Kira, Mohamed Ihab "Estimation of a Firm's Optimal Scale of Operations and Size-Related Analyses" *ACRN Oxford Journal of Finance and Risk Perspectives 6.4 (2017): 1-16.*

ESTIMATION OF A FIRM'S OPTIMAL SCALE OF OPERATIONS AND SIZE-RELATED ANALYSES

MOHAMED IHAB KIRA
Department of Financial Management and Accounting Arab Academy for Science, Technology and Maritime Transport (Dokki Branch)

Received: 13 November 2017
Revised: 10 December 2017
Accepted: 30 December 2017

***Abstract**: The paper proposes a framework that can be utilized for deciding upon the optimal scale of operations for a firm. The basic premise of the paper is that a firm that operates in a stable economic environment and a stable mode of business operations experiences a varying elasticity of scale that, after some scale level, becomes less than one and continues to decline with further increases in scale. A firm operating in such a stable system can search for its optimal scale of operations. The paper's premise does not refute the possibility that a firm can experience a shift from a lower elasticity of scale to a higher one; but this shift is not part of the mechanics of a stable system. Rather, it is due to some shock to the economic environment and/or the firm's mode of operations. Firms that operate under the assumption that they face a constant elasticity of scale, when in fact the elasticity varies, expose themselves to detrimental consequences. The model is in congruence with the Tobin's q criteria and can shed some light on why average q is a poor estimate for marginal q. The model sheds some light on the small-firm effect and on a major difference between internal and external growth.*

***Keywords**: optimal scale, elasticity of scale, Tobin's q, small-firm effect, internal growth, external growth*

JEL: G30, L1

Introduction

It is important for firms to fine tune their scale-of-operations decisions to avoid squandering valuable resources on assets that quickly become redundant; this is especially important in the current economic environment that experiences frequent innovative and 'disruptive' technologies that unsettle the status-quo (El-Erian, 2016). This fine-tuning should take place whether the firm operates in a stable economic environment with a stable mode of operations or in an unstable environment that forces it to change its mode of operations. The paper, however, is mainly concerned with the first situation in which it is reasonable to assume that the elasticity of scale for a firm varies and, beyond some scale level, becomes less than one and continues to decline with further increases in scale. In the second situation, the shocks to the system can cause a firm to change from a lower elasticity of scale to a higher one (Diewert et al, 2011) and the fine-tuning is more difficult but, nevertheless, should be undertaken as the system stabilizes.

Firms that operate under the assumption that they face a constant elasticity of scale, when in fact the elasticity varies, expose themselves to detrimental consequences. A firm needs to estimate as accurately as possible the underlying returns to scale regime it faces.

For tractability considerations, the paper does not delve into the topic of the growth rate with which a firm should approach its optimal scale of operations. There is a huge literature that analyzes the factors that have an impact on, and the distribution of, firms' growth rates. However, on a qualitative basis, one can argue that changes in scale generate returns to investors. If a firm reaches its optimal scale during a short period then the market return during this period might be stellar; in the long subsequent periods, however, during which there is no significant change in scale, the return might be mediocre. A firm might choose to average 'stellar' and 'mediocre' to achieve 'attractive' returns over a prolonged period of time during which it gradually approaches its optimal scale of operations. This perspective might provide a partial explanation for the empirical finding that smaller firms are characterized by faster growth rates compared to larger firms; even with faster growth rates they still have a long time to reach their optimal scale of operations, during which they can generate attractive returns. Several papers indicate that smaller firm sizes are associated with higher expected returns, the 'small-firm effect', (Banz, 1981; Reinganum, 1981; Keim, 1983; Chan et al, 1985; Fama and French, 1993)[1].

The model is used to shed some light on this 'effect'. It is also used to draw a distinction between scale optimality with internal growth, where marginal benefits are the overriding consideration, and scale optimality with external growth where both marginal and average benefit considerations are important.

The model is in congruence with the Tobin's q criteria and can shed some light on why average q is a poor estimate for marginal q. Expressions for average q and marginal q are generated and used, in a numerical example, to show why the former can be a very poor estimate for the latter. Ang and Beck (2000) report that average and marginal q ratios show significant differences and using the average measure as a decision variable can lead to erroneous investment decisions "…approximately fifty percent of the time."

