utilization into the estimation of prices of production (Chilcote, 1997). Direct prices represent a set of monetary prices that are proportional to labor values. The sum of direct prices equals the sum of market prices. The inclusion of depreciation lowers the dispersion for direct prices and the inclusion of fixed capital lowers the dispersion for prices of production.

Marx

In classical economics thoughts, value has been discussed as for underlying prices with emphasis on the social importance of values and prices. The classical economists argue that market phenomena are regulated by underlying forces, which give rise to centers of gravity. Ricardo (1817/1951) acknowledges that a theory of relative prices cannot be established as both the technical conditions of production and the distribution of income jointly affect relative prices when the turnover of capitals differs across industries (Chilcote, 1997).

Karl Marx is the first to refer to price-value deviations (Shaikh, 1984). Marx links values to market prices by introducing the intervening category of prices of production, where prices are compatible with equal rates of profit across industries. Marx and Keynes tie prices to periodic recurrence of boom and bust in Capital and the General Theory (Sherman, 1967). In Marx, every object has a price and there is an agency involved just in setting the price. Marx identified values, or direct prices, to directly regulate both prices of production and market prices; and prices of production to serve as the center for gravity regulating market prices (Chilcote, 1997). While Marx argues that prices of production are dominated and governed by direct prices, he also expects market prices to gravitate more closely around prices of production than direct prices (Chilcote, 1997). In Marx prices are derived magnitudes from values through the redistribution of surplus value according to the logic of a capitalist system (Pasinetti, 1977). In this context, the labor theory of value is criticized for uniform rates of profit and prices of production assumptions. Deviations of market prices from prices of production mostly reflect imbalances in supply and demand conditions. According to Marx, there are expectations of outcomes in financial markets in the interest rate. Expectations and outcomes differ.

In contrast to the modern classical emphasis on distribution, Smith, Ricardo and Marx focus on relative prices. Technology appears to be the central force governing the motion of relative prices through time. Movements of relative vertically integrated labor coefficients are essential in determining prices. Smith, Ricardo, and Marx stress differential rates of growth in vertically integrated labor coefficients as the central force governing the movement in relative prices through time (Chilcote, 1997). Movements in relative vertically integrated labor coefficients are positively related to movements in relative prices (Chilcote, 1997). In Marx the price level is derived as the product of the price of commodities relative to gold, which is determined by the monetary regime. Competition between industries makes relative prices of individual commodities gravitate around prices of production and prices reflect equal profit rates. According to Sraffa (1960), the price of commodities is a joint result of the price of the commodity itself, the costs of the commodity production as responding to changes in the prices of those commodities, which enter into its production.

Marx is the first to analyze the capitalist's reactions to changes in the rate of profit in investment spending, which is seen as major cause of fluctuations in the total level of economic activity. Marx finds the causes of fluctuations in economic activity and periodic depressions inherent in a capitalist economy, which marks a definite departure from classical models, such as in Say's Law. In Marx theory of a simple barter economy, people produce goods either for use value of consuming these commodities or bartering the produced goods, which drives the economy. Money can be introduced as medium of exchange that facilitates the division of labor and trade. Capitalism drives the orientation of the economy from the production of use value to exchange value. The capitalist directs the production process in seeking to make profits.

For Marx, the economy is based on the production of commodities to be sold in markets, which are prone to crises. Business capital profitability has a tendency to fall, therefore unemployment rises, businesses fail and the remaining capital centralizes until profitability recovers. Marx believes that crises tend to become more severe and the market system approach will eventually fail. Marx's critique of capitalism is the intellectual offspring of business cycle theories formulized by Richard M. Goodwin (1967) and John Maynard Keynes (1936/2003).

According to Marx, all firms have market prices. Firms constanty try to reduce prices in order to get an edge over competitors. Competition is driven by competitors trying to sell cheaply to drive out competitors based on technological change lowering costs and cheapening commodities. In Marx competition tends to equalized wage rates and profit rates, around which market prices tend to gravitate, but remain different from natural prices of production (Shaikh, 2016). Marx emphasizes the "anarchic" character of gravitational fluctuations. He generalizes Ricardo's argument that only certain conditions of production regulate the market price by extending the notion from agriculture to all industry. Marx argues that competitive firms are active price-setters and aggressive cost-cutters. The creation of techniques with lower production costs generally requires greater investment in fixed capital per unit. Marx holds that capital must participate in profit rate equalization (Hunt & Lautzenheiser, 2001). Marx establishes that prices will be proportional to labor values when all money value added takes the form of labor income, selling prices are equalized by competition, and incomes per unit labor are equalized by the mobility of labor. In Goodwin's (1967) economic model, recession is caused by increased bargaining power of workers as a result of high employment in boom periods, pushing up wage shares of national income and suppressing profits, which will eventually lead to a breakdown in capital accumulation. For Marx crises are inherent to the system and the government can change the timing of economic crises.

Keynes

John Maynard Keynes, who saw capitalism as a system worth maintaining and susceptible to efficient regulation, opened the door of governments to control business cycles. Keynes became prominent in the 1930ies when studying high unemployment across the capitalist world. By capturing the economic fallout of the Great Depression, he presented a solution based on governmental and in particular central bank support of the economy. Keynes (1936/2003) broke apart with the orthodoxy of his time and the belief that real wage automatically moves to bring full employment. By observing the problems of the 1920ies and the end of the World War I period, Keynes saw persistent poverty in the capitalist world. Keynes blocked the claim that if unemployment caused the rate of interest to fall, then the net rate of return would expand and

make investment rise to bring back full employment. Keynes (1936/2003) finds a negative relationship between the rate of growth and unemployment during World War II, when there was a huge deficit in output and unemployment rising.

Keynes' business cycles reflect the possibility that the economy may reach short-run equilibria below or above full employment opening the stage for monetary and fiscal policy in smoothing fluctuations of business cycles. According to Keynes, fluctuations in aggregate demand cause the economy to come to short run equilibria above or below full employment. The implications are that endogenous causes of crises demand for governmental market regulation, for which Keynes advocates. Capitalism was meant to be ruled by governments to create viable employment levels thanks to governmental intervention to adjust for cyclical changes. The state should pump up the economy by intervening to lead to full employment. Keynes' policy foundation in the capitalist world would bring full employment to markets.

Keynes (1936/2003) broke with the traditions in the General Theory proposing that governments should play a role in the economy. Recessions caused by inadequate aggregate demand call for the government increasing aggregate demand to bring the economy back into equilibrium. The government can do so by either increasing money supply through expansionary monetary policy or by increasing government spending or cutting tax by expansionary fiscal policy in order to curb volatile long-term profit expectation fluctuations due to the "tides of irrational optimism and pessimism" (Keynes, in Shaikh, 2016, p. 33).

According to Keynes (1936/2003), firms do not set prices, competition does. Keynes also acknowledges that the actual price level of investments is the result of the sentiment of the public and the behavior of the banking system (Kurz, 2010). Prices depend on confidence in the future, whereby a collapse of confidence precipitates capital flight. The government can step in by reducing the interest rate to override the fall in confidence through investment stimulus.

Already Piero Sraffa (1960) introduced the difference between expected and actual prices. Later work talks about the rational and adapted preferences based on market expectation corrections (Shaikh, in speech). The expected becomes distinguished from the real behavior, whereby corrections portray falsification. In Keynes (1936/2003), the interest rate is determined by subjective preferences and the price of a commodity is based on costs. In Keynes prices and the interest rate are based on expectations. The lowest cost technology is the most competitive as capital intensity becomes higher. A higher profit margin makes firms more capital intensive. To make investment decision, corporate actors rely on animal spirits. If a lot of people are feeling good about the economy, there will be more investment and more growth. When people feel bad about economic outlooks, they stop investing and growth crashes. When the economy is doing well, banks lend more and create more money. When the economy slumps, banks lend less and create less money. This means the money supply automatically moves up and down based on how the economy is doing, which can be a problem for governments that want to control the economy by money supply. Neoclassical monetarists (such as Friedman, Phelps and Lucas) argue for prices not being determined by costs but by the quantity of money, which becomes leveraged as a means to guide the economy. In the classical theory of money, price wages and levels of production determine the long-run growth rate.

Whereas Keynes' (1936/2003) theory of the expected rate of profit is describing short run fluctuations, Hyman Minsky (1975/2008) proposes fluctuations in credit, interest rates and financial instability as underlying basis of long-run business cycles. Economic credit cycle theories featuring an expansion of credit (e.g., increase in private credit or debt as percentage of GDP) yields economic expansions, while the net contraction causes recessions, and if persistent, depressions. In the expansion, interest rates are low and companies borrow money to invest. As firms become excessively indebted, the investing stops, so do the bank credits and the economy enters recession. In the Keynesian tradition (1936/2003), deficit spending can pump up the economy again building on the Ricardian equivalence.

Keynes (1936/2003) rests his analysis of aggregate consumption on underlying subjective and objective factors that, in addition to person income, influence individual savings behavior. Subjective factors include the desire to provide for future consumption and contingencies, to use passive and speculative investment to expand future income, to amass wealth, and for some, even to enjoy miserliness. Objective factors include windfall gains or losses, taxation, price controls, expectations and changes in the interest rate (Shaikh, 2016). Institutional and organizational influences shape and channel all such factors that build expectations about markets. Already Keynes described expectations as volatile moods swing as output moves up and down. Keynes is also the first one to attribute monetary authorities an active role in the determination of distributive variables based on historical and institutional factors affecting the division of surplus value produced (Cozzi & Marchionatti, 2001). Keynes advocacy for governmental influence in shaping the economy is in contrast to the Austrian School of Economics.

Austrian School

In the mid-20th century, Joseph Schumpeter (1949) describes four cycle stages of (1) expansion as the increase in production and prices, low interest-rates, (2) crisis featuring stock exchange crashes and multiple bankruptcies, (3) recession such as drops in prices and output, high interest-rates, (4) recovery of stocks because of the fall in prices and incomes, increase in productivity, consumer confidence, aggregate demand and prices.

Austrian business cycle theory attributes the excessive issuance of credit by banks as driver of crises in banking systems if central bank monetary policy sets interest rates too low. The resulting expansion of money supply causes a boom, in which resources are misallocated because of artificially low interest rates. This unsustainable boom is followed by a bust, in which malinvestments are liquidated below their original costs and the money supply contracts.

Schumpeter (1949) adopts Walras' model of price-taking firms and maximizing agents but adds the constant creation of innovations adding external perturbations (Kauder, 1965). Schumpeter falls short on explaining where innovations come from and why some get picked up and other vanish. The Austrian emphasis on competition as a process that bids away excess profits is close to the core of Shaikh's (2016) real competition, except for its explicit assumption of rapid profit rate equalization and the lack of a distinction between regulating and non-regulating capitals. Austrian economics also share the neoclassical vision that firms are efficient servants of consumers and that union activity and government intervention are unwarranted intrusions into market processes.

George Soros` Theory of Reflexivity

In the classic tradition, the expected rate of profit is tied to the actual rate of profit similar to Soros' Theory of Reflexivity, which connects expected, actual and fundamental values of economic market systems (Shaikh, 2010). One of the most elaborate real-world relevant explanations of business cycles is outlined in George Soros' Theory of Reflexivity. As anticipations of an expected future value in an uncertain world, expectations are a central part of value calculations (Soros, 2000). Expectations are a thought or a belief about anticipated gains or losses, whereby the idea of time is fundamental to rational expectations and outcomes. The subjective expected utility is derived from the value of an outcome and the probability that it will occur. Expectation calculus is used in many domains, including health, wealth, survival and educational outcomes, but likely most prominent in financial realms. While we know that consumers with more optimistic stock market expectations tend to more likely hold riskier assets and acquire stocks in the near future; we still are in need to explore what information and factors drive financial market expectations.

According to Soros (2003), expectations determine prices. To every good and service there is a (1) fundamental value f, what a good is actually worth, when considering the direct market price and price of production and (2) expectations e. When information in markets builds up expectations, expectations start overpromising beyond the fundamental value, therefore e > f. The difference between the expected price and the actual price, Shaikh (2010, 2013) formalizes as the *degree of bullishness* in the market. People believe a good or service is worth more than its actual fundamental value f. The difference between the actual price and the fundamental, Shaikh (2010; 2013) refers to as the degree of *overvaluation of the market*.

Shaikh (2010) concludes that the actual price will rise if the expected price is greater than the actual price, so the actual price changes in response to the degree of bullishness in the market. Fundamentals are affected by market prices, so the fundamental price rises if the actual price is above the fundamental price, so the fundamental prices are driven by the degree of overvaluation in the market (Shaikh, 2010). Key economic variables are thus historically contingent and path dependent (Shaikh, 2010, 2013).

Expectations thereby become central to the formation of economic cycles. In an economic cycle, during booms and busts expectations rise or fall, which are triggering cyclical ups and downs. In the beginning, expectations are gravitated closely to the fundamental value of a good *f*, outlined in blue in Graph 1.

The nature of the collective soul of booms and busts is outline in Graph 1.



Graph 1: The collective soul of booms and busts

If enough market actors think that the stock will go up, expectations inflate the price. If these anticipations activate enough others to act upon, the expectations rise high above the fundamentals. As information about expectations reaches market actors, it builds up inflated surreal exceptional capital. Positive information serves as self-fulfilling prophecy and social reinforcement lets market investors dream about future anticipated growth. Signalling and positive reputation thereby begin an announced overinflated expectational price, which gets perpetuated in the wake of herd mentality and swarm behavior. The number of buyers rises in expectation of high profits and so the further the expectations get inflated over the real price. Yet as time proceeds and information gets shared, the expectations e, as outlined in purple in Graph 1, build up to the point where the spread between e and f gets so large that leaders notice this not being sustainable anymore and e having become too detached from the actual true value f. Effects of moods and expectations change if people realize that assumptions are not sustainable. The boom then reaches its peak and first market leaders start selling the good or service off. If e gets largely inflated and majorly detached from f, first market experts start realizing the unsustainable inflation of a good's or services' expected value.

product and service, hence trying to sell it. If the amount of sale bids reaches a certain threshold, by the law of supply and demand, the price will eventually decline and then, if animal spirits bleed into a run and more investors follow, eventually inflated prices will collapse. The more the number of people rises that believe that the system is not sustainable as the further one gets away from the real price, the more information gets shared, the more people realize that the bubble bursts. With them setting a vibe in the market, the collective soul of booms and busts naturally triggers herd and swarm behavior of followers, who then start selling the good or service. The price expectations collapse and re-adjust to the actual fundamental value. The economy has then gone from boom (in which *e* starts building up over *f*), peak condition (in which there is the largest spread between *e* over *f*) to bust (in which *e* gets re-adjusted and gravitated back towards *f*).

Expectations play a vital role in this theory. In the collective soul of booms and busts framework, individual decisions are made by subjective considerations and finance is a process of social outcomes. Confidence in the overall economy is built on overall expectations that can cause higher outputs. Fundamentals are affected by the historically contingent path, which makes the equilibrating process turbulent and path dependent (Shaikh, 2010). Expectations affect actual prices, actual prices affect fundamentals and expectations influence the behavior of actual prices and fundamental prices in a turbulent process, in which actual prices oscillate turbulently around the gravitational values (Shaikh, 2010). Soros' Theory of Reflexivity addresses the process of turbulent meta-physical equalization. Soros' notion of reflexivity holds that if investors think that the price of stock goes up influences others to think the same. Therefore, expectations will be that there is a rising of price that will activate others to follow. If enough people act on it, that alone can drive up the price above its fundamentals. When prices are above fundamentals, people start to realize that this is not sustainable and pull out. Single investors reap a benefit by leaving markets earlier when representative agents stay in markets. As people are betting against each other, some become winners and others losers in financial bubbles. Those lose who expect the process to be long term eternally growing.

A rise in business confidence based on animal spirits raises investment and causes the positive multiplier to rule. Yet unmet expectations are triggering the market to go under. Prices get out of balance not being justified by fundamentals during the building of the bubble. Soros explains how expectations can create an output but also a collapse if they are not validated anytime and are too removed from the fundamentals below them. Shaikh's (2010, 2013) formalizes Soros' reflexivity mathematically. Similar feedback systems can be found in Goodwin's (1967) predator-prey model.

Soros' (2003) innovative contribution is that expectations can change the fundamentals in creating bubbles and herd mentality and swarm behavior of investors to offset a path to cyclicality (Shaikh, 2013). The actual responds to fundamentals can higher gravity. If the gap between the fundamentals and the actual orbit will grow larger, there will be a reverse effect on the fundamentals. Profit rates may change expectations. If a sufficient number of people expects the price to go up, the price will really go up and create a boom in which expectations are dominating price movements. If the market price is above sustainable levels, more and more people will realize that this is not sustainable, which will lead to short deviations from the fundamentals in contrast to long deviations. Animal spirits will lead markets in disproportion with reality. Expectations' collapse causes a crunch in output and falling employment rates. Depressions cause real wages to rise, moves in direction of free markets making things worse through rising inequality (Shaikh, 1979, 1980). Problem of crisis stagnate action and send countermovement signals to other industries and markets. In the age of globalization, a crunch in one country leaps over to other territories. Especially if either the EU or US is stuck in depression, the whole world joins. Over the decades, George Soros offered many example applications of his Reflexivity

Theory explaining occurrences in world finance, history, and political sciences, which become a function of cognitive and system manipulative tendencies (Kwong, 2008).

Over the business cycle, one can by now predict certain behavior in the wake of dynamic cyclical fluctuation shocks of continuous equilibrating processes. Yet price expectations and economic cycles – to this day – are hard to predict. The actual temporal flow of these recurrent patterns is hard to measure quantitatively. One knows that it will happen but can only guess when it will occur. While there is a pattern of the economy, problems of agents remain to distinguish between noise and information in shocks and change. In the ever-changing nature of capitalism, having evolved from industrialization to post-World War booms to globalization and the new media revolution, the cycles have become more fuzzy and unpredictable.

Professor Anwar Shaikh is the first to formalize reflexivity, with the expected and actual variables gravitating around a possibly moving fundamental value (Shaikh, 2013). As expectations and aspirations of thinking market actors shape their actions based on novel information, they influence the system of which they are a part of (Shaikh, 2013). The actions of participants based on the various views will generally change the reality in which they operate (Shaikh, 2013). So variables only fluctuate around the levels determined by fundamentals over time, which Soros (2013) refers to as equilibration of a turbulent and reflexive gravitational process. Path dependence is a natural consequence of reflexivity (Shaikh, 2013). Market events become a form of history (Soros, 2009). Market information about the past therefore not only reports past performance, it may shape future performance and affect the fundamentals and hence future market outcomes (Shaikh, 2013; Soros, 1994).

Already Dobb (1929) commented on the tendency of modern economics to downplay the psychological and qualitative aspects of value creation, yet information on attributed value based on expectations is to this day mainly missing in neoclassical economics (Sen, 2003). While we have insights about the mathematical calculus of the expected value of a discrete random variable being the probability-weighted average of all possible values, mainly estimated as probabilities by frequencies; we lack an understanding of the economic correlates in the wake of sociopsychological phenomena regarding outperformed or burst expectations - leading to feelings of hybris or disappointment. While behavioral economics has established the notion of loss aversion and losses looming larger than gains (Kahneman & Tversky, 1979); we have no stringent framework on reference point dependence regarding backtesting-based expectations under- or outperformance. To this day we have no sound information how expectations are built up in markets and how the over- or underperformance creates social norms around markets that may inflate to herd behavior or swarm phenomena. In general, we can assert the higher a person's expectation and the lower the performance, the higher the frustration and trigger of cessation of activity and guitting behavior – yet what this means for financial market decisions, we have no concrete information about. As people develop favorable and unfavorable expectations that guide their choices and behavior on a daily basis, we need to unravel how these expectations influence economic market actors' decisions bleeding into the collective soul of booms and busts.

Expectations play a vital role in economic calculus. Since firms have to produce for the future, any production is based on expectations on future outcomes (Keynes, 1936/2003). Expectations are formed in a dynamic process. Anticipations are skewed and dependent on other people. Expectations influence what happens in the process, as people do not correctly anticipate what will be going on. Soros' (2003) financial market theory holds that some people will belief that market prices will rise forever. They do not know that turn will happen and other do not have a clue when turn will happen. If a sufficient number of people beliefs that the system is not sustainable, the bubble bursts. According to Friedman, surprises in markets matter more than expectations. Every deviation besides surprises are temporary. Information on the expectation performances underlying economic ups and downs could help bridge between rational behavior and societal collective action (Soskice, 1978). Unraveling the impact of expectations and the role

of information in building and destroying expectations in financial markets could aid serving the greater goal of fostering sustainable finance – a mandate primarily addressed by central banks and governing financial authorities. Central Banks and market prospects play a crucial role in shaping expectations and correcting over- or undershooting in market predictions.

Central banks' information and crowd control

In society, language is used to embody theories of reality. Discourse is inseparable from social and economic factors; yet more needs to be known about the concrete effect of information communicated in markets echoing in economic fundamentals, such as price and equilibrium formation (Fowler, Hodge, Kress & Trew, 1979). Different social strata and groups but also different institutions and media have different varieties of language available to them. Linguistic variations reflect and actively express the structured social differences, which give rise to inequality and economic dispersion (Orwell, 1949).

Access to information about markets is distributed unequally within society. Certain market actors have faster and more access to aggregate information about prices than others (Benabou & Laroque, 1992). Access to a different level of aggregate data allows for competitive market advantages. Governments and in particular central banks hold more aggregate information than regular market actors. It is on these actors to utilize information appropriately and their technocrats are obliged to grant the populace access to information. Banking and finance information is usually communicated in the realm of central banks.

Central Banks or Reserve Banks are monetary authority institutions that manage a state's currency, money supply and interest rates. Started from ancient Egypt forerunners, the 1609 founded Amsterdam Wisselbank accounts for the first central bank, which soon led to similar banking institutions in central Europe – featuring notable banks in Hamburg, Venice, Nuremberg, Sweden and England – during the 17th century. The 20th century featured an explosion of central banks around the world, for instance, in 1913 the U.S. Congress enacted the U.S. Federal Reserve, followed by central banks in Australia, Peru, Colombia, Mexico, Chile, Brazil, Canada, India, New Zealand and African and Asian countries. The European Central Bank hosted in Frankfurt Germany was established in 1998.

Historically set up to ensure the efficiency of markets and safeguard monetary stability and international trade, central banks oversee the commercial banking system of respective countries or overseen territories. Central banks primarily implement monetary policies by setting the interest rate; manage inflation and exchange rates and controlling a nation's money supply. The main monetary policy instruments available to central banks are open market operations to influence the money supply of an economy, bank reserve requirement, interest rate policy, relending and re-discount via the repurchase market, and credit policy often coordinated with trade policy. As lenders of last resort and manager of foreign exchange, central banks are holders of gold reserves. In cooperation with the legislative and executives bodies of a nation, central banks regulate and supervise the banking industry. Holding the monopoly on controlling the monetary base of a country, central banks are printing and issuing their national currency due to their exclusive Right of Issuance and therefore control the fiat monetary supply. Central banks foster stability of interest rates, financial markets and the foreign exchange. Central banks also aim to control economic growth via the interest rate. Especially during economic downturns lowering the interest rate incentivizes industry borrowing more money to invest to encourage economic growth. Raising the interest rate is often used as means to curb inflated economic growth as a counter-cyclical move to keep the economy from overheating and to avoid market bubbles.

Central banks also have supervisory and regulatory powers to ensure the solvency of member institutions, prevent bank runs and reckless or fraudulent behavior by member banks. Some

central banks' mandates include unemployment control. In most developed nations, central banks are institutionally designed to be independent in the operational management from political interference, yet accountability and control by the executive and legislative bodies are usually enacted. As holders of aggregate information about a country's economic status and outlook, central banks play a vital role in the world's economy as for issuing information about markets, primarily focused on economic growth and economic outlook prospect.

People have a right to know in access to information mandates and regulatory discretion that require central banks to publish data about economic forecasts. In doing this, central banks have an elevated role to set incentives and constraints but also build and destroy expectations in market performance in backtesting reporting. In the operationalization of market communication, central bank standards differ. While some banks publish re-adjustments of their forecasting errors, others refrain to do so. These differences even persist in the harmonized and standardized Euro-zone territory, which – to this day – features national central banks besides the European Central Bank.

While there are writings on the role of information to control behavior, little is known about the unintended consequences of information in expectation building and hence bubble creation. Economic uncertainties may partially stem from fluctuations of discourse on outcomes and prospects. Hardly any information exists on the role of central banks in informing citizens about the status of the economy. The paper therefore sets out to study what linguistic structures exist to create realities about prices and markets. Information will be portrayed as a means to regulate the ideas and behaviors of social masses. Contrary to other linguistic analysis of communication, this article does not aim at unveiling the misuse of language by institutional agents to control society but rather seeks to shed light on unknown dangers of communication echoing in economic correlates. Information may implicitly aid in building expectations and hence bubbles to add to economic fluctuations. Central bank information is thereby revealed as powerful mode of language and thought. Central bank announcements and economic forecasting mass communication is viewed as a tool to unconsciously drive the economic engine but may also cause disturbance in financial markets. Unprecedentedly described implicit economic fluctuations built by central bank information will become the focus of attention. The media will be portrayed as means to control society's relationship to material reality (Fowler et al., 1979).

What is the driver of equilibrium and what communication causes fluctuations in the economy? The article will show how central bank communication about economic prospects is related to economic outcomes. Contemporary writings address political, institutional and social processes as makers of crises, yet the role of information for the building of economic moods is mainly neglected. Economic prospect information released by central banks is argued to shape individual's expectations and social processes that influence behavior and group outcomes in economic markets bleeding into the collective soul of booms and busts. This article thereby focuses on the unintended and unforeseeable consequences of information on economic forecast corrections in their relation to actual outcomes following the greater goal to gain more certainty about the economic situation in order to breed financial market stability and economic prosperity.

6. Research question

With Shaikh's Capitalism (2016, p. XXXV) granting "a genealogy of the tenets of classical economics; and the repair, refinement, and application of these to modern capitalism," a gate has been opened to allow a trenchant heterodox analysis of information sharing in financial markets. Contrary to neoclassic economics "supreme optimality of the market" argument of the "ever-perfect invisible hand" and "representative agents;" this paper targets at a heterodox economic perspective of business cycles innovatively also shedding light on imbalances based on over- and undershooting expectation corrections (Shaikh, 2016, p. 4). While neoclassical economics begins from a perfectionist base and introduces imperfections as appropriate modifications to the underlying theory, the following paper will highlight the role of information for real competition

in order to argue for a democratization of information flows. Introducing emergent risk mitigation strategies within globalized economic markets may thus help avert future socioeconomic crises and imbue public trust in open market economies through improved economic market stability and societal welfare stemming from universal access to equally shared benefits of global economies.

The following part focuses on representing connections of expectation corrections in backtesting and actual market performance. The paper thereby targets at opening the black box of deliberately future-oriented market prospect reporting and the stylized linear time scale in neoclassical economics, which will be challenged to be disrupted by seasonal information shocks. The paper will show the problem with the neoclassical assumption of perfect information and feature inconsistent representations of information by shedding light on imperfections that produce certain types of outcomes in consumption, equilibrium and price. The paper will thereby outline how the market responds to central bank market communications and how market corrections are related to actual market performance in the near future but also backlash to cyclical tendencies in the more distant future. Attention will be paid to temporal heterogeneity, the information blast moment differing from a linear time scale. On a wider scale, the article will paint a picture of markets behaving in line with corrected market predictions. Acknowledging that agents make choices under social constraints, a dependence of past performance on current actions will be unraveled. Studying information on market prospects will allow constructing a framework of socially structured market fundamentals and derive conclusions how expectation outcomes echo in economics. The importance of historical conditions but also social and cultural Thereby an opening of time consistent predictions will allow structures will be outlined. contributing to non-linear models of predictions and market outcomes paying tribute to the idea of turbulent real economies. Business cycles will be shown to obey some kind of natural complexity, they are whimsical based on socio-historic and political trends as well as follow the occasional madness of actual human behavior. The paper thereby embraces diversity in granting heterodox perspectives of our contemporary knowledge on the formation of business cycles. Overall, the article will connect micro-economic information flows with macro-economic fundamentals and address the emergent properties of heterogeneous agents through the wide varieties of constructions of expectations.

Departing from classical economics addressing exogenous causes for economic cycles, the following paper will focus on unraveling endogenous – thus system-inherent – business cycle drivers. As an alternative to this debate, the following article innovatively paints a novel picture of the mass psychological underpinnings of business cycles in order to recommend certain communication strategies counterweighting the building of disastrous financial market mass movements. As business cycles are a collective phenomenon, group interactions' potential contribution towards business cycles will innovatively be outlined and the role of information flows among groups unraveled. Studying the role of information in communicating crises appears logic when considering that business cycles are fluctuations found in the aggregate economic activity of nations that organize their work. Information will also set the tact on if and how prices become an abstraction determined by ideology shifts.

The following reflections on the inside of economic ups and downs will introduce psychological elements into economic debates. While economics seem to give clear guidelines on how economic correlates of competition, wage pressure, deregulation and repression of real wage growth play out in economic terms, the irrational exuberance that leads to purchasing and investment decisions in the overestimation of future profit perspectives cannot be explained by orthodox economics. Booms and busts will be portrayed as phenomena that are built by the collective decision making within society as market actors anticipate and panic together. The following empirical part will shed light on communication of economic prospects and information representation of estimates of central banks and capture a novel relation of past market predictions correction with near and distant future market performance.

Method

In order to test for the relation of information and expectations on markets shaping prices, information about market projections of a central European central bank was retrieved online.¹ This central bank bi-annually publishes the real Gross National Product (GNP) in percentage changes to the previous year (GNP, *Bruttoinlandsprodukt Veränderung zum Vorjahr in % (real)* in the original) of a central European country for four years – 1 year retroactively, the year in which the report is issued and the prospect of the following year and the 2-year prospect. The report includes the correction of the former expectation of the GNP in the former calendar year t - 1, the same calendar year t as the report is issued and the prospect about the future year t + 1 and two years t - 2 in advance. The information contains the backtesting previous year market performance correction t - 1, the current year t, the following year t + 1 and the prospect of market performance in two years t + 2. Information on such market performance and prospects is available since 1998 until now at the central bank's homepage accessible for anyone with access to the internet. These reports get published twice a year, in June and in December of each year.

The reports offer information about expectations and market prediction corrections. Growth forecast errors have leveraged into a macroeconomic tool to draw inferences about forecasting model weaknesses in order to make better predictions about future market performances (Blanchard & Leigh, 2013). In the wider sense, forecasting error corrections offer an opportunity to study whether markets move in the direction of announced market corrections.

Results

In the data analysis, at first the market corrections for each data point representing the real Gross National Product (GNP) of a central European national economy in percentage changes to the previous year were calculated. The market corrections were then plotted in relation to the GNP market performance after the correction in the next half year t + 1 and in the next year t + 2 period. Graph 2 holds the GNP market performance prediction corrections c and GNP market performance *GDP* after correction derived from a central European central bank for the years 1999 to 2017.

¹ https://www.oenb.at/Geldpolitik/Konjunktur/prognosen-fuer-oesterreich/gesamtwirtschaftliche-prognose.html



J. M. Puaschunder / Finance, Risk and Accounting Perspectives Proceedings (2019) 224-252

Graph 2: GNP market performance prediction corrections and GNP market performance after correction (years 1999-2018)

The data set-up accounted for the corrections c being enacted retroactively in backtesting – so after the time had passed, hence for example a correction for 1999 as t - 1 that was released in t is the year 2000 – but the market performance trend being real time, so 2000 market performance trends in 2000 as t + 1. We can thus infer the relation of a correction for the past performance and the future market performance in the half year (t + 1) and the year (t + 2) thereafter. Market corrections are represented by variations in estimations over time. Relations between variables are investigated by correlation studies (Shaikh, 1984).

A highly significant positive correlation between the overall correction in *t* and subsequent market performance in the next bi-annual period t + 1 ($r_{Pearson(130)}=.417$, p<.000) was found for the years 1999 to 2017. A highly significant negative correlation ($r_{Pearson(126)}=.303$, p<.001) between the overall correction in *t* and subsequent market performance in the period t + 2 ranging from half a year after to a year after the initial correction was found for the years 1999 to 2017.

In order to determine whether past period t - 1 corrections or current state t corrections or future prediction corrections t + 1 and t + 2 are associated with a certain market trend, correlations were calculated for the association between (1) t - 1 corrections and t + 1 actual market performance as well as (2) t - 1 corrections and t + 2 actual market performance (3) t corrections and t + 1 actual market performance. A significant positive correlation between the correction in t - 1 and subsequent market performance in the next bi-annual period t + 1 ($r_{Pearson(37)}=.353$, p<.032) was found for the years 1999 to 2017. A highly significant negative correlation between corrections in t and subsequent market performance in the next period starting after half a year up to a year after the announcement t + 2 ($r_{Pearson(37)}=.633$, p<.00) was found for the years 1999 to 2017.

Overall, retroactive backtesting market corrections are highly significantly positively correlated with the following market performance. Corrections of past periods are positively correlated with current performances in the near future but negatively correlated with future performances in the more distant future. Current period announcements are associated with positive trends in the subsequent period up to half a year after the announcement and negative trends in the period following more than half a year later to a year later. Past market performance shapes the future prospect and expectations about markets.

In order to consolidate the findings and determine the interaction of variables and length of the effect, two regressions were calculated based on the following equation 1 and 2:

$$\Delta Y_{i,t+1} = \alpha + \beta_1 F_{i,t-1} + \beta_2 F_{i,t} + \varepsilon_{i,t},$$

(Equation 1)

whereby *Y* represents *GDP* and $Y_{i,t:t+1}$ GDP growth in percent from the current period in the following half year periods t + 1, representing a half year interval between *t* and t + 1.

Growth forecast corrections are calculated by the difference between actual real GDP percentage changes during one period *t*, based on the latest data, minus the forecast prepared for the period under scrutiny presented in the previous period t - 1. $F_{i,t}$ is the forecast error correction in GDP growth for the given periods of *t*. The associated forecast error correction is $F_{i,t:t} = \{Y_{i,t} | \Omega_t\}$, where $Y_{i,t}$ denotes the market performance in GDP conditional on Ω_t , the information set available in the current *t* period.

Positives values of $F_{i,t}$ represent an unpredicted outperformance of the market in GDP terms in the next half year. Negative values of $F_{i,t}$ an unexpected underperformance correction of the

market predictions in GDP terms. $F_{i,t:t}$ denotes the forecast error change in GDP growth for the periods *t*. The standard error term of the regression is $\varepsilon_{i,t}$.

Under the null hypothesis H_0 that forecast expectation corrections have no impact on or relation with actual future market performance, the β -coefficient would be zero.

Data was retrieved from a European central bank's bi-annual GDP growth prospects. A regression to describe the relation of forecast error correction and actual market performance within the next half year (t + 1) over 35 data points reveals an overall fit with R square .147 and adjusted R square .096 of the model. The regression coefficient β -value of .348 for explaining the market performance in the 1st period after the announcement of the correction is significant at the 5 percent one-sided *t*-testing level *t*=1.148, with a *p*-value of 0.039. A one percentage point change in forecasting corrections is associated with a .348 percentage point change in GDP growth in the same direction in the first half year following the announcement.

The cyclical nature of markets become apparent in a second regression considering the longterm market performance based on equation 2. The second regression was calculated based on the following equation 2:

$$\Delta Y_{i,t+2} = \alpha + \beta_1 F_{i,t-1} + \beta_2 F_{i,t} + \varepsilon_{i,t}, \qquad (\text{Equation 2})$$

whereby *Y* represents *GDP* and $Y_{i,t:t+2}$ GDP growth in percent from the period t + 2 starting from half a year until one year after the announcement.

Growth forecast corrections are calculated by the difference between actual real GDP percentage changes during one period *t*, based on the latest data, minus the forecast prepared for the period under scrutiny presented in the previous period t - 1. $F_{i,t}$ is the forecast error correction in GDP growth for the given periods of *t*. The associated forecast error correction is $F_{i,t:t} = \{Y_{i,t} | \Omega_t\}$, where $Y_{i,t}$ denotes the market performance in GDP conditional on Ω_t , the information set available in the current *t* period.

Positives values of $F_{i,t}$ represent an unpredicted outperformance of the market in GDP terms in the next half year. Negative values of $F_{i,t}$ an unexpected underperformance correction of the market predictions in GDP terms. $F_{i,t:t}$ denotes the forecast error change in GDP growth for the periods *t*. The standard error term of the regression is $\varepsilon_{i,t}$.

Under the null hypothesis H_0 that forecast expectation corrections have no impact on or relation with actual future market performance, the β -coefficient would be zero.

A regression to describe the relation of forecast error correction and actual market performance starting from half a year after the announcement up to a year after the announcement was calculated over 35 data points that reveals an overall fit with R square .414 and adjusted R square .378 of the model. The regression coefficient β -value of -.648 for explaining the market performance in the 2nd period after the announcement of the correction is significant at the 5 percent one-sided *t*-testing level *t*=-4.821, with a *p*-value of 0.000. A one percentage point change in forecasting corrections is associated with a -.648 percentage point change in GDP growth in
the same direction starting after the first half year after the announcement up to a year after the announcement.

Overall, the history of past predictions and necessary corrections are positively correlated with future market performance for a half year and negatively associated for the period from half a year to a year after the announcement of a market prediction correction.

Discussion

In the literature, information is attributed to social power and societal status in social relations (Fowler et al., 1979). Language can change attitudes and information expression. Social exchange based on information was shown to be related to economic investment and transaction decisions shaping markets. Capitalist economies are characterized by some powerful long-term patterns in which order and disorder appear hand in hand. An economy's growth is expressed through recurrent fluctuations, punctuated by period depressions. Expectations about future market performances dominates these deeply rooted system dynamics. Dynamic expectations change substantially over time. While there is a wealth of knowledge on future discounting and market performance, hardly any information exists on retroactive expectation corrections' influence on market performance. This article is a first introduction to the idea of an influence of past market expectation corrections on market performance. Past prediction corrections were found to be highly significantly correlated in directionality and strength with aggregate market performance patterns. Past market performance comments were found to be systemically related to aggregate patterns of capitalist economies. The presented results are evidence for the equilibrating process in markets being inherently turbulent. Market outcomes are thereby portrayed as to be more than the sum of its parts, influenced by noise of corrections of past The interdependence of market actors in prices was shown to be biased by expectations. communication creating expectations but also expectation corrections.

The dataset presented is unique insofar as back-testing and reporting of market prediction corrections is not common with some key central banks. For instance, the German Deutsche Bank refrains from back-testing reporting. Yet the results presented offer invaluable insights into the relation of backwards corrections of expectations being systemically related to future performance. In the sense of Kahneman and Tversky's (1979) prospect theory, past losses loom in future performance.

As for future implications, now that we have outlined that our past predictions' corrections are associated with actual performance, we need to better understand what specific market information can influence market performance. In the future, the concrete use of language as market control may be unraveled. A qualitative study of media contents could enlighten on the concrete and qualitative contents that make a market go up or under. Future prospective research could apply emergent risk theory onto economic fluctuations, which could serve as an innovative way to explain how and what information represented in the media creates economic fluctuations (Centeno et al., 2013). Linguistic analyses of newspaper articles about the economy could then shed light on how media representations and temporal foci echo in economic correlates and shape market outcomes. As business cycles are a collective phenomenon, group interactions' potential contribution to create business cycles could innovatively be captured in laboratory and field experiments on the role of information flows among groups in creating price expectations. Social discourse forms a social representation. Mapping out the systematic patterns of information flows' impact on economic correlates would allow predictions about processing and classifications of communication and aid explaining and interpreting economic transformation over time.

Newspaper are sites in which the views of various combinations of social forces and practices are articulated. Newspaper directly speak to the groups and organizations to which the readers belong, the institutions, movements and sections of society they identify. Information thereby shapes social perception. Using linguistic analysis as a way of uncovering the making of economic booms and busts will affect the general consciousness about language as an implicit economic correlate and basic of economic fluctuations. This could also lead to a richer understanding of the echo of market reporting on aggregate properties that give rise to stable aggregate patterns. Mass media providing a platform for social discourse to debate ideologies that presents information about what is happening, which gives rise to reinterpretation and an expectation in the market should become the study object of a follow up paper (Fowler et al., 1979). To determine the lexical variation and meanings embedded within different linguistic systems of expressing different ideologies or theories will aid to connect the linguistic structure with the social correlates of economic booms and busts. Evaluation of discourse could thereby enlighten on the process of economic booms and busts.

Future work may study the linguistic processes to formulate the relation between economic fundamentals and economic outcomes through price. As a configuration of ideas and systems of reality, discourse is a pattern of categorization of complex information. Processes like interpretation, selection, abstraction are yet shunned from orthodox economic analyses. Further, we do not have information on correspondence and linguistic and theoretical processes in the price formation or how linguistic changes can manipulate economic outcomes. Linguistic transformation through focus in time appears important yet is to this day undescribed. The sequence of changes that create booms and busts but also the selection of wording of booms, busts and crises should therefore become subject to scrutiny. Linguistic changes that determinate theoretical and ideological significance could be studied, whereby linguistic discourse will become part of an economic analysis. All these endeavors would allow deriving cheap and easily implementable information nudges as countercyclical alleviation of economic frictions.

Since the production of text and the reception in terms of economic correlates are rather unstudied, we need an understanding how the nature of communicative events influences economic correlates, which may determine differences in economic power and status. Finding the linguistic structures that socialize market reactions will elucidate the role of information in the turbulent construction of the economy. Products of prevailing forms of economic and social organization will help reflect on how social processes and structures are related to material conditions.

The paper acknowledges the fact that there are different frameworks of interpretation in explaining economic long-term cycles. The societal influences of language on economic correlates should be studied in relation to communication's influence on society's ideological impress. How text embodies interpretations of subjects, evaluations of prices and relationships between the real economy and the financial community could be described in a qualitative study. As interpretative meanings are created uniquely in time, the systematic use of linguistic structures that is connected with the text's placement should become subject to scrutiny. As in each socio-economic system there is a social meaning to the natural language and economic communication, which is distinguished in its lexical and syntactic structures articulated; historical data but also cross-country datasets may reveal what communication facets are particular for a certain economic trend or economic market system. The linguistic structure of the economy and linguistic variations throughout different economic times will help describe how different economic outlooks or circumstances are portrayed and potentially perpetuated by the media.

How individuals' perceptions of the future and the state of the economy influence individuals' spending and investment choices may also have wider implications around the world. A future study on international differences could consolidate a global validity of the results but may also highlight different scenarios around the globe. In some countries, people may end up in an

economically unfavorable situation through a self-fulfilling prophecy or self-enforcing mechanism. Developing nations with less fiscal space may face a vicious cycle transmitted through financial markets, where financial stress and macroeconomic self-enforcing feedback mechanisms eliminate the positive impacts of automatic market stabilizers (Semmler, 2013). Contractionary multipliers resulting from a reduction in fiscal spending, which recently gained attention of EU policymakers in the aftermath of the 2008/09 World Financial Crisis, may in particular imply negative effects in post-crisis economies (European Commission, 2014). Regime-dependent multiplies weaken economically already left-behind regions even more (Mittnik & Semmler, 2012). How expectations and expectation corrections influence international trade around the globe could be another interesting extension of the first preliminary results.

Future studies may also investigate the temporal foci of communication strategies' impact on current decisions in order to unravel concrete strategies how market communication should be regulated. A future studies should outline the frequency of specific economic term references and distribution of active or passive, forward looking or backward thinking in relation to economic cycles. Since temporal foci were found to play a significant role in tax allocation preferences over time (Puaschunder & Schwarz, 2012), temporal perspectives and temporal bundling strategies for the information sharing of important market prospects should be considered. Neutrality of communication contents should ensure a degree of certainty, continuity and universality to lower the perturbation biases as well as ideological governance decisions can create. The role of signaling and reputation information for the formation of expectation is another qualitative linguistic research area that could be addressed.

In the wider application, the results mirror the stylized fantasy of market predictions' reality being far from rational. Communicating expectation corrections gets re-packaged at the recipient, whose experience may influence their decision making and guide actions. Information on corrections may thus shape market outcomes in the sense of reflexivity. For linguistics, the study adds to the meaning and implicit meta-meanings of market prospects on economic correlates which are built upon receipt of messages. Unveiling the reality of communication as driver of economic ups and downs will aid deriving communication recommendations to ease economic fluctuations. Central planners but also journalists should be enabled to understand the economic ethos of words and the moral imperative of their economic coverage.

In this sense, access to market information determines the distribution of power within society. Central banker's privileged access to information but also governmental insights, technocrats' knowledge and media moguls control through language draw a hierarchy of those who hold information before it reaches the general populace. The implicit meaning of the results triggers us to re-evaluate economic systems for a democratization of information and derive concrete communication strategy recommendations. Central bankers could embrace a culture of an information fiduciary to responsibly communicate market information and control the potential fall out pro-actively. Leakage of sensitive information and fraud through insider trading based on information may be an essential boundary condition that deserves closer scrutiny in future work in this domain. Fake news and misleading information's emergent risk potential should become integrated into macroeconomic frameworks and legislative control should be adapted to the potential threat of new social media tools creating expectations. Due diligence of information provision and accuracy should become part of corporate governance frameworks and artificial intelligence ethics should embrace information quality control mechanisms in Fintech solutions. Nations around the world may imbue in the hallmarks of democracy and free markets the democratization of information meaning a fair and free access to accurate information to all presented in a cautious, informed and forward-thinking way.

References

- Arestis, Ph. & Eatwell, J. (2008). Issues in finance and industry: Essays in honour of Ajit Singh. New York, NY: Palgrave Macmillan.
- Armstrong, Ph., Glyn, A. & Harrison, J. (1991). Capitalism since 1945. Cambridge, MA: Basil Blackwell.
- Benabou, R. & Laroque, G. (1992). Using privileged information to manipulate markets: Insiders, gurus, and credibility. The Quarterly Journal of Economics, 107, 3, 921-958.
- Bienenfeld, M. (1988). Regularity in price changes as an effect of changes in distribution. Cambridge Journal of Economics, 12, 247-255.
- Blanchard, O. & Leigh, D. (2013). Growth and forecast errors and fiscal multipliers. Washington, D.C.: IMF Working Paper.
- Brenner, R. (2002). The boom and the bubble: The US in the world economy. New York, NY: Verso.

Brenner, R. (2006). The economics of global turbulence: The advanced capitalist economies from Long boom to long downturn, 1945-2005. New York, NY: Verso.

- Centeno, M.A., Cinlar, E., Cloud, D., Creager, A.N., DiMaggio, P.J., Dixit, A.K., Elga, A.N., Felten, E.W., James, H., Katz, St.N., Keohane, R.O., Leonard, Th.C., Massey, W.A., Mian, A.R., Mian, Z., Oppenheimer, M., Shafir, E. & Shapiro, J.N. (2013). Global systemic risk. Unpublished manuscript for research community. Princeton, NJ: Princeton Institute for International and Regional Studies, Princeton University, April 2013.
- Centeno, M.A. & Tham, A. (2012). The emergence of risk in the global system. Working paper, Princeton, NJ: Princeton University.
- Chilcote, E.B. (1997). Interindustry structure, relative prices, and productivity: An input-output study of the U.S. and O.E.C.D countries. Doctoral Thesis. New York, NY: The New School for Social Research Department of Economics.

Cockshott, W.P. & Cottrell, A. (1993). Towards a new socialism. Nottingham, UK: Spokesman.

- Cozzi, T. & Marchionatti, R. (2001). Piero Sraffa's political economy: A centenary estimate. London, UK: Routledge.
- Dobb, M.H. (1929). A skeptical view of the theory of wages. Economic Journal, 39, 156, 506-519.
- European Commission (2014). Financial crisis: Causes, policy responses, future challenges: Outcomes of EU-funded research. EUR 26554 EN, Research and Innovation.
- Fowler, R., Hodge, B., Kress, G. & Trew, T. (1979). Language and control. London, UK: Routledge.
- Gilpin, R. (2010). Global political economy: Understanding the international economic order. Princeton, NJ: Princeton University Press.
- Goodwin, R.M. (1967). A growth cycle. In C.H. Feinstein (Ed.), Socialism, Capitalism and Economic Growth, pp. 54-58. Cambridge, UK: Cambridge University Press.
- Held, D. & McGrew, A.G. (2007). A new world economic order? Global markets and state power: Beyond globalization/anti-globalization: Beyond the great divide. New York, NY: Polity.
- Hunt, E.K. & Lautzenheiser, M. (2001). History of economic thought: A critical perspective. London, UK: M.E. Sharpe.
- Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47, 2, 263-291.
- Kalecki, M. (1954/2009). Theory of economic dynamics: An essay on cyclical and long-run changes in capitalist economy. New York, NY: Monthly Review Press.
- Kauder, E. (1965). History of marginal utility theory. Princeton, NJ: Princeton University Press.
- Keynes, J.M. (1936/2003). The general theory of employment, interest and money. Cambridge, MA: Harvard University Press.
- Kurz, H.D. (2010). Keynes, Sraffa and the latter's "secret scepticism." In B. Bateman, T. Hirai & M.C. Marcuzzo (Eds.), The Return to Keynes, pp. 51-67. Cambridge, MA: Harvard University Press.
- Kwong, C.P. (2008). Mathematical analysis of Soros's theory of reflexivity. Retrieved at https://www.researchgate.net/publication/23951999_Mathematical_analysis_of_Soros%27s_theory_of_reflexivity
- Leonhardt, J.M., Keller, L.R. & Pechmann, C. (2011). Avoiding the risk of responsibility by seeking uncertainty: Responsibility aversion and preference for indirect agency when choosing for others. Journal of Consumer Psychology, 21, 4, 405-413.
- Leontief, W. (1997). Input-output economics. Oxford, UK: Oxford University Press.
- Marx, K. (1867/1995). Capital: A critique of political economy. Moscow, RU: Progress.
- Miller, G., & Rosenfeld, G. (2010). Intellectual hazard: How conceptual biases in complex organizations contributed to the crisis of 2008. Harvard Journal of Law and Public Policy, 33, 2, 807-840.
- Minsky, H. (1975/2008). John Maynard Keynes. New York, NY: McGraw-Hill.

- Mittnik, St. & Semmler, W. (2012). Regime dependence of the multiplier. Journal of Economic Behavior and Organization, 83, 3, 502-522.
- Ochoa, E.M. (1989). Values, prices, and wage-profit curves in the US economy. Cambridge Journal of Economics, 13, 3, 413-429.
- Okamoto, K.S. (2009). After the bailout: Regulating systemic moral hazard. UCLA Law Review, 57, 1, 183-236.
- Orwell, G. (1949). Nineteen eighty-four. New York, NY: Penguin.
- Pasinetti, L.L. (1977). Lectures on the theory of production. New York, NY: Columbia University Press.
- Puaschunder, J.M. & Schwarz, G. (2012). The future is now: How joint decision making curbs hyperbolic discounting but blurs social responsibility in the intergenerational equity public policy domain. Cambridge, MA: Harvard University Situationist Law and Mind Sciences Working Paper.
- Ricardo, D. (1817/1951). On the principles of political economy and taxation. In P. Sraffa (Ed.), The Works and Correspondence of David Ricardo, pp. 1951-1973. Cambridge, UK: Cambridge University Press.
- Schumpeter, J.A. (1949). Economic theory and entrepreneurial history. Cambridge, MA: Harvard University Press.
- Semmler, W. (2013). The macroeconomics of austerity in the European Union. Social Research, 80, 3, 883-914.
- Sen, A. (2003). Sraffa, Wittgenstein, and Gramsci. Journal of Economic Literature, 16, 1240-1255.
- Shaikh, A.M. (1979). Foreign trade and the law of value. Part I. Science and Society, 43, 281-302.
- Shaikh, A.M. (1980). The laws of international exchange. In E.J. Nell (Ed.), Growth, Profits and Property: Essays in the Revival of Political Economy, pp. 204-235. Cambridge, UK: Cambridge University Press.
- Shaikh, A.M. (1984). The transformation from Marx to Sraffa. In E. Mandel & A. Freeman (Eds.), Ricardo, Marx, Sraffa: The Langston Memorial Volume, pp. 43-270. London, UK: Verso.
- Shaikh, A.M. (1998). The empirical strength of the labor theory of value. In R. Bellofiore (Ed.), Conference Proceedings of Marxian Economics: A Centenary Appraisal, pp. 225-251. London, UK: Macmillan.
- Shaikh, A.M. (2010). Reflexivity, path dependence, and disequilibrium dynamics. Journal of Post Keynesian Economics, 33, 1, 3-16.
- Shaikh, A.M. (2013). On the role of reflexivity in economic analysis. Journal of Economic Methodology, 20, 4, 439-445.
- Shaikh, A.M. (2016). Capitalism: Competition, conflict, crises. Oxford, UK: Oxford University Press.
- Sherman, H.J. (1967). Marx and the business cycle. Science & Society, 31, 4, 486-504.
- Soros, G. (1994). The theory of reflexivity. The MIT Department of Economics World Economy Laboratory Conference, Washington D.C. Retrieved at http://mertsahinoglu.com/research/the-theory-of-reflexivity-by-george-soros/
- Soros, G. (2000). Open society: Reforming global capitalism. New York, NY: Public Affairs.
- Soros, G. (2003). The alchemy of finance. Hoboken, NJ: Wiley Finance.
- Soros, G. (2009). The crash of 2008 and with is means. New York, NY: Public Affairs.
- Soros, G. (2013). Fallibility, reflexivity and the human uncertainty principle. Journal of Economic Methodology, 20, 309-329.
- Soskice, D. (1978) Strike waves and wage explosions, 1968-1970: An economic interpretation, In C. Crouch & A. Pizzorno (Eds.), The Resurgence of Class Conflict in Western Europe since 1968, pp. 221-245, London, UK: Palgrave Macmillan.
- Sraffa, P. (1960). Production of commodities by means of commodities. Cambridge, UK: Cambridge University Press.
- Stiglitz, J.E. (2006). Making globalization work. New York, NY: Norton.
- Summers, L. H., & Pritchett, L. (2012). Societies of the world: The future of globalization: Issues, actors, and decisions. Cambridge, MA: Harvard University Faculty of Arts and Sciences.
- Urry, J. (2012). Complex systems and crises of energy. Working paper. Lancaster, UK: University of Lancaster.
- von Weizsäcker, C.Ch. (2010). Cost-benefit analysis with adaptive preferences. Bonn, Germany: Max Planck Institute for Research on Collective Goods. Retrieved at https://www.coll.mpg.de/download/Weizsaecker/Adaptive%20Preferences.pdf



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings



homepage: http://www.acrn.eu/frap

How Blockchain Technology can Monetize New Music Ventures: An Examination of New Business Models

Marcus O'Dair, Robyn Owen*

Centre for Enterprise and Economic Development, Middlesex University Business School

Abstract

The paper examines how blockchain technology is disrupting business models for new venture finance. The role of blockchain technology in the evolution of new business models to monetize the creative economy is explored, by means of a case study approach. The focus is on the recorded music industry, which is in the vanguard of new forms of intermediation and financialization. There is a particular focus on emerging artists. The paper develops an evolutionary tiered theory of technology-driven business models which apply on the one hand to new forms of financial intermediaries, more correctly referred to as 'infomediaries', and on the other hand to new forms of direct monetization by artists.

Keywords: Blockchain; Business Models; Infomediaries; Music Industry; New Venture Finance

Introduction

The paper examines how blockchain, as a disruptive technology, is leading to new business models for new venture finance, focusing on the roles of new emerging financial intermediaries. These are more correctly referred to as 'infomediaries' (Hagel and Rayport, 1997), since they utilise new digital technologies to improve financing and payment alongside a range of associated support services. The focus is on emerging new recored music artists, since this industry is in the vanguard of fundamental changes taking place in the creative sector.

Disruptive technology has long been recognised as a key driver of new business models (Tumasjan and Beutel, 2018; Baden-Fuller and Haefliger, 2013; Chesbrough, 2003, 2010). That is not to say that the business model is universally defined and readily understood (George and Bock, 2011). Zott et al's (2011) extensive review of business model theory and application demonstrated that, since emerging strongly in the mid-1990s (Teece, 2010), thousands of references present a picture of an evolving system theory unit of analysis which is dynamic and encapsulates elements of business networks and connectivity (holistic approaches), venture (individualistic) and value creation (monetization) approaches, which are influenced in various ways by technology and socio-economic conditioning (institutional factors). Zott et al. (2011) criticise the silo-restricted views of many studies and present the case for more rigorous, and more integrated, theoretical development. This paper develops a holistic technological perspective to emerging infomediary business models as value creators which incorporates wider supply-chain systems (Sitonio and Nucciarelli, 2018); in this case including music recording artists, record labels, pubishers and distributors such as music streaming platforms. Crucially, the paper focuses on how these new infomediary business models are seeking to create value and do business more effectively and equitably.

Since this is a newly emerging infomediary activity, with few cases, we adopt a case study approach to examine the role of blockchain technology in the evolution of new business models to monetize recorded music. The industry has experienced an uplift in recent years and is in the vanguard of new forms of intermediation and financialization – notably through Initial Coin Offerings (ICOs) by artists such as DJ Grammatik, and related projects by high-profile artists such as Bjork. A novel evolutionary tiered disruptive technology business model theoretical framework is applied to analyse new forms of financial infomediaries support services to music creators (artists, composers, engineers and producers) and on the other hand to new forms of monetization for new and emerging recorded music artists.

The paper provides the general context, outlining the impact of blockchain, and digital technologies more broadly, on the record music industry. A review of relevant business model literature assists construction of a disruptive technology business analytical framework. The qualitative case study research methodology is then outlined, before presenting key findings, discussion of research implications and conclusions.

Digital and Blockchain Technology Impacts on Recorded Music Industry Intermediaries

Blockchain, as a new and disruptive technology, can provide new ways for creative ventures to do business (O'Dair, 2019). This is particularly the case for the recorded music sector, as Elder (2017) proclaimed:

"Twenty years after peer-to-peer (P2P) file sharing decimated the music industry, blockchain is emerging as a new P2P technology that could rip the industry anew. But this time the revolution promises to be different!"

From a technological perspective, the digital era, notably during the last decade, can be viewed as a key driver of the sharing economy, leading to the emergence of disruptive and innovative intermediary business service model platforms such as Uber and Airbnb. Tumasjan and Beutel (2018) note how these centralised internet-based models represent an expanded digital realisation of former advertising and business promotion models, such as newspaper and journal advertising, with added advantages including wider market exposure and real time connectivity. Within the recorded music industry, the recent rise of YouTube, Spotify, Bandcamp, AppleMusic and Soundcloud provide examples of a burgeoning, and highly centralised, internet-based music streaming industry. Whilst such digital models offer new value creation opportunities for artists through potentially global exposure, virtual/real-time performance and artist-fan contact, O'Dair and Owen (2019) find that the do-it-yourself potential in the creative economy is hard to realise, due to the increased competition that is the corollary of lower barriers to entry.

Emerging musicians still struggle to gain attention without access to established distribution channels and large marketing budgets; indeed, ostensibly 'independent' artists still frequent major label networks to distribute their records. Furthermore, whilst recent growth in online music streaming has halted the decline in global recorded music revenue evident since the turn of the millennium (IFPI, 2018), 99% of streaming income reportedly goes to just 10% of tracks (Krukowski, 2018). Successful financing of new music ventures (here referring to new recordings of existing or emerging artists) has typically favoured well-known artists with a track record of commercial success. Although digital technologies ostensibly empower artists, the music industry remains largely centralised and pyramidal; the vast majority of income goes to a relatively small number of established artists, usually linked to the three major record labels (Universal, Sony, Warner). Potential revenue streams do exist beyond recorded music, including live performance and direct-to-fan opportunities. However, since these revenue streams depend on a significant fanbase, they are rarely available to less established artists. Despite claims of digital technology

ushering in a revolution in the music industry, then, it can also be argued that there has not in fact been a significant shift of power (Rogers, 2013). Blockchain business models, potential offer more equitable solutions.

Blockchain technology has been hailed as decentralising the digital economy (Nakamoto, 2008; Tumasjan and Beutel, 2018), offering disintermediation and opportunities for new forms of peer to peer (P2P) direct sales business models. Blockchains (and related distributed ledger technologies) offer secure, real-time transactions with immutable recording of activity between parties. Crucially, blockchains also offer a way for parties who may not know or trust one another to reach consensus. Tumasjan and Beutel (2018) suggest that blockchain potentially offers a 'true' sharing economy without intermediaries. They cite OpenBazaar as an example of a blockchain-enabled P2P marketplace with no platform transaction costs, unlike centralised sharing platforms such as eBay and Amazon. However, it should be noted that payments are required in cryptocurrency, such as Bitcoin or the Ethereum blockchain currency, which in itself requires intermediary exchanges that typically charge for fiat currency exchange.

Within the recorded music industry, blockchain technology is groundbreaking (Tapscott and Tapscott 2018) in reducing the role of third parties. Famously, blockchain technology has facilitated new forms of crowdfunding in the form of token sales or ICOs, including for new music and creative economy ventures. Breaker, a company (formerly known as SingularDTV) that aims to decentralise the entertainment industry to empower creators, raised the equivalent of \$7.5m using the Ethereum blockchain in just 15 minutes in 2016. They did this by means of a token sale or ICO, in which cryptographic tokens are sold to raise finance. DJ Grammatik, a Slovenian electronic music producer and DJ, raised over \$2 million by the same means in 2017, in less than 24 hours. Bjork has not launched an ICO but she did issue a cryptocurrency to fans who purchased her album, Utopia, in 2017; Imogen Heap, a singer, songwriter, producer and musician, also made headlines in 2015 by using the Ethereum blockchain to pay all contributors to a track, 'Tiny Human', via a 'smart contract' or programmable transaction executed on the Ethereum blockchain. Arguably more important than ICOs, then, is the fact that blockchains can be used to secure artists' rights, and to automate both royalty payments and licensing through smart contracts (O'Dair, 2019). Blokur, one of the start-ups interviewed for this paper, recently announced a project with Massive Attack that used blockchain technology to track the intellectual property in remixes.

Broadly speaking, music copyright can be split into songs and compositions (publishing rights, typically exploited via a music publisher) and recordings (recording or master rights, typically exploited via a record label). Labels and publishers, then, maintain important databases of copyright ownership – and, when a song is written by more than one person, each represented by a different publisher, these databases can potentially disagree. The problem is compounded by the fact that databases are also maintained by collection societies, responsible for collective or 'blanket' licensing, in countries across the world, and these databases, too, can conflict. Blockchains could transform this landscape of centralised, potentially conflicting copyright databases into a more inclusive world of distributed, networked databases – providing a single source of truth, such as who should receive royalties and respond to licensing requests, for a particular song (O'Dair, 2019). This has the potential to provide a growing range of innovative and important funding streams for artists. It includes income from music streams, but also further commission on future sales, similar to sell-on rights in football transfer agreements by smaller feeder clubs (AMG, 2018) or droit de suite in the fine art world. This secondary rights market could exist alongside spillover internet-based activities such as subscription access to artists and micro-metering sales of music segments (Takahashi, 2017).

However, the operation of a fully P2P, decentralised business model system is not without problems, most notably around issues of usability, trust, transparency and regulation – and particularly where market scale-up is envisaged (Beck et al. 2018; Seidel 2017; Voshmgir 2017).

Thus, Pinna and Ruttenberg (2016) forecast the need for new intermediation approaches to benefit from blockchain's reduced reconciliation costs, streamlined post-trade value chain and more efficient use of collateral and regulatory capital. Pinna and Ruttenberg (2016) propose three business system operational models: (i) Clusters, which facilitate maintenance of the largely corporate status quo, whilst improving efficiencies and reducing costs (e.g. inter-bank activity); (ii) Collectives, where new intermediary brokers can enter a regulated market operating in 'commons' with smart contracts (potentially new music industry intermediaries such as Breaker); (iii) P2P, where individuals can operate directly with each other within an open market (e.g. OpenBazaar). Pinna and Ruttenberg perceive P2P as complex, invoking regulatory and privacy issues. Therefore, given the necessity for emerging artists to receive adequate promotion through established, trusted platforms, the 'Collective' approach appears most suitable and attractive.

Thus, whilst blockchain technology may be inherently disintermediating, blockchains are unlikely to remove third parties altogether. A thinning of intermediaries may occur; their revenue share may decrease and their position in the value chain may shift. For example, Sitonio and Nucciarelli (2018) point to the rise in digital service aggregators, such as Ditto (established, 2007), which link digital music to customer facing streaming platforms. They suggest that aggregators add another tier of intermediary service costs to label artists, whilst yielding little revenue to new independent artists (since the typically value of a single play is 0.003p), and that blockchain will reduce their role and costs. We can also predict the emergence of new intermediaries. These new financial intermediaries can be understood as infomediaries: that is, 'custodians, agents, and brokers of customer information' (Hagel and Rayport 1997). A type of infomediary, playing multiple roles in supporting, promoting and generating revenue streams for new ventures, would appear to be at the heart of the new blockchain financing axis. There is the potential, then, for increased artistic freedom and greater remuneration for artists, especially those artists that are less well established.

Literature Review – Forming a Robust Business Model Research Structure

Since coming to the fore in the mid-1990s, business models have formed a growing theoretical framework to explain how technology influences and impacts on business activity (Teece, 2010). Here, the business model is simply defined as how to do business, relating to the strategic act of an individual venture within a system (Zott et al, 2011). Our perspective is necessarily systems-based, since it requires consideration of the role of intermediaries (or more precisely, infomediaries) within the business system and specifically focuses on the role of technology (Tumasjan and Beutel, 2018; Baden-Fuller and Haefliger, 2013; Chesbrough, 2003, 2010) – notably recent blockchain technology developments – as a fundamental driver in the evolution of intermediary business services in the recorded music industry (O'Dair, 2019).

Sitonio and Nucciarelli (2018) examine the evolving supply chain linkages in the recorded music industry. They recognise that paradoxically, digitisation led to a more horizontal distribution channel for artists to post (e.g. MP3 downloads) and more latterly stream music either directly from their own websites or through promotional platforms. However, this offered opportunities for unlicensed file sharing and difficult to trace use, which led to the formation of a new tier of digital aggregator services to track down and collect and deliver payments to publichers, labels and artists. Sitonio and Nucciarelli (2018) recognise that blockchain technology can eliminate the unwieldy, inefficient role of aggregators. This will be achieved through combining digital rights with blockchain technology, potentially creating a more horizontal P2P music distribution and payment process for artists. However, they speculate that this will require the new emerging types of infomediaries examined in this paper.

Scholtens and van Wensveen (2003) convincingly argue that if intermediaries are perceived to continue to add value to their clients, they will have a continuing role. A business model system can demonstrate how businesses can develop more effectively through the key evolving role for blockchain-based infomediaries (Hagel and Rayport, 1997). Additionally, sustainability and equitability over socio-environmental rights are increasing concerns, brought about by the need to address climate change and persistent poverty, for institutional business models (Boons and Ludeke-Freund, 2013; Grath, 2010; Morris et al, 2005). Within the recorded music industry, where currently established digital platforms (e.g. YouTube) have effectively undervalued new music by making it widely freely available, there are strong arguments for more equitable business models (O'Dair and Owen, 2019).

Taking technology as our driver of business model evolution, Baden-Fuller and Haefliger (2013) focus on value creation, relating to customer engagement (McGrath and MacMillan, 2009), and value capture, relating to how value is delivered and monetized (Teece, 2010). Baden-Fuller and Haefliger (2013) develop this to present a framework which examines: (i) Customer identification and whether different customer types pay or have free access to services (Teece, 2010); (ii) Customer engagement, essentially whether the customer service is one-size-fits-all (mass market) or customised for individual clients (consultative); (iii) Value delivery and linkages relating to how the service adds value to particular customer groups; and (iv) Monetization, or value capture, relating to how payment is secured. Additionally, Baden-Fuller and Haefliger (2013) recognise a complex systems connection within the evolving digital (blockchain) paradigm in terms of complementary service capture between businesses that influence business models. Within the music industry this clearly relates to the use of blockchain technology to facilitate the financial transactions of new music service providers such as Breaker.

Developing a Blockchain Infomediary Business Model for the Recorded Music Industry

Having established a broad theoretical base for a blockchain technology business model framework, we now turn our attention to the crucial role of financial intermediaries within the recorded music industry and their evolution through adoption of new technology.

Digital technology is driving the evolution of new business models, for instance in financial services (see Schueffel et al. 2016 review of 200 articles). Advancing Dorfleitner's (2017) categorisation of fintech's digital internet and app activities, four broad groups are evident: (i) financing (e.g. crowdfunding), (ii) asset management and insurance, (iii) payments, blockchain and cryptocurrencies, and (iv) search and artificial intelligence (AI) activities. The latter two groups are particularly relevant to the recorded music industry, because they are already being trialled (e.g. Sodatone's algorithmic search and selection of new artists from streaming, touring and socialmedia data, bought out by Warner's A&R division), and could enhance blockchain technology approaches to contractual ownership and payments for recorded music artists and creators (Rose, 2019).

Financial intermediation theory (Leland and Pyle, 1977; Allen and Santomero, 1997) explains that financial intermediaries are institutions or individuals operating as conduits for diverse parties in order to facilitate financial transactions. Financial intermediation theories are primarily founded in reducing information asymmetries (Leland and Pyle, 1977), reducing transaction costs (Benston and Smith, 1976) and the need for regulation and market confidence (Merton, 1995). Key tenets are that they are justified through adding value and, or, reducing the cost of services for their clients. Here the more appropriate term of 'infomediary' (Hagel and Rayport, 1997) is applied, as this includes crucial specialist non-financial service support, often required by early stage ventures; such as legal IP rights, technical marketing and distribution expertise. Traditionally, this was the role corporate A&R staff ('Artists and Repertoire' staff responsible for spotting and developing artists), but has increasingly been taken on by independent labels, incubators, accelerators (e.g. Ignite in the UK and Marathon Labs in Israel, Sweden, Germany and UK) and online network organisations (e.g. The Rattle in London) in the music industry (O'Dair and Owen, 2019).

Combining the technology-driven business model literature with that relating to infomediaries, a novel framework is formulated to facilitate analysis of how blockchain technology is driving the evolution of new business models of financial intermediation in the recorded music industry. Table 1 presents this dynamic framework which incorporates the three key stages of technological evolution from pre-2000 analogue and post-2000 digitisation, downloading and streaming and post-2015 blockchain developments. The framework also incorporates the elements of operating market system and institutional/regulatory external factors, as well as the venture-specific customer and value creation aspects which comprise the business model theory. Figure 1 provides a conceptual flow-chart (adapting and simplifying Sitonio and Nucciarelli, 2018) for the recorded music industry's distribution and remuneration pre-blockchain (pre-digital and and early digital) and post-blockchain technology, demonstrating blockchain's potential P2P-related simpler path flow, which will still require some forms of new blockchain infomediary facilitators (see findings and discussion).

Technology/	Pre-digital/traditional	Digital/Internet (2000	Blockchain (2015 -	
Business model	(pre-2000)	-))	
Operating system	Major / indie labels Copyright information held in silos	Major/indie labels, growing number of distributors	Major/indie labels, growing number of distributors	
	Physical distribution (retail)	Rise of aggregators in digital distribution	Copyright information networked /	
		Copyright information held in silos	possibility of automated payments	
		Digital distribution through multiple centralised download and streaming platforms	and licensing through smart contracts	
			Possibility of decentralised streaming platforms /P2P	
Venture/customer	Recording artist signed to record label and music publisher	Recording artist probably signed to record label or distributor and music publisher Growing range of independent options, especially for	Recording artist signed to record label or distributor and music publisher – although boundaries blurring Range of independent options continues to	
		established acts	grow, especially for established artists	

Table 1: Technological and Business Model Component Evolution - Recorded Music Industry Infomediaries

Technology/	Pre-digital/traditional	Digital/Internet (2000	Blockchain (2015 -
Business model	(pre-2000)	-))
Value creation recording and add- ons	Recorded music (sales – i.e. one-off purchases)	Recorded music (shifting from ownership to access model) Increased emphasis on other revenue streams (live performance, merchandise, sync, direct-to-fan)	Recorded music (high volume of micropayments) Continued emphasis on other revenue streams (live performance, merchandise, sync, direct-to-fan)
		Crowdfunding	Crowdfunding/ICO
Institutional / Regulatory	Copyright, advances from labels and publishers, royalty collection (relatively low numbers of relatively high- value transactions) via collection societies, tax	Copyright, advances from labels (and some distributors) / publishers, digital royalty collection via collection societies, tax	Copyright, advances from labels/distributors and publishers, digital royalty collection (smart contracts arguably reduce role of collection societies and arguably allow real-time micropayments), secondary rights market, tax



Figure 1: Pre and Post Blockchain Distribution and Revenue Model Framework Source: Adapated from Sitonio and Nucciarelli (2018)

Research Methodology

Taking new infomediaries as our unit of analysis for this study of evolving business models of financial intermediation within the recorded music industry, we adopted a literature review and case study approach (Creswell, 2003). Step one required a search of academic and grey literature, the former revealing little evidence and the latter informed mainly by internet-based reviews and evidence from the infomediaries' websites. This provided a shortlist of 16 music-related infomediaries (Annex Table) adopting blockchain technology including networking, streaming, digital rights and payments services. These represent the most established and high-profile cases globally on the internet, and therefore most accessible to find for music artists. After reviewing these infomediaries, the case studies selected were those that were most aligned to assisting new and emerging recording artists and which use, or plan to use, blockchain technology to develop funding streams for artists. This encouraged purposively focus in this paper on the digital rights and payment services which are already engaged with blockchain technology and in some cases operating Beta pilot testing in the market with recorded music artists (Table 2). Four case study infomediary businesses were interviewed (Mycelia, Blokur, Jaak and Dot Blockchain Media). Additionally, as a contextual industry control, a parallel key informant interview was undertaken with a former Vice President of Strategy from a major record label who is now a private music industry consultant.

Case studies were undertaken in February 2019, involving in-depth qualitative Skype interviews with senior staff developing and delivering services. Interviews typically took at least one hour and utilised a topic guide approach which ensured a consistent approach whilst also facilitating opportunity for exploratory lines of questioning. The topic guide was designed, based upon the literature review, to capture information in relation to the evolving business models and contained the following sections: (i) profile of the respondent and the infomediary organisation; (ii) services offered to new recording artists; (iii) the role of new technology and blockchain in the current and planned service offer to monetize new recorded music; (iv) how blockchain and new technologies are impacting on their business models and industry structure; (v) other external factors impacting on the industry (e.g. institutional and regulatory); (vi) future visions for the industry.

Interviews were recorded, transcribed and checked for factual details and clarifications with the interviewees and supplemented with each infomediary's online web-based and published information and data. The clean and correctly transcribed content was analysed using Excel spreadsheets to derive key phrases and words and inductive findings of trends and approaches which underpin the new business models (Eisenhardt, 1989; Yin, 2003). All findings and interpretations were double-blind checked to avoid individual researcher bias.

Findings

A profile of the four small infomediary business case studies is presented in Table 2. The study's key findings (summarised in Table 3) derived from the interview content analysis gave rise to a set of four headline themes within the technology-driven infomediary business model framework: new blockchain services; customer and monetization; blockchain-based business models; external governance and regulation.

Table 2: Case study summary

	Services	Number of employees	Development stage of blockchain service	Business model
Blokur	Accurate source of global music publishing data	6; further 7 roles just advertised	Beta	SaaS subscription
Dot Blockchain Media	'Collective truth' concerning song and recording copyright; songs as bundles of rights and code	8	Beta	SaaS + services + access
Jaak	Aggregate music rights, low- value/high- volume licensing	16	Beta	Rights network with token; SaaS products
Mycelia	Creative passport: P2P verification for music creators (and other artists)	3-4	Alpha	Services pay to access data; money distributed to creators after maintenance costs deducted

Table 3 Summary of key findings

	Services offered	Blockchain contribution	Broader context	Barriers to adoption	Future scenarios
Blokur	Helps songwriters and publishers get paid what they should, when they should	Distributed database approach more likely to succeed than GRD; genuinely global infrastructure	Copyright law opening up; growth of remix culture requires record of 'granular' use of IP	Small number of players have large amount of influence. GDPR a challenge for some companies (not Blokur)	Ongoing transition from low- volume/high- value to high- volume/low- value necessitates efficient collection
Dot Blockchain Media	New substrate on which multiple parties can share and own information, including 'persistent ownership layer'	Songs as 'bundles of rights and code'; information stored collectively & immutably; provenance; licensing	Biggest threat posed by blockchains is to music distributors	Music industry slow to adopt new technologies, little investment in R&D. Blockchain difficult to understand	Expects one dominant music copyright blockchain, others for licensing. Potential: 'the most powerful system for media that the world has ever known'
Jaak	Aggregated view of	Centralised databases have	More and more services	Premature regulation	Trend towards smaller

	Services offered	Blockchain contribution	Broader context	Barriers to adoption	Future scenarios
	music rights; low- value/high- volume licensing	failed; needs consensus / substrate linking existing databases	outsourced by record labels	(although in favour in the longer term). Consolidation of value around three major labels	organisations. Blurring of content types (music, games, video) and roles (labels, publishers, managers)
Mycelia	Creative passport for all involved in creative process	Blockchain shed light upon problems. Now tech- agnostic: blockchain 'one of the best solutions'	Direct-to-fan potential of digital era is exciting but only goes so far	International politics – nationalism etc – as running counter to global potential of blockchain. GDPR also a factor	In general, intermediaries that do not add value may not survive
Consultant (contextual views)	At major Label, helped with artist 'story', early- stage positioning	n/a	Streaming is about scale, and revenue over time	Problems not technological but human: blockchain perceived as threat. Also, how does this work retrospectively?	Major labels will suffer if do not adapt cost model to fit streaming

New services which require a blockchain operating platform

All of the infomediary business respondents are developing new innovative services which require a blockchain technology platform. The services are clustered around digital rights and the transparent and immutable record of the ownership of those rights – which, in turn, provides a more efficient and effective path to licensing and payment for artists. The focus of this work is around recorded music, primarily because streaming has become the mainstream distribution channel for music and because this channel offers opportunities for new, more efficient approaches to ownership and payment. There is widespread acknowledgment that the Global Repertoire Database (GRD), an unsuccessful attempt to create a centralised aggregate of music rights, was simply too unwieldy, with too many competing interests, whereas blockchain potentially enables a more decentralised, networked solution.

Mycelia, spearheaded by Imogen Heap, are developing a 'Creative Passport' that will provide peer-to-peer verification for all involved in music creation. The alpha stage, which did not use blockchain technology, has been made available to only a small number of music-makers; the beta, focused on decentralised self-sovereign identity, is expected in Q2/3 2019. The passport is intended as a single place for music-makers to post verified data about themselves; services are expected to pay to access that data, that money being distributed to music-makers after the deduction of maintenance costs:

"It will be a digital container for music makers of IDs, skills, verified information, stored with blockchain technology. It's literally going to be your digital self. The whole point of creating this platform is to help musicians be ready to do more business with other services or third-party services... At the moment the process is very convoluted... The artist needs to upload a biography in 250 different places. It's a time-consuming task – and technology would allow the artist to do a hard job [just] one time."

Dot Blockchain Media is predicated on the belief that the file formats currently used for recorded music are outdated, not least because it is too easy to change the metadata relating to copyright ownership. The Dot Blockchain Media model, instead, is to reconceptualise the song as a 'bundle of rights and code', thereby creating the 'substrate' for a more modern music industry. The appeal of blockchain technology for Dot Blockchain Media lies in the fact that it cannot be owned by any one party, and the fact that it provides an audit trail: Dot Blockchain Media aims to immutably stamp ownership in a given piece of music, such that a 'chain of custody' will persist even, for instance, if a song is remixed. Dot Blockchain Media is currently in beta ('stealth mode'), and is operating a hybrid business model: software as a service (SaaS) + services + access.

"For artists, it speeds up licensing immensely, it allows them to prove cases of theft, and it also means that you can express rights into it. You can say this song is remixable by anybody as long as it doesn't contain X Y Z. When I first proposed this, people said, isn't this DRM [digital rights management]? I said no, it's digital rights expression... You can build legally binding contracts into the code, so I can assign you my rights for six months on a non-exclusive basis, and it's machines doing it, not humans pushing paper. That's how you get this industry to be hundreds of billions a year: it has to work with code, not telegraph cables."

JAAK are using blockchain technology to build an aggregated view of music copyright ownership, and to use that aggregated view to make music licensing (especially at the low-value/high-volume end) easier and more scalable. Jaak have been through beta testing with rightsholders and are currently beta testing with developers. In terms of their business model, they are building a rights network, which has a token, and are also building SaaS products on top. Blockchain technology is fundamental to the JAAK vision, because it offers a way to achieve consensus between multiple parties. Without blockchain, copyright databases will be centralised, and attempts to build such a database in the past have failed – the most notorious example being the Global Repertoire Database.

"In the finance sector, if you have six people who want to trade, traditionally you would have a bank in the middle – but in bitcoin, they transact in a network and have confidence that things are being done in the right way. Take that model and apply it to rights. The way you'd be confident in the traditional world is you'd have... a central database that becomes a version of truth. We're saying, actually, we can get that version of truth without having one single database. Those databases that already exist contribute towards it, but we create the substrate that links them all together and provides that communication and helps consensus to be found across them all."

Blokur was founded by the person who, in his previous role at Ujo, was responsible for the pioneering Imogen Heap track, 'Tiny Human'. Blokur use blockchain technology to build an accurate source of global music publishing data. Any given song might have multiple writers, each with their own publishing company, leading to a complex picture in which numerous, low-value payments are collected from across the world and routed back to the relevant writers. Blokur bring together different pieces of that complex jigsaw of publishing rights to ensure that

songwriters are properly credited – and properly paid – when their songs are used. They are currently in beta, and their business model is SaaS subscription.

"What Blokur is doing now, in terms of offering this as a software platform, [is broadly similar to] the way that other industries have been transformed into software platforms instead of agencies – like taxis and travel, with Airbnb and Uber. That is new for the music industry. It's possible for us to make an impact because we are not requiring everybody, from the beginning, to formally sign a bit of paper saying this is how the music industry is going to organise itself in the future, which is what the GRD required. That failed because of the competing interests of everyone round the table. And from a technical perspective, the music industry has not previously had an effective way to do enough automation to make it possible for technology to be the driver of the solution. It's almost always been driven by people-power when it comes to matching data, resolving conflicts, all those kind of things."

All of this must be contextualised by the fact that these services are at various stages of alpha and beta testing. As was observed by another respondent, formerly VP of strategy at a major label and currently a consultant:

"I don't think blockchain is changing the industry structure at the moment. Artists like Imogen [Heap] are in a position to do things but that's individuals, not at scale It will take a few breakout hits for people to take notice."

However, it was also recognised that the shift to music streaming has resulted in an industry that privileges the individual track, rather than entire albums, and that the high-volume, low-value return nature of music streaming requires a completely different model to that which dominates the record industry today. Therefore, inevitably, it will create an industry shake-down, in a drive to leaner, more efficient and value-added infomediary services.

Customer and Monetization Model

All of the featured case study businesses are oriented towards providing blockchain-related services for the recorded music industry (and beyond). The key focus is on ensuring that new recorded music creators and artists (including performers, songwriters and, in the case of Mycelia, producers and engineers) can immutably record their authorship of a given piece of music. In some cases, such as Jaak, the vision extends to creating the infrastructure for rights and licensing across the creative economy (e.g. film and photography) where high-volume, low-value transactions take place. Indeed, Jaak's clients include music publishers and record labels, as well as individual artists, whilst Blokur focuses on music publishers, but both retain the aim to use blockchain technology to ensure that artistic ownership is far more readily traceable and therefore that remuneration is accurately and appropriately apportioned. Dot Blockchain Media is working with distributors and collection societies as well as offering services for individual artists, while Mycelia is focused on music creators themselves. As the respondent notes, there are important roles that 'got lost' in the shift to streaming, since liner notes and credits did not survive the transition: Mycelia's Creative Passport is intended for all contributors to a given piece of music, not only featured artists and songwriters, but also session musicians, engineers and producers. Dot Blockchain Media, meanwhile, seek to digitally imprint ownership and rights expression in every song. There is also recognition that the industry faces the challenge of retro-fitting in terms of existing digitised music being re-imaged to enable blockchain tracking.

In terms of how artists get paid, there is universal agreement that the industry has shifted considerably towards an online streaming model that is driven by single tracks and a high volume of low-value (fractions of a cent per play) payments, where income is accrued over time. This initially posed a challenge to record label business models, which were based around the sale of physical vinyl and CDs, yet labels with large catalogues are now seeing the benefit of the shift to streaming. For artists and songwriters, however, streaming is a winner-takes-all scenario in which only the likes of Adele, Drake and Ed Sheeran earn significant sums while everyone else is obliged to seek out additional revenue streams in order to make a living.

There was widespread scepticism regarding ICOs, dismissed as a 'flash in the pan' and a distraction from the true value of blockchain technology, which lies in its ability to provide that single source of truth. One respondent suggested that blockchain start-ups were already turning to more traditional sources of finance, offering equity in return for capital. The problem with every artist issuing their token to be traded on an open market, one respondent suggested, is not technological but economic. Local currencies such as the Oxford Pound, the respondent pointed out, typically fail because 'a pound that you can only spend in Oxford is obviously worth less than a pound that you can spend anywhere in the UK.' Artist tokens, he suggested, are likely to fail for the same reason.

Respondents also agreed that real-time payments from music streams, though technically possible, are impractical, since the value of payments per stream is so small: 'no artist wakes up and says, if I don't have that tenth of a penny soon, I'm done for.' And as one respondent stated, faced with a choice between fast payments and accurate payments, creators will choose accuracy every time. Respondents were sceptical of blockchain start-ups claiming to allow trading of song rights, and sceptical of those who claim to put smart contracts into songs: 'smart contracts are only intelligent if ownership is secured.' What is exciting, then, is not the ability to make micropayments as a track is streamed but the ability to let streaming platforms know instantly if and when ownership information is changed – and for that information to be accessible to all, so that siloed databases are replaced by a single source of aggregated truth. There is also a belief that blockchain technology offers opportunities for derivative works, particularly for micro-metering, where small sections of music might be sampled and re-used (for instance in the Massive Attack IP tracking of digital re-mixes project being run by Blokur).

New emerging blockchain-based business models

The interviewees acknowledged that blockchain technology has yet to change the structure of the recorded music industry but respondents from Jaak, Blokur and Dot Blockchain Music all believed that it would begin to change that structure in time, leading to new business models. The four case study businesses are themselves at the forefront of these new business models developing the new infomediary services that blockchain both facilitates and necessitates. Yet, as the industry moves from vertical integration to a more horizonal ecosystem, the demand for value-creating infomediaries is increasing. Dot Blockchain Media anticipate consolidation into a preferred single blockchain for core music rights information. There is also a need for improved licensing to address the evolving high-volume, low-value nature of streaming, with several respondents working on blockchain technologies for this purpose. The Mycelia respondent, together with the consultant, was less wedded to blockchain itself, although both still believed it had some potential. Here there is a view that advancing digital technology in the field of big data analysis and AI (as already noted in the case of Sodatone) will impact on new infomediary service models, alongside blockchain. Here the interest is less with the creation of AI music and more around enabling better distribution and end user experience (BPI, 2016). For example Resonate's embedded playlist approach which combines blockchain with AI algorythms to target streaming sites and provide tailored playlists of their artists to streaming end users, whilst offering improved options to own track downloads to those offered by the streaming sites.