The paper is organized as follows. The second section sets up the model, derives a formula for the elasticity of scale and shows how the optimal scale of operations can be estimated. The third section clarifies some technical details of the model. The fourth section shows the congruence of the model with the Tobin's q criteria. In the fifth section some light is shed on the small-firm effect and on a difference between internal and external growth. The last section concludes.

The Model

Assume that V is a reference level of the market value of a given firm (this could be the average current value). The firm's management wants to decide upon the scaling factor β, $-1 \le \beta \le \infty$, such that the market value becomes $V(1+\beta)$.

At the level V the total (replacement) financing cost is K. At the level $V(1+\beta)$ the total (replacement) financing cost is given by

[1] Other papers report that the evidence is mixed (Handa et al, 1989; Horowitz et al, 2000). The model might help to provide a partial explanation for these opposed arguments.

$$C = K\left((1-f) + \left(f^{\frac{1}{\alpha}} + \beta \right)^{\alpha} \right)^{\eta} \tag{1}$$

Note that when $\alpha = 1$,

$$C = K(1+\beta)^{\eta} \tag{2}$$

From equation (1):

$$\beta = \left(\left(\frac{C}{K} \right)^{\frac{1}{\eta}} - (1-f) \right)^{\frac{1}{\alpha}} - (f)^{\frac{1}{\alpha}} \tag{3}$$

$$V(1+\beta) = V\left(1 + \left(\left(\frac{C}{K} \right)^{\frac{1}{\eta}} - (1-f) \right)^{\frac{1}{\alpha}} - (f)^{\frac{1}{\alpha}} \right) \tag{4}$$

An expression for the elasticity of scale ε can be derived as follows:

$$\begin{aligned} \varepsilon &= \frac{dV(1+\beta)}{dC} \frac{C}{V(1+\beta)} \\ &= \frac{C}{\alpha\eta K((1+\beta))} \left(\left(\frac{C}{K} \right)^{\frac{1}{\eta}} - (1-f) \right)^{\frac{1}{\alpha}-1} \left(\frac{C}{K} \right)^{\frac{1}{\eta}-1} \end{aligned} \tag{5}$$

Replacing C with the right hand side of equation (1), this transforms (after some algebraic manipulation) to:

$$\varepsilon = \frac{\left(f^{\frac{1}{\alpha}} + \beta \right)^{(1-\alpha)} (1-f) + f^{\frac{1}{\alpha}} + \beta}{\alpha\eta(1+\beta)} \tag{6}$$

Equation (6) provides a tractable way to model a varying elasticity of scale.

At β = zero,

$$\varepsilon = \frac{f^{\frac{1-\alpha}{\alpha}}}{\alpha\eta} \tag{7}$$

Note that when α = 1, equations (6) and (7) reduce to[2]

$$\varepsilon = \frac{1}{\eta} \tag{8}$$

This is a constant elasticity of scale. It implies decreasing returns to scale when $\eta > 1$, increasing returns to scale when $\eta < 1$, and constant returns to scale when $\eta = 1$.

As a special case, when $f = 1/2$ equations (6) and (7) take the forms:

$$\varepsilon = \frac{\left(2^{-\frac{1}{\alpha}} + \beta\right)^{(1-\alpha)} + 2^{\frac{-1+\alpha}{\alpha}} + 2\beta}{2\alpha\eta(1+\beta)} \quad (6') \qquad \text{and} \qquad \varepsilon = \frac{2^{\frac{-1+\alpha}{\alpha}}}{\alpha\eta} \quad (7')$$

Example:
Let f = 0.5, α = 1.8 and η = 0.7
Using equation (7'), at β = zero ε = 1.08.
Using equation (6'), at β = 0.1 ε = 1.003, at β = 0.2 ε = 0.948, at β = 0.3 ε = 0.909
This example illustrates the peril of assuming the elasticity of scale is constant at the current level (at β = zero) when actually it varies and declines with scale.
Every effort should be made to estimate the parameters f, α and η as accurately as possible and, of course, sound economic sense should guide the estimation.
For example, an ε = 1.08 at β = zero can also be obtained with the following two combinations of α and η (f remaining at 0.5):

- η = 0.8 and α = 1.4215, which gives: at β = 0.1 ε = 1.03, at β = 0.2 ε = 0.996,

  at β = 0.3 ε = 0.9696 , a situation also of an ε that declines with scale but at a slower rate compared to the situation above.