The financing models emerging will vary, but a common theme is that these early emerging services are B2B, typically serving publishers, record labels or collection societies as well as artists and songwriters directly.

External: governance and regulation

The significance of external institutional and regulatory factors upon business models has already been noted: no technology exists in a vacuum.

Interview findings on external factors were relatively limited, but what was striking was the positive attitude towards regulation: these are not bitcoin maximalists, typically anarcholibertarian in outlook, but attempts to build businesses on top of the technology, committed to discouraging bad actors. The prevailing sense, however, was that regulation should be light-touch (these are small businesses, with no wish to become bogged down in bureaucracy) and should wait until a clearer picture emerges. Premature regulation, the respondent from JAAK suggests, could be problematic. Historically, he suggests, successful business models are usually clear by the time a bubble bursts; regulation, coming in at that point, then provides 'the rails on which true innovation can take place'. The respondent sees the ICO bubble, however, as anomalous, because access to finance was so easy; clear business models have yet to emerge, and we should not introduce regulation until after their emergence. Copyright law is also an important factor, although the respondent from Blokur saw recent changes in European copyright as positive. The respondent also suggested that some blockchain companies may struggle to comply with GDPR, because of the difficulty of deleting data, although Blokur do not put personal information on the blockchain.

Respondents pointed out that the industry has traditionally underinvested in R&D and struggled to keep pace with technological change. There is also the fact that the recorded music industry is dominated by a small number of big players who may resist change. As the consultant suggested, the industry is built around labels that issue advances and managers who depend on a cut of those advances:

"I have no doubt that there are microfinancing options that would be amazing for artists, but to the extent that they have a manager whose livelihood is taking 25% of a big number then how on Earth are they going to be monetised in this new world? And, therefore, they resist it. The music industry, if it has shown one thing consistently over the last fifty years, it's that the host organism always tries to reject the invader species... There is a whole ecosystem that is very firmly rooted to a way of financing talent which is linked to their own remuneration, so there is a hesitation to explore things that are going to change that."

The consultant also pointed out the challenge that we are not starting this afresh: 'It's fine if you're starting out from scratch but how are you going to retrospectively deal with everything?'

Further reflections on external factors are beyond the scope of this paper. For more detail on political, economic, social, legal and environmental barriers to, and risks of, adopting blockchain technology in the creative economy context, see O'Dair (2019).

Discussion

The general trend in the recorded music industry, respondents agreed, is from small numbers of high-value units (vinyl, CDs) to much larger volumes of low-value uses (music streaming). Streaming, as the consultant pointed out, is a high-volume/low-value business in which revenue accrues over time. For individual artists and rightsholders alike, income in the streaming era depends upon being able to efficiently collect micropayments from multiple streams all over the world. It seems very likely that technology (including blockchain and AI) will play a key part of the solution. AI tracking, facilitatated by new recorded music metadata, will enhance streaming audience benefits alongside the promotion of new artists (Lyons et al, 2019).

The respondent from Blokur points out that the general trend over recent decades has been for the recorded music industry to become less vertically integrated, as everything from recording studios and manufacturing plants to PR and radio plugging has been gradually outsourced: 'Over time, it has basically been unbundled into an ecosystem.' JAAK expect this trend to continue, with record labels focusing increasingly on areas in which they can add most value, such as A&R, and outsourcing other services: 'My vision for the music industry, is one of small organisations competing with each other over actually providing value to artists'. Again, the decentralising, disintermediating nature of blockchain technology appears aligned to broader industry trends. In the short-term, however, the artists who enjoy the associated freedoms are typically either already established or content to remain relatively underground; those emerging artists who desire commercial success are likely to be obliged to work with major labels for the foreseeable future.

As regards blockchain technology itself, the correspondent from Blokur stated that the technological barriers have largely been overcome, due to a combination of improvements in the technology itself and a better understanding of workarounds such as 'half-on half-off' processing. Blockchain cannot solve all issues overnight, but it can operate highly effectively for those wishing to opt in, particularly for new recorded music which can adopt facilitating metadata. This has to address issues of standard identifiers and the interoperability of platforms, which is achievable, given desirability for efficiencies and reduced operational costs for streaming platforms (Lyons et al, 2019). Other barriers to adoption were identified, ranging from the dominance of major record labels to the danger of premature regulation (although respondents were typically in favour of regulation in the longer term). One suggested that some blockchain companies would struggle to adhere to to GDPR; another pointed to the music industry's historic neglect of new technologies and the difficulties of explaining blockchain technology. The consultant argued that blockchain was likely to be perceived as 'technobabble' by a music industry 'jaded by utopians'. The failure of the GRD, a previous attempt to get various parties to collaborate on building a shared music database, was mentioned by more than one respondent. Yet these barriers, as the respondent from Blokur pointed out, are not offputting: 'For an entrepreneur, barriers to entry are attractive... They are the thing that will create value if you succeed.'

Blockchain technology, the correspondent from Blokur conceded, has so far had 'zero impact' on the structure of the music industry, although he expected this to change as the sector matures, anticipating a 'fundamental upgrade of the music industry's infrastructure'. He saw the overall impact as positive, although 'the role of the individual players obviously is going to be subject to change.' Other respondents agreed that blockchain technology had yet to significantly affect the structure of the recorded music industry, yet one predicted that, within the next three or four years, we would see one blockchain initiative establish itself as the ultimate source of music copyright data, with other blockchains taking care of licensing.

The respondent from Mycelia suggested that she was ultimately tech-agnostic – but that blockchain, at the very least, was an 'earthquake' that has introduced a focus on challenges (not least that 'sometimes there are too many intermediaries not bringing value') that are hugely

important – regardless of whether or not these challenges are ultimately solved by blockchain technology. This is similar to the argument proposed by Baym et al (2019), that blockchain is less a technological solution for the music industry than a 'convening force' that brings together disparate actors.

Several respondents pointed out that the potential extends far beyond recorded music, right across the creative economy: 'This whole thing starts with music and it goes into every other file format.'

Conclusions

The paper has explored the new emerging infomediary services at the vanguard of blockchain technology's potential transformation of the recorded music industry. In doing so, it has developed a unique tiered business model approach to analyse these changes and provide insights into the future of the industry and wider repercussions for the creative economy. Four themes emerge from the case studies. First, that blockchain is enabling new infomediary services to develop. The case study evidence suggests that the most advanced of these new services, currently at alpha and beta testing stages, aim to ensure the rights and licensing of recorded music, incorporating artist creator details, smart contract payments and artistic expression into the digitised online streaming and wider artistic distribution mechanisms facilitated by the internet and blockchain. They not only offer vastly improved, secure payment mechanisms for artists, but also contribute to new forms of service (e.g. micrometering of third-party music extracts for film, television and sampling) artistic control over the increasingly fragmented distribution process (e.g. expression of how the music may be used) - and above all an immutable forever ownership and payment stream. Second, blockchain is enabling new infomediary customer services to monetize new recorded music far more efficiently through the new online streaming music distribution services. The main customers will be the artist creators, publishers and labels, and the online music streaming services, with the value creation in terms of greater efficiency (speed, accuracy and cost reduction) in establishing artists rights and payments by using blockchain technology, digital encoding and AI search mechanisms to replace physical collection society activities. Third, the new business models are B2B services, with little cost directed towards the artist/creators. The expectations are that the publishers and labels and streaming distributors will pay for these enhanced services, either through licensing rights or small percentage service charges on transactions. Additionally, the streamlining services also have the potential to cover costs by third party advertising and listener memberships (e.g. for premium advertising free services). Finally, all new service providers are in favour of regulation as a means of gaining public recognition, trust and acceptance of the blockchain services – the main question here is the extent to which industry regulations can be adopted globally and the desire for a blockchain facilitating light touch approach. So, whilst blockchain is inherently disintermediating, it is also necessarily in need of legal enforcement of global rights and infomediary activities to ensure the promotion and more equitable payment of artists.

The paper is necessarily limited by the lack of data and operational case studies (at least to beta testing) in a fast-developing blockchain service sector. The examples provided therefore represent just the initial vanguard of infomediary services that will most likely develop in the next few years around the digital creative economy – with recorded music and its associated digital add-on services at the forefront.

References

Allen F and Santomero AM (1997) The Theory of Financial Intermediation. Journal of Banking and Finance 21(11-12): 1461-85

Baden-Fuller C, and Haefliger S. (2013) Business Models and Technological Innovation. Long Range Planning Vol.46 pp. 419-426

Baym, N, Sqwartz, L and Alarcon A. (2019) Convening Technologies: Blockchain and the Music Industry. International Journal of Communication. 13, 402–421 1932–8036/20190005.

- Beck, R., C. Müller-Bloch, and J.L. King. (2018) Governance in the blockchain economy a framework and research agenda. Journal of the Association for Information Systems 19(3): Forthcoming.
- Benston, G.J., and Smith Jr., C.W. (1976). A transactions cost approach to the theory of financial intermediation, Journal of Finance 31, 215-231

Boons, F. and Lüdeke-Freund, F. (2013): Business Models for Sustainable Innovation:

State of the Art and Steps Towards a Research Agenda, Journal of Cleaner Production, Vol. 45: 9

-19.

BPI (2016) Music's Smart Future: How will AI impact the music industry? Music Ally report for the British Publishing Industry (BPI), accessed 26/03/2019

https://www.musictank.co.uk/wp-content/uploads/2018/03/bpi-ai-report.pdf

Carpenter, R., & Petersen, B. (2002). Is the growth of small firms constrained by internal finance? The Review of Economics and Statistics, 84: 298–309

- Chesbrough HW, (2013) Open Innovation: the New Imperative for Creating and Profiting from Technology. Harvard University Press.
- Chesbrough HW, (2010) Business model innovation: opportunities and barriers. Long Range Planning Vol.43(2-3): 354-363
- Dorfleitner, G, Hornuf, L., Schmitt, M., Weber, M. (2017) FinTech in Germany. Springer International, Basel, Switzerland
- Eisenhardt, K. M. (1989) Building Theories from Case Study Research. Academy of Management Review, 14 (4): 532-550.
- Elder R (2017) Blockchain tokens could transform the music industry. Business Insider UK posting 08/11, sourced 11/05/18
- http://uk.businessinsider.com/blockchain-could-transform-the-music-industry-2017-11

Fraser S (2011) Access to finance for creative industry businesses. Final Report with IFF research for DCMS & BIS George G, and Bock AJ. (2011) The Business Model in Practice and its Implications for Entrepreneurship Research. Entrepreneurship Theory and Practice Vol.35(1):83-111

- Gompers, P.A. (1995). Optimal Investment, Monitoring, and the Staging of Venture Capital. Journal of Finance, 50, 1461-1490
- Grasmayer B (2016) Are accelerators, incubators really helping founders and creators to build new music businesses? Posted on Hypebot 24/10 http://www.hypebot.com/hypebot/2016/10/projecting-trends-accelerating-musics-future.html
- Grath R.G (2010) Business Models: A Discovery Driven Approach. Long Range Planning Vol.43(2-3):247-261

Hagel III, J., and Rayport, J.F. (1997) "The Coming Battle for Customer Information". Harvard Business Review, January

- IFPI. (2018) Global Music Report 2018. Available at: http://www.ifpi.org/downloads/GMR2018.pdf. Last accessed 23.07.18
- Krukowski, D. (2018) How to Be a Responsible Music Fan in the Age of Streaming. Pitchfork. Accessed online 02.03.18,
- https://pitchfork.com/features/oped/how-to-be-a-responsible-music-fan-in-the-age-of-streaming/

Leland, H.E., and Pyle, D.H. (1977). Informational asymmetries, financial structure, and financial intermediation, Journal of Finance 32, 371-387

Lyons F, Un H, Collopy D, O'Hagan P, and Currant K. (2019) Music 2025. Ulster University report to UK Intellectual Property Office DPS-006417, Newport, June file:///C:/Users/Robyn/Documents/Robyns%20Current%20Documents/Robyns%20Work/academic%20pape rs%202017/Blockchain%20articles%20drafts/Music-2025%20Ulster%20University%20June%202019.pdf

McGrath R.G, and MacMillan I.C. (2009) Discovery Driven Growth: A Breakthrough Process to Reduce

Risk and Seize Opportunity, Harvard Business Publishing, Boston, Mass.

- Mason C (2017) Financing Entrepreneurial Ventures. Chapter 17 (Ed) Blackburn R. De Clercq D and Heinonen J. The Sage Handbook of Small Business Entrepreneurship. Sage, London
- Mason C and Harrison R (2015) Business angel investment activity in the financial crisis: UK evidence and policy implications. Environment and Planning C: Government and Policy 33(1): 43–60.

- Merton, R.C. (1995). Financial innovation and the management and regulation of financial institutions, Journal of Banking and Finance 19, 461-481.
- Morris, M, Schindehutte, M, and Allen, J. (2005) The entrepreneur's business model: toward a unified perspective. Journal of Business Research Vol.58: 726-735
- Musicoin (2017) Musicoin: A decentralised platform revolutionizing creation, consuption and distribution. White paper v2.0, October https://drive.google.com/file/d/1KVvcwPKUngMNffgWW65k1p4UvKg5QG0u/view
- Nakamoto, S (2008) "Bitcoin: A Peer-to-Peer Electronic Cash System". https://bitcoin.org/bitcoin.pdf. Accessed 24 March 2018.
- Nesta, (2017) Incubators and Accelerators: An Updated Directory for the UK. Report by the National Endowment for Science and Technology Agency (Nesta), April. Accessed 24/6/2018. https://www.nesta.org.uk/blog/incubators-and-accelerators-an-updated-directory-for-the-uk/
- O'Dair M. (2019) Distributed Creativity: How Blockchain Technology Will Transform the Creative Economy. Cham, Switzerland: Palgrave.
- O'Dair M and Owen R (2019) Monetizing New Music on Blockchain: Four Possible Futures? International Journal of Entrepreneurship and Innovation https://doi.org/10.1177%2F1465750319829731
- Pinna A and Ruttenberg W (2016) Distributive ledger technologies in securities post trading: Revolution or Evolution? European Central Bank, Occasional Paper Series No.172, April
- Rogers, J. (2013) The Death and Life of the Music Industry in the Digital Age. London: Bloomsbury
- Rose A (2019) Three Ways AI is Reshaping the Music Business. Entrepreneur.com, article 11/02/2019, accessed 15/03/2019. https://www.entrepreneur.com/article/327781
- Scholtens B and Wensveen D van (2003) The Theory of Financial Intermediation: An Essay on what it does (not) explain. SUERF The European Money and Finance Forum, Vienna https://www.suerf.org/docx/s 903ce9225fca3e988c2af215d4e544d3 143 suerf.pdf
- Schueffel P.(2016) Taming the beast: A Scientific Definition of Fintech. Journal of Innovation Management Vol.4(4):32-54
- Seidel, M.D.L. (2017) Questioning centralized organizations in a time of distributed trust." Journal of Management Inquiry 19(1): 1-5.
- Siklos, P (2001). Money, Banking, and Financial Institutions: Canada in the Global Environment. Toronto: McGraw-Hill Ryerson
- Sitonio C, and Nucciarelli A. (2018) The Impact of Blockchain on the Music Industry. Paper to the R&D Management Conference, Milan 30/06 to 04/07.
- Statista (2019) Global crowdfunding data for 2017, accessed 21/01/18
- https://www.statista.com/statistics/946668/global-crowdfunding-volume-worldwide-by-type/
- Takahashi R (2017) How can creative industries benefit from Blockchain? McKinsey & Co, August https://www.mckinsey.com/industries/media-and-entertainment/our-insights/how-can-creative-industries-benefit-from-blockchain
- Tapscott D. and Tapscott, A. (2018) Blockchain Revolution: How the Technology Behind Bitcoin and Other Cryptocurrencies is Changing the World, updated edition. Penguin Random House: New York.
- Teece, DJ. (2010) Business models, business strategy and innovation. Long Range Planning, 43: 172-194.
- Tumasjan, A. and Beutel, T. (2018) Blockchain-based Decentralized Business Models in the Sharing Economy: A Technology Adoption Perspective, in Treiblmaier, H. and Beck, R. Business Transformation through Blockchain, Volume 2, Cham, Switzerland: Palgrave Macmillan
- Voise (2017) Voise: A Fair Platform of Musicians. White Paper, Ottawa https://cryptorating.eu/whitepapers/VOISE/whitepaper-voise.pdf
- Voshmgir, S. (2017) Disrupting governance with blockchains and smart contracts. Strategic Change 26(5): 499–509.
- Xiao L. & Owen (Baldock) R. (2016). The role of trust in influencing the equity crowdfunding platform development path: an exploratory study of AngelCrunch in China. Paper to the Institute of Small Business and Entrepreneurship, Paris 10-11th November 2016.
- Yin, R.K. (2003) Case Study Research. London: Sage Publications, Thousand Oaks, California
- Zott C, Amit R, and Massa L. (2011) The Business Model: Recent Developments and Future Research. Journal of Management Vol.37(4): 1019-1042



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings



homepage: http://www.acrn.eu/frap

Social Impact Investment for Housing Projects in Russian Regions

Leonid A. Shafirov*

Southern Federal University

Abstract

This article is submitted as the full paper for the 17th FRAP Finance, Risk and Accounting Perspectives Conference, Track 7. Social and Sustainable Finance, Entrepreneurship and Impact Investing.

The paper deals with the problem of affordable housing. It is a vital part of the Sustainable Development Framework. The research problem is how to improve living conditions for Russian households by enhancing individual housing construction. The rationale for the impact investing in localized individual housing construction projects, with sustainable consumer credit as one of the financial sources, is provided. Such projects are treated as one of the key priorities for the regional socio-economic policy, the financing priorities (funded targets), and the instrument to make financing more effective and efficient. In this study, the evidence is provided by a series of surveys conducted in three federal districts of the Russian Federation (n = 1050), and interviews with the community leaders of a small depressed Russian mono-industrial town (n = 23), to identify:

- prospective lenders, investors and other stakeholders of localized household housing projects,
- household activities related to home construction and home improvements (as the prospective households' budgets allocations).
- prospective ways of improving mechanisms of interactions between the parties involved in housing household projects (as the conditions to make the investing effective and efficient).

Overall, the results of this study show that impact investing in individual housing construction has potential utility as a part of the socio-economic policy for improving the living conditions and economic well-being of Russian households.

Keywords: impact investing; local economic development policies; affordable housing; sustainable credit; household projects

Background and Research Problem

Recent qualitative and quantitative research indicated that activity in affordable housing is relatively high (Czischke & van Bortel, 2018; Elmer, 2017; Pomeroy & Marquis-Bissonnette, 2016; Sharam et al., 2018a; 2018b; Meen, 2018; Gibb, 2018).

Being an essential part of the Sustainable Development Agenda¹, affordable housing appears as a significant part of a diversified impact investment portfolio (Elmer, 2017: 13).

Social impact is often treated as a key motivation to invest in affordable housing and a key investment driver for the investors (Elmer, 2017: 10).

¹ See Sustainable Development Goals. Goal 11: Make cities inclusive, safe, resilient and sustainable

⁽https://www.un.org/sustainabledevelopment/cities/); Affordable housing key for development and social equality, UN says on World Habitat Day, 02 Oct 2017, Goal 11: Sustainable Cities, News

⁽https://www.un.org/sustainabledevelopment/blog/2017/10/affordable-housing-key-for-development-and-social-equality-un-says-on-world-habitat-day/).

E-Mail address: <u>shafirov@mail.ru</u>

Empirical research studies on identifying the linkages between housing and a broad range of non-housing outcomes suggest housing conditions as an essential factor for health and family stability (Pomeroy & Marquis-Bissonnette, 2016). Thus, "improved affordability and housing condition may be an important mediating factor in the transmission of intergenerational and neighbourhood disadvantage that might otherwise exert negative influences on outcomes such as health or opportunities to secure earnings" (Ibid: i).

Among meaningful social performance metrics mentioned by the leading impact investors from Europe and the USA in emerging and frontier markets, are actual improvements of the community life (the quality of life for the broader community) due to investments for affordable housing (Elmer, 2017: 10).

Affordable housing is nowadays among the priorities within the local economic development policies framework in Russia's regions (Analytical Center for the Government of the Russian Federation, 2016: 27)².

Overview of the empirical studies on the programs aimed at assistance in housing improvements (see Polterovich, 2015; Shapiro et al., 2007) shows that joint efforts of the firms based on the establishment of consortiums, partnerships and foundations, result in:

- the lower costs for corporate housing programs due to scale effect;
- the lower interest rate on a mortgage loan;
- the lower price of housing as a response to guaranteed demand for housing;
- establishment of a foundation to diminish risks to provide guarantees for payments to creditors and employees.

Presidential Address to the Federal Assembly 2019³ emphasizes the need to develop affordable financial instruments to assist the households in their private housing endeavors; thus, improving affordability of housing is an integral part of the socio-economic discourse in Russia.

However, the official documents on the tools for territorial development do not refer to proposed mechanisms for housing construction and renovation as the possible tools for local socio-economic development (see also Shafirov, 2017).

As the data from the Federal State Statistics Service and the Ministry of Construction of the Russian Federation suggest, in the past three years, housing construction in Russia has dropped. In 2018, housing construction as a whole amounted to 75.3 million square metres (95.1% versus 2017), in 2017 – 79.2 million square metres (98,7 versus 2016). Meanwhile, the Housing and Urban Environment national project prescribes increase in housing construction up to 120 million square metres by 2024.

There is evidence from the Russian researchers shows that there are sharp intraregional and interregional differences in indicators of living area per person (in particular, half of the poor households reported in 2013 about their crowded housing conditions, which were less than 15 square meters for one family member (Burdyak, 2015)). Although official statistics show no sharp differences in indicators of living area per person, more detailed assessments show that the problem exists. (Ibid., pp. 276–277). Recent data by an integrated housing development institution, the state-owned company DOM.RF (previously – the Russian Agency for Housing Mortgage Lending, AHML) demonstrate that 44% of the Russian households (32% of the total housing stock, which is equal to 1.2 billion square metres, or about 15 times higher than the current housing construction per year) report their need in the living conditions improvements.

² See also program documents by the RF Government: State programme: Affordable and Quality Housing and Utilities for 2011–2017 (http://government.ru/en/docs/3361/); document related to the implementation of the federal target program "Housing" for 2015–2020 years; "Providing affordable and comfortable housing and communal services to the citizens of the Russian Federation" (http://www.minstroyrf.ru/en/trades/realizaciya-gosudarstvennyh-programm/29/), etc. Furthermore, a program for supporting private housing construction is currently being drafted

program for supporting private housing construction is currently being drafted. ³ Presidential Address to the Federal Assembly 2019, February 20 (http://kremlin.ru/events/president/news/page/4).

It should be noted that in 2018, a new record for the mortgage market was set: 1.5 million loans, which equals to 3 trillion rubles, were issued (+ 50% by 2017), DOM.RF reports . Besides, the average loan amount for all mortgage loans has increased – in 2018 it was 2 million rubles, against 1.86 million in 2017. Against the background of a decline in the housing construction, these data show a tendency when, due to a decrease in living standards, the households are increasingly trying to solve the housing problem through the use of credit financing, rather than by savings.

Previous author's studies also have shown that housing problem is an urgent issue for the Russian mono-towns and that this problem may be addressed relying on the sustainable credit concept (see in details Shafirov, 2017; Shafirov, 2014b). In particular, the author has provided rationale for the fact that the local authority and other prospective stakeholders could be interested in housing improvements of the local households in Russian regions, with the housing affordability as one of the critical points promoting local economic development (Shafirov, 2013; 2014a). In this research, due to their prospective role as the favorable catalyst for the local development, housing projects are suggested as one of the key priorities for the regional development. concept of a "rational", or sustainable, consumer credit is introduced (see for example Shafirov, 2014b; 2017) as one of the financial sources to ensure effective and efficient investments in housing construction.

Conceptual Framework

Relying on the previously conducted studies (Shafirov & Tanaka, 2017; Shafirov, 2017; Volchik & Shafirov, 2017), the author applies the research approach which uses the elements of the Project management theory to treat the household housing asset-building activities as the project activities, that is, as a temporary social organization to which resources are assigned to do work to deliver beneficial change (Turner et al., 2010). Thus, the household project is seen as a temporary organization owned by the household, established on the initiative of household or project stakeholder/stakeholders interested in the full or partial project delivery (project beneficiaries). In turn, the household localized project is viewed as the household project which output and/or outcome are consistent with the local economic development goals.

Using the concept of the affordable credit by Collard & Kempson (2005), affordable housing refers in this paper to such housing that is affordable and appropriate for people on low and minimum incomes. This approach (see also: Deeming, Collard & Hayes, 2011; Shafirov, 2014b; 2014a) involves both affordability and appropriateness of the housing for people on low and minimum incomes, implying:

1) availability of housing, offered in the housing market, suited to the livelihood needs of the poor, low-income and excluded families;

2) adequacy, correspondence of price and non-price terms and characteristics of the housing with the financial and non-financial resources of the poor, low-income and excluded families.

The author's concept of a sustainable consumer loan (see Shafirov, 2014b) suggests affordable consumer loans as one of the sources to finance the localized housing projects. These loans include for example loans for housing construction and renovation, and other asset-building activities, such as agricultural development loans, loans for collective purchasing of fuel, clothes, and essentials; loans to pay for medical services, education loans, etc. (see also Shafirov, 2014a).

In turn, research perspective of the Original Institutional Economics (Gruchy, 1947; Stanfield, 1999; Hodgson, 2000; 2003; 2004; 2013; Mirowski, 1987; Koslowski, 2000; Volchik & Oganesyan, 2014) is combined with the perspectives on projects as social process, or social

organization, or, more broadly, a social construct (see in detail Gareis, 2005; Zwikael & Smyrk, 2012; Small & Walker, 2012; Turner & Müller, 2003; Müller et al., 2015; Morris, Pinto & Soderlund, 2012) to describe household project in terms of converging (but also sometimes conflicting) interests in dynamic interactions of a range of stakeholders (see also Shafirov, 2017; 2014b; Shafirov & Oganesyan, 2013).

To provide the rationale for the social impact investing in affordable housing as one of the instruments for local and regional socio-economic development, the author focuses on the prospective role of the local authorities. Their role in organizing housing projects is viewed as one of the key priorities for the regional socio-economic policy. The institutionalization of sustainable lending practices to enhance housing affordability is suggested therefore as the tool to harmonize the interests of prospective stakeholders of the localized housing projects. This harmonization implies the creation and habitualization of rules and practices promoting rational spending of credit resources. In turn, localized housing projects as the financing priority (funded targets) for the prospective investors (local and regional banking and non-banking credit providers) are viewed as the instrument to make financing more effective and efficient.

Methodology

In this study, the mixed-method framework (Wheeldon, 2010), with the emphasis on the qualitative methodology (Richards & Morse, 2013; Maxwell, 2013), is combined with the literature on theory and practice of community-based studies, on the role of community leaders as the key informants (see Veldman et al., 1996; Eyler et al., 1999; Tumiel-Berhalter et al., 2005; NRCS, 2000). This study uses the common practice of identifying the local community leaders trough a brief preliminary survey of the local people (on common practices of identifying the community leaders see for example Warner (Sanagorski) & Galindo-Gonzalez, 2014), and subsequent qualitative research interviews.

The evidence is provided by:

1) interviews with the community leaders of a small depressed Russian mono-industrial town Gukovo (Southern Federal District, n = 23) organized and conducted by the author in 2015, and 2) a series of surveys organized and conducted by the author in H1 2019 in three federal districts of the Russian Federation (n = 1050) – Far Eastern Federal District, North Caucasian Federal District, and Northwestern Federal District).

Interview guidance had been designed to identify:

- household activities related to home construction and home improvements (as the prospective households' budgets allocations);
- prospective lenders, investors and other stakeholders of localized household housing projects;
- prospective ways of improving mechanisms of interactions between the parties involved in housing household projects (as the conditions to make the investing effective and efficient).

The interview questions were focused on understanding the past, current and prospective practices of the local households on housing construction and renovation through the informants' narratives.

The sample survey questionnaire was designed to support the evidence from the qualitative interviews; the sample survey aimed at identifying the most typical problems the households face during their housing endeavours as well.

Findings

Through the expert interviews with the community leaders, a list of household activities related to home construction and home improvements had been identified, showing the most relevant ones, in the context of the sustainable development agenda, taking into account the specific characteristics of the territory:

- 3) to acquire (purchase) dwelling (house/flat);
- 4) to build a house;
- 5) to buy a plot of land for the house construction;
- 6) to buy a dacha or a plot of land for agriculture/gardening;
- 7) significantly reconstruct a dwelling, to redesign a flat;
- 8) to build an annex to a home or extent the house;
- 9) to insulate housing and/or otherwise improve its energy efficiency (e.g., through reducing heat and consumption);
- 10) to connect to the gas trunk pipeline;
- 11) to connect to the water drain;
- 12) to construct a septic tank;
- 13) to repair dwelling (without reconstruction or redesign);
- 14) to improve the landscape in a dacha or a plot of land for agriculture/gardening;
- 15) to improve the landscape around the house;
- 16) to scale up a private farm.

As the interviews have demonstrated, local people usually associate related activities with the "projectized" ways of implementation. In other words, such activities are typically carried out through creation of the temporary social organization and imply different kinds of works and various resources to be assigned, with a range of stakeholders (beneficiaries) involved. In this context, a list of prospective stakeholders of localized household housing projects is specified:

- households and its members;
- local authority (local administration as the governmental institution);
- local employers;
- non-profit organizations;
- local authority's representatives, officials from local government body responsible for economic development;
- local administration employees (technical expert or consultant assigned by the local administration);
- bank staff experts (estimators and credit consultants);
- outside specialists and technical experts employed (foremen);
- members of the household (as a rule, the head of the family);
- specialists assigned by the local administration;
- banking and non-banking organizations;
- suppliers and contractors.

The survey results have demonstrated the typical problems which occur when the respondents implement these kinds of asset-building activities (see Table 1), with the lack of financial resources at the top (about 61% of the respondents).

Table 1: Typical problems related to housing construction and renovation

Difficulties related to housing construction and renovation	% of respondents
Lack of financial resources	60.5
The poor infrastructure surrounding the construction site	37.4
Lack of appropriate plot of land	22.3
Lack of reliable suppliers to purchase high-quality materials	12
Insufficient engineering communications and difficulties to connect to the utility networks	11.7
Lack of appropriate skills	7.7
Lack of reliable contractors to do the work on the construction and renovation	6.3
Lack of time, skills and motivation	5.9
Lack of legal knowledge	5.6
Lack of standardized projects for housing construction/connection to the engineering communications	4.4

It worth noting that among the respondents from the regions with the negative migration balance (Far Eastern District), 67.5% reported that the construction of a private house would deter them from going to permanent residence in another region or city. 68% of the respondents said that they were more likely to decide on the birth of a child if they had a private house in their city or region. 55% of respondents are ready to get the loan to finance a private house construction in their city or region.

Related positive effects for the local communities treated as the actual improvements of the community life in the context of the Sustainable Development Goals Agenda are to result in:

- increased financial and legal literacy of the local population;
- improved quality of life of local households and increased financial stability;
- household's empowerment, increased bargaining power of the households;

- reduced risks of the high turnover rate of employees among local employers, increased labor warranty;
- reduced risks for the lenders (banks and other credit providers);
- increased loyalty and trust of the local people toward the local administration;
- increased perceived performance of local government in the eyes of the local population.

Each of these positive effects was analyzed through the theoretical lens of the Project Management theory, as described in Shafirov (2017). This approach provided a fruitful perspective to treat the prospective role of local administration's representatives as involving the following functions (operationalized as the "measures of assistance") through application of related methods of project management:

"1. To facilitate coordination between the members of the local households (or their groups) and other parties involved (creditors, suppliers and contractors, bureaucracy agencies, including facilitating access to the objects of engineering infrastructure).

2. To provide informational support for the parties involved.

3. To control the fulfillment of the credit obligations by the households.

4. To implement monitoring of completion and efficiency of housing projects.

5. To promote and make efforts to replicate best practices of the use of credit resources for housing construction and renovation by the borrowing households to generate future incomes, ensure cost savings – i.e., to promote rationalisation of borrowing and consumer behaviour and prevent wasteful spending." (Ibid: 84–85).

Further, this idea was elaborated to gain an understanding of the particular stage of the household project activity to which a specific measure of assistance may refer. Relevant classifications are given below:

I. Preliminary "pre-project" stage.

- 1. Free or preferential provision of land for the construction of a residential house
- 2. Free or preferential execution of the following works (services):
 - project design;
 - preliminary cost estimate (incl. utility connection services);
 - technical specifications for connecting to engineering communications (incl. heat,
 - water, power supply; sewage; gas pipeline)
- 3. Drawing up construction documentation

4. Scheduling construction and related financial costs

5. Drawing up a list of materials, equipment, kind of works required for housing construction, and delivery/implementation schedule

6. Providing a typical construction project

7. Conclusions on feasibility of contracts with suppliers, contractors and creditors based on the property and financial position of the household for the nearest perspective

8. Mediation in the search for reliable suppliers and contractors

9. Negotiations with the neighboring households, suppliers and contractors

10. Advisory, coordinative and expert assistance to the household

11. Provision of templates for contracts with suppliers and contractors, creditors etc., explanation of the procedure for their use

12. Coordination of the interaction of groups of households to solve common problems, to share experience via a special databank

II. Project implementation stage

1. Expert and technical advice on:

- the engineering issues;
- the legal issues

2. Monitoring by the expert engineers of the quality and quantity of work performed by the contractors

3. Financial assistance to the households and low-income neighboring households in connecting to the: gas networks; electrical networks; sewage; central heating

4. Free connection of the land plot of the household to the engineering networks

5. The organization of the interactions of households with utility providers to connect households to: gas networks; electrical networks; sewage; central heating

6. Coordination of the interactions of groups of households to solve common problems, to share cooperation experience via a special databank

7. Assistance in determining and imposing penalties on unscrupulous suppliers and contractors

III. Stage of completion of the housing project

- 1. Assistance in assessing the success of the implemented project
- 2. Promoting the best supplying and contracting practices
- 3. Promoting the best household project practices
- 4. Analysis and synthesis of negative experience of household developers

5. Assistance in determining and imposing penalties on unscrupulous suppliers and contractors

6. Checking the quality terms of work performed by contractors by the expert engineers

Recommendations for Socio-Economic Policy

The following tasks are suggested as a part of coherent economic and social policies that address the issues of impact investments in housing affordability for the households.

1. To overcome the land shortage for individual housing construction and renovation.

This issue might be addressed in several ways:

- to amend the existing law to prevent regional and municipal authorities from confiscation of the plots of land, previously leased to agricultural producers, if such plots are near the settlements lacking land plots for individual housing construction;

- to delegate authority (and responsibility) for the use and disposal of federally owned land plots, from the Federal Agency for State Property Management (Rosimushchestvo) – to the regional authorities. This can enable regional authorities together with the tenant farmers to decide whether the plots should be used for individual housing construction purposes;

- to amend the existing federal law to extended the categories of citizens for whom the plots of lands for private housing construction purposes are provided free of charge or at heavily subsidized prices (either for private ownership, without the auction as a phase, or for rent-free use);

- to revise the existing federal law so as the priority for the provision of land plots for individual housing construction shall be established for low-income and socially vulnerable groups under long-term lease or rent-free use rather than under private ownership (to prevent resale, any misuse, misappropriation or improper exploitation of the land);

- to agree on the rules on granting subsidies from the federal budget on the provision of land for individual housing construction for low-income and socially vulnerable groups – either for purchase or for lease – in cases if local authorities cannot provide such groups of people with plots of land for free;

– to amend the existing federal law so as the possibility shall be provided for regional and local authorities to offer land plots for individual housing construction for up to 20 years (instead of a fixed term of 20 years), with the minimum duration of 5-7 years. This measure can positively affect the discipline of the individual developers through enhancing the opportunities for the most interested and well-resourced ones;

- to create a supportive legislative and regulatory framework for providing rent-free land plots for private housing construction on a competitive basis for those citizens who implement cost-effective energy-efficient and renewable energy housing projects, environmentally sound and sustainable emission reduction housing projects, using products and services of the local providers and contractors.

2. To provide the households with an opportunity to create (build/construct) private farms on the agricultural lands owned by them.

3. To ensure the development of engineering infrastructure.

In particular, participative budgeting can be used for financing independent power supply, water supply, autonomous heating supply and sewer, in the regions experiencing a severe budget deficit. As a result, in those regions where local resources can cover the necessities of the whole settlements, infrastructure barriers can be overcome by mobilizing the local communities to transform their finance into the source of the regional economic development. Additional address support for the low-income and socially vulnerable groups is required, at that.

4. To provide support for the "dacha" owners and inhabitants in small towns and settlements.

Housing projects should be co-financed and supported through non-financial measures in small towns and settlements, including organizational support in infrastructure development, such as heat, water, electricity supply, construction and rehabilitation of roads. L cal and regional authorities should support the readiness of the local communities (inhabitants of the towns and settlements) to participate financially in covering the part of related costs.

5. To reduce administrative and bureaucratic burden through the implementation of standardized housing project design.

This measure can increase the quality of housing, reduce associated risks, and minimize related transaction costs. Besides, this allows replication of successful housing projects through the dissemination of knowledge on positive and negative households' experiences. Furthermore, if designed according to the standardized project, houses can serve as collateral for the loans, which can positively affect the secondary housing market of the in the long-term perspective. The use of advanced technologies, materials and equipment, in turn, can provide cost savings, with shorter construction periods.

6. To provide financial and organizational assistance for the local people planning housing construction and renovation.

In particular, additional support has to be provided for the owners of private farms, with the cooperative distribution and processing as the expected factor of its profitability. Increased

profitability of private farms, in turn, would result in increased household incomes, and enlarge the financial basis to cover the house maintenance costs. Thus, this measure may provide further impetus for individual housing construction development.

Discussion and Directions for Further Research

The idea of this research emerged from a suggestion that "If somebody gets value from looking upon what they are doing as a project, and wants to use some project management tools to help manage it, we would encourage them to do so. There is a spectrum of endeavors from the totally routine to the totally unique. ... That is really the point – which perspective gives people value to help them manage what they are doing better" (Turner et al., 2010: 107).

Thus, the current study presents the researcher's attempt to view the household housing assetbuilding activity as the project activity using the elements of the Project Management theory as the fruitful theoretical perspective. The potential usefulness of this perspective may be challenged; however, we need better insight into the socio-economic and financial situation in the Russian regions, with the particular emphasis on housing construction and renovation, as well as the expected social impact of related projects.

It seems reasonable to suggest that other kinds of household asset-building activities be the subject of further research from the Project Management theory perspective, either qualitative or quantitative. The focus on assessing the expected household demand for the measures of assistance from the local authorities in implementing household asset building may be an appropriate measurable objective for such a research.

The potential conflict of interests between the prospective stakeholders of the household projects may provide a good focus for qualitative research. Attempts to identify and measure the social impact of investments in housing projects should rely on both qualitative and quantitative methodologies as well.

Conclusions

The research findings allow demonstrating how the elements of Project Management theory can contribute to the study of household asset-building activities, and to view related activities as the project.

Furthermore, research approaches within the Project Management theory have provided valuable insight into household project activities.

In particular, the suggested approach views the household asset-building activities as prospective interactions between the potential stakeholders of household housing projects. The author's perspective provides the rationale for the expected measures of assistance from the local authorities, which can act as a trigger for impact investments in affordable housing and housing improvements in the Russian regions. Nevertheless, further research is needed to gain insights into the specificities of various regions and municipalities of the Russian Federation, and to justify the recommendation and provide supporting evidence for the pilot localized household projects as one of the instruments for the local economic development.

References

Analytical Center for the Government of the Russian Federation (2016). Human Development Report in the Russian Federation. UN Sustainable Development Goals and Russia. Brief version, 44 p. (http://ac.gov.ru/files/publication/a/11138.pdf). (In Russian.)

Burdyak, A. Ya. (2015). Housing in post-Soviet Russia: Inequality and the problem of generation. Journal of Social Policy Studies, 13(2), 273–288. (In Russian.)

Collard, S., & Kempson, E. (2005). Affordable credit: The way forward. York: Joseph Rowntree Foundation.

- Czischke, D. & van Bortel, G. (2018). An exploration of concepts and polices on 'affordable housing' in England, Italy, Poland and The Netherlands. J Hous and the Built Environ, pp 1–21, https://doi.org/10.1007/s10901-018-9598-1
- Deeming, C., Collard, S., & Hayes, D. (2011). Affordable credit: Lessons from overseas, A report prepared for Consumer Focus by the Personal Finance Research Centre (PFRC) University of Bristol (http://www.bris.ac.uk/media-library/sites/geography/migrated/documents/pfrc1109.pdf).
- Elmer, P. (2017). State of Investment in Affordable Housing. Habitat for Humanity International, Terwilliger Center for Innovation in Shelter, 21 p. (https://www.findevgateway.org/sites/default/files/publication_files/state_of_investment_in_affordable_housin g report-published-11152017.pdf).
- Eyler, A. A., Mayer, J., Rafii, R., Housemann, R., Brownson, R. C. & King1, A. C. (1999). Key informant surveys as a tool to implement and evaluate physical activity interventions in the community. Health Educ. Res., 14(2), 289–298. doi: 10.1093/her/14.2.289

Gareis, R. (2005). Happy Projects! Vienna: Manz Velog.

Gibb, K. (2018). Funding new social and affordable housing: Ideas, evidence and options. Social Housing Policy Working Group Paper, November 2. The UK Collaborative Centre for Housing Evidence (https://housingevidence.ac.uk/wp-

content/uploads/2018/11/R2018_SHPWG_funding_models_Ken_Gibb_final.pdf).

Gruchy, A. G. (1947). Modern Economic Thought. The American Contribution. New York: Prentice-Hall, Inc.

Hodgson, G. M. (2000). What is the essence of institutional economics? Journal of Economic Issues, 34(2).

Hodgson, G. M. (2003). The hidden persuaders: institutions and individuals in economic theory. Cambridge Journal of Economics, 27, 159–175.

- Hodgson, G. M. (2004). Reclaiming habit for institutional economics. Journal of Economic Psychology, (25), 651–660.
- Hodgson, G. M. (2013). From Pleasure Machines to Moral Communities: An Evolutionary Economics without Homo economicus. Chicago and London: The Chicago University Press.
- Koslowski, P. (2000). The Theory of Ethical Economy as a Cultural, Ethical, and Historical Economics: Economic Ethics and the Historist Challenge, pp. 3–15 / In: P. Koslowski (ed.) (2000). Contemporary Economic Ethics and Business Ethics. Springer-Verlag Berlin Heidelberg, 266 p.
- Maxwell, J. A. (2013). Qualitative Research Design: An Interactive Approach. USA: Sage Publications, Inc.
- Meen, G. (2018). Policy approaches for improving affordability. Report, September 3. The UK Collaborative Centre for Housing Evidence (https://housingevidence.ac.uk/wp-content/uploads/2018/09/R2018_02_02_policy-approaches-for-improving-affordability.pdf).
- Mirowski, Ph. (1987). The philosophical bases of institutionalist economics. Journal of Economic Issues, 21, 1001–1038. Reprinted in: Lavoie, D. (Ed.) (2005). Economics and Hermeneutics (pp. 74–110). Taylor & Francis e-Library. [Russian ed.: Mirowski, Ph. (2013). The philosophical bases of institutionalist economics. Terra Economicus, 11(3), 72–88.]
- Morris, P. W. G., Pinto, J. K., & Soderlund, J. (Eds.) (2012). The Oxford Handbook of Project Management. Oxford, UK: Oxford University Press.
- Müller, R., Pemsel, S., & Shao, J. (2015). Organisational enablers for project governance and governmentality in project-based organizations. Int. J. Proj. Manag., 33(4), 839–851. http://dx.doi.org/10.1016/j.ijproman.2014.07.008
- Natural Resources Conservation Service (NRCS) of the U.S. Department of Agriculture (USDA). Social Science Team (2000). Working with Community Leaders. People, Partnerships, and Communities, 43 (http://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1045582.pdf).
- Polterovich, V. (ed.) (2015). Home Purchase Assistance Programs in Corporations: A Problem of Institutional Design. Moscow: CEMI RAS, 81 p. MPRA Paper No. 66981 (https://mpra.ub.uni-muenchen.de/66981/). (In Russian.)
- Pomeroy, S. & Marquis-Bissonnette, M. (2016). Non-Housing Outcomes of Affordable Housing, Research Report, March. Focus Consulting Inc. and Carleton University Centre for Urban Research and Education (CURE) (https://carleton.ca/cure/wp-content/uploads/Non_Housing_Outcomes_of_Affordable_Housing.pdf).
- Richards, L., & Morse, J. M. (2013). Readme first for a user's guide to qualitative methods. USA: SAGE Publications, Inc.
- Shafirov, L. A. & Oganesyan, A. A. (2013). Consumer lending rationalization for the local development benefits: an
institutional approach.TerraEconomicus,11(4.3),27–42(http://te.sfedu.ru/evjur/data/2013/journal11_4_3.pdf). (In Russian.)

- Shafirov, L. A. & Tanaka, H. (2017). A Conceptual Scientific Approach to Local Economic Development by Way of Portfolios of Projectised Household Asset Building Activities a Case of Depressed Russian Monotown, pp. 222–228 / In: Proceedings for 30th IPMA Congress, Kazakhstan, Astana, 4th–7th of September, 2017. Published on behalf of 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), 5–8 Sept. 2017. doi: 10.1109/STC-CSIT.2017.8099452
- Shafirov, L. A. (2013). Russian monotowns residents' living conditions as the basis for estimation of households' sustainable lending potential. Terra Economicus, 11(4.3), 208–217 (http://te.sfedu.ru/evjur/data/2013/journal11_4_3.pdf). (In Russian.)
- Shafirov, L. A. (2014a). Potential of sustainable lending to individuals: Russian small monotowns' population peculiarities. Journal of Economic Regulation, 5(3), 124–138 (http://www.hjournal.ru/files/JER_5_3/jer_5_3-12.pdf). (In Russian.)
- Shafirov, L. A. (2014b). Sustainable Consumer Lending: Local Communities, Rationality, and Economic Policy. Journal of Institutional Studies, 6(3), 67–82 (http://www.hjournal.ru/files/JIS_6_3/jis_6_3-5.pdf).
- Shafirov, L. A. (2017). Sustainable lending for housing projects: Is affordable housing a myth for depressed Russian monotowns? ACRN Oxford Journal of Finance and Risk Perspectives, 6.3 Special Issue FRAP/SSFII Conference, 2017, 71–88 (http://www.acrn-journals.eu/iframe-6/jofrp/jofrp0603.html).
- Shapiro, A. L., Reeves, W. E., Quigley, M., Wides, B., & Williams, J. L. (2007). Understanding Employer-Assisted Mortgage Programs: A Primer for National Banks. US Department of the Treasury (https://www.occ.gov/topics/community-affairs/publications/insights/insights-employer-assisted-mortgage-programs.pdf).
- Sharam, A., Moran, M., Mason, C., Stone, W. & Findlay, S. (2018a). Understanding opportunities for social impact investment in the development of affordable housing, AHURI Final Report No. 294. Melbourne: Australian Housing and Urban Research Institute Limited (https://www.ahuri.edu.au/research/final-reports/294, doi:10.18408/ahuri-5310202).
- Sharam, A., Moran, M., Mason, C., Stone, W. & Findlay, S. (2018b). Social impact investing for Australian affordable housing, AHURI Final Report No. 294. Melbourne: Australian Housing and Urban Research Institute (https://ahuri.edu.au/research/final-reports/294).
- Small, J., & Walker, D. (2012). Chapter 12. The project entity as human activity system and social process: Providing the structural openness to connect with context / In: H. Linger and J. Owen (eds.) The Project as a Social System: Asia-Pacific Perspectives on Project Management. Monash University Publishing (http://books.publishing.monash.edu/apps/bookworm/view/The+Project+as+a+Social+System%3A+Asia-Pacific+Perspectives+on+Project+Management/171/OEBPS/tp.htm).
- Stanfield, J. R. (1999). The Scope, Method, and Significance of Original Institutional Economics. Journal of Economic Issues, XXXIII(2), 231–255.
- Tumiel-Berhalter, L., Watkins, R., & Crespo, C. J. (2005). Community-Based Participatory Research: Defining
CommunityCommunityStakeholders.MetropolitanUniversities,16(1),93–106(http://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1025&context=commhealthfac).100100100
- Turner, J. R. ,& Müller, R. (2003). On the nature of the project as a temporary organisation. Int. J. Proj. Manag., 21 (1), 1–7. [Op. cit.: Mueller et al., 2015].
- Turner, R., Huemann, M., Anbari, F. & Bredillet, C. (2010). Perspectives on Projects. New York: Routledge.
- Veldman, C., Zuckerbraun, S. & Schneiderman, S. (1996). Methodology for a community-based key informant survey. Proceedings of The National Opinion Research Center (http://www.amstat.org/sections/srms/proceedings/papers/1996_156.pdf).
- Volchik, V. V. & Oganesyan, A. A. (2014). What is OIE needed for, if there is NIE? Terra Economicus, 12(4), 79–90 (http://te.sfedu.ru/en/journals/2014/60-no-4/503-what-is-oie-needed-for-if-there-is-nie.html). (In Russian.)
- Volchik, V. V. & Shafirov, L. A. (2017). Prospects for institution localized household credit projects: an evidence from a small Russian monotown. Economics Science of Modern Russia, 2, 29 46. (In Russian.)
- Warner (Sanagorski), L. A. & Galindo-Gonzalez, S. (2014). Identifying Key Community Leaders to Assess Extension Programming Needs. Agricultural Education and Communication Department, University of Florida IFAS Extension, WC164 (http://edis.ifas.ufl.edu/wc164).
- Wheeldon, J. (2010). Mapping Mixed Methods Research: Methods, Measures, and Meaning. Journal of Mixed Methods Research, 4(2), 87–102. DOI: 10.1177/1558689809358755
- Zwikael, O., & Smyrk, J. (2012). Chapter 14. The project as a social system: Identifying key players to ensure outcome realization / In: H. Linger and J. Owen (eds.) The Project as a Social System: Asia-Pacific Perspectives on Project Management. Monash University Publishing (http://books.publishing.monash.edu/apps/bookworm/view/The+Project+as+a+Social+System%3A+Asia-

(http://books.publishing.monash.edu/apps/bookworm/view/The+Project+as+a+Social+System%3A+Asia-Pacific+Perspectives+on+Project+Management/171/OEBPS/c14.htm).



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings



homepage: http://www.acrn.eu/frap

The Influence of EU Non-Financial Reporting Directive on CSR Disclosure: Empirical Evidence from Finland

Hanna Silvola^{*,1}, Eija Vinnari², Suvi Viljaranta³

¹ Hanken School of Economics ² University of Tampere ³ KPMG

Abstract

To improve the disclosure of non-financial information, on 26th June 2017, the European Commission introduced the Directive on the disclosure of non-financial and diversity information (Directive 2014/95/EU). From 2018 onwards, large public-interest entities have required to include in their annual reports a statement on environmental, social, employee-related, anti-corruption and bribery matters, respect for human rights, and diversity (European Commission, 2017). The purpose of this study is to examine the effects of the directive on corporate social responsibility (CSR) reporting. Our research question is formulated as follows: How, if at all, has the implementation of the non-financial reporting Directive influenced the CSR disclosures of organizations falling under the scope of the Directive? This topic is examined in the empirical context of Finland because of two major reasons. First, with the exception of a dozen state-owned companies, CSR reporting has thus far been voluntary, thus enabling the analysis of the effects of mandatory regulation on the practice. Second, the small size of the Country makes it possible to analyze the annual and CSR disclosures of all organizations falling under the scope of the Directive. The empirical analysis of all the relevant CSR reports disclosed during the years 2016-2018; and a thematic analysis of interviews conducted with the representatives of 11 companies, both before and after the implementation of the Directive

Key words: Non-Financial Information (NFI), Corporate Social Responsibility (CSR), Sustainability Reporting, Assurance, EU directive, Regulation

Introduction

To improve the disclosure of non-financial information (NFI), on 26th June 2017, the European Commission introduced the Directive on the disclosure of non-financial and diversity information (Directive 2014/95/EU) on 26th June 2017. The purpose of the directive is to improve and standardize the disclosure of non-financial information in large public-interest entities (PIEs), i.e. listed companies, banks and credit institutions, insurance companies and other organizations determined as public interest entities by a government. Large companies are defined to employ on average 500 or more employees during a fiscal year. Over 6000 companies across the European Union's (EU) Member States are estimated to belong within the scope of the Directive. (EC, 2017.)

From 2018 onwards, large public-interest entities have required to include in their annual reports a statement on environmental, social, employee-related, anti-corruption and bribery matters, respect for human rights, and diversity (European Commission, 2017). The directive requires companies to disclose information on the above mentioned sustainability aspects but
companies have a lot of flexibility to choose the relevant issue and exact measure for disclosing that information. To illustrate, the issue of the required environmental aspect can be, for example, water, CO2 emissions or energy. Moreover, 'water' can be measured and disclosed by a consumption per produced tons or per a product. As a result, the six aspects of non-financial information are required to disclose but it up to the company on which issue and measure it choose to disclosure because the different issues are relevant within different industries.

The EU Directive on the disclosure of non-financial and diversity information was launched to improve transparency, standardize CSR reporting and increase the importance of CSR in firms. Member states implement the directive in their national legislation. They were allowed a considerable degree of discretion in deciding on reporting framework, format and content. The Member States are allowed to set state-specific requirements on companies. For example, a company can choose whether to apply any framework in CSR reporting and the non-financial information can be provided with an integrated way as a part of a management report or separately within the certain time frame. The Member States vary as well in the ways on how they determine a large undertaking and which organizations to recognize as PIE. Moreover, the Member States can determine by themselves if the non-financial information bust be externally assured and if any penalties will be imposed upon organizations which are not fulfilling the reporting requirements. (CSR Europe & GRI, 2017.)

Companies' voluntary disclosure of social and environmental information (also known as CSR reporting or sustainability reporting) has fascinated accounting scholars for several decades. Especially the drivers and motivations for such disclosures have been studied extensively, from various theoretical perspectives. First, it has been proposed that in such a situation, voluntary disclosure might be used as a mechanism to signal that a firm has "nothing to hide", thus avoiding an adverse market reaction (Brammer & Pavelin, 2004) that might have eventuated from non-disclosure. Alternative perspective suggests that firms voluntarily disclose social and environmental information in order to ensure their legitimacy, or their social license to operate. In other words, CSR reports are generated in response to pressures exerted by diverse stakeholders (Sinclair-Desgagne & Gozlan, 2003).

Critical accounting literature also claim that it is somewhat naïve to expect corporate voluntary self- reporting to question the very premises on which the business world and capitalist market economy is built (Gray et al., 2014; Cho et al., 2015). Thus, it has been argued (e.g. Braam et al., 2016) that mandatory reporting is required to prevent the use of social and environmental disclosures as an instrument of greenwashing. This parallels the finding from financial reporting literature, where mandatory disclosure is considered necessary in instances where managers do not voluntarily disclose information that would increase social welfare (Beyer et al., 2010).

Rational perspective points out the strategic competitiveness of sustainable companies and CSR reports as a tool to communicate that information to the markets. It has also been reported that voluntary front-runners of CSR reporting have been reported to benefit strategically more than followers (Bhimani, Silvola & Sivabalan, 2016). It remains an open question whether that competitive edge disappear in a situation where all companies are required to disclose the same information.

Research arguing for/against mandatory CSR reporting (e.g. Habek, 2013; Gatti et al., 2018) does not clearly indicate how mandatory reporting requirements would affect CSR reporting. In order to fill that gap in the existing literature, we investigate the effects of the NFI directive on corporate social responsibility (CSR) reporting. In our study, we examine how, if at all, has the implementation of the non-financial reporting Directive influenced the CSR disclosures of organizations falling under the scope of the Directive? This topic is examined in the empirical context of Finland because of two major reasons. First, except for a dozen state-owned companies, CSR reporting has thus far been voluntary, thus enabling the analysis of the effects of mandatory regulation on the practice. Second, the small size of the country makes it possible to

analyze the annual and CSR disclosures of all organizations falling under the scope of the Directive. The country has relatively long CSR reporting history and the listed companies in Helsinki Stock Exchange have been the most active in disclosing the sustainability information in the international comparison of stock exchanges. The empirical analysis undertaken consists of two parts: a content analysis of all the relevant CSR reports (88 companies, 226 reports) disclosed during the years 2016-2018; and a thematic analysis of interviews conducted with the representatives of 11 companies, both before and after the implementation of the Directive.

Literature Review on Voluntary Versus Mandatory Disclosure of Non-Financial Information

Firms' voluntary disclosure of social and environmental information (also known as CSR reporting or sustainability reporting) has fascinated accounting scholars for several decades. Especially the drivers and motivations for such disclosures have been studied extensively, from various theoretical perspectives. More generally, accounting disclosure theory posits that reporting regimes emerge endogenously in order to alleviate the problem of information asymmetry between managers and financers (Healy & Palepu, 2001). In the case of social and environmental reporting, managers are considered to possess private information that is unavailable to, but relevant for the decision-making of, financers as well as other stakeholder groups such as regulators, suppliers, customers, analysts, non-governmental organizations and the broader society. It has been proposed that in such a situation, voluntary disclosure might be used as a mechanism to signal that a firm has "nothing to hide", thus avoiding an adverse market reaction (Brammer & Pavelin, 2004) that might have eventuated from non-disclosure. In some studies, CSR disclosure has been found to correlate with a firm's social and environmental record (e.g. Mahoney et al., 2012) or future financial performance (Lys, Naughton & Wang, 2015).

Alternatively, it has been suggested that firms voluntarily disclose social and environmental information in order to ensure their legitimacy, or their social license to operate. In other words, CSR reports are generated in response to pressures exerted by diverse stakeholders (Sinclair-Desgagne & Gozlan, 2003). Research applying the legitimacy lens to the study of social and environmental reporting strongly suggests that voluntary disclosure is insufficient for tackling the major sustainability challenges facing contemporary societies. Numerous empirical studies suggest that companies undertaking CSR reporting are motivated merely by the wish to secure their own private interests (e.g. Aerts & Cormier, 2009; Milne & Gray, 2013) or by reputational/legitimacy concerns (Cho, 2009; Michelon et al., 2015). Thus, the main aim of social and environmental reporting is not to provide signals of forward-looking plans for addressing substantive concerns (Patten, 2012) but to obfuscate readers and deflect criticism (Cho, Freedman & Patten, 2012). Gray, Brennan and Malpas (2014) claim that it is somewhat naïve to expect corporate voluntary self-reporting to question the very premises on which the business world and capitalist market economy is built (see Cho et al., 2015). Thus, it has been argued (e.g. Braam et al., 2016) that mandatory reporting is required to prevent the use of social and environmental disclosures as an instrument of greenwashing. This parallels the finding from financial reporting literature, where mandatory disclosure is considered necessary in instances where managers do not voluntarily disclose information that would increase social welfare (Beyer et al., 2010). On the other hand, voluntary front-runners of CSR reporting have been reported to benefit strategically more than followers (Bhimani, Silvola & Sivabalan, 2016).

Despite these insightful contributions, research arguing for/against mandatory social and environmental reporting (e.g Gatti et al., 2018) does not clearly indicate how mandatory reporting requirements would affect sustainability reporting, especially in previously voluntary reporting environments. The social and environmental impacts of voluntary non-financial disclosure have

been addressed in only a few empirical studies, including both functionalist and more interpretive ones.

In the functionalist stream, Chen, Hung & Wang (2018) analyze the effect of mandatory CSR disclosure in China. More specifically, they examine the impact of the new reporting requirement on firms' economic performance and social externalities, the latter operationalized as the amount of industrial wastewater discharge and sulphur dioxide (SO2) emissions. Their results indicate that in cities with a high proportion of firms falling under the remit of the mandatory disclosure requirement, the post-implementation levels of both wastewater discharge and SO2 emissions were significantly lower than pre-implementation. Moreover, they found that the introduction of mandatory CSR reporting correlated with an increase in CSR related spending. Reflecting on the prominent role of state-owned enterprises in China, Chen et al. (2018) draw the more general conclusion that the institutional environment plays a marked role in shaping the outcomes of accounting regulation.

Moving on to interpretive studies, Bebbington (1999) examined mandatory environmental reporting in Denmark based on the Danish government's evaluation of this exercise. According to her, the Danish experience points out the potential benefits and pitfalls of compulsory environmental reporting. Although it generated notable environmental, economic and organizational benefits, the experiment also fell short in engaging stakeholders in the reporting process. Bebbington (1999) explains this dual outcome through the conceptualization of administrative versus institutional reform. Making environmental reporting mandatory constitutes an administrative reform that can prove successful as such but does not yet result in a transition towards sustainability. In order to environmental reporting to yield far-reaching societal changes, Bebbington (1999) argues, a deeper institutional or cultural reform is required.

To summarize the above review of relevant literature, there is relatively little research on the impacts of mandatory CSR reporting in the sense of advancing a transition towards a more sustainable state of affairs. Yet, the small number of such studies that exist as well as corresponding research from financial reporting all point towards the significant role of a country's legal, political and cultural environment in shaping the effects of mandatory disclosure requirements. Our framework for studying these aspects is presented in the following sub-section.

Document Analysis

Based on guidelines developed by the European Commission (2017), we identified 18 aspects to analyze in the NFI disclosures. We identified 88 firms which filled the criteria of the NFI directive to disclose non-financial information and they all were included in the sample of document analysis. According to PwC (2018) there were in total 165 CSR reporters in Finland, meaning that most of the sustainability reporters need to follow NFI requirements. Our field study reports that the majority of the reports were disclosed by old reporters but as much as 19% of the reports (n17) were published by the first-time reporters. As a first indication of the NFI disclosure directive, we can report that that the *absolute number of the reporters have increased*. However, we can not analyze to what extent the NFI disclosure has been the reason to start the CSR reporting or whether these companies would have started it in any case.

The non-financial information is represented the most often in annual report. *Only two of the new reporters provided a minimum information in a separate NFI disclosure* indicating that the first-time reporters also integrate their NFI either in the annual report or have started to publish a more comprehensive sustainability report. About the quarter of the old reporters provided a separate CSR disclosure (this increased from the previous year). GRI is the most popular framework. Increase in *materiality assessments have increased from 66 % in 2016 to 80 % in 2018*. The number of a third- party CSR assurance remained on the same level than before the NFI disclosure

requirement. Our data shows that 29 companies have assured their NFI information while PwC (2018) shows that 39 Finnish companies assured their CSR information in total.

The use of historical data on different NFI aspects has increased from 77% (2015) to 82% (2017) even though in 2016 the frequency of historical data in CSR disclosures was higher than in 2017. The high number of the first-time reporters in 2018 explains the decrease from 2017 to 2018 in that respect.

Based on the document analysis, the *frequency of numerical targets has increased* during the research period. In 2015, around 56% of the CSR disclosures included 'at least some numerical targets', while the frequency in 2017 disclosures was 60%. To be more specific, especially CSR disclosures measuring 'more than three issues with numerical targets' have increased from 38% to 46% during the research period.

The NFI directive requires to disclose information on environmental, social, employee-related, anti- corruption and bribery matters, respect for human rights, and diversity (European Commission, 2017). The document analysis show that the companies *have disclosed more information on human rights, anti-corruption and bribery issues than before*. Traditionally these issues have not been experienced to be the most problematic not relevant in the Finnish corporation culture and therefore they may have not been reported intensively before.

Almost half of the sample companies (42/88) companies indicate either in the *CEO's report or Board's report of activities on the actions caused by NFI disclosure requirement*. This is a signal that the issue has become to the agenda of the top management of companies.

In sum, the Directive seems to *have had no dramatic impacts on CSR disclosures while incremental changes can be identified*. First, the number of CSR reporters has increased by 19%. Second, the increase in materiality assessments increased from 66 % in 2016 to 80 % in 2018. Third, amount of a quantitative historical data and numerical targets have increased. Fourth, more information on human rights, anti-corruption and bribery issues have been disclosure. Finally, NFI requirements have brought up to the top management's agenda.

Next, we analyze the interviews because the document analyses tell us only what can be observed from outside the company. Interviews will give us deeper understanding how the NFI disclosure requirement have influenced inside the organizations, including actions, practices and behavior that is not (yet) seen from outside the company.

Preliminary Findings from the Interviews

Even though the document analysis did not show dramatic changes in the outputs of the CSR reporting due to the NFI directive implementation, our interviews indicate that it has had influence on the internal development work within companies. We interviewed experts from 11 companies before and after the NFI implementation as described in Appendix 1. In the interviews, we identified several internal development processes which are related to the following aspects of CSR reporting: organizational level and structure of reporting, data collection processes, risk management practices and systems, and role of CSR assurance.

Reporting obligation is considered as a driver for improvement:

I'm not saying we had acted irresponsibly earlier, but it was this [reporting] obligation which really made us consider things. So who's responsible, are these issues a part of finance or risk management or what. (Interviewee #1, Finance Development Manager)

NFI reporting as part of Board of Directors' report of activities considered a good thing because it forced **to think about the responsibilities of the CSR** in a big picture:

It (CSR) became part of the Board's agenda, they looked really closely at it, which I think is excellent because ultimately, they are the ones responsible [for what the firm does]. (Interviewee #1, Finance Development Manager)

"Well, that have been the change in practice. The NFI text now goes through the CFO on the CEO's desk before it goes to the Board. That's a kind of a new thing, and it increases attention to this issue. But how does it affect the work and actions of the Executive Board, I can't say based on this first round, you see it maybe a couple of years from now." (Interviewee #9, Responsibility Consultant)

After the NFI implementation, **CSR is discussed more on the top level** of organizations instead of being only the responsibility of a sustainability manager and a separate sustainability team. The NFI directive has brought the CSR on the top of the organizations, at least formally, due to compliance requirement. As a result, **the audit committees in addition to the Boards are involved in CSR.**

"Well, definitely management is now more interested in reporting than before, because this is now such a compliance requirement and it's one good way to get things up to the top. Yes, it definitely has increased interest." (Interviewee #7, Corporate responsibility data and reporting specialist)

"Even if there would be two identical reports, before and after the directive, the later one must be approved by the Board because it is included in the Board of Directors' report of activities. And I know that in many companies the audit committee has taken this on their agenda. In particular, the AC chairs drive this, and it is really advanced in some companies." (Interviewee #9, Responsibility Consultant)

In many companies, the NFI disclosure requirement also **increased cooperation between the finance department and the sustainability team** which has usually been responsible for sustainability reporting. As the finance department has typically provided some numbers to GRI reporting framework, for example, now they were more intensively thinking about the relevant NFI indicators with the sustainability team.

"Both the finance people and the sustainability team, which has traditionally produced the reports, had to face this new issue of NFI and look it from the other perspective than usually and see it in a new format." (Interviewee #9, Responsibility Consultant)

Since the Board's signature is required, they indeed **have become more careful on what do they sign-in** as the following quote by Responsibility Advisor illustrates:

"The Board's signature is now required so it did affect the accuracy of reading the text of sustainability report that was given for them to read and comment. So, we got more comments on the report than ever before. " (Responsibility Advisor)

It seems that the NFI reporting is taken seriously not only as a formal document, but it has increased internal communication of how to communicate on NFI performance to stakeholders. More importantly, the NFI disclosure requirement **have caused changes in organizational structure** harmonizing practices on the organizational level and bringing sustainability issues upper in the organization.

"We now have to form our own working groups around the NFI subject areas. Now we'll have the environmental working group on the whole group level while in the past that used to be only on business unit level, and then we now have a new working group of responsible sourcing, and then responsibility will become to the management team's agenda on every quarter." (Interviewee #8, Development Manager)

The NFI disclosure requirement have caused some companies to invest in their data collection systems and processes quite heavily.

"My role and why I have been hired here is not only related to this directive, but to develop those processes out there for data collection that they can then be utilized in this reporting. But yes, we have had to start completely new processes because of this directive, and those are the very important things that we want to develop further, because of course we want to report the correct information. So it's the biggest exercise here that we have internally done here during this year and on what we'll be focusing in the future. That's why we're not taking any huge leap forward in this year's report because we want first to have the information that we need and then report that when it is correct. It is the starting point for our reporting that those internal processes are in order and that we can guarantee the accuracy of the information." (Corporate responsibility data and reporting specialist)

As the companies want to be sure about the quality and accuracy of the reported information, these beginners have reported only the minimal information this year. Based on the external document analyses, one could claim that the directive did not cause big change. However, what is happening internally in these companies, seems to be significant development in data collection and process development which can be seen as an improved quality of NFI reporting in the future.

NFI directive has also integrated sustainability activities in some companies. Even though the risk management would have been existing as a separate function, it has now integrated to reporting process:

"Risk management has had any role in sustainability reporting but was now involved with the NFI report. It was certainly the most essential change, to incorporate risk management into the NFI reporting process, and there was a new person who came in and said that in order to be able to take a good stance on this, he has to make a proper investigation on this documenting the description and results of an investigation." (Responsibility Advisor)

National regulator in Finland remained sustainability assurance as a voluntary activity when the NFI directive was implemented. 29/88 sample companies voluntarily assured their CSR information, including NFI, and the number of the companies remained somewhat similar compared to the previous year. Even though the NFI disclosure requirement **did not made assurance more popular, it cause a lot of debate in the companies, especially on the top level** to whom this issue was relatively new in many companies.

> "Yeah, this has been a debate for some companies, especially in the big ones. Some have hoped for or wanted to have the NFI information assured before they agree to sign-in the report. But that has been challenging due to the fact that not all figures have been completed before the assurance would take place. But yeah there have been a number of cases where the Board has come up with a message saying they somehow want to verify that information is correct before they put it under their name." (Responsibility Consultant)

Based on the comment above, it could be assumed that the NFI assurance might become more popular in the future when on-going development on data collection systems and processes are completed. Some of the companies hired sustainability assurance providers to write an expert option to the Board to calm them regarding the NFI reporting if the assurance was impossible to implement:

"We had no assurance, but the only thing we did - and that came from the Board and the audit committee - we hired an expert who is more knowledgeable about this [NFI assurance] and who does it to other companies and we asked him to write a statement that has eased the Board's pain. He states that he thinks our [NFI] reporting meets all those requirements, but that's not assured. It's pretty free-form statement, more like a memo-type. That he believes, and according to his experience and view, our report fulfills these requirements of the law, though it is not in itself an assurance. In a way, it does not commit him to anything, but he has watched it and worked with us and want to reassure our Board by saying that he thinks everything is okay with what he has seen." (Interviewee #8, Finance Development Manager)

In sum, the interview findings indicate the Directive gave a stronger mandate (legitimation) for CSR in companies. The results of the study indicated that the changes in the disclosure information have not yet been extremely significant but the NFI directive has done sustainability more important in the organizations and it has been taken more seriously beyond the sustainability team. Several internal changes in the organizational structure and data collection processes are not realized yet but the consequences in external reporting can be expected to see in near future.

Discussion and Conclusions

Major motivation to this study is to examine on how mandatory reporting requirements affect CSR reporting. Our study builds on the research arguing for/against mandatory CSR reporting (e.g. Habek, 2013; Gatti et al., 2018) which does not clearly indicate how mandatory reporting requirements affect CSR reporting. In order to fill that gap in the existing literature, we examine the effects of the NFI directive on corporate social responsibility (CSR) reporting. In our study, we examine how, if at all, has the implementation of the non-financial reporting Directive influenced the CSR disclosures of organizations falling under the scope of the Directive? This topic is examined in the empirical context of Finland which has relatively long history of voluntary CSR reporting. Our empirical findings highlight the importance of data triangulation. The document analyses based on disclosed NFI reports indicate relatively limited changes in disclosed information. However, in-depth interviews in companies reveal several organizational changes and on-going development projects that are not yet visible to external stakeholders. Thus, the examination of published NFI reports on the first year after the regulation took place will not give a comprehensive view on the consequences of mandatory reporting. The outcomes can expect to be seen in near future. The interview data show that NFI disclosure requirement has had influence on the internal development work within companies. We identified several internal development processes which are related to the following aspects of CSR reporting including organizational level and structure of reporting, data collection processes, risk management practices and systems, and role of CSR assurance.

The earlier literature has claimed CSR reporting to ensure companies' legitimacy, or their social license to operate. In other words, CSR reports are generated in response to pressures exerted by diverse stakeholders (Sinclair-Desgagne & Gozlan, 2003). Our study shows that mandatory CSR requirement will legitimate the sustainability itself among the top management in Finnish companies. Boards and Audit Committees are now responsible to sign-in the NFI disclosure and be responsible of the disclosed information. While the earlier research strongly

suggests that voluntary disclosure is insufficient for tackling the major sustainability challenges facing contemporary societies, our case shows that the mandatory disclosure works in Finland as the regulation aimed for. Since top management is now involved in CSR reporting, they have to secure their own private interests (e.g. Aerts & Cormier, 2009; Milne & Gray, 2013) and be concerned by their reputational/legitimacy concerns (Cho, 2009; Michelon et al., 2015) as the earlier literature suggest. In sum, the interview findings indicate the NFI Directive gave a stronger mandate (legitimation) for CSR within companies in which the top management had not voluntarily been involved earlier.

References

- Aerts, W. & Cormier, D. (2009). Media legitimacy and corporate environmental communication.
- Accounting, Organizations and Society, 34 (1), 1-27.
- Archel, P., Husillos, J., Larringa, C. & Spence, C. (2009). Social disclosure, legitimacy theory and the role of the state. Accounting, Auditing and Accountability Journal, 22(8), 1284-1307.
- Bebbington, J. (1999) Compulsory environmental reporting in Denmark: an evaluation. Social and Environmental Accounting, 19(2): 2–4.
- Beyer, A., Cohen, D.A., Lys, T.Z. & Walther, B.R. (2010). The financial reporting environment: Review of the recent literature. Journal of Accounting and Economics, 50, 296–343.
- Bhimani A, Silvola H., & Sivabalan P. (2016). Voluntary Corporate Social Responsibility Reporting: A Study of Early and Late Reporter Motivations and Outcomes. Journal of Management Accounting Research 28(2),77-101.
- Braam, G., Weerd, L., Hauck, M., Huijbregts, M., 2016. Determinants of corporate environmental reporting: the importance of environmental performance and assurance. Journal of Cleaner Production, 129, 724-734.
- Brammer, S., & Pavelin, S. (2004). Voluntary social disclosures by large UK companies. Business Ethics: A European Review, 13(2-3), 86-99.
- Chen, Y.-C, Hung, M. & Wang, Y. (2018). The effect of mandatory CSR disclosure on firm profitability and social externalities: Evidence from China. Journal of Accounting and Economics, 65, 169–190.
- Cho, C. (2009). Legitimation Strategies Used in Response to Environmental Disaster: A French Case Study of Total SA's Erika and AZF Incidents. European Accounting Review, 18(1), 33-62.
- Cho, C. H., Freedman, M., & Patten, D. M. (2012). Corporate disclosure of environmental capital expenditures: A test of alternative theories. Accounting, Auditing and Accountability Journal, 25(3), 486–507.
- Cho C.H., Laine M., Roberts R.W. & Rodrigue M. (2015). Organized Hypocrisy, Organizational Façades, and Sustainability Reporting. Accounting, Organizations and Society, 40(1), 78-94.
- Gatti, L. Vishwanath, B., Cottier, B. & Seele, P. (2018). Are We Moving Beyond Voluntary CSR? Exploring Theoretical and Managerial Implications of Mandatory CSR Resulting from the New Indian Companies Act. Journal of Business Ethics (in Press).
- Gray, R. (2010). Is accounting for sustainability actually accounting for sustainability...and how would we know? An exploration of narratives of organisations and the planet. Accounting, Organizations and Society 35, 47–62.
- Gray R., Adams, C.A. & Owen, D. (2014). Accountability, Social Responsibility and Sustainability.
- Accounting for Society and the Environment. Harlow, UK: Pearson.
- Gray, R., Brennan, A. & Malpas, J. (2014). New accounts: Towards a reframing of social accounting.

Accounting Forum, 38(4), 258-273.

- Healy, P.M., & Palepu, K.G., (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. Journal of Accounting and Economics 31, 405–440.
- Lys, T., Naughton, P. & Wang, C. (2015). Signaling through corporate accountability reporting.
- Journal of Accounting and Economics, 60(1), 56-72.
- Mahoney, L.M., Thorne, L., Cecil, L. & LaGore, W. (2013). A research note on standalone corporate social responsibility reports: Signaling or greenwashing? Critical Perspectives on Accounting, 24(4–5), 350-359.
- Michelon, G., Pilonato, S. & Ricceri, F. (2015). CSR reporting practices and the quality of disclosure: An empirical analysis. Critical Perspectives on Accounting, 33, 59-78.
- Milne, M. J., & Gray, R. (2013). W(h)ither ecology? The triple bottom line, the Global Reporting Initiative and corporate sustainability reporting. Journal of Business Ethics, 118(1), 13–29.
- Patten, D. M. (2012). White tigers, zoos and sustainability reporting: A cynical reflection. Social and Environmental Accountability Journal, 32(1), 17–25.
- PwC (2018). Yritysvastuubarometri. https://www.pwc.fi/fi/julkaisut/tiedostot/pwcn- yritysvastuubarometri-2018.pdf
- Sinclair-Desgagne, B. & Gozlan, E. (2003). A Theory of Environmental Risk Disclosure. Journal of Environmental Economics and Management 45(2), 377-393.

APPENDIX 1. INTERVIEWS

The names of the interviewees and their companies are not disclosed in this study to keep the answers anonymous. The companies have been given specific acronyms that are used in the text.

#	Date	Company acronym	Type of company / interviewee	Record length
#0	20.9.2016		University lecturer/ researcher (background interviews)	
#1	19.10.2016	Company Mercury	Reporting company	60 minutes
#2	3.11.2016	Company Sirius	Consulting company	42 minutes
#3	4.11.2016	Company Cassiopeia	Consulting company	60 minutes
#4	15.11.2016	Company Mars	Reporting company	47 minutes
#5	24.11.2016	Company Jupiter	Reporting company	97 minutes
#6	30.11.2016	Company Andromeda	Consulting company	61 minutes
#7	2.12.2016	Company Saturn	Reporting company	55 minutes
#8	8.12.2016	Company Uranus	Reporting company	86 minutes
#9	16.12.2016	Company Vega	Consulting company	40 minutes
#10	20.12.2016	Company Neptune	Reporting company	44 minutes
#11	21.12.2016	Company Pluto	Reporting company	54 minutes
#12	2.5.2017		Law firm (background interview)	

Appendix 1a. Interviews conducted before the law came into force in Finland

	Date	Company acronym	Type of company	Record length
#5	14.3.2018	Company Jupiter	Reporting company	55 minutes
#8	26.3.2018	Company Uranus	Reporting company	83 minutes
#10	4.4.2018	Company Neptune	Reporting company	50 minutes
#1	6.6.2018	Company Mercury	Reporting company	81 minutes
#4	17.9.2018	Company Mars	Reporting company	42 minutes
#9	5.10.2018	Company Vega	Consulting company	37 minutes

Appendix 1b. Interviews conducted after the law had come into force in Finland



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings





Designing visualizations to identify and assess correlations and trends: An experimental study based on price developments

Lisa Perkhofer*, Conny Walchshofer, Peter Hofer University of Applied Sciences Upper Austria

Abstract

Alongside the increase in available data, long histories, and the need to look at unconventional investment strategies (high risk and low risk by focusing on parallel or opposing stock price developments), multiple visualization options have emerged. This is caused by the ability of visualizations to provide insights such as an accurate and efficient assessment of possible correlations and trends. This study focuses on an optimal way to visualize correlations between two officially listed price developments (stock prices, indices, and commodity goods). In this regard, the choice and the design of the visualization used can influence decision accuracy substantially, however, explicit effects on visualization use and design choices are mostly lacking. To fill this gap, this study tests two highly recommended visualization types (a scatterplot and a parallel coordinates plot) and three concrete design features (regression line – yes vs. no; color – mono vs. multi; interaction – filter vs. select). Although the results indicate that scatterplots outperform parallel coordinates plots in all design conditions, parallel coordinates plots are less affected by deviations from a normal distribution (measured by kurtosis and skewness) and with increasing experience they might be equally effective.

Introduction

An increasing trend when visualizing large amounts of data is the utilization of visualizations that allow a fast and accurate detection and assessment of correlations between two or more variables (e.g., an optimal portfolio depends on the understanding of correlations between various financial assets) (Lehmann, Kemmler, Zhyhalava, Kirschke, & Theisel, 2015). Visualizations therefore help in gaining an abstract overview and allow for an enhanced amount of data to be analyzed at a given time (Federico, Heimerl, Koch, & Miksch, 2017). In other words, the higher the amount of data, the more important it is to present the data in a visual form (Perkhofer, L., Hofer, P., Walchshofer, C., Plank, T., & Jetter, H.-C., 2019). Nevertheless, choosing and designing specific visualizations for specific purposes is indispensable (Falschlunger, Lehner, & Treiblmaier, 2016; Speier, 2006; Vessey & Galletta, 1991). Only if the right visualization types are used to foster the recognition of patterns can a minimum cognitive load on working memory and an enhanced performance be the consequence (Falschlunger, Treiblmaier, & Lehner, 2015; Lehmann et al., 2010).

In particular when dealing with financial data and assessing them with (un-)supervised machine-learning algorithms, visualizations that graphically display correlations might support the user in a more efficient and effective decision-making process (Keim et al., 2008; Li, Martens, & van Wijk, 2010). One of the oldest and simplest forms of visualizing correlations is the scatterplot (SP) visualization (Lewandowsky & Spence, 1989; Sarikaya & Gleicher, 2018). SPs

display data as a collection of points forming a pattern; they are a popular and widespread visual representation not only because of their simplicity, but also because of their familiarity and visual clarity (Abi Akle, Yannou, & Minel, 2019; Urribarri & Castro, 2016). Within the data mining process, SPs are considered to be an ideal early visualization to scan new datasets (Elmqvist, Dragicevic, & Fekete, 2008; Sarikaya & Gleicher, 2018) as they not only indicate correlation but also highlight the distribution and clearly depict outliers (Abi Akle et al., 2019). However, as datasets grow in complexity and scalability in a big data related environment (becoming multidimensional), SPs are said to rapidly become ineffective (Sarikaya & Gleicher, 2018). "Even if we employ 3D graphics as well as point color, shape, and size as graphical properties, a standard scatterplot diagram can only visually represent a handful of data dimensions at a time" (Elmqvist et al., 2008, p. 1141). Alternatives which are more effective in terms of dealing with multiple dimensions include the dense pixel display, dimensional stacking, stacked charts (stacked line, area, bar, or column charts) and the parallel coordinates plot (PCP) (Harrison, Yang, Franconeri, & Chang, 2014).

The PCP has been created with the intention to display multiple dimensions within one chart (Inselberg, 1985) and is the most frequently mentioned and applied alternative, especially in the Information Visualization community (Claessen & van Wijk, 2011; Elmqvist et al., 2008; Harrison et al., 2014; Johansson & Forsell, 2016; Kuang, Zhang, Zhao, & McGuffin, 2012). The design is based on perception studies, which indicate that the human eye is highly sensitive to identification of horizontal and/or crossing lines. While horizontal lines indicate a positive correlation, crossing lines stand for a negative one (Diehl, Beck, & Burch, 2010). The emerging pattern can only be analyzed between neighboring axes (Johansson & Forsell, 2016). Although PCPs can be used with only two or three attributes, comparable to SPs, they can easily be extended to a considerably larger number by adding additional parallel placed axes which is one of the major advantages of this visualization type (Harrison et al., 2014).

Despite their usefulness in larger and multidimensional datasets, research on PCPs is still limited while the claim of SPs becoming ineffective and the push for alternatives persists in literature (Harrison et al., 2014; Rensink, 2017). Current research does not give a clear indication of whether one visualization type is better than the other and almost no recommendations can be found concerning their optimal design (Johansson & Forsell, 2016; Kanjanabose, Abdul-Rahman, & Chen, 2015). Further, studies are mostly focused on value retrieval tasks (Abi Akle et al., 2019; Kanjanabose et al., 2015; Kuang et al., 2012; Netzel et al., 2017), while limited attention has been paid to correlation identification and assessment (Harrison et al., 2014). Experiments are also frequently based on normal distributed and computer generated data as opposed to real-world datasets, limiting the external validity of study results (Harrison et al., 2014).

This study contributes to the body of literature by testing specific design features (integration of a regression line, use of different color schemas, use of different interaction techniques) for both mentioned visualization types and comparing their respective performance based on estimation accuracy. Thus, the fundamental goal of this study is to test whether PCPs can – if designed optimally – be seen as a promising alternative to SPs, in particular when performing correlation identification tasks with financial data under varying but real-life conditions (positive or negative correlations, strong or weak correlations, high or low amount of data). During the analysis we ascertained that not only the design but also the amount of data being displayed and the distribution of the variables under investigation strongly influence a decision-maker's ability to detect and assess correlations correctly. In particular, the latter (effect of distribution) is a new but very important finding, influencing the usefulness of PCPs as they seem to be rather unaffected by deviations from normal distribution while SPs are affected significantly by excess kurtosis and skewness. While excess kurtosis negatively affects estimate accuracy, skewness enhances performance.

The remainder of this paper is structured as follows: first, the two visualization types are introduced in detail before a discussion of previous research comparing PCPs and SPs. Based on these findings, hypotheses for our empirical investigation are deduced. In the Method chapter, the method and the stimuli material are described in detail, with the results presented in the Results chapter. Finally, in the chapters Discussion and Conclusion and Limitations and further Research Opportunities the results are examined and further research opportunities as well as the limitations of our experimental design are laid out.

Theoretical background

Before being able to compare the two visualization types, namely SP and PCP, we analyze them and their areas of application in more detail in this chapter. Further, their respective strategies for identifying correlations are explained.

The scatterplot

An SP depicts two variables by presenting a collection of points (Elmqvist et al., 2008) on two continuous, orthogonal dimensions and was designed to emphasize the spatial distribution of the data presented (Sarikaya & Gleicher, 2018). The visualization uses positions to encode the respective values of two variables in a two dimensional space and allows for patterns and correlations to become visible (Abi Akle et al., 2019). This means one variable is presented on the vertical (also called the Y-axis) and the other on the horizontal axis (also called the X-axis). A data point is drawn for each pair of values. The collection of points forms a pattern (or visual guidance rules), which indicates the kind and degree of correlation that might exist between the two presented variables (Lehmann et al., 2015). The degree of correlation is described by the correlation coefficient (r), which ranges from -1 to +1. A perfectly negative correlation has a coefficient of -1 (see Figure 1 on the left-hand side) and a perfectly positive one a coefficient of +1 (see Figure 1 presented on the right-hand side). Therefore, three main correlation types have to be distinguished:

- a) **Strong negative correlation**: If two variables show a strong negative correlation, it means that if an increase is visible in the first variable a decrease is visible in the second one. The visible pattern starts in the upper left corner and ends in the bottom right.
- b) **No correlation**: If two variables do not move in the same or the opposite direction, no clear connection between them can be associated. Therefore, the data points form the shape of a cloud.
- c) **Strong positive correlation**: If two variables show a strong positive correlation, the increase in one variable is also visible in the second one. If both variables increase, a diagonal pattern starting at the bottom left corner and ending at the upper right corner will appear.