- η = 0.7 and α =0.33582, which gives: at β = 0.1 ε = 1.6, at β = 0.2 ε = 2,

  at β = 0.3 ε = 2.3267 , an absurd situation of an ε that increases with scale.

The firm's management makes the scale of operations decision (i.e. decides upon β) in order to maximize the appreciation of current shareholders' wealth.

The firm's management decision program can be stated as:

$$\textit{Maximize} \quad \frac{u}{U}P^{AR} - S \quad \textit{or} \quad \textit{Maximize} \quad \frac{u}{U}P^{AR} - P^{BR} \quad \textit{or} \quad \textit{Maximize} \quad \frac{u}{U}P^{AR}$$

$$\textit{subject to the financing identity} \quad C = S + D + \frac{U-u}{u}P^{BR}$$

---

[2] Also, if $f = 1$ then $\varepsilon = \dfrac{1}{\alpha\eta}$. Thus when either f or α (or both) equal one then this indicates a constant elasticity of scale. If $f = 0$ then $\varepsilon = \dfrac{\beta^{(1-\alpha)} + \beta}{\alpha\eta(1+\beta)}$.

It does not make any difference which form of the objective function is used.

Where

- u is the number of shares that current shareholders own
- U is the total number of shares (including u) after rescaling operations with the scale parameter β
- P^{AR} is market value of the equity of the firm after rescaling operations
- P^{BR} is market value of the equity of the firm before rescaling operations which belongs to current shareholders
- S is the investment amount that current shareholders paid to the firm
- D is total debt of the firm after rescaling operations

Note that

- P^{AR} equals $V(1+\beta)-D$, total market value of the firm after rescaling operations minus total debt of the firm after rescaling operations
- $\frac{u}{U}P^{AR}$ is the part of the market value of the equity of the firm after rescaling operations that belongs to current shareholders
- $\frac{U-u}{u}P^{BR}$ is the amount of investment funds that the firm receives from new shareholders (equals: number of new shares to be issued ($U-u$) times current price per share (P^{BR}/u)

The decision program can be restated as:

$$\textit{Maximize}\quad \frac{u}{U}\left((1+\beta)V-K\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\beta\right)^{\alpha}\right)^{\eta}+S+\frac{U-u}{u}P^{BR}\right)$$

The first derivative of the objective function with respect to β is

$$\frac{u}{U}V-\frac{u}{U}\eta K\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\beta\right)^{\alpha}\right)^{\eta-1}\alpha\left((f)^{\frac{1}{\alpha}}+\beta\right)^{\alpha-1}$$

Equating this first derivative to zero leads to the following First Order Condition (FOC)[3]

$$V=\alpha\eta K\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\beta\right)^{\alpha}\right)^{\eta-1}\left((f)^{\frac{1}{\alpha}}+\beta\right)^{\alpha-1} \qquad (9)$$

[3] New prospective shareholders reach the same equation because their objective function differs from that of current shareholders in that (U-u)/U replaces u/U at the beginning of the function, which is immaterial.

This condition reflects $\dfrac{dV(1+\beta)}{d\beta} = \dfrac{dC}{d\beta}$

To evaluate whether the FOC reflects a maximum or a minimum the second derivative of the objective function with respect to β is needed (the second order condition SOC)

$$-\frac{u}{U}\eta K\left(\begin{array}{l}(\eta-1)\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\beta\right)^{\alpha}\right)^{\eta-2}\left(\alpha\left((f)^{\frac{1}{\alpha}}+\beta\right)^{\alpha-1}\right)^{2}\\ +\alpha(\alpha-1)\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\beta\right)^{\alpha}\right)^{\eta-1}\left((f)^{\frac{1}{\alpha}}+\beta\right)^{\alpha-2}\end{array}\right)$$

No closed form expression for optimal β can be derived from equation (9). However, for specific values for V/K, f, α and η, the equation can be solved for optimal β using mathematical programs; the calculated optimal β can then be plugged in the expression for the SOC to find if it represents a maximum (if the SOC is negative) or a minimum (if the SOC is positive).

Example (cont'd):
Continuing with the numerical example started in this section (f = 0.5, α = 1.8 and η = 0.7), adding that V/K= 1.2 , equation (9) yields an optimal β = 0.508 at which the SOC is negative (the firm needs to increase its scale by about 50%). ε at this level (calculated using equation (6)) equals 0.85.