Figure 1: Kind of correlation and correlation coefficient of scatterplots by Li et al. (2008) Left: Negative correlation (r=-1); Middle: No correlation (r=0); Right: Positive correlation (r=1)

The SP is a very popular and familiar visualization technique (Elmqvist et al., 2008) as it is used very early on in elementary school textbooks to support information transmission; as a result, users usually show extensive experience when interpreting this particular chart type (Harrison et al., 2014). Further, SPs are widely available in commercial visual analytics software (e.g., Tableau, Microsoft PowerBI, Qlik) as well as in Microsoft Excel (Elmqvist et al., 2008), which is still the most common tool for (primary) data analysis as well as for visualizing financial data (Perkhofer, L. M., Hofer, P., Walchshofer, C., Plank, T., & Jetter, H.-C., 2019).

SPs mostly depict bi-dimensional datasets (Sarikaya & Gleicher, 2018), however, they can also be extended to multidimensional and larger ones (Urribarri & Castro, 2016). Admittedly, this extension to multidimensionality is limited and the number of data points visible within one chart is also restricted (Elmqvist et al., 2008). These two aspects – the limited number of distinguishable attributes and the reduced number of data points – are the two major drawbacks of the SP visualization (Urribarri & Castro, 2016).

The parallel coordinates plot (PCP)

A PCP depicts two or more variables and links them on two (or more) vertically aligned axes using polylines (Abi Akle et al., 2019; Johansson & Forsell, 2016; Netzel et al., 2017). Each axis represents one variable and the PCP can therefore be seen as a tool to provide a continuous view on multivariate datasets (Inselberg & Dimsdale, 1990; Kuang et al., 2012). "It avoids the limits of orthogonal coordinate systems by placing each axis of coordinates in parallel" (Abi Akle et al., 2019, p. 234). By doing so, a PCP can be employed to detect outliers, trends, clusters and correlations. However, the greatest strength lies in its ability to quickly provide an overview of multivariate datasets (Johansson & Forsell, 2016). Nonetheless, interpretation can be limited as only neighboring axes give clear indications about correlations and trends (Lehmann et al., 2010). If more axes are displayed it is therefore of great importance to use a flexible layout where the participant can modify the displayed order of the axes, include or exclude variables, and highlight specific aspects of the dataset (Hofer, Walchshofer, Eisl, Mayr, & Perkhofer, 2018; Johansson & Forsell, 2016; Netzel et al., 2017).

Further, and more importantly in the context of this study, correlations "can quickly be identified at a glance" (Abi Akle et al., 2019, p. 234). Each data pair (e.g., choosing the value of the same day for both variables, the same person, the same sale...) is represented by a line between these two axes. The collection of lines representing each data pair forms a pattern, which can be informative in terms of detecting the kind and the degree of correlation that might exist between the two variables (see Figure 2 on the right-hand side) (Lehmann et al., 2015; Li et al., 2008; Netzel et al., 2017).

- a) **Strong negative correlation**: If two variables show a strong negative correlation, an increase in the first variable and a decrease in the second one is visible. The visible pattern is represented by significantly crossing lines.
- b) **No correlation**: If two variables do not move in the same or the opposite direction, no clear connection between them can be associated. A mix of horizontal and crossing lines is visible.
- c) **Strong positive correlation**: If two variables show a strong positive correlation, an increase in the first as well as in the second variable is identifiable. If both variables increase, all lines (or the majority of lines) are drawn horizontally.



Figure 2: Kind of correlation and correlation coefficient of PCPs by Li et al (2008) Left: Negative correlation (r=-1); Middle: No correlation (r=0); Right: Positive correlation (r=1)

The PCP was introduced to the InfoVis community by Inselberg in 1985 and has been used almost exclusively within visualization research (Kanjanabose et al., 2015). The community hypes the visualization because of its versatility and the easy detection of correlations (visual assessment of horizontal and crossing lines) (Diehl et al., 2010). Despite encouraging findings for multidimensional datasets and complex tasks (clustering, outlier detection, cluster analysis) (Kanjanabose et al., 2015), the visualization has not fully convinced other research disciplines. As the complexity of datasets and tasks has become more challenging for financial performance management, it might be time to investigate their use in more detail.

Hypotheses development

Some studies have already investigated the difference between SPs and PCPs depending on different scenarios. However, only a few have used a correlation estimation task within their experimental design (Abi Akle et al., 2019; Harrison et al., 2014; Li et al., 2008), only one study has employed real-life data (Henley, Hagen, & Bergeron, 2007), and none of the studies have tested effects on different design features (such as color use or the use of integrated interaction techniques). An overview of identified studies comparing SPs and PCPs, their experimental tasks and results are presented in the following table. The findings are discussed in more detail in the subsection below, which provides the basis for our hypotheses and consequently the empirical investigation.

Author	Experimental task	Result
(Henley et al., 2007)	Genome identification task and preference survey (small usability test, 7 participants)	PCPs outperform SPs; SPs are subjectively preferred by users
(Li et al., 2008)	Correlation estimation task (lab-experiment, 25 participants)	SPs outperform PCPs in all conditions (strength of correlation, amount of data)
(Holten & van Wijk, 2010)	Cluster identification task	The combined layout PCP including SPs outperforms other PCP layouts
(Kuang et al., 2012)	Value retrieval task	Depending on the dimensionality
	(two small lab experiments with think aloud protocols for triangulation, 12 and 18 participants)	 and data density: low dimensions and low data density: PCPs outperform SPs; high dimensions and high data density: SPs outperform PCPs
(Harrison et al., 2014)	Correlation estimation task	Overall SPs outperform PCPs; for
	(large scale online experiment using MTurk, 1,687 participants)	positive correlations SPs show better results, for negative correlations PCPs do better (especially if correlations are weak)
(Kanjanabose et al., 2015)	4 different task types: value retrieval, clustering, outlier detection, and change detection; no change in data dimension or number of data points	PCPs outperform SPs in clustering, outlier detection and change detection; no difference in value retrieval
	Constant: 4 data dimensions, 8 data	
	(lab-experiment; 42 participants)	
(Netzel et al., 2017)	Value retrieval task; judging distances between points	Results depend on data dimensions (2, 4, 6, 8):
	(Eye tracking lab-experiment, 24 participants)	 SPs (matrix) in low dimensions (2, 4, 6) PCPs with high dimensions (8)
(Abi Akle et al., 2019)	 Data exploration (first assessment, outliers) Data discovery (detailed analysis, correlations, clusters) 	For data exploration and data discovery, the SP or respectively the SP matrix outperforms the PCP
	(lab experiment, 42 participants)	

Table 1: Previous studies investigating SP vs. PCP

The visualization type effect

An optimal visualization clusters information or supports inferences, which are helpful for the decision-making process (2014). For this purpose, multiple options for visualizing data are available; however, each type has the potential to uncover and present a different type of insight to its audience while at the same time hiding another (Perkhofer, L. et al., 2019). To shed light on the special insight a user wants to uncover, the right visualization needs to be used. Generally the dimensionality, the data type and the task type are indicators of an optimal visualization use (Harrison et al., 2014). As highlighted in the introduction, for the display of correlations a few

visualization types are recommended (Harrison et al., 2014), however, the most common ones cited in literature which are also used in practice are SPs and PCPs (Johansson & Forsell, 2016). Both visualization types enable the researcher to look for visual patterns to uncover correlations and their respective strength (Johansson & Forsell, 2016). They also use a coordinate system allowing them to be "ordered" by default (Harrison et al., 2014). This means values are positioned on the axis using a numerical scale in ascending order starting at the bottom (or left if we talk about the x-axis in a scatterplot) with low values which increase towards the end. The range of the scale primarily depends on the range of the dataset and therefore ideally starts with the minimum and ends at its maximum.

As indicated in our analysis in Table 1, SPs seem to outperform PCPs when estimating correlation coefficients (Abi Akle et al., 2019; Harrison et al., 2014; Li et al., 2008). Nonetheless, taking a closer look the effectiveness of the used visualizations, performance also depends on whether the visualization depicts a strong or a weak as well as a positive or a negative correlation (Harrison et al., 2014; Li et al., 2008). Results from a large quantitative study indicate that estimation accuracy is more precise in PCPs when negative correlations are illustrated (especially if correlation coefficients are below -0.5), while for positive correlation estimation accuracy deteriorates (Harrison et al., 2014). An explanation for this effect can be found when analyzing the visual pattern that appears in a PCP depending on the direction of the correlation coefficient; the intersecting lines of a negative correlation are easier to assess than the horizontal lines from a positive one as crossing lines tend to stand out. Even a low number of intersecting lines negatively affects a participant's decision-making when positive correlations are presented. This bias towards an underestimation of correlation coefficients in PCPs has been called the "diablo effect" (Li et al., 2008, p. 17). In contrast, in an SP the visual form remains the same: a horizontal dot cloud generally pointing upwards (positive correlations) or downwards (negative correlations), however, outliers do not irritate or stand out as much. Therefore, the "diablo effect" is not applicable for the SP visualization (Harrison et al., 2014).

- H1. Overall, SPs show a higher estimation accuracy than PCPs.
- H2. SPs show a higher estimation accuracy than PCPs for positive correlations.
- H3. PCPs show a higher estimation accuracy than SPs for negative correlations.

The effect of design

The more complex the dataset that is represented, the simpler the design of the display should be to avoid distracting the user and drawing attention to unnecessary details (Sarikaya & Gleicher, 2018). Nonetheless, if an additional design feature (e.g., the use of a regression line, the use of supplementary interaction techniques, the use of color) supports the user in executing the task, it can be beneficial (Falschlunger, Eisl, Losbichler, & Greil, 2014). We therefore test three commonly used features to determine whether they can be considered non-data ink (Tufte, 1983), or if they hold meaning and support information processing by supporting the task:

1. Regression line: Annotations, such as a linear regression trendline, can help the decisionmaker to navigate through the dots in an SP as they help order the data and provide additional information. For a correlation estimation task in particular, a linear regression line is helpful as it highlights possible correlations within the dataset (Sarikaya & Gleicher, 2018). Further, the angle of the regression line is indicative of the strength of the correlation if the decision-maker is familiar with this concept. Although frequently used in textbooks and figures, no experimental research investigating the benefits of a linear regression, used for judging correlations in an SP has been identified. In addition, no reference concerning trendlines has been found for PCPs. For a fair comparison, a trend is included in our experimental study for both visualization types (for instructions on how to include a trend in a PCP, we refer to Chapter 4.2 explaining the stimuli material – Table 2, Design 2).

H4. The use of a regression line has a positive effect on estimate accuracy.

2. Color: Color has a huge impact on decision-making accuracy as specific features can strongly influence judgment (Yoo & Smith-Jackson, 2011). For example, the use of different sizes and colors allow for certain visual aspects to "stand out and thus increase speed of identification" (Parsons & Sedig, 2014, p. 464). In particular, color is the dominant channel for humans in order to distinguish between multivariate information. The use of color therefore has an advantage over using different symbols (Li et al., 2010). Nonetheless, colors should not be used too extensively as the advantage diminishes if the number of different dimensions/attributes they are supposed to represent increases (Claessen & van Wijk, 2011; Liu, Maljovec, Wang, Bremer, & Pascucci, 2017; Yoo & Smith-Jackson, 2011). This is attributed to the fact that the decision-maker might already be overwhelmed by making sense of the complex dataset, leaving no processing capacity to comprehend further features or information (Perkhofer & Lehner, 2019). In such scenarios, color blending (the use of different shades of the same color) might be more useful as it has been demonstrated to have a positive effect on the perception of order and structure (Liu et al., 2017).

H5. An increased amount of colors used (to distinguish between dimensions) has a negative effect on estimate accuracy.

3. Interaction technique: Although more dimensions can be included by using different colors and/or shapes, the identification of patterns and clusters in such a cluttered visualization is difficult (Abi Akle et al., 2019). In order to ensure a useful interpretation, the dimension of interest should be isolated to focus the user's concentration on this particular subset of information. Two different possibilities of interaction allow for such an approach: filter or select (Keim, 2001; Perkhofer, L., Hofer, P., & Walchshofer, C., 2019; Pike, Stasko, Chang, & O'Connell, 2009). With filter, the visualization and all axes displayed are resized (depending on the remaining range of the dataset – the minimum and the maximum) and only the subset of data chosen is displayed. With select, the entire dataset is presented, however, deselected items are sent to the background by using a uniform light grey color (for example please see Table 2, Design 4) (Perkhofer, L., Hofer, P., & Walchshofer, C., 2019). Especially in a PCP, the use of filters has presented itself as a very helpful option to isolate specific parts of the dataset (Hofer et al., 2018). Further, from an information-processing related perspective it seems favorable to reduce the amount of data presented as fewer cognitive resources are needed for processing (Perkhofer & Lehner, 2019) and the decision-maker does not have to suppress the data presented in the background during interpretation (Perkhofer, L., Hofer, P., & Walchshofer, C., 2019).

H6. The interaction technique filter shows a higher estimation accuracy than the interaction effect select.

The effect of large datasets

It has long been shown that too much information overburdens a decision-maker (Miller, 1956; Sweller, 2010); however, a certain amount is necessary in order to identify relevant patterns and trends (Falschlunger, Lehner, & Treiblmaier, 2016). The same is true for SPs and PCPs. On the one hand, a certain number of dots in the SP or respectively lines in the PCP might be necessary to detect a pattern (Li et al., 2008; Rensink, 2017). On the other hand, too many dots or lines

might overlap and distort the pattern, making it difficult to detect (Li et al., 2008; Liu et al., 2017; Sarikaya & Gleicher, 2018; Urribarri & Castro, 2016). Overall, there seems to be a trend indicating that large amounts of data negatively affect performance. Despite this, Kuang et al. (2012) indicate that in a data retrieval task, performance is affected more when displayed in a PCP. Li et al. (2008) found similar results in a correlation estimation task. Both studies link this effect to the visual shape of the used charts: dots are more space efficient than lines and therefore are less affected by an enriched data volume. This is a well-known problem when working with PCPs as the visualization becomes cluttered if the dataset is too large, thus hindering the discovery of patterns (Johansson & Forsell, 2016; Liu et al., 2017).

H7. An increased amount of data used has a negative effect on estimation accuracy. H8. SPs show a higher estimation accuracy than PCPs if large amounts of data are visualized.

The dimensionality of the dataset can lead to further challenges in the user's perception. It is proposed that PCPs should be able to depict more attributes simultaneously. Still, the results of empirical studies are inconsistent (Kuang et al., 2012). For our study, although the topic of dimensionality is the logical next step, it is important to investigate the effects of data amount and design first and use optimized visualization types for the investigation of dimensionality. This particular topic is therefore not part of this experiment but is a further research endeavor.

The effect of experience

With experience, the decision-maker can rely on mental representations, which relieve the processing load of the user. Learned schemata on how to work with and read a visualization can be applied and a faster but more importantly more accurate analysis of the data is the consequence (Chandler & Sweller, 1991; Perkhofer & Lehner, 2019; Sweller, 2010). A novice not only needs to put more effort into the understanding of the presented dataset but also into the ways a particular visualization needs to be analyzed in order to retrieve the necessary insights. Many of the visualizations designed for multivariate datasets require extensive practice and initial effort in order for the decision-maker to perform a predefined task (Lehmann et al., 2015). An expert, on the other hand, experiences less "cognitive work", allowing him/her to concentrate solely on locating interconnections and correlations and therefore making him/her superior in a task comparison analysis (Abi Akle et al., 2019).

With respect to the two visualizations used in this study, the SP has clear advantages in terms of experience. An SP has a long history of being used in linear algebra at school, indicating that early training exists. This training is the foundation of a strong developed intuition and knowledge when interpreting data presented visually (Abi Akle et al., 2019). Thus, users appear to have fundamental knowledge regarding correlation perception and estimation in SPs (Rensink, 2017). Judging correlations in an SP is often seen as common knowledge and the participant's accuracy is therefore predicted to be high (Lewandowsky & Spence, 1989). PCPs, on the other hand, are said to be less intuitive to novices (Kanjanabose et al., 2015; Lehmann et al., 2010) and unfortunately they are endorsed in the InfoVis community in situations with high dimensionality, however, their use is still limited (Johansson & Forsell, 2016; Perkhofer, L. M. et al., 2019). PCP are therefore widely unknown to users and "an unskilled user is prone to make wrong decisions" (Urribarri & Castro, 2016, p. 1).

H9. A high level of experience has a positive effect on estimation accuracy.

Method

The purpose of this study is the evaluation of two different visualization types (SPs and PCPs) combined with various design features to test their suitability for correlation assessment (correlation types used: strong positive, weak positive, weak negative, strong negative) (for details see Section The Visualization Type Effect). For each visualization-design combination, a separate but identical experiment was created; as a result, every participant evaluated only one visualization-design combination, but had to assess all correlation types. Further, a variation in the number of data points was introduced to test the influence of low and high data volumes (by using not only the whole dataset but also a subset for analysis; for more details we refer to Section The Dataset). As previous research strongly suggests a connection between the level of experience with specific visualization types and a participant's performance (Falschlunger, Lehner, Treiblmaier, & Eisl, 2016), data on experience has been collected and used as a control variable.

The experimental study was conducted online between April and June 2019 using LimeSurvey and the crowdsourcing platform Amazon Mechanical Turk (MTurk). The results have been shown to be congruent with lab experiments in the context of visual analytics (Harrison et al., 2014). These results convinced us to believe that MTurk is a proper fit for this kind of experimentation. Our research model, including all independent and control variables used within the online experiment, is summarized in Figure 3 and explained in detail in the following sections.



Figure 3: Research model

The dataset

The dataset represents the stock price development of indices or companies as well as price developments of real goods such as coffee, milk, or gold. Each trading day is represented by one data point within the given time period of 3.25 years (2015-2018; 2018 only 3 months). The amount of data presented was manipulated by looking at the whole dataset vs. looking at one particular year within the given time period. Depending on the stock exchange, trading days vary (largest n=833 using 3.25 years; smallest n=53 using 3 months).

The use of real data is a unique feature of our study as most experiments use normally distributed data, which is randomly calculated via software depending on the correlation coefficient under investigation. In this experiment, however, we focused on the use of visualizations for finance analyses of concrete enterprises and goods and therefore real datasets instead of demo data have been introduced. Further, previous experiments mostly focus on smaller datasets (dataset equivalent to one year) (Harrison et al., 2014; Li et al., 2008), while

knowledge regarding larger ones is scarce. However, because the amount of data is of special interest in a big data related context, we also tested with an increased amount of data.

The stimuli material

Two different visualization types and four different designs were tested, resulting in eight separate experimental surveys to reduce each participant's overall time requirement and ensure high levels of concentration. To reduce visual clutter, rather small dots (radius 1.5) and respectively thin lines (width 0.5) with slight transparency (opacity 0.1) were used. The visualization itself always shows two dimensions/attributes, namely the two-stock price or commodity price developments on the respective axis. The third dimension/attribute – years, ranging from 2015-2018 – was included by using different colors (Designs 1, 2, and 4) or different shades of a single color (Design 3). All visualizations had the same physical dimensions (size and position of SVG). The design features are listed in the following table (all visualizations used in the experiment, including their correlations and distribution specifications, are illustrated in the Appendix).

Table 2: Visualization types and designs used within the experiment (example: negative correlation r = -0.62)



L.Perkhofer, C.Walchshofer, P. Hofer / Finance, Risk and Accounting Perspectives Proceedings (2019) 294-340



The experimental task and assessment of the dependent variable

The purpose of the task was to identify the kind and the degree of correlation that might exist between the two variables presented on the x- and y-axes or the left (first) and the right (second) axis, respectively. The actual linear correlation between x and y is indicated by the Pearson Correlation coefficient, which is used for evaluation but naturally invisible to the participants during the experiment. To test which visualization type and design works best to identify correlations, an experimental task targeted to assess the degree of correlation was used (see Figure 4).





For the assessment of accuracy, the deviation of the participants' estimates from the true score (or the actual correlation coefficient) was used. Over or underestimations are treated identically (Harrison et al., 2014), resulting in absolute values for evaluation. Ideally, the deviation from the true score should be close to zero as this would mean the participants correctly estimated the calculated correlation coefficient.

The procedure

To avoid possible confusion due to the use of mobile devices, participants were instructed to access the experimental survey solely from their computers. The experimental questionnaires started with a tutorial explaining the respective visualization types as well as demonstrating examples to illustrate the visual patterns of all relevant correlation types used within the study (obviously, none of the introductory examples were included in the actual experiment afterwards). If further explanation was necessary, the user had the opportunity to watch an instructional video information from a textual description on to retrieve demand (for SPs: or http://www.usivis.org/en/explaining-correlations-in-a-scatterplot-visualization/; for PCPs: http://www.usivis.org/en/explaining-correlations-in-a-parallel-coordinates-plot/). Before starting the experiment, short review questions were implemented to check whether the tutorial had been studied carefully. If more than four out of the six stated questions were answered incorrectly, answers were not used for evaluation. This exclusion criteria mitigates "clickthrough" responses, which often impact crowdsourced experiments.

Further, as participants were recruited on a crowdsourcing platform and because the SPs and more importantly the PCPs might not be understood by all participants, a training session including feedback on the performance was used for all visualization types and designs to provide a warm-up (12 individual judgments). This procedure is also necessary to guarantee a familiar procedure during the main trials. For each experimental session, four positive (r > 0.6), four negative (r < -0.6), and four weak correlations (-0.6 < r <= -0.2; 0.6 > r >= 0.2) were used and presented in random order. A difference between high and low correlations is important as correlations above 0.6 are easier to estimate than lower correlations irrespective of the visualization type (Harrison et al., 2014). No time limit was imposed on the participants. After the experimental tasks were finished, demographic questions on age, gender, and experience were posed.

The participants

To ensure high quality, qualification levels for participation on MTurk were specified: participants had to have at least a Bachelor's degree. For their time, participants were compensated with \$5 per experimental questionnaire which took participants approximately 30 minutes (average: SP 36 min, average PC: 26 min) to finish. We recruited n=388 participants in total. One participant could only complete one experimental survey, which means that the user had to work with one design-visualization combination; this left us with roughly 50 participants per group. Details on demographics are presented in Table 3. A total of 4,692 correlations were assessed and used for analysis.

Designs	SP participants	PCP participants
Design 1: default	Total: 50	Total: 52
gender (% female)	44.9%	40.4%
Ø age	37.06	36.63
Ø experience	1.14	0.54
Design 2: regression-line	Total: 45	Total 49
gender (% female)	42.2%	44.9%
Øage	34.93	37.84
Ø experience	1.22	0.18
Design 3: color	Total: 52	Total: 43
gender (% female)	48.1%	51.2%
Øage	37.33	37.67
Ø experience	1.40	0.40
Design 4: interaction technique	Total: 50	Total: 47
gender (% female)	42.0%	40.4%
Ø age	35.28	36.21
Ø experience	1.42	0.64

Table 1: Response rate per experimental survey and demographic information

Unfortunately, in our experiment no change in experience can be observed with respect to previous research, which suggests very low levels of familiarity with PCP (Harrison et al., 2014; Kuang et al., 2012; Li et al., 2008). While experience with SPs (scale from 0-4) is mediocre, experience with PCPs is almost not existent (see Table 4; the difference in experience is significant: Students t-test p=0.000, T=26.23). This low level of experience with PCPs needs to be considered in data analysis and interpretation.

Table 4: Details on experience with the visualization types used

Experience w	vith Scatterplots	Experience wi	th parallel coordinates plots
0	19%	0	62%
1	43%	1	33%
2	29%	2	3%
3	7%	3	2%
4	2%	4	0%

4.6. Necessary adaptions to the experimental design

After a preliminary and quick analysis of the obtained results based on the correlation estimation task of the whole dataset (4 years), it became clear that other influences besides the intended manipulations (visualization type, design, number of data points and correlation types) introduced did influence results. For this initial and quick analysis, estimated values within a range of +/-15% of the true score (calculated correlation coefficient) are coded with one (for correct), while values outside this tolerance interval are coded with zero (for incorrect).

This quick analysis demonstrates that SPs outperform PCPs for the correlation estimation task (overall accurate estimates SP: 49% vs. PCP: 42%). Additionally, it is evident that the use of a linear regression line might be beneficial not only in an SP but also in PCPs. Moreover, performance in SPs is substantially better when confronted with positive correlations (accuracy: 54%) in comparison to negative correlations (accuracy: 43%) while this difference is not visible for PCPs (accuracy: 41% or respectively 42%). The following tables summarize the estimation accuracy results based on this quick initial analysis.

Table 5: Estimates within +/- 15% of actual correlation using SPs (4 years)

	r <= -0.8	r <=- 0.6	r <= -0.2	r > 0.2	r > 0.6	r > 0.8	Overall
Design 1	66%	18%	8%	44%	52%	86%	46%
Design 2	69%	44%	31%	60%	64%	56%	54%
Design 3	48%	33%	48%	27%	56%	67%	46%
Design 4	80%	38%	34%	24%	46%	70%	49%
Ø neg. vs. Ø pos.		43%					49%

Table 6: Estimates within +/- 15% of actual correlation using PCPs (4 years)

	r <= -0.8	r <=- 0.6	r <= -0.2	r > 0.2	r > 0.6	r > 0.8	Overall
Design 1	50%	38%	27%	19%	37%	38%	35%
Design 2	80%	49%	24%	31%	35%	65%	47%
Design 3	63%	26%	21%	30%	35%	70%	41%
Design 4	64%	38%	26%	34%	43%	57%	44%
Ø neg. vs. Ø pos.		42%			41%		42%

However, paying closer attention to the obtained results, differences between Design 1 and Design 4 should not exist in the 4 year data condition period for either the SP or the PCP. Design 4 should only show substantially different results when a subset of the overall information – one selected year – is presented to the user because all other factors remain constant compared to Design 1. Therefore, in the presented initial analysis the exact same visual design and data density is presented to the user. Thus, it is not only the correlation type (positive or negative, weak or strong) and the design that influences the appearance of the visual pattern, but on closer inspection there also appears to be a variance attributed to the distribution of each variable being displayed.

A detailed analysis of the dataset revealed that the further the distribution deviates from normality, the harder it becomes to visually assess the correlation coefficient. Therefore, common measures describing the distribution of a variable – excess kurtosis (heavy tails and peakedness) and skewness (peak shifted left or right) – have been introduced as controls, resulting in the following and adapted research model for the analysis presented in the Results Section.



308

Measurement models

For a better understanding, the following table summarizes all variables and their measurement models. Details on the distribution of the used variables per visual stimuli, including histograms and values for both additional control variables, are presented in the Appendix.

Table 7: Summary of variables and measurement models and coding Variable Description Measurement and coding Visualization Two discrete variables are used to distinguish SP = 1between SPs and PCPs PCP = 2type independent Variable Trend: Three discrete variables are used to measure the Design No regression line included = effect of the tested designs: linear regression line, independent 0 color, and interaction technique Regression line included = 1Variable Color: Monochrome = 1Multi-color use = 2Variable Interaction: Filter = 0Select = 1Continuous scale; depending on the stock and the Min = 52Number of data amount of years used, different values for trading points Max = 833 days are the result. independent +Quadratic effect Research always suggests that both extremely small and large amounts of data are bad for decision-making. To account for this influence, we also test for a quadratic effect. Correlation Two discrete indicators are used to measure the Indicator direction: negative = 1correlation type: one to account for a negative or a types positive correlation and one to account for the positive = 2independent degree of correlation. Indicator magnitude: • weak correlation $(-0.6 < r \le -0.2; 0.6 > r \ge -0.2; 0.6 > -0.2; 0.2 > -0.2$ 0.2) = 1• medium correlation (-0.8 < r <= -0.6; 0.8 > r >= 0.6) = 2• strong correlation $(r \le -0.8; r > 0.8) = 3$ **Deviation from** Continuous scale: compares the estimate of the Min = 0true score participant with its true score on a scale from -100 Max = 181 (someone entered (perfect negative correlation) to +100 (perfect a positive instead a negative dependent positive correlation) correlation or vice versa) No difference is made if the estimate is below or above the actual value (values are absolute!) Excess kurtosis Continues scale: for each of the variables used in Min = 0.11the correlation analysis, excess kurtosis is Max = 5.47control calculated. Kurtosis deviates from normality if the +Quadratic effect curve is flat (below 0) or steep (above 0). Therefore, absolute values are used to indicate a

deviation from normality; from the two values

Variable	Description	Measurement and coding
	calculated per correlation, the higher one is used for further analysis.	
	Further, we observe a severe degradation of results with high kurtosis levels, indicating that this effect might not be linear. As a result, we also test for a quadratic effect.	
Skewness	Continues scale: skewness is calculated for each of	Min = 0.06
control	the variables used in the correlation analysis. If	Max = 2.42
	peak is shifted to the left (above 0) or to the right (below 0); as a result, absolute values are used to indicate a deviation from normality. From the two value calculated per correlation, the higher one is used for further analysis.	<u>+Quadratic effect</u>
	Also for skewness, we test for a quadratic effect to see whether an increase in skewness influences estimation accuracy to a higher degree than a linear effect would suggest.	
Experience	Discrete values: entered level of experience by the	Mean level of experience for
control	user (self-assessment) on a 5-point Likert scale (0: no experience at all – 4: high experience)	SP = 1.30 Mean level of experience for PCP = 0.44
		Also see Table 4

Results

Analysis is based on estimate accuracy. Visualization comparison is based on the Student's t-test (see Section Comparison between SPs and PCPs), while for the evaluation of the proposed research model structural equation modelling (SEM) is used. SEM allows for the evaluation of all the influences proposed simultaneously (PLS-SEM) as well as for a direct comparison of the different effects depending on the two visualization types (using multi-group analysis). To analyze the model as well as the effects of the two different visualizations, first results are presented for the SP (see Section Model Evaluation: Results of the SP), then for the PCP (see Section Model Evaluation: Results of the PCP), and finally the differences between the two visualizations are analyzed in detail (see Section Model Evaluation: Comparison between SPs and PSPs).

Comparison between SPs and PCPs

To compare results based on correlation estimations between the SP and the PCP, multiple t-tests depending on the correlation types have been conducted (Table 8).

Tuble 0: 0 ferview Ebtili		teste, acpent	ing on conclut	ion type
Comparison	Years	p-Val ue	T-Statisics	Mean difference (ED PCP - ED SP)
All correlation types	all data	0,000	-10,726	8,829
	only high amount (4 years)	0,000	-8,328	8,896
Positive correlations	all data	0,000	-8,331	12,104
	only high amount (4 years)	0,000	-10,411	10,942
Only > 0.8	all data	0,000	-9,035	19,648
	only high amount (4 years)	0,000	-5,231	15,229
Only > 0.6	all data	0,000	-4,701	10,286
	only high amount (4 years)	0,000	-4,953	12,960
Only > 0.2	all data	0,000	-3,961	5,237
	only high amount (4 years)	0,000	-4,310	8,124
Negative correlations	all data	0,000	-4,426	5,828
	only high amount (4 years)	0,000	-3,645	5,687
Only < -0.8	all data	0,005	-2,832	8,004
	only high amount (4 years)	0,029	-2,189	7,189
Only < -0.6	all data	0,074	-1,792	3,794
	only high amount (4 years)	0,685	-0,407	1,120
Only < -0.2	all data	0,000	-4,317	8,057
	only high amount (4 years)	0,000	-4,863	8,751

Table 8: Overview Estimate Deviation (ED) SP vs. PCP (t-tests) depending on correlation type

Our analysis reveals that overall the SP outperforms the PCP. Further, in both separate analyses based on the direction of the presented correlation – positive and negative – the SP seems to be the better visual support. Nonetheless, as correlations become weak or negative, the difference between PCPs and SPs is reduced and even nonexistent in the r < -0.6 condition.

The results of this analysis allow the first three hypotheses to be answered:

- ✓ Accept: Overall, SPs show a higher estimation accuracy than PCPs. (H1).
- ✓ Accept: SPs show a higher estimation accuracy than PCPs for positive correlations. (H2).
- * Reject: PCPs show a higher estimation accuracy than SPs for negative correlations. (H3).

It remains to be seen whether a variance in experience or the deviation from normal distribution can reverse or strengthen these findings.

Model evaluation: Results of the SP

Before analyzing the structural model based on our independent and control variables, the model itself needs to be evaluated. We want to find out how much of the deviation of the true score can be explained by our variables and how much this deviation influences the participants' estimates. Therefore, it is necessary to include the actual or true score within the structural model. This analysis is presented in the following figure.



Figure 6: Explained variability of the participants' estimates - SP

From this analysis, we can see that the true score has a higher influence on the estimate than the deviation from the true score (which can be expected), but nonetheless, our intended manipulations affect the estimate to a high extent (-0.314). Overall 45% of the variability of each participant's assessment can be attributed to these two variables, which indicates a high explainability and a robust model. In addition, the quality criteria for the proposed model are within the recommended limits (SRMR=0.066, NFI=0.864 and Chi-Square 2,365.753).

Figure 7 shows the results of the proposed research model, focusing on the deviation from the true score. Results show that a more accurate assessment of the independent variables as presented in the previous section (in Table 5 and Table 6) is possible by using the structural model and SmartPLS for assessment. This is due to the fact that the indicated influences of the independent variables have been corrected by simultaneously assessing the influence of the proposed control variables. Although the independent variables show a significant influence (the level of significance is presented in Table 9), the results deviate slightly from our observations in Section Necessary Adaptions to the Experimental Design.

From the proposed variables the strongest effect on the deviation from the true score can be attributed to the amount of data presented within the SP ($f^2=0.057$ further boosted the quadratic effect $f^2=0.028$), followed by kurtosis ($f^2=0.046$; additionally the quadratic effect of $f^2=0.017$) and skewness ($f^2=0.022$; additionally the quadratic effect of $f^2=0.035$). Details on all variables, namely design choice, amount of data, correlation type, excess kurtosis, skewness and experience, are described in the following bullet points.



Figure 7: PLS model for the SP (x1: results on quadratic effects)

Independent variables:

• **Design choices**: In contrast to our previous analysis, the display of a linear regression line has no effect on estimation error when evaluated using the model and therefore does not influence judgment of correlations within an SP. Further, the use of multiple colors reduces the deviation,

indicating that it is better to use different colors to indicate different years than to use one color with different shades. With respect to the interaction technique, using select (only highlighting the data points within the whole dataset instead of solely focusing on the selected year) increases the deviation from the true score and therefore should be avoided.

- Amount of data used: The higher the amount of data presented within the SP, the lower the deviation from the true score. Using a limited amount of data can negatively affect a decision-maker's ability to detect a correlation. As expected, this effect is also not linear but quadratic. A limited amount of data shows severely distorted results while assessments based on a larger amount of data become more accurate. The amount of data used is the strongest of the tested independent variables and supports the application of this chart type, especially in situations with high data density.
- **Correlation type**: As already presumed in our quick analysis in Section The Participants, both positive correlations and stronger effects (in both directions, negative and positive) result in estimates that are more accurate. This is also indicated by the results based on our structural model. The strength of the correlation is the stronger indicator when compared to the direction of the correlation (positive or negative).

Control variables:

- Excess kurtosis: The higher excess kurtosis (in our case values are high irrespective of whether the distribution is flat or steep as absolute values are used see Table 7), the higher the negative effect on the deviation from the true score becomes. Further, the additional intensifying quadratic effect indicates that results are even worse if values for kurtosis are at the higher end of the spectrum. Consequently, we can conclude that high excess kurtosis severely distorts results when estimating correlation in an SP visualization.
- **Skewness**: Skewness, on the other hand, increases estimate accuracy. A concentration of data points on either the left or right sides of the chart supports participants in their evaluation. This effect is also significantly stronger if skewness becomes higher. The effect is therefore not linear but quadratic, intensifying the reducing effect on the deviation of the true score. As a result, we can support the statement that a high skewness (the concentration of data points on one specific side of the chart) increases estimate accuracy in an SP visualization.
- **Experience**: Although the indicated experience was of only medium strength, the effect of experience is clearly visible. As predicted, the higher the experience with the SP the lower the deviation of the true score, or in other words the better estimate accuracy.

L.Perkhofer, C.Walchshofer, P. Hofer	/ Finance, Risk and Accounting	Perspectives Proceedings	(2019) 294-340
--------------------------------------	--------------------------------	--------------------------	----------------

Table 9. Bootstrapping results for the SP (significan	t lesuits ale l	nuicateu. p >	0.01°°°, p	≥ 0.05 ····, p	(× 0.1 °)
Scatterplots	Original	Sample	STDEV	Т	Р
	Sample	Mean (M)		Statistics	Values
	(0)				
TrueScore -> Estimate	0.562	0.561	0.014	39.307	0.000
DeviationFromTrueScore -> Estimate	-0.314	-0.316	0.027	11867	0.000
Trend -> DeviationFromTrueScore	-0.007	-0.007	0.024	0.291	0.771
Color -> DeviationFromTrueScore	-0.102	-0.102	0.034	2.971	0.003
Interaction -> DeviationFromTrueScore	0.147	0.147	0.028	5.232	0.000
DataPoints -> DeviationFromTrueScore	-0.255	-0.255	0.026	9.965	0.000
Quadratic Data -> DeviationFromTrueScore	-0.389	-0.388	0.047	8.197	0.000
Correlation type -> DeviationFromTrueScore	-0.113	-0.115	0.018	6.208	0.000
Kurtosis -> DeviationFromTrueScore	0.951	0.953	0.089	10.709	0.000
Quadratic Kurtosis ->	0.248	0.249	0.034	7.230	0.000
DeviationFromTrueScore					
Skewness -> DeviationFromTrueScore	-0.454	-0.455	0.062	7.291	0.000
Quadratic Skewness ->	-0.585	-0.587	0.061	9.651	0.000
DeviationFromTrueScore					
Experience -> DeviationFromTrueScore	-0.108	-0.107	0.019	5.704	0.000

Model evaluation: Results of the PCP

Figures 8 and 9 show the structural model based on the estimates made for the PCP. In Figure 8, we can again see that the true score has a higher influence on the estimate than the deviation. However, slightly less variability can be explained by the model, namely only 32%, which still indicates medium to high explainability. Quality criteria for the proposed model are slightly above the recommended limits but still sufficient for model assessment (SRMR=0.117, NFI=0.655 and Chi-Square 5,137.055).



Figure 8: Explained variability of the participants' estimates - PCP

Once again, the true score accounts for more variability of the estimate than the deviation from the true score. However, this difference is only marginal. The strength of the influence of the deviation from the true score is stronger for PCPs (-0.383) compared to SPs (-0.314), yet whether this difference is also significant is considered in the next step of the analysis, which is presented in the next section.

Details on the influences of our proposed research model are presented in the Figure 9. It is evident that the independent variables have a significant influence on the deviation from the true score (the level of significance is presented in Table 11). Nonetheless, with respect to the introduced controls on the distribution (excess kurtosis and skewness), no significant effects can be obtained when analyzing only PCPs. Consequently, these results indicate that PCPs are harder to read but estimation accuracy is not influenced by the distribution of the presented variables. From the proposed variables, the strongest effect on the deviation from the true score can be attributed to the amount of data presented within the PCP ($f^2=0.030$); for all other variables f^2 is below 0.02, indicating minimal effects. It is notable that the reducing effect of an initial increasing amount of data is turned into a positive effect if the dataset becomes too large. This negative effect turning positive indicates a u-shape-like effect of the amount of data when analyzing PCPs.



Figure 9: PLS model for the PCP (x1: results on quadratic effects)

Independent variables:

- **Design choices**: The use of a linear regression function, which highlights the data pair representing the median from the left axis (for details on the regression line please see Table 2, Design 2), increases estimate accuracy. This effect was also visible in our initial analysis in Section Necessary Adaptions to the Experimental Design. On the other hand, the use of multiple colors worsens the results. Again, this effect was indicated before. It is better to highlight different dimensions (in our case, the different years) with different but distinguishable shades of the same color rather than utilizing multiple colors (one representing each year within the dataset). With respect to the interaction techniques used, the highlighting of the selected information within the whole dataset allows better correlation estimates as opposed to solely focusing on the filtered data. To ascertain why it is also better to represent the whole dataset for analyzing only a subset of data, more research on data retrieval when working with PCPs might be necessary.
- Amount of data used: As has already been stated, the amount of data used within the PCP follows a u-shaped curve. Initially, the higher the amount of data, the better the results are. However, if too much data is being displayed the results are negatively affected. This corresponds with the often mentioned influence of visual clutter within a PCP. For an excessively large amount of data, data reduction methods might be necessary so that valuable information can be retrieved from a PCP.
- **Correlation type**: No significant effect can be found for correlation type. The correlation strength and direction have no effect on the deviation of the true score. This is also congruent with our initial analysis presented in Section Comparison between SPs and PSPs. Control variables:
- **Deviation from normality**: Neither excess kurtosis nor skewness have an effect on estimate accuracy. The PCP seems to be unaffected by the distribution of the variables it is representing, which could represent a huge advantage over an SP.
- **Experience**: Surprisingly, with PCPs it seems the higher the experience, the lower estimate accuracy. Admittedly, the indicated overall experience is very low (or nonexistent) and therefore results might be biased. Further, we did ask for participants' experience with the visualization

type but not their experience with performing a correlation estimation task. Further research is necessary to investigate the negative effect of experience in more detail.

Parallel coordinates plot	Original	Sample	STDEV	Т	Р
	Sample	Mean		Statistics	Values
	(0)	(M)			
TrueScore -> Estimate	0.434	0.434	0.018	24353	0.000
DeviationFromTrueScore -> Estimate	-0.383	-0.383	0.024	16.177	0.000
Trend -> DeviationFromTrueScore	-0.109	-0.106	0.029	3824	0.000
Color -> DeviationFromTrueScore	0.125	0.124	0.027	4.657	0.000
Interaction -> DeviationFromTrueScore	-0.124	-0.123	0.030	4.123	0.000
DataPoints -> DeviationFromTrueScore	-0.184	-0.183	0.022	8.367	0.000
Quadratic Data -> DeviationFromTrueScore	0.122	0.115	0.056	2.155	0.031
Correlation type -> DeviationFromTrueScore	0.006	0.017	0.024	0.249	0.803
Kurtosis -> DeviationFromTrueScore	0.076	0.085	0.072	1.042	0.298
Quadratic Kurtosis ->	0.002	0.005	0.039	0.051	0.959
Skewness -> DeviationFromTrueScore	-0.021	-0.025	0.051	0.412	0.681
Quadratic Skewness -> DeviationFromTrueScore	0.013	0.007	0.039	0.323	0.747
Experience -> DeviationFromTrueScore	0.052	0.052	0.023	2.257	0.024

Table 10: Bootstrapping results for the PCP (significant results are indicated: $p > 0.01^{***}$; $p > 0.05^{**}$; $p > 0.1^{*}$)

Model Evaluation: Comparison between SPs and PCPs

When comparing the results of the SP and the PCP, the following results can be obtained (results are based on multi-group analysis and statistical significance is based on the Welch-Satterthwait Test, both presented in Table 11):

Estimate:

- Effect of the true score: The effect of the true score on the estimate is stronger for SPs than for PCPs. This corresponds strongly with the observation that estimate accuracy is higher with SPs. This difference is significant if tested between the two visualization types. Thus, overall, we can state that SPs outperform PCPs.
 - ✓ Accept: Overall, SPs show a higher estimation accuracy than PCPs (H1, based on model assessment).
- **Deviation of the true score**: The difference between the true score and the given estimate demonstrates a higher effect for PCPs in comparison to SPs, however, this difference is only significant at a p < 0.1 level. Consequently, the design and the amount of data represented account for slightly more variance in a PCP than in an SP. A good design is consequently of higher importance for PCPs than for SPs.

Independent variables:

• **Design choices**: For SPs a linear regression indicating the trend does not influence results, while highlighting the data pair representing the median of the left axis in the PCP demonstrates a

positive influence. This difference is significant (p < 0.01). Further, SPs should be displayed using multiple colors to show different dimensions, while for parallel coordinates it is preferable to use a monochrome color palate. Again, this difference between visualization types is significant (p < 0.01), indicating the needed difference in design. Further, concerning the optimal interaction technique, different results can be obtained (p < 0.01): when focusing on one dimension within the dataset in an SP it is best to filter and rescale both axes according to the selected dimension. With PCPs, on the other hand, it is best to show the whole dataset and highlight the selected dimension with low opacity. All differences mentioned are significant between groups.

~ **Possibly reject**: The use of a regression line has a positive effect on estimate accuracy. (H4, based on model assessment true for PCPs but not for SPs).

~ **Possibly reject**: An increased amount of colors used (to distinguish between dimensions) has a negative effect on estimate accuracy. (**H5**, based on model assessment true for PCPs but not for SPs).

~ **Possibly reject**: The interaction technique filter shows a higher estimation accuracy than the interaction effect select. (**H6**, based on model assessment true for SPs but not for PCPs).

- Amount of data used: For both SPs and PCPs, an increase in the represented amount of data from the lowest amount used (n=52) to a medium amount (n=250) is beneficial. However, as soon as the amount increases to its maximum (n=833) estimate accuracy increases when represented in an SP while it decreases in a PCP (p < 0.01). This difference gives SPs a large disadvantage when analyzing large datasets, indicating better use for questions relevant to big data. The benefits of data reduction methods as proposed for the use of PCPs need to be investigated further to indicate congruent results.
 - Reject: An increased amount of data used has a negative effect on estimation accuracy (H7, based on model assessment dismissed for SPs for all data density levels; dismissed for PCPs in low density scenarios but accepted for high density levels).
 - ✓ Accept: SPs show a higher estimation accuracy than PCPs if large amounts of data are visualized (H8, based on model assessment).
- **Correlation type**: With SPs the identification of strong positive effects is easiest, while with PCPs the strength and the direction of the correlation seems to have no effect. This difference is also significant in the multi-group analysis (p < 0.01).
 - ✓ Accept: SPs show a higher estimation accuracy than PCPs for positive correlations. (H2, based on model assessment).
 - ✗ Reject: SPs show a higher estimation accuracy than PCPs for negative correlations. (H3, based on model estimation).

Control variables:

- Excess kurtosis: High kurtosis (flat or steep curves) negatively influences results in SPs, while it has no effect when working with a PCP.
- **Skewness**: High skewness (data points crowded on the left or right side of the chart) positively influences results in an SP while there is no effect on the PCP.
- Experience: Overall, the level of experience is higher with SPs in comparison to PCPs. For SPs, the higher the experience the lower the deviation from the true score. This relationship is

reversed for PCPs. While the difference is significant (p < 0.01), more research is needed to identify the reason why a higher experience causes distorted results when judging correlations.

~ **Possibly reject**: A high level of experience has a positive effect on estimation accuracy (H9, based on the model assessment true for SPs but not for PCPs).

	Path	T-Value	p-Value
	Coefficients		F · ·····
	Difference		
TrueScore -> Estimate	0.127	5.650	0.000
DeviationFromTrueScore -> Estimate	0.068	1.930	0.054
Trend -> DeviationFromTrueScore	0.102	2721	0.007
Color -> DeviationFromTrueScore	0.227	5.173	0.000
Interaction -> DeviationFromTrueScore	0.272	6.495	0.000
DataPoints -> DeviationFromTrueScore	0.071	2.122	0.034
Quadratic Data -> DeviationFromTrueScore	0.510	6.842	0.000
Correlation type -> DeviationFromTrueScore	0.119	4.025	0.000
Experience -> DeviationFromTrueScore	0.159	5.421	0.000
Kurtosis -> DeviationFromTrueScore	0.875	7.655	0.000
Quadratic Kurtosis -> DeviationFromTrueScore	0.246	4.821	0.000
Skewness -> DeviationFromTrueScore	0.434	5.471	0.000
Quadratic Skewness ->	0.598	8.324	0.000
Experience -> DeviationFromTrueScore	0.159	5.323	0.000

Table 11: Multi-group analysis (significant results are indicated)	n >	<u> 00</u>	1 * * * .	n >	0.05	**. *	$\gamma > 0$	1 *)	
Table 11. Walti-group analysis (significant results are indicated.	P -	- 0.0	ı,	P -	0.05	· · ·	j > 0	.1)	
	_			_				_		_

Discussion and Conclusion

Increasing datasets require new forms of data analysis. The use of visualizations in situations where algorithms cannot be used, are too expensive to implement or too complex to understand, is one option to enhance data processing and decision quality (Janvrin, Raschke, & Dilla, 2014; Keim et al., 2008). This realization has led to the development of visual analytics tools and also to the creation of new visualization types (besides the widely used business visualizations) (Perkhofer, L. M. et al., 2019) that allow for larger and more complicated datasets to be represented (Kehrer & Hauser, 2013). Visualizations facilitate an intuitive and fast understanding of the underlying data structure. This is the reason why many different visualization types have been developed to tackle specific analytics problems in different fields of application, including risk management (Dilla & Raschke, 2015; Jones, Melis, Gaia, & Aresu, 2018) and management accounting (Appelbaum, Kogan, Vasarhelyi, & Yan, 2017; Kokina, Pachamanova, & Corbett, 2017; Ohlert & Weißenberger, 2015; Perkhofer, L. M. et al., 2019). Unfortunately, evidence from experimental research on the correct use and design of a visualization is largely missing but would help further improve their use (Abi Akle et al., 2019; Urribarri & Castro, 2016).

In this paper, we concentrate on visual representations that allow for the detection and assessment of correlations. Correlation detection is an important task in portfolio optimization, risk assessment or sensor monitoring and therefore important for accounting and risk professionals (Appelbaum et al., 2017). Two visualization types have been identified to be suitable

for this task in previous research, namely the SP and the PCP (Harrison et al., 2014; Kanjanabose et al., 2015; Kuang et al., 2012). While the SP is the most frequently used visualization type to present multivariate datasets for correlation analysis (Claessen & van Wijk, 2011), the PCP has gained increasing attention during the last decade, especially in the Information Visualization community (Inselberg & Dimsdale, 1990; Kanjanabose et al., 2015). This paper has attempted to investigate which of the proposed visualizations is the most appropriate aid for correlation analysis based on a real-life financial dataset. To do so, the two visualizations have been analyzed in detail depending on:

- their particular design and possible characteristics that might support correlation estimation tasks,
- their ability to visualize large datasets,
- their dependency on the previous experience of the data analyst (or decision-maker),
- and their dependency on normal distribution.

Overall, we can support previous results indicating that SPs outperform PCPs; however, PCPs do have advantages over SPs if designed optimally and under specific circumstances regarding the dataset. The results obtained are presented and discussed in the bullet points below and a summary is presented in Table 13 (recommendations on design) and Table 12 (characteristics of the dataset). When reading these results, please keep in mind that experience with the PCP is drastically low, which could distort results as this factor has been shown to have a negative influence. This is the case as based on previous experience and knowledge structures a user's brain develops certain cognitive expectations. These expectations are used to predict an outcome of certain problems faced by the user. Thus, PCPs are still harder to work with when addressed by an unexperienced user (Lehmann et al., 2010). Thus, even though "scatterplots remain one of the oldest and simplest yet most flexible and widely used visual representations" (Elmqvist et al., 2008, p. 1141), their increased estimation accuracy might be attributed to the enhanced experience with this particular visualization type. More research is needed to ascertain whether an increased experience with the PCP could enhance estimation accuracy.

For the design, almost in all conditions opposite results can be obtained for the two visualization types under investigation (these results are based on model evaluation).

- **Design regression line**: Annotations can either be seen as helpful or as non-data ink (Sarikaya & Gleicher, 2018). They are helpful if they support the user in performing the task, while they can be accounted non-data ink if the display of the specific annotation distracts the user. With respect to the use of a regression line, we can find no evidence of support nor distraction within the SP. However, we have identified a supporting role within the PCP where a display of the regression line resulted in superior results. Overall, we would therefore argue for the display of a regression line in both visual aids.
- **Design color use**: With respect to color-use, it is best to present different dimensions with multiple and clearly distinguishable colors within an SP, whereas it is best to use one color in different but distinguishable shades within a PCP. This might be attributed to the fact that an enhanced variability in color use increases cognitive load, which is already higher for PCPs as limited experience is available for data processing (Falschlunger, Lehner, Treiblmaier et al., 2016; Perkhofer & Lehner, 2019).
- **Design interaction technique**: For an SP, focusing (=filter) solely on the selected option (in our case one year) should be allowed, while within a PCP it is better to present the whole dataset while highlighting (=select) the selected part with brighter colors. In addition, in this context more research is needed to determine why reducing the dataset does not result in better estimation accuracy in the PCP. More information on the data retrieval process is necessary to answer this question.
| Table 12. Recommendations on visualization | in type design |
|---|--|
| Design an SP to estimate correlations | Design a PCP to estimate correlations |
| with or without a regression line (no indication for positive nor negative influence) | with a data pair that highlights the median of the left axis showing the corresponding data point |
| with multiple colors to represent multiple | using the linear regression
with mono-color (one color and in |
| dimensions.
with the interaction technique filter (if you | distinguishable shades of this particular color)
with the interaction technique select (always |
| want to be able to interactively focus on different dimensions) | show the whole dataset) |

Table 12: Recommendations on visualization type design

With respect to the dataset, three aspects are of major importance: the amount of data that needs to be represented in one comprehensive visualization, the distribution of the variables and the type of correlation that is represented within the chart.

- Number of data points: With respect to the amount of the dataset, we could find an increased estimation accuracy for both visualization types under medium data density levels (roughly 250 data points) with respect to a low one (around 50 data points). However, if the dataset increases further (up to approximately 850 data points), the results deteriorate for PCPs while continuing to improve for SPs. If one has to analyze a truly large dataset, SPs seem to be the better choice. For PCPs, data reduction methods would be needed before visualizing the data.
- Distribution of the variables: The distribution in particular constitutes an aspect of high importance, which we did not foresee when designing the experiment. Strong kurtosis and skewness of index-distributions have an especially strong effect on the estimate accuracy when analyzed with SP visualizations, while no effects are visible for PCPs. High kurtosis either because the distribution is flat or high-peaked is one of the strongest predictors of deviations in the true score and causes results on estimate accuracy to deteriorate. On the other hand, high skewness crowding data points on the left or on the right side of the mean increases estimate accuracy in the SP visualization. These insights are not discussed in current literature as for most of the experiments normally distributed data is used. However, as a lot of datasets in real-life deviate from normality, the usefulness of PCPs might be drastically underestimated.
- **Correlation type**: Previous research has indicated that positive correlations are more accurately estimated when using an SP. This finding can be supported by our analysis. However, previous research also suggests that negative correlations might be better supported in the PCP. This finding cannot be supported as in a PCP the direction of the correlation (positive or negative) has no effect on estimation accuracy. The effect of the magnitude of the correlation is also much stronger in SPs (strong correlations are detected more often), whereas this effect is almost nonexistent in the PCP.

Use an SP to estimate correlations	Use a PCP to estimate correlations
if the dataset has a very high data density (the more	if the dataset has a moderate data density (too
data points – the more crowded and focused they are –	much information reduces estimate accuracy)
the better the estimate)	
if the dataset you need to display is likely to be	if the data you need to display is likely to show
shifted either to the right or to the left indicated by high	many outliers (with fat or flat tails) indicated by high
skewness	kurtosis
if correlations are more likely to be strong	if correlations are more likely to be moderate
if the data is more likely to indicate positive	no indications on the direction can be found,
correlations	positive and negative correlations have the same
	change to be detected

Table 13: Recommendations on visualization type use

From our perspective, and given more training, we think that PCPs can be seen as a proper alternative to SPs, particularly if the dataset is not normally distributed and the likelihood of negative correlations is considerable. Particularly in finance, where the judgment of correlations is of high importance and data is unlikely to be normal, the PCP could be a useful addition to the current pool of visualizations employed. For more widespread use, however, the users need to be educated on the usefulness of this particular visualization type and they need to be taught how to properly read and interact with the visualization (Perkhofer, L. M. et al., 2019).

Limitations and further research opportunities

This chapter discusses limitations as well as future research opportunities. We start by the methodological approach, which shows limitations via the use of MTurk, problems associated with limited experience, especially with the PCP, and the fact that only estimation accuracy is evaluated whereas no indications on the user's information retrieval process is given. Next, we discuss specific features of the visualizations and the dataset used and how they might influence results. Finally, these visualizations are not exclusively designed to visualize correlations. Further tasks important in finance are considered and laid out for future research endeavors.

Use of MTurk: While the use of MTurk comes with the advantage of a large pool of possible participants and fast survey completion rates, there are also some limitations. First, during data collection we as researchers have no control over the process. Participants could be disturbed or interrupted, drawing away the necessary attention that might be needed to successfully fulfill the required tasks. Therefore, special attention needs to be paid in the design of the questionnaire and quality checks need to be implemented to sort "good" from "bad". Second, most workers live in the United States and India, which might introduce a cultural bias. However, workers tend to be more educated than the general population and therefore more complex issues can be posted on MTurk, which was important for our study. Further, specific characteristics of the workers (e.g., the need for a bachelor's degree) can be linked to the posted HIT (human intelligence tasks) in exchange for higher payment. Despite the possible drawbacks, a comparative study which was posted on MTurk and also executed in the lab in the context of visual analytics produced comparable results (Harrison et al., 2014). For this initial stage of our research, we need answers to many manipulations (design, dataset, correlation type...) and we therefore believe MTurk to be an appropriate platform.

No information on information retrieval process: The way information is retrieved from a visual display gives a lot of indications on design problems. For example, in our analysis we could see that the use of the interaction technique filter (restricting the data represented to the current analysis focus) is helpful for SPs, while the opposite is true for PCPs. Eye tracking studies, which have proven to be particularly useful in providing reasons for such contradictions in perception (Falschlunger et al., 2014; Falschlunger, Lehner, & Treiblmaier, 2016), might be able to shed light on the issue.

Increasing dimensionality of the dataset: "A set of N-dimensional data points can be visualized using either N-1 SPs, or a single parallel-coordinates plot" (Netzel et al., 2017, p. 119). When more than three dimensions have to be examined, the expansion from one single SP to an SP matrix and from one PCP to an entire PC matrix might help to detect correlative data at an earlier stage. A good understanding of how to design and utilize interactive visualizations for data exploration is therefore valuable and could potentially increase the use of parallel coordinates and SPs (matrices) in financial dashboards in order to take corrective actions at an early stage. Testing the number of different attributes and the application of said matrices was not part of this study but is a logical next step following the identification of the optimal design. It is not clear how many dimensions should be visualized at the same time in order for the decision-maker to be able

to still comprehend the data. However, some studies indicate a higher accuracy if multiple dimensions (> = 8) are displayed in a PCP rather than a SP matrix (Netzel et al., 2017).

Other visualization tasks: Typically, when confronted with the analysis of a dataset, different tasks need to be performed (Claessen & van Wijk, 2011). In this paper, we concentrate on the aspect of correlation assessment; however, other tasks are also important and the usefulness of PCPs might differ depending on the task. For example, it has been shown that the more complex the task, the better the results of the PCP when compared to the SP (Kanjanabose et al., 2015). Tasks can be classified into different categories depending on their complexity (Claessen & van Wijk, 2011; Kanjanabose et al., 2015; Lehmann et al., 2015). Nonetheless, multiple task types have not been addressed yet and these are highlighted in bold:

- Identifying specific values (value retrieval; *determine range; outlier detection...*)
- Identifying patterns (*characterize distribution*; assess correlations, clusters, *or temporal trends*...)
- Performing predictions/analysis (dependency reasoning; change detection; regression analysis; *multidimensional scaling*...)

References

- Abi Akle, A., Yannou, B., & Minel, S. (2019). Information visualisation for efficient knowledge discovery and informed decision in design by shopping. Journal of Engineering Design, 30(6), 227–253. https://doi.org/10.1080/09544828.2019.1623383
- Appelbaum, D., Kogan, A., Vasarhelyi, M., & Yan, Z. (2017). Impact of business analytics and enterprise systems on managerial accounting. International Journal of Accounting Information Systems, 25, 29–44. https://doi.org/10.1016/j.accinf.2017.03.003
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. Cognition and Instruction, 8(4), 293–332.
- Claessen, J. H. T., & van Wijk, J. J. (2011). Flexible linked axes for multivariate data visualization. IEEE Transactions on Visualization and Computer Graphics, 17(12), 2310–2316. https://doi.org/10.1109/TVCG.2011.201
- Diehl, S., Beck, F., & Burch, M. (2010). Uncovering strengths and weaknesses of radial visualizations--an empirical approach. IEEE Transactions on Visualization and Computer Graphics, 16(6), 935–942. https://doi.org/10.1109/TVCG.2010.209
- Dilla, W. N., & Raschke, R. L. (2015). Data visualization for fraud detection: Practice implications and a call for future research. International Journal of Accounting Information Systems, 16, 1–22. https://doi.org/10.1016/j.accinf.2015.01.001
- Elmqvist, N., Dragicevic, P., & Fekete, J. (2008). Rolling the Dice: Multidimensional Visual Exploration using Scatterplot Matrix Navigation. IEEE Transactions on Visualization and Computer Graphics, 14(6), 1141-1148. https://doi.org/10.1109/TVCG.2008.153
- Falschlunger, L., Eisl, C., Losbichler, H., & Greil, A. (2014). Improving information perception of graphical displays: An experimental study on the display of column graphs. In V. Skala (Chair), Conference on Computer Graphics, Visualization and Computer Vision (WSCG), Pilsen.
- Falschlunger, L., Lehner, O., & Treiblmaier, H. (2016). InfoVis: The Impact of Information Overload on Decision Making Outcome in High Complexity Settings. In Special Interest Group on Human-Computer Interaction, Proceedings of the 15th annual Pre-ICIS Workshop on HCI Research in MIS (1–6, Paper 3). AIS Electronic Library: Association for Information Systems.
- Falschlunger, L., Lehner, O., Treiblmaier, H., & Eisl, C. (2016). Visual Representation of Information as an Antecedent of Perceptive Efficiency: The Effect of Experience. In Proceedings of the 49th Hawaii International Conference on System Sciences (HICSS) (pp. 668–676). IEEE. https://doi.org/10.1109/HICSS.2016.88
- Falschlunger, L., Treiblmaier, H., & Lehner, O. (2015). Cognitive differences and their impact on information processing: An empirical study comparing survey and eye tracking data. In F. D. Davis, R. Riedl, J. Vom Brocke, P.-M. Léger, A. B. Randolph, B. Weber, . . . P. Sachse (Eds.), Lecture Notes in Information Systems and Organization, Information Systems and Neuroscience (pp. 137–144). Springer International Publishing.
- Federico, P., Heimerl, F., Koch, S., & Miksch, S. (2017). A Survey on Visual Approaches for Analyzing Scientific Literature and Patents. IEEE Transactions on Visualization and Computer Graphics, 23(9), 2179–2198. https://doi.org/10.1109/TVCG.2016.2610422

- Harrison, L., Yang, F., Franconeri, S., & Chang, R. (2014). Ranking Visualizations of Correlation Using Weber's Law. IEEE Transactions on Visualization and Computer Graphics, 20(12), 1943–1952. https://doi.org/10.1109/TVCG.2014.2346979
- Henley, M., Hagen, M., & Bergeron, D. (2007). Evaluating two visualization techniques for genome comparison. In E. Banissi (Ed.), 11th [IEEE] International Conference Information Visualization: IV 2007 [proceedings] 4-6 July 2007, Zurich, Switzerland (pp. 1–6). Los Alamitos Calif., Washington D.C.: IEEE Computer Society; Conference Publishing Services.
- Hofer, P., Walchshofer, C., Eisl, C., Mayr, A., & Perkhofer, L. (2018). Sankey, Sunburst & Co: Interactive Big Data Visualizierungen im Usability Test. In L. Nadig & U. Egle (Eds.), Proceedings of CARF 2018: Controlling, Accouting, Risk, and Finance (pp. 92–112). Verlag IFZ.
- Holten, D., & van Wijk, J. J. (2010). Evaluation of Cluster Identification Performance for Different PCP Variants. Computer Graphics Forum, 29(3), 793–802. https://doi.org/10.1111/j.1467-8659.2009.01666.x
- Inselberg, A., & Dimsdale, B. (1990). Parallel coordinates: A tool for visualizing multi-dimensional geometry. In Proceedings of the First IEEE Conference on Visualization: Visualization `90 (pp. 361–378). IEEE Comput. Soc. Press. https://doi.org/10.1109/VISUAL.1990.146402
- Inselberg, A. (1985). The plane with parallel coordinates. The Visual Computer, 1(2), 69–91. https://doi.org/10.1007/BF01898350
- Janvrin, D. J., Raschke, R. L., & Dilla, W. N. (2014). Making sense of complex data using interactive data visualization. Journal of Accounting Education, 32(4), 31–48. https://doi.org/10.1016/j.jaccedu.2014.09.003
- Johansson, J., & Forsell, C. (2016). Evaluation of Parallel Coordinates: Overview, Categorization and Guidelines for Future Research. IEEE Transactions on Visualization and Computer Graphics, 22(1), 579–588. https://doi.org/10.1109/TVCG.2015.2466992
- Jones, M., Melis, A., Gaia, S., & Aresu, S. (2018). Does graphical reporting improve risk disclosure?: Evidence from European banks. Journal of Applied Accounting Research, 19(1), 161–180. https://doi.org/10.1108/JAAR-07-2016-0068
- Kanjanabose, R., Abdul-Rahman, A., & Chen, M. (2015). A Multi-task Comparative Study on Scatter Plots and Parallel Coordinates Plots. Computer Graphics Forum, 34(3), 261–270. https://doi.org/10.1111/cgf.12638
- Kehrer, J., & Hauser, H. (2013). Visualization and visual analysis of multifaceted scientific data: A survey. IEEE Transactions on Visualization and Computer Graphics, 19(3), 495–513. https://doi.org/10.1109/TVCG.2012.110
- Keim, D. A. (2001). Visual exploration of large data sets. Communications of the ACM, 44(8), 38-44. https://doi.org/10.1145/381641.381656
- Keim, D. A., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., & Melancon, G. (2008). Visual analytics: Definition, process, and challenges. In A. Kerren, J. T. Stasko, J. D. Fekete, & C. North (Eds.), Lecture Notes in Computer Science: Vol. 4950. Information Visualization (Vol. 4950, pp. 154–175). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Kokina, J., Pachamanova, D., & Corbett, A. (2017). The role of data visualization and analytics in performance management: Guiding entrepreneurial growth decisions. Journal of Accounting Education, 38, 50–62. https://doi.org/10.1016/j.jaccedu.2016.12.005
- Kuang, X., Zhang, H., Zhao, S., & McGuffin, M. J. (2012). Tracing Tuples Across Dimensions: A Comparison of Scatterplots and Parallel Coordinate Plots. Computer Graphics Forum, 31(3pt4), 1365–1374. https://doi.org/10.1111/j.1467-8659.2012.03129.x
- Lehmann, D. J., Albuquerque, G., Eisemann, M., Tatu, A., Keim, D., Schumann, H., . . . Theisel, H. (2010). Visualisierung und Analyse multidimensionaler Datensätze. Informatik-Spektrum, 6(33), 589–600.
- Lehmann, D. J., Kemmler, F., Zhyhalava, T., Kirschke, M., & Theisel, H. (2015). Visualnostics: Visual guidance pictograms for analyzing projections of high-dimensional data. Computer Graphics Forum, 34(3), 291–300. https://doi.org/10.1111/cgf.12641
- Lewandowsky, S., & Spence, I. (1989). Discriminating Strata in Scatterplots. Journal of the American Statistical Association, 84(407), 682–688. https://doi.org/10.1080/01621459.1989.10478821
- Li, J., Martens, J.-B., & van Wijk, J. J. (2008). Judging correlation from scatterplots and parallel coordinate plots. Information Visualization, 1, 69. https://doi.org/10.1057/palgrave.ivs.9500179
- Li, J., Martens, J.-B., & van Wijk, J. J. (2010). A model of symol size discrimination in scatterplots. In E. D. Mynatt, S. E. Hudson, & G. Fitzpatrick (Eds.), CHI Conference (pp. 2553–2562). New York N.Y.: Association for Computing Machinery.
- Liu, S., Maljovec, D., Wang, B., Bremer, P.-T., & Pascucci, V. (2017). Visualizing High-Dimensional Data: Advances in the Past Decade. IEEE Transactions on Visualization and Computer Graphics, 23(3), 1249–1268. https://doi.org/10.1109/TVCG.2016.2640960
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 101(2), 343–352.
- Munzner, T. (2014). Visualization analysis and design (1st Edition). AK Peters Visualization Series: CRC Press.

- Netzel, R., Vuong, J., Engelke, U., O'Donoghue, S., Weiskopf, D., & Heinrich, J. (2017). Comparative eye-tracking evaluation of scatterplots and parallel coordinates. Visual Informatics, 1(2), 118–131. https://doi.org/10.1016/j.visinf.2017.11.001
- Ohlert, C. R., & Weißenberger, B. E. (2015). Beating the base-rate fallacy: an experimental approach on the effectiveness of different information presentation formats. Journal of Management Control, 26(1), 51–80. https://doi.org/10.1007/s00187-015-0205-2
- Parsons, P., & Sedig, K. (2014). Adjustable properties of visual representations: improving the quality of humaninformation interaction. Journal of the Association for Information Science & Technology, 65(3), 455–482.
- Perkhofer, L., Hofer, P., & Walchshofer, C. (2019). BIG Data Visualisierungen 2.0: Optimale Gestaltung und Einsatz neuartiger Visualisierungsmöglichkeiten. In L. Nadig & U. Egle (Chairs), CARF. Symposium conducted at the meeting of Verlag IFZ Hochschule Luzern, Luzern.
- Perkhofer, L., Hofer, P., Walchshofer, C., Plank, T., & Jetter, H.-C. (2019). Interactive visualization of Big Data in the field of accounting: A survey of current practice and potential barriers for adoption, in Print.
- Perkhofer, L., & Lehner, O. (2019). Using gaze behavior to measure cognitive load. In F. Davis, R. Riedl, J. Vom Brocke, P.-M. Léger, & A. Randolph (Eds.), Lecture Notes in Information Systems and Organisation: NeuroIS Retreat 2018. Information Systems and Neuroscience: NeuroIS Retreat 2018 (1st ed., pp. 73–83). Springer International Publishing.
- Perkhofer, L. M., Hofer, P., Walchshofer, C., Plank, T., & Jetter, H.-C. (2019). Interactive visualization of big data in the field of accounting. Journal of Applied Accounting Research, 5(1), 78. https://doi.org/10.1108/JAAR-10-2017-0114
- Pike, W. A., Stasko, J., Chang, R., & O'Connell, T. A. (2009). The Science of Interaction. Information Visualization, 8(4), 263–274. https://doi.org/10.1057/ivs.2009.22
- Rensink, R. A. (2017). The nature of correlation perception in scatterplots. Psychonomic Bulletin & Review, 24(3), 776–797. https://doi.org/10.3758/s13423-016-1174-7
- Sarikaya, A., & Gleicher, M. (2018). Scatterplots: Tasks, Data, and Designs. IEEE Transactions on Visualization and Computer Graphics, 24(1), 402–412. https://doi.org/10.1109/TVCG.2017.2744184
- Speier, C. (2006). The influence of information presentation formats on complex task decision-making performance. International Journal of Human-Computer Studies, 64(11), 1115–1131.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane gognitive load. Educational Psychology Review, 22(2), 123–138.
- Tufte, E. R. (1983). The visual display of quantitative information (1st). Connecticut: Graphics Press.
- Urribarri, D. K., & Castro, S. M. (2016). Prediction of data visibility in two-dimensional scatterplots. Information Visualization, 16(2), 113–125. https://doi.org/10.1177/1473871616638892
- Vessey, I., & Galletta, D. (1991). Cognitive fit: an empirical study of information acquisition. Information Systems Research, 2(1), 63–84.
- Yoo, H. S., & Smith-Jackson, T. L. (2011). Colour size illusion on liquid crystal displays and design guidelines for bioinformatics tools. Behaviour & Information Technology, 30(6), 775–785.

Appendix

Normal distribution and correlations used within the study: SP design 1: 4 years – 46% correct answers

Correlation	SP	Histogram (y-axis)	Histogram (x-axis)
Negative correlatio n r > -0.8 Answers correct +/- 15PP: 66%	Company A & Company B	Company A	Company B
	Correlation coefficient: -0.84	Kurtosis: -0.96 Skewness: 0.62	Kurtosis: -1.49 Skewness: -0.20
Negative correlatio n r > -0.6	Company B & Company C	Company B	Company C
+/- 15PP: 18%	Correlation coefficient: -0.62	Kurtosis: -1.49 Skewness: -0.20	Kurtosis: 5.47 Skewness: 2.42
Negative correlatio n r > -0.2 Answers correct	Company B & Company N	Company B	Company N
+/- 15PP: 8%	Correlation coefficient: -0.26	Kurtosis: -1.49 Skewness: -0.20	Kurtosis: -0.66 Skewness: -0.66
Positive correlation r < 0.2 Answers correct +/- 15PP: 44%	Company E & Company L	Company E	Company L
	Correlation coefficient: 0.25	Kurtosis: -1.12 Skewness: 0.05	Kurtosis: -0.53 Skewness: -0.45
Positive correlation r < 0.6 Answers correct +/- 15PP: 52%	Company A & Company O	Company A	Company O
	Correlation coefficient: 0.62	Kurtosis: -0.96 Skewness: 0.62	Kurtosis: -0.15 Skewness: -0.15
Positive correlation r < 0.8 Answers correct +/- 15PP: 86%	Company A & Company J	Company A	Company J
	Correlation coefficient: 0.84	Kurtosis: -0.96 Skewness: 0.62	Kurtosis: -0.39 Skewness: 0.29

Completion			
Correlation	SP Recent	Histogram (y-axis)	Histogram (x-axis)
Negative	Company F & Company K	Company F	Company K
correlatio	1.00 G200 G200 G1000 G100 G100 G100 G100 G100 G100 G	\frown –	
n r > -0.8	244		
	and the second		
Answers correct	and the second s	· ·	
+/- 15PP: 69%	- A Contraction		
	van van dan dan	Kurtosis: -1 16	Kurtosis [.] -0.88
	Correlation coefficient: -0.85	Skewness: -0.32	Skewness: 0.42
Negative	Company E & Company H	Company E	Company H
correlatio			company m
n r > -0.6	B Rogenson (no		
Answers correct			
+/- 15PP· 44%	and the second se		
• / 1011.11/0			TT (1.40
	Correlation coefficient: -0.62	Kurtosis: -1.12	Kurtosis: -1.42
27.1		Skewness: 0.05	Skewness: 0.85
Negative	Company F & Company N	Company F	Company N
correlatio	Biogrammon Lies	\frown –	\frown
n r > -0.2			
	A STATE OF		
Answers correct			
+/- 15PP: 31%			
	and the set	Kurtosis: -1.16	Kurtosis: -0.66
	Correlation coefficient: -0.29	Skewness: -0.32	Skewness: -0.66
Positive	Company L & Company O	Company L	Company O
correlation $r < 0.2$			_
	e Regesor Lie		
Answers correct			
+/- 15PP: 60%			
	2 	Kurtosis: -0.53	Kurtosis: -0 15
	Correlation coefficient: 0.25	Skewness: -0.45	Skewness: -0.15
Positive	Company E & Company O	Company E	Company O
correlation $r < 0.6$		Company 12	company c
	a Togranio Lie		
Answers correct			
+/- 15PP: 64%			
		7	
		110	
	Correlation coefficient: 0.67	Kurtosis: -1.12	Kurtosis: -0.15
D iii		Skewness: 0.05	Skewness: -0.1
Positive	Company E & Company G	Company E	Company G
correlation $r < 0.8$	and Sold Sold Sold Sold Sold Sold Sold Sol	-	\frown
A	and the second se		
Allsweis correct	and the second se		
+/- 13PP: 30%			
	an en internet		
	Correlation coefficient: 0.86	Kurtosis: -1.12	Kurtosis: -0.16
		Skewness: 0.05	Skewness: 0.92

SP design 2: 4 years – 54% correct answers

r design 5. 4 years – 2			
Correlation	SP	Histogram (y-axis)	Histogram (x-axis)
Negative	Company B & Company G	Company B	Company G
correlatio	2000 C 1019		\frown
n r > -0.8	and the second sec		
	- And Same and a		
Answers correct	and the same		
+/- 15PP: 48%	The second se		
	1988 - 1986 - 1986 - 1986 - 1986 - 1986 - 1986 - 1986 - 1986 - 1986 - 1986 - 1986 - 1986 - 1986 - 1986 - 1986 -	Kurtosis: -1.49	Kurtosis: -0.16
	Correlation coefficient: -0.86	Skewness: -0.20	Skewness: 0.92
Negative	Company C & Company H	Company C	Company H
correlatio	8 7987 9 797 6 499		_
n r > -0.6	later a		
Answers correct			
+/- 15PP: 33%	and the second s		
		Kurtosis: 5.47	Kurtosis: -1.49
	Correlation coefficient: -0 64	Skewness: 2.42	Skewness: -0.20
Negative	Company C & Company M	Company C	Company M
correlatio		company c	
n r > -0.2	E CONTRACTOR DE		\square
	14		
Answers correct	- de		
+/- 15PP: 48%			
	and the second	Vurtagis: 5 47	Kurtosis: 0.54
	Correlation coefficient: 0.23	Skewness: 2.42	Strownoss: 0.01
Desitive	Company L& Company N	Composer I	Skewiiess. 0.91
Positive $r < 0.2$	Company J & Company N	Company J	Company N
	and the second sec		
Answers correct			
+/- 15PP 27%			
1/-1311.21/0	· · · · · · · · · · · · · · · · · · ·		
	and the group of the second		
	10 5 10 10 10 10 10 10 10 10 10 10 10 10 10	Kurtosis: -0.39	Kurtosis: -0.66
	Correlation coefficient: 0.23	Skewness: 0.29	Skewness: -0.66
Positive	Company C & Company E	Company C	Company E
correlation $r < 0.6$	and the second		- ^
	ner i i i i i i i i i i i i i i i i i i i		
Answers correct			
+/- 15PP: 56%	1		
	1.000 1.000 1.000 1.000 1.000 1.000	Kurtosis: 5.47	Kurtosis: -1.12
	Correlation coefficient: 0.65	Skewness: 2.42	Skewness: 0.05
Positive	Company C & Company G	Company C	Company G
correlation $r < 0.8$	nom 8000		
	-		
Answers correct			
+/- 15PP: 67%			
		Kurtosis: 5.47	Kurtosis: -0.16
	Correlation coefficient: 0.84	Skewness: 2.42	Skewness: 0.92

SP design 3: 4 years – 46% correct answers

Correlation Histogram (y-axis) Histogram (x-axis) SP Company H Negative Company H & Company I Company I correlatio n r > -0.8 Answers correct +/-15PP:80% Kurtosis: -1.49 Kurtosis: -0.41 Correlation coefficient: -0.84 Skewness: -0.20 Skewness: 0.85 Negative Company H & Company L Company H Company L C 2018 C 2017 C 2016 correlatio n r > -0.6 Answers correct +/-15PP:38% Kurtosis: -1.49 Kurtosis: -0.53 Skewness: -0.20 Skewness: -0.45 Correlation coefficient: -0.63 Negative Company J & Company M Company J Company M 5 2018 5 2017 6 2010 correlatio n r > -0.2 Answers correct +/-15PP:34% Kurtosis: 0.54 Kurtosis: -0.39 Skewness: 0.29 Skewness: 0.91 Correlation coefficient: -0.33 Positive Company N Company O Company N & Company O 2518
2517
2516
2516 correlation r < 0.2the set Answers correct +/-15PP:24% Kurtosis: -0.66 Kurtosis: -0.15 Skewness: -0.15 Skewness: -0.66 Correlation coefficient: 0.30 Positive Company G Company O Company G & Company O correlation r < 0.6Answers correct +/-15PP:46% Kurtosis: -0.16 Kurtosis: -0.15 Skewness: 0.92 Skewness: -0.15 Correlation coefficient: 0.60 Positive Company I & Company K Company I Company K correlation r < 0.8Answers correct +/-15PP:70% Kurtosis: -0.41 Kurtosis: -0.88 Skewness: 0.85 Correlation coefficient: 0.88 Skewness: 0.42

SP design 4: 4 years – 49% correct answers

CF design 1. 4 years -	- 55% correct answers	TT • / • \	TT• / • \
Correlation	PCP	Histogram (y-axis)	Histogram (x-axis)
Negative	Company B & Company G	Company B	Company G
correlatio	74.05 C 2015		
n r > -0.8			
	1111 Lui		
Answers correct	10		
+/- 15PP: 50%	10		
	4.15	Kurtosis 1.40	Kurtosis: 0.16
		Sharry and 0.20	Shown and 0.02
	Correlation coefficient: -0.86	Skewness: -0.20	Skewness: 0.92
Negative	Company C & Company H	Company C	Company H
correlatio	200 Collin		
n r > -0.6	100 dan		
	100		
Answers correct			
+/- 15PP: 38%			
	40	Vurtagia 5 47	Kurtosia 140
	40.	Strong 2, 42	Strown age, 0.20
	Correlation coefficient: -0.64	Skewness. 2.42	Skewness0.20
Negative	Company C & Company M	Company C	Company M
correlatio	10 0 4000 0 2000 0 2000		\frown
n r > -0.2	11		
	-10		
Answers correct			
+/- 15PP: 27%	u		
	III.	Vurtagia 5 47	Vertecies 0.54
		Kultosis. 5.47	Kurtosis: 0.54
	Correlation coefficient: -0.23	Skewness: 2.42	Skewness: 0.91
Positive	Company J & Company N	Company J	Company N
correlation $r < 0.2$	2 Min		
Answers correct			
+/- 15PP: 19%			
		-	
		Kurtosis: -0.39	Kurtosis: -0.66
	Correlation coefficient: 0.23	Skewness: 0.29	Skewness: -0.66
Positive	Company C & Company E	Company C	Company E
correlation $r < 0.6$	1.44 UNIX 100 UNIX 10		
	Contract of the second s		
Answers correct	449		
+/- 15PP: 37%			
	11		-
		Vartagia 5 47	Variation 1.10
		Kurtosis: 5.47	Kurtosis: -1.12
	Correlation coefficient: 0.65	Skewness: 2.42	Skewness: 0.05
Positive	Company C & Company G	Company C	Company G
correlation $r < 0.8$	2007 2005 2005		
	A01		
Answers correct	am		
+/- 15PP: 38%			
	140	Vurtosis: 5 47	Kurtogia: 0.16
	4.44 -	Kultosis, 3.47	
	Correlation coefficient: 0.84	Skewness: 2.42	Skewness: 0.92

PCP design 1: 4 years – 35% correct answers

CP design 2: 4 years –	- 47% correct answers	TT . ()	TT
Correlation		Histogram (y-axis)	Histogram (axis)
Negative	Company H & Company I	Company H	Company I
correlatio	In Contraction (inclusion)		
II I > -0.8	1.0		
Answers correct			
$\pm /_{-} 15 PP \cdot 80\%$			
17-1511.0070	Us Not	TT	
	Correlation coefficient: -0.84	Kurtosis: -1.49	Kurtosis: -0.41
Nagating	Company II & Company I	Skewness: -0.20	Skewness: 0.85
Negative	Company H & Company L	Company H	Company L
n r > -0.6	a 2014 S Degression Line		
111 > -0.0			
Answers correct			
+/- 15PP: 49%	0m		
	Um Water	Kurtosis: 1 10	Vurtagia: 0.52
	Correlation coefficient: -0.63	Skewness: -0.20	Skewness: -0.55
Negative	Company J & Company M	Company J	Company M
correlatio	The second secon	y	•••••••
n r > -0.2	and the second s		
	11 Mar 10 Mar		
Answers correct			
+/- 15PP: 24%			
	Completion of figure 0.22	Kurtosis: -0.39	Kurtosis: 0.54
	Correlation coefficient: -0.33	Skewness: 0.29	Skewness: 0.91
Positive	Company N & Company O	Company N	Company O
correlation $r < 0.2$	T Bitty Start Le	\frown	1 A A A A A A A A A A A A A A A A A A A
	a construction of the second sec		\frown
Answers correct			
T/-15FF. 5170	Dar we derive a		
	Correlation coefficient: 0.30	Kurtosis: -0.66	Kurtosis: -0.15
Docitive	Company G & Company O	Skewness: -0.66	Skewness: -0.15
$\frac{1}{100} r = 0.6$	Company & & Company O	Company G	Company O
	a Contemportante de la contemp	\frown	
Answers correct	-		
+/- 15PP: 35%			
	14	Kurtosis [.] -0.16	Kurtosis: -0.15
	Correlation coefficient: 0.60	Skewness: 0.92	Skewness: -0.15
Positive	Company I & Company K	Company I	Company K
correlation r < 0.8			
	Comparison to a		
Answers correct			
+/- 15PP: 65%	10		
	Los Fax		
	Correlation coefficient: 0.99	Kurtosis: -0.41	Kurtosis: -0.88
		Skewness: 0.85	Skewness: 0.42

P

Correlation	PCP	Histogram (v-axis)	Histogram (y-avis)
Negative	Company A & Company B	Company A	Company B
correlatio	Company A & Company D	Company A	
n r > 0.8	ter		
11 - 0.0	1.00		
Answers correct			
\pm / 15DD· 630/	10 2.57		
+/-1511.05/0	100 (40) 100		
	100	Kurtosis: -0.96	Kurtosis: -1.49
	Correlation coefficient: -0.84	Skewness: 0.62	Skewness: -0.20
Negative	Company B & Company C	Company B	Company C
correlatio			\mathbf{N}
n r > -0.6			
Answers correct			
+/- 15PP: 26%			
	Correlation coefficient: 0.62	Kurtosis: -1.49	Kurtosis: 5.47
		Skewness: -0.20	Skewness: 2.42
Negative	Company B & Company N	Company B	Company N
correlatio	17 Mar 2019 6019		\frown
n r > -0.2	14 A		
A			
Allsweis collect	10		
+/-15PP. 21 %	10 (en 220		
	и 101	Kurtosis: -1.49	Kurtosis: -0.66
	Correlation coefficient: -0.26	Skewness: -0.20	Skewness: -0.66
Positive	Company E & Company L	Company E	Company L
correlation $r < 0.2$	1.14 2.37 1.47 2.37 2.37	\sim	
	Contraction of the second seco		
Answers correct			
+/- 15PP: 30%	Un		
	1.0		
	101	Kurtosis: -1.12	Kurtosis: -0.53
	Correlation coefficient: 0.25	Skewness: 0.05	Skewness: -0.45
Positive	Company A & Company O	Company A	Company O
correlation $r < 0.6$			
A reasonable contract	a (1)		
Answers correct $\pm / 15 \text{ DD} \cdot 250/$	20 20 20 20 20 20 20 20 20 20 20 20 20 2		
+/-1511.55/0	10 ⁻		
		Kuntania 0.00	TT
		Skownoss: 0.62	Kurtosis: -0.15
	Correlation coefficient: 0.62	SKCW11C55. U.02	Skewness: -0.15
Positive	Company A & Company J	Company A	Company J
correlation $r < 0.8$	1 4 4 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5		_
A possiona compact	IIII LUI		
Answeis correct $+ /_{-} 15 \text{PD} \cdot 700/_{-}$	Net 145		
·/- 1511. /U/0	444		
	100 100	Kurtaria 0.06	TK (0.00
		Skowness 0.62	Kurtosis: -0.39
	Correlation coefficient: 0.84	SKEWHESS. U.02	Skewness: 0.29