If the parameters are erroneously estimated as η = 0.8 and α = 1.4215, the optimal β turns out to be 1.31, reflecting an increase of scale of 131%. It is perilous to make a decision based on such an estimate. But it is even more perilous to make a decision based on the assumption that the ε is constant at 1.08 (its β = 0 level) as is illustrated below.

Note that when α = 1 equation (9) reduces to:

$$V = \eta K(1+\beta)^{\eta-1} \quad (10)$$

which can be solved easily for optimal β

$$\beta_{opt} = \left(\frac{V}{\eta K}\right)^{\frac{1}{\eta-1}} - 1 \quad (11)$$

The SOC reduces to

$-\dfrac{u}{U}(\eta-1)\eta K(1+\beta)^{\eta-2}$ which is negative if η > 1 (reflecting a maximum) and positive if η < 1 (reflecting a minimum).

Example (cont'd):
Assume that, somehow, ε is correctly estimated to equal 1.08 but that it is erroneously expected to remain constant at that level. Based on this erroneous expectation and equation (8), η is estimated as 0.926. The SOC represents a minimum and all that management has to do is move as far as possible from the minimum calculated using equation (11); the sky is the limit, a perilous disposition indeed!

ACRN Oxford Journal of Finance and Risk Perspectives
Vol.6 Issue 4, December 2017, p.1-16
ISSN 2305-7394

The Reference Level is Immaterial

Some readers might be worried that the optimal scale of operations derived using the model is dependent on the reference level *V*. This worry might be due to the observation that the parameters *V/K*, *f*, α and η used in equation (9) are tied to the reference level.

The purpose of this section is both to try to alleviate this worry and to outline various relationships that need to be satisfied to ensure the internal consistency of the model.

Assume that a different reference level V^* is chosen which is related to the original *V* as follows:

$$V^* = V(1+\gamma)$$

$$K^* = K\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\gamma\right)^{\alpha}\right)^{\eta}$$

Where γ is the percentage change from *V* that leads to V^*

For a rescaling *R* made from the original reference level *V*

$$V^R = V(1+\beta^R)$$

$$C^R = K\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\beta^R\right)^{\alpha}\right)^{\eta}$$

If the same rescaling is made from the different reference level V^* the end results have to be the same, that is:

$$V^R = V(1+\beta^R) = V(1+\gamma)(1+\beta^{R*}) \tag{12}$$

β^R is the percentage change from *V* that leads to V^R

β^{R*} is the percentage change from V^* that leads to V^R

$$\beta^{R*} = \frac{(1+\beta^R)}{(1+\gamma)} - 1 \tag{13}$$

$$\begin{aligned} C^R &= K\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\beta^R\right)^{\alpha}\right)^{\eta} = K^*\left((1-e)+\left((e)^{\frac{1}{\delta}}+\beta^{R*}\right)^{\delta}\right)^{\theta} \\ &= K\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\gamma\right)^{\alpha}\right)^{\eta}\left((1-e)+\left((e)^{\frac{1}{\delta}}+\beta^{R*}\right)^{\delta}\right)^{\theta} \end{aligned} \tag{14}$$

e, δ and θ are the parameters tied to V^* that need to be calculated; since these are three unknowns three rescalings are needed to generate three equations of the form (14).

Example (cont'd):
Continuing with the numerical example ($f = 0.5$, $\alpha = 1.8$, $\eta = 0.7$, and $V/K = 1.2$)

Let the different reference level V^* be related to the original V as follows:
$V^* = V(1+\gamma) = V(1+0.3)$ where γ (0.3) is chosen to be less than the optimal β =0.508

It follows that

$$\frac{V^*}{K^*} = \frac{V}{K}\frac{1.3}{\left(\frac{1}{2}+\left(\left(\frac{1}{2}\right)^{\frac{1}{1.8}}+0.3\right)^{1.8}\right)^{0.7}} = 1.1941$$

Assume the first rescaling has a β = 0.45 (chosen to be less than the optimal β = 0.508) from the original reference level. Using equation (13)

$$\beta^{R*} = \frac{(1+0.45)}{(1+0.3)} - 1 = 0.115385$$

Using equation (14):

$$\left(\frac{1}{2}+\left(\left(\frac{1}{2}\right)^{\frac{1}{1.8}}+0.45\right)^{1.8}\right)^{0.7} = \left(\frac{1}{2}+\left(\left(\frac{1}{2}\right)^{\frac{1}{1.8}}+0.3\right)^{1.8}\right)^{0.7}\left((1-e)+\left((e)^{\frac{1}{\delta}}+0.115385\right)^{\delta}\right)^{\theta} \quad (15)$$

Assume the second rescaling has a β = 0.4 (chosen to be less than the optimal β = 0.508) from the original reference level. Using equation (13)

$$\beta^{R*} = \frac{(1+0.4)}{(1+0.3)} - 1 = 0.076923$$

Using equation (14):

$$\left(\frac{1}{2}+\left(\left(\frac{1}{2}\right)^{\frac{1}{1.8}}+0.4\right)^{1.8}\right)^{0.7} = \left(\frac{1}{2}+\left(\left(\frac{1}{2}\right)^{\frac{1}{1.8}}+0.3\right)^{1.8}\right)^{0.7}\left((1-e)+\left((e)^{\frac{1}{\delta}}+0.076923\right)^{\delta}\right)^{\theta} \quad (16)$$

Assume the third rescaling has a β = 0.35. Using equation (13)

$$\beta^{R*} = \frac{(1+0.35)}{(1+0.3)} - 1 = 0.038462$$

Using equation (14):

$$\left(\frac{1}{2}+\left(\left(\frac{1}{2}\right)^{\frac{1}{1.8}}+0.35\right)^{1.8}\right)^{0.7} = \left(\frac{1}{2}+\left(\left(\frac{1}{2}\right)^{\frac{1}{1.8}}+0.3\right)^{1.8}\right)^{0.7}\left((1-e)+\left((e)^{\frac{1}{\delta}}+0.038462\right)^{\delta}\right)^{\theta} \quad (17)$$

Equations (15), (16) and (17) can *theoretically* be solved together for e, δ and θ. This is a relatively complicated system of equations. Thankfully, a mathematics program was capable of solving (an algebraic elaboration on[4]) this relatively complicated system of equations. It turns out that e = 0.751554, δ = 2.21658, and θ = 0.580758 (it was found that these values satisfy the equations for other rescaling values quite well). Optimal β from the perspective of the different reference level, using equation (9), turns out to be 0.16. From equation (13)

$\left(1+\beta^{R}\right)=\left(1+\gamma\right)\left(1+\beta^{R*}\right)=1.3x1.16=1.508$, optimal β from the perspective of the original reference level is the same as calculated before. The two reference levels lead to the same scale of operations decision.

The irrelevance of the reference level can only be illustrated with numerical examples (and powerful mathematics programs) for the general case where α is not equal to one. It can also be shown with numerical examples that if γ = optimal β from the perspective of the original reference level, then the optimal β from the perspective of the different reference level is zero.

For the simpler case of α = 1, where ε and η are constants that do not depend on the reference level, it is easy to show the irrelevance analytically using equation (11):

$$V^{*}\left(1+\beta^{opt*}\right)=\left(\frac{V(1+\gamma)}{\eta K(1+\gamma)^{\eta}}\right)^{\frac{1}{\eta-1}}V(1+\gamma)=\left(\frac{V}{\eta K}\right)^{\frac{1}{\eta-1}}V=V\left(1+\beta^{opt}\right)$$

For the general case, the parameters tied to the different reference level (V^*) should also be estimable by utilizing a number of other relationships that need to be satisfied for the model to be internally consistent. Some of these relationships are listed below.

1) The elasticity at V^* should be the same whether calculated from the standpoint of V or V^*

$$\frac{\left(f^{\frac{1}{\alpha}}+\gamma\right)^{(1-\alpha)}(1-f)+f^{\frac{1}{\alpha}}+\gamma}{\alpha\eta(1+\gamma)}=\frac{e^{\frac{1-\delta}{\delta}}}{\delta\theta} \qquad (18)$$

For the numerical example of this section it can be calculated as follows:

$$\frac{\left(\frac{1}{2}^{\frac{1}{1.8}}+0.3\right)^{(-0.8)}\left(\frac{1}{2}\right)+\frac{1}{2}^{\frac{1}{1.8}}+0.3}{(1.8)(0.7)(1.3)}=0.90866=\frac{e^{\frac{1-\delta}{\delta}}}{\delta\theta} \qquad (19)$$