PCP design 3: 4 years – 41% correct answers

CP design 4: 4 years –	- 44% correct answers		
Correlation	РСР	Histogram (y-axis)	Histogram (x-axis)
Negative	Company F & Company K	Company F	Company K
correlatio	C 2014 C 2017 D 2014	<u>_</u>	
n r > -0.8			\sim
	210		
Answers correct	10		
+/- 15PP· 64%			
., 1011.01/0	100		
	100	Kurtosis: -1.16	Kurtosis: -0.88
	Correlation coefficient: -0.85	Skewness: -0.32	Skewness: 0.42
Negative	Company E & Company H	Company E	Company H
correlatio	2001 2007 2009		_
n r > -0.6	888		
	1.M		
Answers correct			
+/- 15PP: 38%	100		
	4.0	110	Kauta in 140
		Kurtosis: -1.12	Kurtosis: -1.49
	Correlation coefficient: -0.62	Skewness: 0.05	Skewness: -0.20
Negative	Company F & Company N	Company F	Company N
correlatio	8 0017 1000 1000	\frown –	\frown
n r > -0.2			
Answers correct		, i i i i i i i i i i i i i i i i i i i	
+/- 15PP: 26%	1.00		
		Kurtosis: 116	Kurtosis: 0.66
	Correlation coefficient: 0.20	Skowness: 0.32	Strownoos: 0.66
Desitive	Company I & Company O	Skewness0.52	Skewness0.00
$\begin{array}{c} POSITIVE\\ correlation \ r < 0.2 \end{array}$	Company L & Company O		Company O
correlation 1 < 0.2	и 		1. In 1.
A	·		
Answers correct			
+/- 15PP: 34%			
		Kurtosis: -0.53	Kurtosis: -0.15
	Correlation coefficient: 0 25	Skewness: -0.45	Skewness: -0.15
Positive	Company E & Company O	Company E	Company O
correlation $r < 0.6$			
	4 348 Gm11		
Answers correct	a construction of the second sec		
$+/_{-}15PP \cdot 43\%$	10 X.00		
1/-1511.45/0			
	9 (m) 20		
	8	Kurtosis: -1.12	Kurtosis: -0.15
	Correlation coefficient: 0.67	Skewness: 0.05	Skewness: -0.15
Positive	Company E & Company G	Company E	Company G
correlation $r < 0.8$	8 2016 6 2016 6 2017		
	1000 CHI 0 111		
Answers correct	Tool Control C		
+/- 15PP: 57%			
	and the second		
	100	Vurtaria 1 12	Vurtagian 0.16
	430	Kurtosis: -1.12	Kurtosis: -U.16
	Correlation coefficient: 0.86	Skewness: 0.05	Skewness: 0.92

PCP design 4: 4 years – 44% correct answers

Correlation	SP	Histogram (v-avis)	Histogram (v_avis)
Negative	Company A & Company B 2015	Company A	Company B
correlatio	Company A & Company B 2013	Company A	Company D
n r > -0.2	Date Date Date Date Date Date Date Date	\frown	
111 0.2			
Answers correct	24		
+/- 15PP: 24%	TON	-	
	DX	Kurtosis: -0.60	Kurtosis: 1.33
	Correlation coefficient: 0.24	Skewness: -0.36	Skewness: -0.97
Negative	Company B & Company C 2016	Company B	Company C
correlatio	110 Cites		
n r > -0.8	uno Traj		
	a the second sec		
Answers correct	and the second		
+/- 15PP: 32%	and the first state		
		Kurtosis: -1.14	Kurtosis: -0.17
	Correlation coefficient: -0.86	Skewness: -0.48	Skewness: 0.62
Negative	Company B & Company N 2016	Company B	Company N
correlatio	1000 1000 1000 1000 1000 1000 1000		
n r > -0.6	and a set of the set o		
	, and the second se		
Answers correct			
+/- 15PP: 14%	ب ر بر ساند ر		
	27% 10% 25% 10% 100 110 110 110	Kurtosis: -1.14	Kurtosis: -1.47
	Correlation coefficient: -0.75	Skewness: -0.48	Skewness: 0.30
Positive	Company E & Company L 2018	Company E	Company L
correlatio	C Doors C Doors	- ~	
n r < 0.2			
Answers correct			
+/- 15PP 18%			
., 1011.10/0	01	Kautasia 1.52	TZ (1 0.22
	Correlation coefficient: 0.35	Skowposs: 0.26	Kurtosis: -0.33
Positive	$Company \wedge \& Company \cap 2019$	Company A	Company O
correlatio	Company A & Company O 2018		
n r < 0.8	13×6 13×6		
Answers correct	E.		
+/- 15PP: 58%			
		Kurtosis: -1.24	Kurtosis: -1 11
	Correlation coefficient: 0.88	Skewness: 0.39	Skewness: -0.06
Positive	Company A & Company J 2015	Company A	Company J
correlatio	12349 2237 1236		
n r < 0.6	XXX		
Answers correct	62W		
+/- 15PP: 70%			
	100	Kurtosis: -0.60	Kurtosis: -0.93
	Correlation coefficient: 0.78	Skewness: -0.36	Skewness: -0.40

SP design 1: 1 year – 36% correct answers

CorrelationSPPositive correlation $r < 0.2$ Company FAnswers correlatio n $r > -0.6$ Correlation c Negative correlatio n $r > -0.6$ Company EAnswers correlationCompany FAnswers correlationCompany FAnswers correlationCompany FAnswers correlationCompany FAnswers correlationCompany FAnswers correlation $r < 0.2$ Company LAnswers correlation $r < 0.2$ Company EAnswers correlation $r < 0.2$ Company E		TT'	
Positive correlation r < 0.2Company F aAnswers correlatio n r > -0.6Correlation cNegative correlatio n r > -0.6Company E aAnswers correlationCompany F aAnswers correlationCompany F aAnswers correlationCompany F aAnswers correlationCompany F aAnswers correlation r < 0.2		Histogram (y-axis)	Histogram (x-axis)
correlation $r < 0.2$ Answerscorrect+/- 15PP: 31%Correlation ofNegative correlation n $r > -0.6$ Company EAnswerscorrect+/- 15PP: 56%Correlation ofNo correlation Answers correctCompany FAnswers correlation r < 0.2	c Company K 2018	Company F	Company K
Answerscorrect $+/-15PP: 31\%$ Correlation ofNegative correlation n r > -0.6Company EAnswerscorrect $+/-15PP: 56\%$ Correlation ofNo correlation Answers correctCompany FAnswers correlation r < 0.2	CLOVE CLOVE	\sim	
Answerscorrect $+/-15PP: 31\%$ Correlation ofNegative correlation n r > -0.6Company EAnswers $+/-15PP: 56\%$ Correlation ofNo correlation Answers correct $+/-15PP: 87\%$ Company FAnswers correlation r < 0.2			
+/- ISPP: 31%Correlation cNegative correlation n r > -0.6Answers correct +/- 15PP: 56%Correlation Correlation cNo correlation Answers correct +/- 15PP: 87%Correlation r < 0.2 Answers correct +/- 15PP: 56%Positive correlation r < 0.2 Answers correct +/- 15PP: 56%Positive correlation r < 0.2 Answers correct +/- 15PP: 44%Correlation c Correlation cPositive correlation r < 0.2 Correlation c			
Negative correlation n r > -0.6Company E Company E Company E Correlation cAnswers correlationCompany F Correlation cNo correlation Answers correct $+/-15PP: 87\%$ Company F Correlation cPositive correlation r < 0.2			
Negative correlation n r > -0.6Company E Company E Company E Correlation cAnswers correlationCorrelation cNo correlation Answers $+/-15PP: 87\%$ Company F Correlation cNo correlation Correlation r < 0.2			
Negative correlation n r > -0.6Company E Company E Correlation CorrelationAnswers correlationCorrelation c Correlation cNo correlationCompany F Correlation cAnswers correct +/- 15PP: 87%Company L Correlation cPositive correlation r < 0.2		Kurtosis: -0.49	Kurtosis: -1.00
Negative correlatio n r > -0.6Company E correct +/- 15PP: 56%Answers CorrelationCorrelation cNo correlationCompany F Correlation cAnswers correct +/- 15PP: 87%Company L correlation cPositive correlation r < 0.2	Semcient: 0.57	Skewness: -0.38	Skewness: -0.19
correlatio n r > -0.6CorrelationAnswers +/- 15PP: 56%Correlation of CorrelationNo correlation Answers +/- 15PP: 87%Company F Correlation of Correlation	¿ Company H 2017	Company E	Company H
n r > -0.6 Answers correct +/- 15PP: 56% Correlation Answers correct +/- 15PP: 87% Correlation c Positive correlation r < 0.2 Answers correct +/- 15PP: 56% Company L Correlation c Correlation c Positive correlation r < 0.2 Answers correct +/- 15PP: 56% Correlation c Positive correlation r < 0.2 Answers correct +/- 15PP: 44% Correlation c	2077 D2976 C0979 Distriction in c	\frown	\frown
Answers +/- 15PP: 56%CorrelationNo correlationCompany FAnswers +/- 15PP: 87%Company FCorrelation cCorrelation cPositive correlation r < 0.2	an lasters		
Answerscorrect $+/-15PP: 56\%$ Correlation cNo correlationCompany FAnswerscorrect $+/-15PP: 87\%$ Correlation cPositive correlation r < 0.2	A State of the second s		
+/-15PP: 56%No correlationCorrelation ofAnswers correctCompany F $+/-15PP: 87%$ Correlation ofPositive correlation r < 0.2			
No correlationCorrelation ofAnswerscorrect+/- 15PP:87%Correlation ofCorrelation ofPositive correlation r < 0.2	No.		
CorrelationNo correlationAnswers correct+/- 15PP: 87%Correlation cPositive correlation r < 0.2		Kurtosis: -0.96	Kurtosis: 0.18
No correlationCompany FAnswers correct	pefficient: -0.76	Skewness: -0.06	Skewness: -0.28
Answers +/- 15PP: 87%Correlation cPositive correlation r < 0.2	company N 2015	Company F	Company N
Answerscorrect+/- 15PP:87%Correlation cPositive correlation r < 0.2	020% 020% 020%	100 A. 100 A	
+/- 15PP: 87% Positive correlation r < 0.2 Answers correct +/- 15PP: 56% Correlation c Positive correlation r < 0.2 Answers correct +/- 15PP: 44% Correlation c	C Regression Line	~	\frown
Positive correlation r < 0.2Correlation cAnswers +/- 15PP: 56%Company LPositive correlation r < 0.2			
Positive correlation r < 0.2Company L Company L Company L Company L Company L Correlation cPositive correlation r < 0.2			
Positive correlation r < 0.2Company L Company L Company L Company L Correlation cAnswers correlation r < 0.2			
Positive correlation r < 0.2Company L Company L Company L Correlation cAnswers +/- 15PP: 56%Correlation cPositive correlation r < 0.2	10.000 10.000 10.000	Kurtosis: 0.01	Kurtosis: 1.33
Positive correlation r < 0.2Company L Company L Answers +/- 15PP: 56%Positive correlation r < 0.2	pefficient: -0.06	Skewness: -0.67	Skewness: -0.89
correlation r < 0.2	k Company O 2015	Company L	Company O
Answers correct +/- 15PP: 56% Correlation c Positive correlation r < 0.2 Answers correct +/- 15PP: 44% Correlation c	Dans Dans Dans		1 V
Answers correct +/- 15PP: 56% Correlation c Positive correlation r < 0.2 Answers correct +/- 15PP: 44% Correlation c	© Regression Line		
+/- 15PP: 56% Correlation c Positive correlation r < 0.2 Answers correct +/- 15PP: 44% Correlation c Correlation c			
Positive correlation r < 0.2Company EAnswers +/- 15PP: 44%Correlation cPositiveCorrelation c	and the second second		
Positive correlation r < 0.2Company E Company EAnswers +/- 15PP: 44%Correlation cPositiveCorrelation c			
Positive correlation r < 0.2Company E Company E Correlation cAnswers +/- 15PP: 44%Correlation cPositiveCompany E	1.0 111 100	Kurtosis: -0.46	Kurtosis: -0.82
Positive correlation r < 0.2 Answers correct +/- 15PP: 44% Correlation c	pefficient: 0.59	Skewness: 0.08	Skewness: -0.14
correlation r < 0.2 Answers correct +/- 15PP: 44% Correlation of Positive	& Company O 2017	Company E	Company O
Answers correct +/- 15PP: 44% Correlation of Positive	2019 52017 72016	\sim	
Answers correct +/- 15PP: 44% Correlation of Positive	E Represent the		
+/- 15PP: 44% Correlation of Correlation of Company F	1		
Correlation of Control			
Correlation c	1 - 14 - 14 - 14 - 14 - 14 - 14 - 14 -		
Correlation c	100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100	Kurtosis [.] -0.96	Kurtosis: -0.36
Positive Company F	pefficient: 0.38	Skewness: -0.06	Skewness: 0.48
	k Company G 2016	Company E	Company G
correlation r < 0.8	2019	1 ,	1 .
908	an a		
Answers correct	and the second second		
+/- 15PP: 53%			
······································	K. A. w		
and the second			Kurtosis: 0 11
Correlation of	21	Kurtosis: 0.00	
Answers correct +/- 15PP: 53%	Barn Bargaren		Kurtosis: 0.11

SP design 2: 1 year – 45% correct answers

Correlation	SP	Histogram (v-axis)	Histogram (x-axis)
No correlation	Company B & Company G 2018	Company B	Company G
Answers correct +/- 15PP: 94%	an company D & Company C 2010		
	Correlation coefficient: 0.08	Kurtosis: -0.76 Skewness: -0.23	Kurtosis: -0.58 Skewness: 0.35
Negative correlatio	Company C & Company H 2017	Company C	Company H
n r > -0.6 Answers correct +/- 15PP: 46%			
		Kurtosis: 3.72	Kurtosis: 0.18
	Correlation coefficient: -0.76	Skewness: 2.03	Skewness: -0.28
inegative	Company C & Company M 2015	Company C	Company M
r r > -0.2	□2019 ■2019		
Answers correct +/- 15PP: 31%			
	이 이 말했는 것은 것 같은 것 같은 것 같다.	Kurtosis: 1.80	Kurtosis: 0.97
	Correlation coefficient: -0.40	Skewness: 1.60	Skewness: 1.22
Positive	Company J & Company N 2015	Company J	Company N
correlation r < 0.2 Answers correct +/- 15PP: 52%			
	103 	Kurtosis: -0.93	Kurtosis: 1.33
	Correlation coefficient: 0.29	Skewness: -0.40	Skewness: -0.89
Positive	Company C & Company E 2017	Company C	Company E
correlation r < 0.6 Answers correct +/- 15PP: 52%			
	Correlation coefficients 0.67	Kurtosis: 3.72	Kurtosis: -0.96
Docitive	Company C & Company C 2014	Company C	Company C
correlation $r < 0.8$ Answers correct	Company C & Company G 2016		Company G
·/- 1011. 10 /0		Kurtosic: 0.17	Kurtosis: 0, 11
	Correlation coefficient: 0.82	Skewness: 0.62	Skewness: -0.22

S

Correlation	SP	Histogram (y-axis)	Histogram (x-axis)
Positive	Company H & Company I 2015	Company H	Company I
correlation $r < 0.2$		1 5	
	" Change and the second se		
Answers correct	····		
+/- 15PP: 8%	- 44		
	-		
		Kurtosis: 0.78	Kurtosis: -0.32
	Correlation coefficient: 0.44	Skewness: -0.31	Skewness: -0.73
Negative	Company H & Company L 2016	Company H	Company L
correlatio	Land Land Land Land Land Land Land Land	- r J	- r ·· J
n r > -0.2		\wedge	
Answers correct			
+/- 15PP: 22%	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		
	4.4	Kurtosis: -0.36	Kurtosis: -0.33
	Correlation coefficient: -0.41	Skewness: 0.67	Skewness: -0.56
No correlation	Company J & Company M 2016	Company J	Company M
Answers correct	danı		
+/- 15PP: 2%			
		N	
		Kurtosis: 1.06	Kurtosis: -0.91
	Correlation coefficient: 0.17	Skewness: 0.93	Skewness: 0.23
No correlation	Company N & Company O 2018	Company N	Company O
		1 5	1 5
Answers correct			
+/- 15PP: 80%			
		Kurtosis: 5.30	Kurtosis: -1.11
	Correlation coefficient: 0.01	Skewness: 1.81	Skewness: -0.06
Positive	Company G & Company O 2018	Company G	Company O
correlation r < 0.8			
	- 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997 - 1997		
Answers correct			
+/- 15PP: 62%	· · · · · · · · · · · · · · · · · · ·		
	1 (A 100) 2(10) (20) (20)	Kurtosis: -0.58	Kurtosis: -1.11
	Correlation coefficient: 0.82	Skewness: 0.35	Skewness: -0.06
Negative	Company I & Company K 2015	Company I	Company K
correlatio	C 2013 C 2017 C 2017 C 2017		
n r > -0.2			
	an in Al align		
Answers correct	00		-
+/-15PP:0%	-		
		Kurtosis: -0.32	Kurtosis: -1.28
	Correlation coefficient: -0.20	Skewness: -0.73	Skewness: -0.12

SP design 4: 1 year – 42% correct answers

CP design 1: 1 year –	27% correct answers	Histogram (y avis)	Histogram (x axis)
Negative	Company B & Company G 2017	Company B	Company G
correlatio	Company B & Company G 2017	Company D	Company G
n r > -0.6	AVR 181 Date	\frown	\sim
11 2 0.0	108 741		
Answers correct	D/M [81		
+/- 15PP: 21%	10		
	10	Vurtagia: 0.20	Vurtagia: 0.20
	Correlation coefficients 0.68	Skewness: 0.42	Skewness: 0.69
No correlation	Company C & Company H 2018	Company C	Company H
		company c	
Answers correct			
+/- 15PP: 52%	130		
	139		
		Kurtosis: 0.31	Kurtosis: -0.62
	Correlation coefficient: -0.03	Skewness: 0.94	Skewness: -0.61
Positive	Company C & Company M 2018	Company C	Company M
correlation $r < 0.6$	10	I I J I	r r J
	18		
Answers correct	0 0 0		
+/- 15PP: 12%	17		
	00		
	(i)(i)	Kurtosis: 0.31	Kurtosis: 1.11
	Correlation coefficient: 0.63	Skewness: 0.94	Skewness: 0.93
Positive	Company J & Company N 2016	Company J	Company N
correlation $r < 0.2$	1 2007 1 2007 1 2007 1 2007		_
Answers correct			
+/- 15PP: 6%			
	3		
	17 500	Kurtosis: 1.06	Kurtosis: -1.47
	Correlation coefficient: 0.45	Skewness: 0.93	Skewness: 0.30
Positive	Company C & Company E 2016	Company C	Company E
correlation $r < 0.6$	30 B 2200 B2200 B2200		
A norman correct	80 110		\frown
\pm / 15DD 250/			
T/-15FF. 55%			
		W	
		Kurtosis: -0.17	Kurtosis: 0.00
Desitions	Correlation coefficient: U.73	Skewness: 0.62	Skewness: 0.12
$\begin{array}{c} \text{POSITIVe} \\ \text{correlation } r < 0.9 \end{array}$	Company C & Company G 2017	Company C	Company G
	12700 UNIV		
Answers correct	333		
+/- 15PP: 17%	10 M		
	10 A		
	0.00 10.00 10.00	Kurtosis: 3.72	Kurtosis0 39
	Correlation coefficient: 0.91	Skewness: 2.03	Skewness: 0.69

P

CP design 2: 1 year -2	PCP	Histogram (v. avie)	Histogram (avis)
Negative	Company H & Company I 2016	Company H	Company I
correlatio	Company II & Company I 2010		
n r > -0.8	Was BregesknUre	\sim	\sim
	NY		
Answers correct	ux		
+/- 15PP: 39%	50 mm		
	(A)	Kurtosis -0.36	Kurtosis: -0.86
	Correlation coefficient: -0.89	Skewness: 0.67	Skewness: -0.29
No correlation	Company H & Company L 2015	Company H	Company L
	1327 12270 2220 8 Tapacente	-	\sim
Answers correct	····		
+/- 15PP: 45%			
	100		
	08		
	Correlation coefficient: 0.03	Kurtosis: 0.78	Kurtosis: -0.46
Negative	Company I & Company M 2017	Company I	Company M
correlatio	company 5 & company w 2017		
n r > -0.6	and the second s		
Answers correct			
+/- 15PP: 24%	3		-
	*	Kurtosis: -0.26	Kurtosis: -0.61
	Correlation coefficient: -0.65	Skewness: 0.99	Skewness: 0.48
No correlation	Company N & Company O 2017	Company N	Company O
Answers correct	*		
+/- 15PP: 59%			
	0 0		
	Correlation coefficient: 0.15	Kurtosis: -1.31	Kurtosis: -0.36
Positivo	Company G & Company O 2015	Skewness: -0.19	Skewness: 0.48
$\frac{1}{100} \frac{1}{100} \frac{1}$	Company G & Company O 2015	Company G	Company O
conclution 1 × 0.2	■ Use Control Provide Control		
Answers correct	995		
+/- 15PP: 33%			
	4 944		
	4	Kurtosis: 1.19	Kurtosis: -0.82
	Correlation coefficient: 0.55	Skewness: -1.37	Skewness: -0.14
No correlation	Company I & Company K 2018	Company I	Company K
	200 Class 200 Class 200 Class	\frown	
Answers correct	223 23 23 23 23 23 23 23 23 23 23 23 23		
+/-15PP:%	10. 10.		
	1.0		
	Correlation coefficient: -0 18	Kurtosis: 0.04	Kurtosis: -1.00
		Skewness: -0.5/	Skewness: -0.19

PCP design 2: 1 year –28% correct answers

Cr ucsigii 5. r year –			
Correlation		mistogram (y-axis)	THISTOGRAM (X=axis)
Negative	Company A & Company B 2016	Company A	Company B
correlatio	10 Conv 751 Conv Cross	\sim	
n r > -0.8	29		
	50 Sta		
Answers correct	137		
+/- 15PP: 37%	10		
	19 M		
	10	Kurtosis: -0.34	Kurtosis: -1.14
	Correlation coefficient: -0.81	Skewness: 0.38	Skewness: -0.48
Positive	Company B & Company C 2015	Company B	Company C
correlation $r < 0.2$	231/ 231/ 231/ 231/ 231/ 231/		
		A	
Answers correct	M		
+/- 15PP· 14%	238		
• • • • • • • • • • • • • • • • • • • •	*		-
	388	Kurtosis: 1.33	Kurtosis: 1.80
	Correlation coefficient: 0.46	Skewness: -0.97	Skewness: 1.60
Positive	Company B & Company N 2017	Company B	Company N
correlation $r < 0.2$	1 5 1 5	1 5	1 5
			\sim
Answers correct			
\pm / 15DD · 140/	111		
+/-1JII.14/0	10 	/	
	11		
		Kurtosis: -0.30	Kurtosis: -1.31
	Correlation coefficient: 0.20	Skewness: 0.42	Skewness: -0.19
Positive	Company E & Company L 2017	Company E	Company L
correlation $r < 0.2$			
	Cons		\sim
Answers correct	00 CO		
Allsweis collect			
+/-13PP. 19%			
	an		
		Kurtosis: -0.96	Kurtosis: 0.58
	Correlation coefficient: 0 45	Skewness: -0.06	Skewness: -0.32
Positive	Company A & Company O 2015	Company A	Company O
correlation $r < 0.6$		company 11	company c
	a ver anno	\sim	
Answers correct	14 (A)		
Allsweis collect	4 		
+/-13PP. 21 %	·*		
	4 4.0	Kurtosis: -0.60	Kurtosis: -0.82
	Correlation coefficient: 0 56	Skewness: -0.36	Skewness: -0.14
Positive	Company A & Company I 2018	Company A	Company J
correlation $r < 0.8$		company 11	company b
	10200 1021 (1021)		
A	184		
Answers correct	244 244		
+/-13PP:14%			
	###		
	10	Kurtosis: -1.24	Kurtosis: -1.67
	Correlation coefficient: 0 92	Skewness: 0.39	Skewness: 0.25

PCP design 3: 1 year – 26% correct answers

Cruesign 4. 1 year –			
Negeting		Compare F	Compare V
Negative	Company F & Company K 2017	Company F	Company K
correlatio	184 USA	_	\frown
n r > -0.6	244	\sim	
Answers correct	193		
+/- 15PP: 43%	101 (101 (101 (101 (101 (101 (101 (101		
		Kurtosis: -0.03	Kurtosis: -0.79
	Correlation coefficient: -0.64	Skewness: 0.06	Skewness: 0.19
Negative	Company E & Company H 2017	Company E	Company H
correlatio			
n r > -0.6	099 100		
	,AU		
Answers correct	178		
+/- 15PP 45%	N/0		
., 1011.10,0	An Contraction of the Contractio		
		Kurtosis: -0.96	Kurtosis: 0.18
	Correlation coefficient: -0.75	Skewness: -0.06	Skewness: -0.28
Negative	Company F & Company N 2018	Company F	Company N
correlatio			
n r > -0.2			
	LEAD CONTRACTOR OF		
Answers correct			
+/- 15PP: 9%	14		
		Kurtosis: -0.49	Kurtosis: 5 30
	Correlation coefficient: -0.53	Skewness: -0.38	Skewness: 1.81
Positive	Company L & Company O 2015	Company L	Company O
correlation $r < 0.2$			company c
Answers correct			
+/- 15PP: 17%	·		
	· All and a second s		
		Versteader 0.46	Kartasia 0.82
		Sharry and 0.08	Kurtosis: -0.82
	Correlation coefficient: 0.59	Skewness: 0.08	Skewness: -0.14
Positive	Company E & Company O 2016	Company E	Company O
correlation $r < 0.6$		1 5	1 5
Answers correct	9 (14)		
+/- 15PP: 49%	a(A)		
	9		
	3	Kurtosis: 0.00	Kurtosis -0 19
	Correlation as officients 0.72	Skewness: 0.12	Skewness: -0.76
Positive	Company F & Company G 2016	Company F	Company G
$\frac{1}{100} \frac{1}{100} \frac{1}$	Company E & Company G 2010		
	D216 Date		
Answers correct	Turk T	\frown	
+/- 15PP· 32%	Part		
·/~ IJII. J#/U	2007		
	440 00X	TZ / 1 0 00	
	NIN 300	Kurtosis: 0.00	Kurtosis: 0.11
	Correlation coefficient: 0.86	Skewness: 0.12	Skewness: -0.22

PCP design 4: 1 year – 30% correct answers



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings

homepage: http://www.acrn.eu/frap



Business Angel Ability to Accurately Identify Investment Opportunities: The Role of Mindfulness

Catherine Crapsky¹, André Soares², Abdel Malik OLA^{3,*}

¹University of Angers

²University of Coventry

³University of Tours

Abstract

Decision making, when it happens under uncertain conditions, retains the attention of researchers in many fields. Entrepreneurial finance is one of them where we are interested in how investment decision is made in the early stage of an entrepreneurial firm (EF) while we lack reliable objective evidence about its future value (uncertainty). Recent studies have shown that one category of investor, the business angel (BA), dominate early stage financing because they can use decision approaches such as heuristics, intuitive judgment or gut feel to sustain their funds allocation. However, little is known about what are the determinants of such dominance. This study aims to find an answer to this question by anchoring in psychological theory the process underlying angel decision expertise. We first suppose that this expertise is explained by the data collecting and processing mode by the BA (cognitive behavior). Second, we build on the literature about mindfulness, a concept that defines the ability of certain actors than others to behave reliably in situation. The concept emphasizes how an actor's attention can sustain his cognitive functioning and explains efficient behavior. Our development consistently relates mindfulness dimensions to angel cognitive strategy when coping with the uncertainty of EF. We build a theoretical model to demonstrate how mindfulness can explain the BA ability to identify investment opportunity (IO) when other categories of investors are absent. We make some propositions that can be tested empirically. We contribute not only to the literature in entrepreneurial finance but also in decision making under uncertain and complex environment.

Keywords: Mindfulness, cognition, business-angel, investment-opportunity

Introduction

Entrepreneurship plays a crucial role in economic development and understanding the financing of new ventures is crucial (Cassar, 2004). It is well recognized that all types of young and small firms face difficulties accessing financial sources due to asymmetric or incomplete information between the entrepreneur and the potential providers of external capital (Ang, 1992). Concerning nascent or new entrepreneurial firms¹ (EF), they suffer from financing constraints firstly due to the high level of risk or uncertainty² they represent for investors. Among those investors, very

¹ Reading academic articles on entrepreneurship or entrepreneurial finance leads to an evidence: there is no consensus on a definition of entrepreneurial firms. Recently, Morris and al. (2018) highlighted a lack of definition of the concept of EF and some confusion regarding the results of empirical research in entrepreneurship due to not enough consideration of the context of entrepreneurship. In this research, we assume that an EF is an independent nascent or new innovative one with a growth or high-growth potential (Löhrer 2017).

² Uncertainty refers to contexts in which probabilities are unknown (Knight, 1921). In this case risks are known but can't be quantified. Diebold and al. (2010) consider that BA are facing contexts of 'unknown unknows" which means that they face both uncertainty and noise, that is a very high specific risk (quoted by Huang, 2018). Indeed, in the traditional financial theory the

often the first external capital provider is a Business Angel (BA) (Harrison et al., 2010; OECD, 2011).

The environment becoming unpredictable, a very high specific risk may create an incapacity or an intellectual paralysis at the decision-maker level (Milliken, 1987). Indeed, BAs have to identify if the nascent or new venture represents an investment opportunity (IO)³ while dealing with a lack of reliable information to envision objectively the future of the firm. In doing so, they affirm to place an important weight on their gut feel about the entrepreneur and the venture (Hisrich and Jankowicz, 1990; Huang, 2017; Huang and Pearce, 2015; Ola, 2016; Ola et al., 2019). Such behavior could be seen as impulsive and emotional. The behavioral finance framework could conclude to biases and false decision process (Kahneman and Tversky, 1982). On the contrary, it has been demonstrated that the BA uses a creative intuitive decision process based on structural alignment (Ola et al., 2019). Structural alignment means that the BA can make sense in situ because he uses discourse to create new interpretation and new significations before investing. Discovering such an ability is possible through an understanding of the real cognitive strategy of BA (Hisrich and Jankowicz, 1990; Huang, 2017; Huang and Pearce, 2015; Moesel and Fiet, 2001). Despite the increasing importance of angel finance, the literature focusing on the cognitive strategy of BAs is very limited and, as pointed out by Huang and Pearce (2015), we are just beginning to understand the decision-making process of BAs.

We consider that BAs are able to decide to invest under uncertainty and noise because they can make sense in situ: they can face reliably uncertainty and ambiguity. At this point, we want to go further in the explanation of the BA identification process of IO by investigating how to predict such a mental attitude.

The nature of the entrepreneurial idea constrains the cognitive strategy of the decision maker. It impacts not only the kind of decision inputs but also the processing of the information. Decision makers cope with such a context through a mental manipulation of outside information in order to produce new decision constructs. Any study, to our knowledge, focuses on this data processing aspect of the BA investment. Such an approach can explain not only which indicators really count for the investor's final perception of financial return, but also how the investor arrives or builds every component necessary for a final decision. The BA's cognitive strategy is unknown. The literature describes the inputs (different decision factors) and the outputs (trust and numerical measure). The link between these two sides is complex because of uncertainty, noise and lack of understanding of the median process. We consider mindfulness is the missing link. Mindfulness matters because of the high specific risk of the EF. Indeed, mindfulness means stable attention so sustain the active seeking and processing of information in situation (Langer, 1989). It means the constant creation and labelling of new categories when dealing with a lack of reliable information to decide (Sun and Fang, 2010). Mindfulness level can explain how a category of actors can behave more efficiently than others in the same situation by opening the black box underlying decision-making and subsequent behaviors. It supports the idea that behaviors depend upon the attention of the actors that can be influenced by human functioning (Good et al, 2016). Building on this argument and the gap in the angel investment literature, this study aims at answering the following question: can mindfulness explain BA's ability to accurately identify investment opportunities? This paper introduces a conceptual model that explains the impact of mindfulness on BA's cognitive strategy through an impact on attention and human functioning. It is the first step of a wider research whose aim is to measure the mindfulness of BAs regarding the different scales of mindfulness which exist (Baer et al., 2006).

We contribute to the literature on entrepreneurial finance in several ways. First, as pointed out by many scholars we think that entrepreneurs have a poor understanding of the BA's decision-

risk of an asset can be divided in systematic risk and specific risk. The systematic risk is a market risk and concerns all assets.

The specific or unsystematic risk is only due to the specificities of one asset.

³ An opportunity exists from the moment the funder's expected financial return exceeds the cost of the investment.

making process and a tiny proportion of innovative projects are funded by BAs. Our analysis may help entrepreneurs better understanding this important potential first external investor in order to increase the number of valuable projects they can finance. Secondly, to address this topic we are convinced that we need to deeply analyze the cognitive strategy of those investors and then to take advantage of other social science academic results, especially psychology. Then, we contribute to the very small literature looking at the real decision investment process under high risk or uncertainty without focusing on cognitive biases. Thirdly, we take into account the transformation of the BA market by integrating the decision-making at the group level in our theoretical model.

The remainder of the paper is as it follows. In order to link the behavioral characteristics of BAs with the mindfulness dimensions, we have analyzed the academic literature concerning BA investments at early-stage. We have identified three dimensions; the non-compensatory relationship between risk and return (Jeffrey et al., 2016); the judgment approach (Huang, 2017; Ola, 2016) and the group level decision (Mason et al., 2016). The first part explains these three dimensions. In the second part, we expose the dimensions of mindfulness and we discuss the propositions to be tested. In the conclusion, we present our conceptual model and we explain how we plan to develop our empirical analysis.

The three behavioral dimensions of BA decision process

The business angel research is developing for more than thirty years (White and Dumay, 2017). Mason et al (2016), looking at the transformation of the BA market, propose this definition: *"Business angels can be defined as high net worth individuals (HNWIs) who invest their own money, either alone or with others, directly in unquoted businesses in which there is no family connection. They normally invest in the form of equity finance in the hope of achieving a significant financial return through some form of exit. Typically, they will also take an active involvement in their investee businesses (Mason, 2006). Investments by business angels are largely focused on new and early-stage technology ventures (Giudici and Paleari 2000; Mason and Harrison 2010, 2011). They are particularly important from a regional economic development perspective because the majority of their investments are local (Avdeitchikova, 2008; Harrison et al., 2010); hence, they are typically recycling and reinvesting locally created wealth (e.g. from the sale of a business –Mason and Harrison 2006)».*

Compared to Venture Capitalists (VC), BAs are considered as informal investors. Through cluster analysis, Sørheim and Landström (2001) have identified four types of informal investors regarding their investment activity and their competences. Among the four categories BAs are defined as follows: they "are characterized by a very high investment activity level, in addition to possessing high competence. They can contribute both knowledge and skills to the firms in which they invest, and they generally engage in many informal investments". Then the focus of our research is on this type of BA.

The literature review on BA decision process can be summarized in three different dimensions: non-compensatory heuristic, intuitive judgement and collective decision.

No balance between risk and return: the non-compensatory heuristic

Like any financial investment decision, early-stage funding may involve the balance between the risk and the return. Paul et al. (2007) study was among the first studies claiming that a key issue for the BA is such a balance. For a rational investor, early-stage funding is an uninteresting asset because the expected return cannot cover the high cost of due diligence, contracting and monitoring (Bonini et al., 2018). In a traditional financial framework, academics claim that BAs will diversify their risk by being a member of a syndicate. They can have access to more heterogeneous project with different risk level and then diversify their portfolio (Paul et al, 2007).

But, according to more recent analysis (Huang, 2017; Huang & Pearce, 2015), the investment' strategy of BA is different from other categories of investors. BAs have privileged access to private information about the deals and such proximity with EF makes the risks on different projects as manageable, without seeking to establish a formal relationship with expected return. BAs are voluntarily involved in closing the gap between the supply and the demand of early-stage funding. They are risk takers and they adjust their behavior to high-risk contexts like EF funding (Croce et al., 2017; Paul et al., 2007). BAs are wealthy individuals investing their own money and are generally under little pressure to generate income or capital growth. They will not choose to diversify their portfolio (Geibel and Yang, 2018) but focus on a few opportunities which risk seems manageable through their own perception. When VC operate according to risk-return principle, angels pay more attention to innovative firms with less profitability expectation (Pedchenko et al., 2018). Unlike banks, they will prefer equity to debt in their investment approach, which represent a more risky investment mode. Angels acknowledge the high level of risk involved in EF and can consider it like measurable and controllable through numbers (Huang, 2017). As far as the angel is able to assess the risk inherent to a project, he will claim that there is a potential for significant financial return, without establishing a positive correlation with risk (Smith et al., 2010). The risk perceived by the BA is twofold: the business risk and the agency risk (Harrison and Mason, 2017). BAs invest in companies whose business segments are closely related to their own industry experiences, and in proposals that are located close to their residence in order to be able to manage the risk (Geibel and Yang, 2017). By accepting a high level of specific risk, angels know that they can try to complement it with their own background in order to try to avoid huge monetary failure.

Then, to deal with EF, angels will use a non-compensatory heuristic (Jeffrey et al., 2016) that will manifest itself in two different ways. First, during the early screening, the heuristic consists in quickly selecting inside a pool of projects by avoiding spending a lot of cognitive energy. A project can be rejected only because information about one aspect is absent or insufficient (fatal flow) even if all the other aspects have a good score (Maxwell et al., 2011). Some factors are dealmakers and others are a deal killer (Clark, 2008) and most often investors are looking for reasons to reject an opportunity (Carpentier and Suret, 2015; Harrison et al., 1997). This fast heuristic-based decision happened during the pitch of the entrepreneur in the face of many angels or investors. Non-compensatory heuristic then means an elimination-by-aspects. The angel's cognitive process helps to « cut the chase » when initially screening deals. Non- compensatory heuristics at this stage confirm also the fact that angels, even most experienced, do not become obsessed with just one or two investment factors but review very quickly « what is present or absent » in each deal (Smith et al, 2010).

During the due diligence, angels will continue using non-compensatory heuristic but with more time and energy demanding. According to Jeffrey et al (2016), instead of compensating the risk and return in their EF's assessment, the used shortcut consists of weighing separately risk and return by aggregating different components. The risk is calculated through the assessment of the business feasibility (Huang, 2017). Each project risk assessment deserves a considerable time for the BA. According to Huang (2017), business feasibility judgment consists in integrating each factor piece-by-piece (hard data), in a "checklist" manner with the objective to provide an overall "score" on whether the opportunity was a credible one that should be pursued. During this assessment, if the forecasted risk is too high or the anticipated return is too low, BAs do not treat these attributes as compensatory. In other words, if the risk exceeds a certain level, they will consider that the return cannot be high enough to compensate for the high risk; conversely, if the return is below a certain level, the risk is assumed to be not too small to compensate for the low return (Jeffrey et al, 2016).

To summarize, with the high level of uncertainty around the EF, researchers have acknowledged that angels use a heuristic called non-compensatory. However, it manifests itself

in two different ways. By behaving according to the heuristic in a different manner, we assume that angels have the ability to cognitively adapt themselves to a complex context of decisionmaking. The ability to shift between the two different modes of non-compensatory shows cognitive flexibility, which distinguishes these investors from others. BAs also show an ability to voluntarily control one's own thinking with the energy required at each stage. This capacity for control and cognitive flexibility is not innate and we think can distinguish BAs from other investors.

BAs do not follow the recommendation of traditional portfolio theory in that they evaluate risk and return separately, rather than in a compensatory manner as the theory recommends. Angels quickly select a small number of EF according to deal killer and deal maker. Only some EF are picked and deserve assessment attention with the acknowledgement that they can pass through some good opportunities. The more an investor will get itself away from the traditional financial theory principle and use the two modes of non-compensatory models, the best he will be able to navigate in the high-risk environment of early-stage investment.

The combination of soft and hard cues: an intuitive judgment approach

A second distinctive trait in the behavior of BAs is the use of intuitive judgment in their investment decision-making. Intuitive judgment implies the conjoint use of hard and soft data in a non-analytical way during a decision-making process (Ola, 2016; Huang, 2017). Indeed, the lack of explicit reliable information in entrepreneurial finance is a reliable incentive for the usefulness of non-traditional tacit cues such as gut feel or intuition in the angel decision-making process (Wu, 2016). Because of uncertainty and noise, the BA prefers to rely on their personal knowledge or on information collected by their own self than third party sources. Four reasons explain such behaviors according to Harrison and Mason (2017): "(1) it is cheap; (2) one trusts one's own information best - it is richer, more detailed, and known to be accurate; (3) individuals with whom one has a continuing relation have an economic motivation to be trustworthy, so as not to discourage future transactions; and (4) departing from pure economic motives, continuing economic relations often become overlaid with social content that carried strong expectations of trust and abstention from opportunism ". Angels prefer to believe more and more their own gut feel, instinct and hunch and these tacit cues are formed through questioning rather than description (Smith et al, 2010; Harrison et al, 1997). Angels will try to experiment things by themselves during an in-depth assessment including personal information search, non-structural meetings and physical site visits (Carpentier and Suret, 2015). Then a pivotal role of nontraditional tacit and not easily transferable information is a key characteristic of the angel's behavior.

The screening of project by the angels proceeds by "fusion" of both hard and soft data about the opportunity (Paul et al, 2007). Hard data including library research, formal market research, formal information about the entrepreneur and formal technologies are among the least likely elements used during the due diligence process. Sometimes angels are hesitant to invest even if a proposal is highly rated on such cues or criteria inferred from them(Cox et al., 2017). Angels will appreciate more when an entrepreneur use language, symbols and images to invoke disruptive claims that highlight their venture's potential to disrupt a market (Sort and Nielsen, 2018). Angels largely focus on their own perception of the entrepreneur's skills. Preparedness, enthusiasm and commitment when assessed, are related to the fact that they have thought and worked a lot about their ventures (Cardon et al., 2017). The fact of investing a lot of their own funds send a signal about the commitment of the entrepreneur. Through the displayed passion, the entrepreneur shares high energy, pleasant emotion with a positive feeling that helps the angel to anticipate a strong relationship and to gauge the EF (Huang and Knight, 2015). During presentation or pitch, entrepreneurs display skills, such as the ability to be understandable and persuasive, that are highly useful for angels (Argerich et al., 2013; Clark, 2008).