2) Since $V^{*}=V\left(1+\gamma\right)$ it follows that $V=V^{*}\left(1+\left(\frac{1}{1+\gamma}-1\right)\right)$, also since $K^{*}=K\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\gamma\right)^{\alpha}\right)^{\eta}$ then

[4] The elaboration is outlined in Appendix A.

$$\frac{V}{K}=\frac{V^*}{K^*}\frac{\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\gamma\right)^{\alpha}\right)^{\eta}}{(1+\gamma)}=\frac{V^*}{K^*}\frac{\left(\frac{1}{1+\gamma}\right)}{\left((1-e)+\left((e)^{\frac{1}{\delta}}+\left(\frac{1}{1+\gamma}-1\right)\right)^{\delta}\right)^{\theta}} \quad (20)$$

$$\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\gamma\right)^{\alpha}\right)^{\eta}\left((1-e)+\left((e)^{\frac{1}{\delta}}+\left(\frac{1}{1+\gamma}-1\right)\right)^{\delta}\right)^{\theta}=1 \quad (21)$$

3) The elasticity at V should be the same whether calculated from the standpoint of V or V^*

$$\frac{\left(e^{\frac{1}{\delta}}+\left(\frac{1}{1+\gamma}-1\right)\right)^{(1-\delta)}(1-e)+e^{\frac{1}{\delta}}+\left(\frac{1}{1+\gamma}-1\right)}{\delta\theta\left(\frac{1}{1+\gamma}\right)}=\frac{f^{\frac{1-\alpha}{\alpha}}}{\alpha\eta} \quad (22)$$

These relationships reflect consistency between the two sets of parameters for the two reference levels.

Congruence of the Model with Tobin's *q* criteria

Average q at a given scale equals:

$$\frac{Total\ Firm Value}{Total\ Financing\ Needs}=\frac{V}{K}\frac{(1+\beta)}{\left(\frac{1}{2}+\left(\left(\frac{1}{2}\right)^{\frac{1}{\alpha}}+\beta\right)^{\alpha}\right)^{\eta}} \quad (23)$$

At β = zero (the reference level), average q equals V/K.

Marginal q equals:

$$\frac{Change\ in\ Total\ Firm\ Value}{Change\ in\ Total\ Financing\ Needs}=\frac{V}{K}\frac{\left(\beta^2-\beta^1\right)}{\left(\frac{1}{2}+\left(\left(\frac{1}{2}\right)^{\frac{1}{\alpha}}+\beta^2\right)^{\alpha}\right)^{\eta}-\left(\frac{1}{2}+\left(\left(\frac{1}{2}\right)^{\frac{1}{\alpha}}+\beta^1\right)^{\alpha}\right)^{\eta}} \quad (24)$$

Example (cont'd):
Continuing with the numerical example (f = 0.5, α = 1.8, η = 0.7, and V/K= 1.2)
At β = zero average q equals V/K = 1.2
At optimal β = 0.508 average q equals (using equation (23)) 1.17
Let β^2 = 0.503 and β^1 = 0.498

Marginal *q* between these two close scale levels can be calculated using equation (24) and is equal to 1.003. Since $q > 1$, the *q* criteria (Tobin, 1969) indicates that the firm should move to $\beta^2 = 0.503$.
Let $\beta^2 = 0.508$ and $\beta^1 = 0.503$
Marginal *q* between these two close scale levels is equal to 1.0009. Since $q > 1$, the *q* criteria indicates that the firm should move to $\beta^2 = 0.508$ (the optimal level).
Let $\beta^2 = 0.5131$ and $\beta^1 = 0.508$
Marginal *q* between these two close scale levels is equal to 0.999. Since $q < 1$, the *q* criteria indicates that the firm should stay at $\beta^1 = 0.508$ (the optimal level).
Let $\beta^2 = 0.51821$ and $\beta^1 = 0.5131$
Marginal *q* between these two close scale levels is equal to 0.997. Since $q < 1$, the *q* criteria indicates that the firm should not go to β^2 (not even to β^1 as shown above)
The example above serves to show that marginal *q* is significantly different from average *q* and that using average *q* as an estimate for marginal *q* can lead to erroneous investment decisions as indicated by Ang and Beck (2000).

A Numerical Exploration of the Small Firm Effect and Growth Policy

The purpose of this section is to provide a rationale for the 'small-firm effect' and to show that this effect is mediated by growth rates. It also attempts to draw a distinction between scale optimality with internal growth, where marginal benefits are the overriding consideration, and scale optimality with external growth where both marginal and average benefit considerations are important.

The Small-Firm Effect

The framework developed in this paper abstracts, for tractability reasons, from the dynamics of financing (e.g. retaining earnings, payouts, consecutive equity issuances).

Thus, it does not allow for a parsimonious modeling of the 'rate of return'. As a proxy for the rate of return, the 'rate of change of value added', R_{VA}, is used; it is defined as the rate of change of the excess of value over replacement cost.

$$R_{VA} = \frac{(1+\beta_2)V - K\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\beta_2\right)^{\alpha}\right)^{\eta} - (1+\beta_1)V + K\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\beta_1\right)^{\alpha}\right)^{\eta}}{(1+\beta_1)V - K\left((1-f)+\left((f)^{\frac{1}{\alpha}}+\beta_1\right)^{\alpha}\right)^{\eta}} \quad (25)$$

Simplifying and dividing numerator and denominator by *K*

$$R_{VA} = \frac{\frac{V}{K}(\beta_2-\beta_1) - \left((1-f)+\left((f)^{\frac{1}{\alpha}}+\beta_2\right)^{\alpha}\right)^{\eta} + \left((1-f)+\left((f)^{\frac{1}{\alpha}}+\beta_1\right)^{\alpha}\right)^{\eta}}{\frac{V}{K}(1+\beta_1) - \left((1-f)+\left((f)^{\frac{1}{\alpha}}+\beta_1\right)^{\alpha}\right)^{\eta}} \quad (26)$$

One can use equation (26) in a numerical example to compare this 'quasi rate of return' for a small and a large firm. Such a numerical example follows.

Example:
The parameters for the 'small firm' are: $f = 0.5$, $\alpha = 1.5$, $\eta = 0.7$, and $V/K = 1.07$

Using equation (7), at β = zero $\varepsilon = 1.19992$

Equation (9) yields an optimal $\beta = 2.1$ at which the SOC is negative (the firm needs to increase its scale by about 210 %).

To develop the parameters, for a 'large firm', that are in the same 'regime' as the small firm the methodology of third section is used, thus it is assumed that the optimal scales for both the small and large firms are the same.

$V_{large} = V_{small}(1+1)$ the large firm is twice the size of the small firm.

$$K_{large} = K_{small}\left((1-f)+\left((f)^{\frac{1}{\alpha}}+1\right)^{\alpha}\right)^{\eta}$$

Going through the same procedure as that followed in the third section, it was found that the parameters tied to V_{large} : e, δ, θ, and V^*/K^* are : 0.93917, 2.50451, 0.4307 and 1.10196 respectively. For V_{large}, $\varepsilon = 0.96267$ and optimal $\beta = 0.55$; as stated above both the small and large firms have the same optimal scale: $(1+2.1)=(1+1)(1+0.55)$

First Exploration

Assume that both the small and large firms grow 10% each period, then the percentage change in their scale (the β*'s* to be used in equation (26)) at the end of each of the first four periods is as follows: 0.1, 0.21, 0.331, and 0.4641.

Plugging a pair of consecutive β*'s* in equation (26) together with the parameters for one of the two firms results in the following consecutive quasi returns

Small firm: 0.314, 0.206, 0.191, 0.136

Large firm: 0.054, 0.038, 0.031, 0.015

The calculation of the quasi return for the small firm in period one is shown below:

$$\frac{1.07(0.1-0)-\left(0.5+\left((0.5)^{\frac{1}{1.5}}+0.1\right)^{1.5}\right)^{0.7}+\left(0.5+\left((0.5)^{\frac{1}{1.5}}+0\right)^{1.5}\right)^{0.7}}{1.07(1+0)-\left(0.5+\left((0.5)^{\frac{1}{1.5}}+0\right)^{1.5}\right)^{0.7}}=0.314$$

Second Exploration

Assume the large firm grows as in the first exploration. The small firm grows at a very small pace such that the percentage change in its scale at the end of each of the first four periods is as follows: 0.01, 0.03, 0.05, 0.07. The small firm's quasi returns are: 0.034, 0.063, 0.058, 0.053.

In the first exploration there is a huge difference between the quasi returns of the small firm and the large firm. In the second exploration the difference is rather insignificant.

It is proposed that the small firm effect is mediated by the difference in growth rates.