Angel investors look for tenacious and passionate entrepreneurs because these traits show that the entrepreneur identifies himself with the business idea and he will be able to overcome adversity (Murnieks et al., 2016). Preference will be high for the entrepreneur who displays both tenacity and passion, as well as inspirational leadership. Angels want to build a good future positive relationship with the entrepreneurs which means good reciprocal resources or information exchange over time. Foreseeing a good future post-investment relationship is then a key issue in the angel's investment decision. For Maxwell and Lévesque (2014), the BA's decision-making process consists in « intuitively auditing positive and negative displays of trustbased behaviors to determine the level of relationship risk before making an investment offer ». The more trust damaging and violating behaviors will be observed, the less the chance to get financing by angels. BA needs the cognitive ability to distinguish behaviors related to trust, to perceive passion and tenacity. The ability to perceive social resources become a crucial component of the angel's decision behavior because they invest their own money. It is also important for a BA to accept the fact that he needs to experiment the venture by himself with personal information search, non-structural meetings and physical site visits. Assuming a good and qualitative future relationship is a mean for the angel, contrary to the VC, to manage the agency risk during the post-investment interaction (Kelly and Hay, 2003).

For Huang (2017), BAs proceed through an intuiting process through which they justify and substantiate investment action. Intuition is opposed to the analytical and rational thinking mode and is more suitable to manage complex and uncertain context. According to this author, more than the hard data, angels will gauge the executional capability based on the interpersonal interactions with the entrepreneur (the perceived risk is then reduced by signals and cues that they picked up on as entrepreneurs gave pitch presentations, and appraisals of the entrepreneur that were more emotionally based). The « exceptionalism and the distinctivism » of the entrepreneur or the team is assessed in relation with a focal opportunity. Finally, the intuition of the angel will yield a fit between the perception of the founding entrepreneur and the business viability. This is the combination of hard and soft data to produce a subjective feeling that can be trusted. By intuition, angels actively process tacit and hard data in a new way in order to infer the distinctive ability of the entrepreneurial team. The ability to identify opportunities is then hidden in their information processing approach in a context where reliable data are missing (Ola, 2016). The reliance on intuition by the BA denotes the subjective nature of their investment decision and is justified by the complexity of the assessment of EF (Jeffrey et al, 2016). Angels are unaware about the influence the presentational or the tacit cues have on their investment decision and they are sometimes reluctant to acknowledge this influence (Clark, 2008).

It is largely accepted that angels use their judgment when dealing with EF. Such approach manifests itself in the way information is selected and process in situ to make sense that will sustain the investment decision. Intuitive judgment represents a distinctive skill in angel's investing. Trusting one own feeling is an angel's investing requirement. Intuition when described like non-efficient and biased decision-making approach (Tversky and Kahneman, 1974), however, is prioritized by angels and will distinguish them from other categories of investors. We think that angels are good in identifying good deals in EF because they are able to use reliably their intuition. They can more trust their gut feeling based on tacit information produced by themselves and integrate them into a relevant decision-making scheme to create meaning despite uncertainty. The more an angel will be able to use their intuition and to trust their personal feeling, the more he can identify investment opportunity in a focal EF.

The collective decision making in angel groups

The growing interest of networks and syndicates with communities of practices is another characteristic of the angels' activities (Paul et al, 2007). The experience and the behaviors of other colleagues are important in making such an investment. Syndication gives the opportunity for business angels to benefit from the combined experiences of a group of individuals (Kelly and Hay, 2003). Many eyes seeing a deal can help gauge and as a consequence lower the inherent risk. Angels' group is a specific efficient organization with an added-value for their decision. Indeed, investing in EF requires tacit, difficult to transfer information such as the investor's trust in the entrepreneur's character and intuition about future market trends, which are acquired through direct interaction with the entrepreneur or long-term personal experience respectively (Wu, 2015). In fact, some of the agents in a group may experiment difficulty with the tacitness of information, because the information acquisition process for their decision is individual and costly. Such agents will still participate in the collective choice of deals, making their vote worse than useless. They dilute the quality of the group decision that would have occurred without them. Angels' groups or networks are recognized as effective in identifying investment opportunities, which means that their internal working rule (explicit or not) is useful in overcoming the difficult transfer of tacit cues. The stronger relationship between deal-makers helps mitigate concerns about the opportunistic use of information, and supported more reciprocal and fair exchanges of complex forms of information (tacit and not formally codifiable) (Scarbrough et al., 2013).

According to Smith et al (2010), angels identified other colleagues inside their group as a key source of their learning: « you get to the stage where you have people that you respect and take opinion from who you listen to their views who are far more experienced than me. He and I have various involvements together so we use each other as a sounding board. I bring in someone else for a second opinion. So if I look at a proposition and think it is good I bring in another person to rip the proposition apart for me to see if I have missed anything. So that is how I operate now. » The above quote summarizes very well the importance of the social dimension or group aspect of the BAs decision to invest. These angels note that learning from other, more experienced angels, both members of the same investment syndicate and investment partners, have had a significant impact on how a group overcomes the difficulty in assessing EF. It has been also demonstrated that the number of investors starting evaluating a startup is a predictor of the final success in the investment process (Argerich et al, 2013).

Angel group can be conceived like a temporary system where trust needs to be developed in order to ease behaviors or decision making. The temporary situation represents « situation where the individuals have a limited history of working together, have limited prospects of working together in the future, and are involved in tasks that are "often complex and involve independent work..., have a deadline... are non-routine and not well-understood...[and] are consequential...[requiring] continuous interrelating [with others in the group] to produce an outcome" (Harrison et al, 1997). Inside a temporary system, trust must be conferred ex-ante of experience telling an individual that another is trustworthy. Trusting other and basing one's decision making on other views need a specific cognitive aptitude. It seems that angels have the secret in placing considerable reliance on their group's members who will provide them with reliable information for their individual judgments.

According to Scarbrough et al (2013), inside the temporary system formed by angels, their vulnerability is not a matter of « whom to trust » but rather their ability to contribute to the collaborative environment and to sustain the reliable sharing of information in order to assess EF. According to Ola et al. (2017), groups of angels have developed aptitude of delegating, collective thinking and sharing of knowledge. Delegating means that group members trust enough each other and they intentionally delegate to their colleague more close to a deal the task to assess

more in-depth the potential of a new project. Collective thinking means that the assessment of deal is always group-based with a committee in charge of interacting with and challenging the entrepreneurial team. Sharing inside the group finally means that investment decision is always based on the collective acceptance that an EF needs and is ready for the support provided by the angel group.

Investment decision of angels seems to be largely group-based with an efficient collaborative environment. Each member acknowledges the influence of colleagues on his decision making. The social capital developed helps to ease the transfer of tacit information and their integration in personal view to build up a final individual judgment. The collective nature of the decision is very important for analysis as judgment and intuition are more conceived like an individual and unjustifiable decision approach (Mintzberg, 1976). A reliable and efficient organization is formed with interesting collaborative characteristics. Angel's groups represent organizations where individual and collective level interoperate reliably and efficiently. A trusting environment has been developed to sustain cognitive operating. More importantly, members are open-mind and able to use a decision-making approach based on tacit cues and social resources collectively constructed. Then, the more the angel will be able to use such social resources and participate in the collective environment the better will be their ability to identify investment opportunity in EF.

These three dimensions justify considering the cognitive strategy of BAs in order to deeply analyze their ability to identify investment opportunities represented by the EF in which they decide to invest. But, at this stage, there is a missing link to prove that BAs are specific investors who are able to face reliably uncertainty and ambiguity. We think that the mindfulness of BAs can explain such capabilities.

How mindfulness dimensions explain the cognitive strategy of BAs

Definition of mindfulness

The term mindfulness is rooted in the Buddhist tradition where it provided a framework for mental training for centuries. Mindfulness is commonly defined "as the state of being attentive to and aware of what is taking place in the present" (Brown and Ryan, 2003). The key aspects of this practice that have been brought into business and management literature are attention and awareness (e.g. Brown and Ryan, 2003; Petchsawang and McLean, 2017; Good et al., 2016; Leroy et al., 2013).

Experiential processing emerges in the literature as being a fundamental process of mindfulness, which involves attention to the internal (e.g. thought, emotion) or external stimulus and registering observable facts (Brown et al., 2007). In experiential processing, common psychological content such as mental images, self-talk and emotions are consciously processed by individuals (Good et al., 2016). This mode of processing has been referred to as "decentralization" (Bishop et al., 2004; Brown et al., 2007; Good et al., 2016) as it requires a high level of awareness when attending experiences. Most mindfulness interventions teach the clients to become aware of their thoughts and feelings in a decentered way rather than as reflections of the self (Bishop et al., 2004).

Like many other psychological constructs, mindfulness has been studied either as being a trait (e.g. Brown and Ryan, 2003; Baer et al., 2006; Desbordes et al., 2012; Ostafin and Kassman, 2012) or dispositional (e.g. Beckman et al., 2012; Mrazer et al., 2012; Ruocco and Direkoglu, 2013; Beach et al., 2013). Research on mindfulness as traits assume that some people have a higher tendency to be mindful than others. For instance, this stream of research established a positive relationship between the tendency of being mindful and insight problem solving (Ostafin

and Kassman, 2012), job performance (Dane and Brummel, 2014) and empathy (Dekeyser et al., 2008).

In turn, research on dispositional mindfulness focuses on the ability of individuals to adopt a mindful state. The literature on mindfulness as a state found, for instance, a relationship between mindful states and reduced mind wandering (Mrazer et al., 2012), improved communication quality (Beckman et al., 2012), improved relations quality between leaders and subordinates (Reb et al., 2014). Mindfulness practices have also attracted the attention of researchers (e.g. Lutz et al., 2009; Maclean et al., 2010; Desbordes et al., 2012; Roeser et al., 2013). These practices include both those that focus attention to perceptual stimuli and those that create awareness of the various stimuli (Good et al., 2016). For instance, research in this has found a positive impact of mindfulness practice on sustained attention (Maclean et al., 2010) and attentional stability (Lutz et al., 2009).

Throughout this article, we refer to mindful individuals or mindful Business Angels to describe those high in mindfulness, independently of considering this mindfulness as a trait or a disposition. We also refer to mindfulness practice to refer to training and interventions that aim to increase mindfulness. Therefore, our focus is the cognitive phenomenon itself which can be facilitated or not by mindfulness practice. This follows the same line of reasoning introduced by Good et al. (2016).

Two key aspects of mindfulness are awareness and attention. Awareness refers to the conscious registration of stimuli, using physical senses, providing direct and most immediate contact with reality (Brown et al. 2007). Before awareness to occur, a stimulus must be sufficiently strong to engage attention. In turn, attention guides awareness to a specific element of the experienced reality, which is manifest as an initial "taking notice of, or "turning toward" that element (Bishop et al., 2004). Our argument is that awareness and attention will make it easier for BAs to identify "fatal flows" and to better assess risk and return. By being aware of and paying attention to their own reactions to new investment opportunities BA will be able to maintain a degree of mental distance or disengagement from self-relevant evaluations. This build on the fact that high awareness and attention increases the capacity to witness events, thoughts and emotions and judge them without being biased by personal memories, learned associations, or future projections (Good et al., 2016). Awareness and attention explain the open mind style of the angels. They can notice and integrate reliably different kind of information, tacit and explicit ones. Tacit and not transferable information are produced actively in situ by focusing one's attention on the present and being aware in the context. It is largely accepted that soft highly detailed cues are privileged by BA (Smith et al, 2010; Harrison et al, 1997). BAs need to actively assess trustworthiness, passion, tenacity and involvement. These qualities will be accurately perceived only if awareness can guide the attention of the mindful angel to those critical elements present in the experienced reality that is the personal information search, non- structural meetings and physical site visits (Carpentier and Suret, 2015). The cost of soft information is very high due to its specific nature and the context in which it is processed.

Mindfulness also means the conscious registration and processing of outside stimuli. Combined with intuitive judgment, mindfulness will decrease the level of unconsciousness during this decision making process. With consciousness, BAs will use more efficiently their intuition when making decisions about IO. They acknowledge the added-value of an intuitive judgment process and choose voluntarily to proceed according to this cognitive strategy. Mindfulness will boost a kind of "conscious intuitive judgment". Their willingness to use heuristics underlying intuitive judgment is high. The recognition of the usefulness of intuition by the BA is observable in the green literature and many interviews reported by researchers. They will be more aware of the flaws and bias that can emerge during intuitive judgement and therefore will be more confident on their own judgement.

Mindfulness impact on Human Functioning

Mindfulness has an impact on major domains of human function such as cognition, emotion, behavior, and physiology through an increase in attention (Good et al., 2016).

Attention

According to previous research, mindfulness improves three qualities of attention – stability, control and efficiency (Good et al., 2016). Attention stability can be defined as "ability to attend to a task without distraction" (Mrazer et al., 2012 pp. 776). Mindfulness training (Mrazek et al., 2012) and dispositional mindfulness (Mrazek et al., 2013) have shown to reduce mind wandering— defined "as a shift of attention from a task to unrelated concerns" (Mrazer et al., 2012 pp. 776) — and therefore increasing attention stability.

Evidence suggests that attentional processes can be trained, for example for emotion regulation, suggesting that attention processes are plastic (i.e. malleable) and can become more efficient with practice (Wadlinger and Isaacowitz, 2011). Most of mindfulness training and practice focus precisely on training attentional processes and increasing awareness, developing attentional control and reducing attention to distracting information (Good et al., 2016). For instance, Tang et al. (2007) found that meditators tend to be less distractible.

Mindfulness increases attentional control and lessens attention to off-task thoughts or activities, attention becomes more efficient (Good et al., 2016). When applied to business angels, mindfulness increases attention by sustaining their attention when evaluating IO (attention stability), by selecting appropriate risk and return indicator by controlling their attention (attention control) and by making economical use of attentional resources (attention efficiency). This leads to the following proposition:

Proposition 1: Mindfulness increases the attention stability, control and efficiency of business angels.

Angel needs to maintain attention during entrepreneur's pitch and when personally experiencing the venture with the entrepreneurial team. Non-compensatory heuristic means the rejection of an idea because only one element is missing, even if the score is high on all the other aspects (Maxwell et al, 2011). It means that IO is perceived like a hole where each part is useful to construct the full image. Pitch is, most of the time, the starting point of the long interaction between entrepreneurial team and BAs. Investors have to spend a long time analyzing investment opportunities from initial screening to the final agreement. There will be many meetings and negotiations with the entrepreneurial team. So, whatever is the stage during which the BA will experience the venture, they need to have stable attention and control that attention. This allows them to attend their task of building the right image of the project without distraction and to increase attentional efficiency. They need to be able to relate each new information to an old cue already presented or displayed by the team members. When assessing the risk level, they will continue to search for new deal killer or fatal that can stop the process at any time (Clark, 2008). This leads to the following proposition:

Proposition 2: Attention stability, control and efficiency have a positive impact on business angels' cognitive strategy underlying the identification of IO by BA.

Cognition

Working memory is a malleable aspect of cognitive capacity (Good et al., 2016). "The term working memory refers to a brain system that provides temporary storage and manipulation of the information necessary for such complex cognitive tasks as language comprehension, learning and reasoning." (Baddeley, 1992 pp. 556).). For instance, when listening to a pitch BAs must retain a large amount of information to support their decision-making, information that will be discarded (i.e. forgotten) after the decision being made. Mindfulness widens the attentional breadth (e.g., seeing more peripheral stimuli and being less focused on a specific target) of the actor in situ and he will need their full working memory in order to withdraw sense from all these information (Good et al., 2016). Mindfulness has been linked to increasing working memory capacity (e.g. Roeser et al., 2013). Its training can also positively impact fluid intelligence as the "ability to process and respond to novel information by assessing patterns and relationships" (Good et al., 2016 pp. 119). Fluid intelligence might be of paramount importance for BAs, especially when dealing with highly innovative and novel investment opportunities.

Mindfulness has also been associated with flexible cognition (Good et al., 2016). Flexible cognition allows individuals to generate novel perspectives and responses what can be advantageous in contexts of high informational complexity (Walsh, 1995). Therefore, flexible cognition is advantageous to Business Angels as it supports them making decisions under uncertainty and noise. Mindfulness also plays an important role in flexible cognition as it showed to be associated with better insight and problem-solving (Ostafin and Kassman, 2012), which are important aspects of BAs activity. Research suggests that brief mindfulness training has a positive impact on problem-solving ability by facilitating the switch of attention from dominant but incorrect representations to less frequent but correct ones (Ding et al., 2015).

According to these studies, it is possible to consider that mindfulness increases cognitive capacity and flexibility via its impact on attention (Good et al., 2016) what will have a positive impact on BAs functioning. We can, therefore, draw the following proposition:

Proposition 3: Mindfulness has a positive impact on cognitive capacity and cognitive flexibility of Business Angels, via its effects on attention.

The due diligence aims at sense making before investing (Weick and Robert, 1993). The attention of the BA needs to take advantage of their working memory resources. Cognitive energy is important at this stage to facilitate the "fusion" of both hard and soft data about the opportunity (Paul et al, 2007). During due diligence, risk and return have to be assessed separately by aggregating many different components (Jeffrey et al, 2016) (Huang, 2017; Huang and Pearce, 2016), which requires high cognitive capacity. Each BA needs a large loading capacity as many of this information is tacit and not transferable (tenacity, passion, inspirational leadership, trust displays, interpersonal relationship qualities, presentational skills...) (Maxwell and Levesque, 2014; Murniecks et al, 2016). Pro-social factors are dominant in the BA's perception. Syndications or investors' groups give the opportunity for business angels to benefit from the combined experiences of others (Kelly and Hay, 2003; Scarbrough et al, 2013; Smith et al, 2010). According to Smith et al (2010), angels identify other colleagues inside their group as a key source of their learning and their decision. The involvement of others, their experience and knowledge are used to weight the BA own opinion before an investment decision. So the efficient and collaborative environment formed by investors' groups eases the transfer of tacit information and then increases the quantity of information for all members. The heuristic approach and the intuitive judgment process privileged by BAs can support the integration of diverse families of factors. Their cognitive capacity will be useful in the loading and the in situ processing of cues. It is largely recognized that the intuitive judgment of the BA involves the subjective construction of patterns and relationships between factors (Hisrich and Jankowicz, 1990). The competence of BAs implies a holistic perception with the subjective non-visible link created in situ (Ola, 2016). So a high cognitive capacity with working memory and fluid intelligence will influence an angel's cognitive strategy.

Their intuitive judgment ability is supposed to support an ability to perceive the value creation of an entrepreneurial idea that will disrupt the market (authors). They need to envision the future of the venture through the speech of the entrepreneurial team without sufficient objective evidence (Ola, 2016, Huang and Pearce, 2016). More known BA always affirms that their role is not to fund and coach an already existing business idea but to identify the ones that will succeed because of the unique and innovative business model. Then following Ding et al (2015), angel's attention will raise their cognitive flexibility in order to identify the uniqueness of an entrepreneurial idea.

During the assessment of IO by BA, cognition is in function to help deal with different contents (objective, subjective and social indicators) and to help in searching for the newness and uniqueness in the entrepreneurial idea at the early stage. Proposition 4 is as follows:

Proposition 4: Cognitive capacity and cognitive flexibility have a positive impact on Business Angels' cognitive strategy underlying the identification of IO by BA.

Emotion

Emotions result from appraisals of the person-environment relationship that are elicited by affective events (Conroy et al., 2017) – i.e. by incidents that result in a reaction from the individual. Different individuals have different emotional recovery times from those incidents (Davidson, 1998) and mindfulness appears to shorten that recovery by reducing the emotional life cycle (Good et al., 2016). For instance, Goldin and Gross (2010) found that mindfulness training reduces the emotional reactions life cycle of patients with social anxiety. Mindfulness also seems to influence the reactivity to emotional stimuli (Good et al., 2016). Research found that mindful individuals tend to display reduced negative emotions after being exposed to stressors (Arch and Craske, 2010). This may occur due to a shift in emotional appraisal promoting a more neutral evaluation because of mindful-experiential processing (Good et al., 2016).

Emotional tone or valence gives an indication if the emotional response is positive or negative. The high level of attention that can be found in mindful states may inhibit the access to the perceived past or future that can lead to negative emotions (e.g. regret about the past, anxiety about the future) (Good et al., 2016). Indeed, a meta-analysis about the effects of mindfulness training found that mindfulness practices are associated with less negative and more positive emotional tone (Eberth and SedImeier, 2012).

Mindfulness seems to impact emotions via attention as it will influence the selection of stimuli for observation (Good et al., 2016). Ultimately, it will influence how the stimuli are evaluated and appraised and consequently the emotional reactions (Wadlinger and Isaacowitz, 2011). This leads to the following proposition:

Proposition 5: Attention stability, control and efficiency have an impact on the emotion life cycle, reactivity and valence of Business Angels

BAs are experienced entrepreneur and/or manager. They have stored cognitive schemes of entrepreneur that can be activated at the pitch session or during the interactions. Activated schemes are most often attached with emotional cues or somatic marker (Jeffrey et al., 2016).

Among the preferred indicators of angels are passion, enthusiasm, trust that are all subject to emotional reaction (Cardon et al, 2017; Huang and Knight, 2015). These factors are displayed through language, image and symbols (Sort and Nielson, 2017) and such media can easily activate the link between stores cognitive schemes and emotional responses. So, to avoid the selection bias or quick reactivity due to slow emotional recovery, mindful attention of BA can be useful for a more neutral assessment of IO without being disturbed by initial non-controlled responses.

Positive feeling vis-à-vis entrepreneurship is necessary to motivate BA in their investment activity. The personal feeling that entrepreneurship is at the heart of development and that it might be supported is a leitmotiv for many interested BAs (more than a mean to make a financial gain)(Avdeitchikova et al., 2008; Hellmann et al., 2015). Positive or negative initial feelings can influence attention and cognition and yield an investment decision without completing a serious assessment. The success rate of the past investees is very low and poor past investment performance or difficulty to exit or opportunist entrepreneurs or non-cooperative business team can weaken their involvement (regret about the past, anxiety about the future). Then, BAs must be able to put in bracket such probable negative personal feeling and mindful attention can moderate the effect of such emotional valence. The above argument about the role of emotion in human functioning can support the proposition below:

Proposition 6: Emotion life cycle, reactivity and valence have a positive impact on Business Angels' cognitive strategy underlying the identification of IO by BA.

Behavior

Changes of behaviour as a result of mindfulness occur through two main mechanisms, self-regulation and reduced automaticity (Good et al., 2016). Mindfulness improves self-regulation of behaviours by creating a separation between the self (e.g., ego, self-esteem, self-concept) and events, emotions, and experiences (Glomb et al., 2011). This may be explained by the process of "decentralization" (Bishop et al., 2004; Brown et al., 2007) through which common psychological content such as mental images, self-talk and emotions are consciously processed by individuals (Good et al., 2016).

The mechanism of self-regulation, as a result of mindfulness, is deeply linked to reduced automaticity, the effect of which is a mental gap between stimulus and behavioral response (Good et al., 2016). Automaticity is a process by which individuals adopt behaviors with no conscious oversight of the operational details, especially when cognitive capacity is constrained (Bargh and Chartrand, 1999). By increasing cognitive capacity, mindfulness reduces automaticity. Mindfulness also reduces automaticity by fostering awareness of automatic operations and behaviors, creating conditions for the individual to choose whether to adopt the automatic response or consciously regulate behavior (Good et al., 2016).

Mindful attention creates a gap between the stimulus (e.g. a negative feature of the IO) and the habitual response (not make an investment), which enables choicefulness and consequently more effective behavioral regulation (make a question to get clarification about the negative feature of the IO) (Good et al., 2016). This leads to the following proposition:

Proposition 7: Attention stability, control and efficiency have an impact on behavior self-regulation and reduced automaticity

For BA, reducing automaticity will lead to the ability to avoid the selection bias when screening project at the first step of the process for example. Experiences can sometimes guide the search

for fatal flows when using a non-compensatory heuristic (Maxwel et al, 2011). The BA must be capable to avoid the influence of negative investment experiences, to wisely stay away from the negative feedback he had from previous interactions with entrepreneurs (successful investment are very scarce in the angel investment activity). A new due diligence based on non- compensatory heuristic and intuitive judgment need to be authentic and focused on the strengths and the weaknesses of each case because every new EF is unique. Experiencing himself entrepreneurial idea in order to form judgment (Carpentier and Suret, 2015) is a reliable approach of decisionmaking because it allows behavioral control of the decision process. Self- regulation and reduced automaticity concern also the relationship between all the members of an angels' group. The behavior observed inside the angel group is more than herding and automatic responses observed for example on the financial market. It is a matter of involvement and contribution to a collaborative environment in order to cope with the uncertainty and the lack of objective evidence (Scarbrough et al, 2013). Internal mechanisms emerge in such group to self-regulate behavior and to reduce automaticity at the individual level (Ola et al, 2017). BA participates to a group to get credible information that can be reliably integrated into their own intuitive judgment before deciding to invest.

Of course, our argumentation in the previous sections points also that the mindful attention of the BA positively influence their cognitive abilities, help control emotion life cycle and the valence effect. He will use more efficiently the non-compensatory heuristic and their intuitive judgment in assessing IO because he is able to self-regulate and limit bias in their investment behavior. We can then formulate the following proposition:

Proposition 8: Behavior self-regulation and reduced automaticity have an impact on Business Angels' cognitive strategy underlying the identification of IO.

All these propositions are represented through the following conceptual model (Figure 1) aiming at explaining the identification of IO by business angels. We believe this is a first important step in understanding why BA can dominate the early stage financing of entrepreneurial firms while such activity is very risky in financial terms. More than the investment decision making, we aim to improve understanding of the decision-making behavior in uncertainty in organizations. BA and their groups are a good example of an efficient organizing activity to sustain business choices.



Figure 1: Theoretical model of the role of mindfulness in explaining the identification of investment opportunity by business angel
Conclusion

Mindfulness, through its impact on human functioning can be linked to the three dimensions of the behavior of BAs. Indeed, the consequences of a mindful state seem to explain the way BAs are taking their decision despite the specific characteristics of EF. These young innovative firms have a high specific risk due to their newness and uniqueness. They constrain the cognitive strategy of the first external investor represented here by an active BA. Linked the important literature that described and the less important one that tried to explain the uncommon decision process of BA, the 8 propositions are, to our knowledge, a first attempt to begin to explain the specificities of the cognitive strategy of BAs. If verified, these propositions could prove that angel's investing require specific cognitive abilities. Moreover, it will be possible to prove that their angels' decision process is suitable in unpredictable contexts represented by uncertainty and noise. Indeed, the mindfulness can explain the BA's attention (stability, control, efficiency). Moreover, mindfulness has an impact on human functioning (cognition, emotion, behaviour). Through increasing of BA attention and higher efficiency of their decision process, mindfulness can explain the three dimensions of BA cognitive strategy in situ. Those three dimensions have been identified through the literature review and are the non-compensatory heuristic, the intuitive judgment and the collective decision. Finally, because they have such a cognitive strategy, explained by their mindful state, these active informal investors are able to identify investment opportunities, when others are not. Our analysis can also contribute to the literature on decision making under uncertainty. We think that mindfulness can definitively useful to move intuition from a biased decision-making approach (Tversky and Kahneman, 1974) to an efficient situated decisionmaking method (Ola, 2016; Huang, 2017).

At this stage, it is necessary to test our propositions. For this purpose, we will need to construct a specific scale regarding the different abilities of BAs that have been identified. We will use the different scales that already exist. We will also contact angels' groups first on the French market and after on several European markets in order to make some international comparison. Even, if there are fewer women BAs, a gender analysis at the European level could be interesting regarding some results in behavioural finance on this specific gender comparison. Finally, it would be very interesting to conduct the same type of analysis with other types of investors, as VC or individuals investing in equity crowdfunding platforms.

Bibliography

- Ang, J., 1992. On the Theory of Finance for Privately Held Firms. The Journal of Entrepreneurial Finance 1, 185–203.
- Argerich, J., Hormiga, E., Valls-Pasola, J., 2013. Financial services support for entrepreneurial projects: key issues in the business angels investment decision process. The Service Industries Journal 33, 806–819.
- Arch, J. J., & Craske, M. G. (2010). Behaviour Research and Therapy Laboratory stressors in clinically anxious and non-anxious individuals: The moderating role of mindfulness. Behaviour Research and Therapy, 48(6), 495–505.
- Avdeitchikova, S., 2008. On the structure of the informal venture capital market in Sweden: developing investment role. venture capital 10, 55–85.
- Avdeitchikova, S., Landström, H., Mansson, N., 2008. What do we mean when we talk about business angels? Some reflections on definitions and sampling. venture capital 10, 371–374.
- Baddeley, A. (1992). Working Memory. Science, 255(5044), 556-559.
- Baer, R.A., Smith, G.T., Hopkins, J., Krietemeyer, J., Toney, L., 2006. Using Self-Report Assessment Methods to Explore Facets of Mindfulness. Assessment 13, 27–45.
- Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. American Psychologist, 54(7), 462–479.
- Beach, M. C., Sankar, A., Roter, D., Korthuis, P. T., Epstein, R. M., Sharp, V., ... Saha, S. (2013). A multicenter study of physician mindfulness and health care quality. Annals of Family Medicine, 11(5), 421–8. https://doi.org/10.1370/afm.1507
- Beckman, H. B., Wendland, M., Mooney, C., Krasner, M. S., Quill, T. E., Suchman, A. L., & Epstein, R. M. (2012). The impact of a program in mindful communication on primary care physicians. Academic Medicine, 87(6), 815–819.
- Bishop, S. R. (2004). Mindfulness: A Proposed Operational Definition. Clinical Psychology: Science and Practice, 11(3), 230–241. https://doi.org/10.1093/clipsy/bph077
- Bonini, S., Capizzi, V., Valletta, M., Zocchi, P., 2018. Angel network affiliation and business angels' investment practices. Journal of Corporate Finance 50, 592–608.
- Brown, K. W., & Ryan, R. M. (2003). The Benefits of Being Present: Mindfulness and Its Role in Psychological Well-Being. Journal of Personality and Social Psychology, 84(4), 822–848. https://doi.org/10.1037/0022-3514.84.4.822
- Brown, K. W., Ryan, R. M., & Creswell, J. D. (2007). Mindfulness: Theoretical foundations and evidence for its salutary effects. Psychological Inquiry, 18(4), 211–237.
- Cardon, M.S., Mitteness, C., Sudek, R., 2017. Motivational Cues and Angel Investing: Interactions Among Enthusiasm, Preparedness, and Commitment. Entrepreneurship Theory and Practice 41, 1057–1085. https://doi.org/10.1111/etap.12255
- Carpentier, C., Suret, J.-M., 2015. Angel group members' decision process and rejection criteria: A longitudinal analysis. Journal of Business Venturing 30, 808–821.
- Cassar, G., 2004. The financing of business start-ups. Journal of business Venturing 19, 261–263.
- Clark, C., 2008. The impact of entrepreneurs' oral "pitch" presentation skills on business angels' initial screening investment decision. Venture capital 10, 257–279.
- Conroy, S. A., Becker, W. J., & Menges, J. I. (2017). The Meaning of My Feelings Depends on Who I Am: Work-Related Identifications Shape Emotion Effects in Organizations. Academy of Management Journal, 60(3), 1071– 1093.
- Cox, K.C., Lortie, J., Gramm, K., 2017. The investment paradox: why attractive new ventures exhibit relatively poor investment potential. Venture Capital 19, 163–181.
- Croce, A., Tenca, F., Ughetto, E., 2017. How business angel groups work: Rejection criteria in investment evaluation. International Small Business Journal 35, 405–426.
- Dane, E., & Brummel, B. J. (2014). Examining workplace mindfulness and its relations to job performance and turnover intention. Human Relations, 67(1), 105–128.
- Davidson, R. J. (1998). Affective Style and Affective Disorders: Perspectives from Affective Neuroscience, 12(3), 307–331.
- Ding, X., Tang, Y. Y., Cao, C., Deng, Y., Wang, Y., Xin, X., & Posner, M. I. (2015). Short- term meditation modulates brain activity of insight evoked with solution cue. Social Cognitive and Affective Neuroscience, 10(1), 43–49.
- Dreyfus, G. (2011). Is mindfulness present-centred and non-judgmental? A discussion of the cognitive, Contemporary Buddhism 12 (1).
- Eberth, J., & Sedlmeier, P. (2012). The Effects of Mindfulness Meditation: A Meta-Analysis. Mindfulness, 3(3), 174–189. https://doi.org/10.1007/s12671-012-

- Geibel, R., Yang, J., 2018. Extended Study on Characterizing Business Angels and Their Impact on Start-up Success. GSTF Journal on Business Review (GBR) 5.
- Glomb, T. M., Duffy, M. K., Bono, J. E., & Yang, T. (2011). Mindfulness at Work. In The Wiley Blackwell Handbook of the Psychology of Positivity and Strengths-Based Approaches at Work (Vol. 7301, pp. 115–157).
- Goldin, P. R., & Gross, J. J. (2010). Effects of Mindfulness-Based Stress Reduction (MBSR) on Emotion Regulation in Social Anxiety Disorder. Emotion, 10(1), 83–91.
- Good, D. J., Lyddy, C. J., Glomb, T. M., Bono, J. E., Brown, K. W., Duffy, M. K., ... Lazar, S.
- W. (2016). Contemplating Mindfulness at Work: An Integrative Review. Journal of Management, 42(1), 114–142. https://doi.org/10.1177/0149206315617003
- Harrison, R., Dibben, M.R., Mason, C.M., 1997. The Role of Trust in the Informal Investor's Investment Decision: An Exploratory analysis. Entrepreneurship Theory and Practice 21.
- Harrison, R.T., Don, G., Johnston, K.G., Greig, M., 2010. The early-stage risk capital market in Scotland since 2000: issues of scale, characteristics and market efficiency. Venture capital 12, 211–239.
- Harrison, R.T., Mason, C.M., 2017. Backing the horse or the jockey? Due diligence, agency costs, information and the evaluation of risk by business angel investors. International Review of Entrepreneurship 15, 269–290.
- Hellmann, T.F., Schure, P., Vo, D., 2015. Angels and Venture Capitalists: Substitutes or Complements? (SSRN Scholarly Paper No. ID 2602739). Social Science Research Network, Rochester, NY.
- Hisrich, R.D., Jankowicz, A.D., 1990. Intuition in venture capital decisions: An exploratory study using a new technique. Journal of Business Venturing 5, 49–62.
- Hodgkinson, G.P., Healey, M.P., 2011. Psychological foundations of dynamic capabilities: reflexion and reflection in strategic management. Strategic Management Journal 32, 1500–1516. https://doi.org/10.1002/smj.964
- Huang, L., 2017. The Role of Investor Gut Feel in Managing Complexity and Extreme Risk. AMJ 61, 1821–1847. https://doi.org/10.5465/amj.2016.1009
- Huang, L., Knight, A.P., 2015. Resources and Relationships in Entrepreneurship: An Exchange Theory of the Development and Effects of the Entrepreneur-Investor Relationship. AMR 42, 80–102. https://doi.org/10.5465/amr.2014.0397
- Huang, L., Pearce, J.L., 2015. Managing the Unknowable The Effectiveness of Early-stage Investor Gut Feel in Entrepreneurial Investment Decisions. Administrative Science Quarterly 60, 634–670. https://doi.org/10.1177/0001839215597270
- Jeffrey, S.A., Lévesque, M., Maxwell, A.L., 2016. The non-compensatory relationship between risk and return in business angel investment decision making. Venture Capital 18, 189–209. https://doi.org/10.1080/13691066.2016.1172748
- Kahneman, D., Tversky, A., 1982. On the study of statistical intuitions. Cognition 11, 123–141. Kelly, P., Hay, M., 2003. Business angel contracts: the influence of context. Venture Capital 5,
- 287–312.
- Langer, E.J., 1989. Minding Matters: The Consequences of Mindlessness–Mindfulness, in: Berkowitz, L. (Ed.), Advances in Experimental Social Psychology. Academic Press, pp. 137–173. https://doi.org/10.1016/S0065-2601(08)60307-X
- Leroy, H., Anseel, F., Dimitrova, N. G., & Sels, L. (2013). Mindfulness, authentic functioning, and work engagement: A growth modelling approach. Journal of Vocational Behavior, 82(3), 238–247. https://doi.org/10.1016/j.jvb.2013.01.012
- Lutz, A., Slagter, H. A., Rawlings, N. B., Francis, A. D., Greischar, L. L., & Davidson, R. J. (2009). Mental Training Enhances Attentional Stability: Neural and Behavioral Evidence. Journal of Neuroscience, 29(42), 13418–13427.
- MacLean, K. A., Ferrer, E., Aichele, S. R., Bridwell, D. A., Zanesco, A. P., Jacobs, T. L., ... Saron, C. D. (2010). Intensive Meditation Training Improves Perceptual Discrimination and Sustained Attention. Psychological Science, 21(6), 829–839.
- Mason, C., Botelho, T., Zygmunt, J., 2016. Why business angels reject investment opportunities: Is it personal? International Small Business Journal 0266242616646622. https://doi.org/10.1177/0266242616646622
- Maxwell, A.L., Jeffrey, S.A., Lévesque, M., 2011. Business angel early-stage decision making. Journal of Business Venturing 26, 212–225.
- Maxwell, A.L., Lévesque, M., 2014. Trustworthiness: A Critical Ingredient for Entrepreneurs Seeking Investors. Entrepreneurship: Theory & Practice 38, 1057–1080.
- Milliken, F.J., 1987. Three Types of Perceived Uncertainty about the Environment: State, Effect, and Response Uncertainty. The Academy of Management Review 12, 133–143. https://doi.org/10.2307/257999
- Mintzberg, H., 1976. Planning on the left side and managing on the right. Harvard Business Review 54, 49–58.
- Moesel, D.D., Fiet, J.O., 2001. Embedded fitness landscapes?part 2: Cognitive representation by venture capitalists. Venture Capital 3, 187–213.
- Murnieks, C.Y., Cardon, M.S., Sudek, R., White, T.D., Brooks, W.T., 2016. Drawn to the fire: The role of passion, tenacity and inspirational leadership in angel investing. Journal of Business Venturing 31, 468–484.

- Mrazek, M. D., Franklin, M. S., Phillips, D. T., Baird, B., & Schooler, J. W. (2013). Mindfulness Training Improves Working Memory Capacity and GRE Performance While Reducing Mind Wandering. Psychological Science, 24(5), 776–781.
- Mrazek, M. D., Smallwood, J., & Schooler, J. W. (2012). Mindfulness and mind-wandering: Finding convergence through opposing constructs. Emotion, 12(3), 442–448.
- Ocasio, W. (2011). Attention to Attention. Organization Science, 22(5), 1286-1296.
- OECD, 2011. Financing High-Growth Firms: The Role of Angel Investors. OECD Publishing. http://dx.doi.org/10.1787/9789264118782-en
- Ola, A.M., 2016. L'identification des opportunités d'investissement en incertitude : : le jugement intuitif des Business Angels dans le financement des firmes entrepreneuriales.
- Ola, A.M., Deffains-Crapsky, C., Dumoulin, R., 2019. Vers une nouvelle approche de l'investissement en amorçage : un raisonnement à travers la théorie de l'alignement structurel. Finance Contrôle Stratégie A paraître.
- Ola, A.M., Deffains-Crapsky, C., Dumoulin, R., 2017. Investissement en amorçage des business angels : la plusvalue du collectif au sein des réseaux. Presented at the deuxième journée de l'Innovation Abbé Grégoire, CNAM, Paris, Frnace.
- Ostafin, B. D., & Kassman, K. T. (2012). Stepping out of history: Mindfulness improves insight problem solving. Consciousness and Cognition, 21(2), 1031–1036.
- Paul, S., Whittam, G., Wyper, J., 2007. Towards a model of the business angel investment process. Venture capital 9, 107–125.
- Pedchenko, N., Strilec, V., Kolisnyk, G.M., Dykha, M.V., Frolov, S., 2018. Business angels as an alternative to financial support at the early stages of small businesses' life cycle.
- Petchsawang, P., & McLean, G. N. (2017). Workplace spirituality, mindfulness meditation, and work engagement. Journal of Management, Spirituality and Religion, 14(3), 216–244.
- Reb, J., Narayanan, J., & Chaturvedi, S. (2014). Leading Mindfully: Two Studies on the Influence of Supervisor Trait Mindfulness on Employee Well-Being and Performance. Mindfulness, 5(1), 36–45. https://doi.org/10.1007/s12671-012-0144-z
- Scarbrough, H., Swan, J., Amaeshi, K., Briggs, T., 2013. Exploring the Role of Trust in the Deal-Making Process for Early-Stage Technology Ventures. Entrepreneurship: Theory & Practice 37, 1203–1228. https://doi.org/10.1111/etap.12031
- Simon, H.A., 1987. Making Management Decisions: the Role of Intuition and Emotion. ACAD MANAGE PERSPECT 1, 57–64. https://doi.org/10.5465/AME.1987.4275905
- Smith, D.J., Mason, C.M., Harrison, R.T., 2010. Angel Investment Decision Making as a Learning Process. Working Paper 10-05, University of Strathclyde, Scotland UK.
- Sørheim, R., Landström, H., 2001. Informal investors A categorization, with policy implications. Entrepreneurship & Regional Development 13, 351–370.
- Sort, J.C., Nielsen, C., 2018. Using the business model canvas to improve investment processes. J of Res in Mrkting & Entrepre 20, 10–33.
- Sun, H., Fang, Y., 2010. Toward a model of mindfulness in Technological Acceptance, in: ICIS 2010 Proceedings. Presented at the International Conference on Information Systems (ICIS).
- Tang, Y.-Y., Ma, Y., Wang, J., Fan, Y., Feng, S., Lu, Q., ... Posner, M. I. (2007). Short-term meditation training improves attention and self-regulation. Proceedings of the National Academy of Sciences, 104(43), 17152–17156.
- Tversky, A., Kahneman, D., 1974. Judgment Under Uncertainty: Heuristics and Biases. Science, New Series 185, 1124–1131.
- Wadlinger, H. A., & Isaacowitz, D. M. (2011). Fixing our focus: Training attention to regulate emotion. Personality and Social Psychology Review, 15(1), 75–102.
- Walsh, J. P. (1995). Managerial and Organizational Cognition: Notes from a Trip Down Memory Lane. Organization Science, 6(3), 280–321.
- Weick, K. E., & Roberts, K. H. (1993). Collective Mind in Organizations: Heedful Interrelating on Flight Decks. Administrative Science Quarterly, 38(3), 357.
- White, B.A., Dumay, J., 2017. Business angels: a research review and new agenda. Venture Capital 19, 183–216. https://doi.org/10.1080/13691066.2017.1290889
- Wu, A., 2016. Organizational Decision-Making and Information: Angel Investments by Venture Capital Partners. Proceedings 2016, 11043.



Contents lists available at SCOPUS

Finance, Risk and Accounting Perspectives Proceedings



homepage: http://www.acrn.eu/frap

About ACRN

The ACRN Oxford Research Center was originally founded as a joint venture by three leading European business schools of renowned universities from Austria, Finland and the UK to promote and host academic conferences and summer-schools in finance, risk and accounting, as well as to advance methodological know-how for doctoral students and young researchers. It is now a distinct (social enterprise) entity with a clear impact mission and a highly respected international academic advisory board overseeing its corporate governance, firmly embedded in the Oxford research and publication landscape.

Our Mission

To provide inspiring courses, conferences and summer schools for academic faculty and doctoral students, with a focus on interdisciplinary research, and multi-perspective approaches in finance, risk and accounting research.

To disperse and translate current academic research into practice through open access publication outlets and advisory services to policy makers.

In our activities and intra-organizational culture, we embrace and advocate heterogeneity and tolerance, and pay attention to context and environment!

We are a social enterprise for higher education, any excess profits are used to further our programs and outreach to the international academic community.