Internal versus External Growth

In the numerical example of the previous sub-section, the large firm is twice the size of the small firm and its optimal $\beta = 0.55$ (said another way the small firm is half the size of the large firm and its optimal $\beta = 2.1$); It might seem that it makes sense that the two firms should merge. This is not correct.

A major difference between internal growth (which is what is discussed throughout the paper up to this point) and external growth (through mergers & acquisitions) is that the overriding consideration in the former is the marginal aspect of benefits (benefits defined as excess of value over replacement cost) whereas in the latter both the marginal and average aspects of benefits are important; the average aspects are important because there are two parties sharing the benefits.

Internal growth

The excess of value over replacement cost for each firm is calculated before and after rescaling and expressed in terms of the small firm's K

Large Firm:

Before: $V^*/K^* = 1.10196$, therefore $V^* - K^* = 0.10196\,K^*$

Since, $K^* = K\left(0.5 + \left((0.5)^{\frac{1}{1.5}} + 1\right)^{1.5}\right)^{0.7} = 1.942\,K$ then $V^* - K^* = 0.198\,K$

After: the firm expands with a $\beta = 0.5$

$$(1+0.5)V^* - K^*\left((1-0.93917) + \left((0.93917)^{\frac{1}{2.50451}} + 0.5\right)^{2.50451}\right)^{0.4307}$$
$$= (1+0.5)\,1.10196\,K^* - 1.53604\,K^* = 0.1169\,K^* = 0.227\,K$$

The benefits increase from 0.198 K to 0.227 K; the firm should pursue internal growth.

Small Firm:

Before: $V/K = 1.07$, therefore $V - K = 0.07\,K$

After: the firm expands with a $\beta = 2$

$$(1+2)V - K\left(0.5 + \left((0.5)^{\frac{1}{1.5}} + 2\right)^{1.5}\right)^{0.7}$$
$$= (1+2)1.07\,K - 2.98296\,K = 0.227\,K$$

The benefits increase from 0.07 K to 0.227 K; the firm should pursue internal growth.

External growth

The sum of the excess of value over replacement cost for the two firms before rescaling, expressed in terms of the small firm's K, can be calculated from the information above and equals $0.198\,K + 0.07\,K = 0.268\,K$. For the merger to be beneficial for both firms the combined firm should have an excess of value over replacement cost greater than 0.268 K. As shown above, this is not the case, the combined firm has an excess of value over replacement cost of just 0.227 K[5]. The merger has a detrimental effect on both firms.

However, mergers between small firms can be beneficial. Consider two exactly similar small firms each having the same characteristics as before.

The sum of the excess of value over replacement cost for the two firms before rescaling, expressed in terms of K, can be calculated from the information above and equals
$0.07\,K + 0.07\,K = 0.14\,K$

If the firms merge, the excess of value over replacement cost for the combined firm is:

$$(1+1)V - K\left(0.5 + \left((0.5)^{\frac{1}{1.5}} + 1\right)^{1.5}\right)^{0.7}$$

$$= (1+1)1.07\,K - 1.942\,K = 0.198\,K$$

The merger is beneficial for both firms because the combined firm has an excess of value over replacement cost of 0.198 K which is greater than the sum of the excess of value over replacement cost for the two firms before rescaling (0.14 K).

Optimality from the perspective of internal growth is very different from optimality from the perspective of external growth.

Closing Remarks

The paper presents a model for the determination of the optimal scale of operations for a firm. The model allows for varying elasticity of scale, for which a formula is derived. While the parameters of the model are associated with a specific reference scale of operations, the estimated optimal scale is independent of this reference level when the parameters are changed correctly for different reference levels. The model is in congruence with the Tobin's q criteria. The paper illustrates why average q is a poor estimate for marginal q. The model is useful for firms seeking to optimize their scale decisions in order not to squander valuable resources. The model is used to shed some light on the 'small-firm effect' and the results from numerical illustrations indicate that this effect is mediated by differences in growth rates for firms of different sizes. The model provides some guidance to mergers/acquisitions decisions and the results from numerical illustrations indicate that optimality from the perspective of internal growth is quite different from optimality from the perspective of external growth.

[5] If the two firms are already combined then a separation makes sense. The value of the separated parts is greater than the value of the combined firm.

Appendix A

The purpose of this brief appendix is to outline the approach taken to simplify and solve the system of equations (15), (16), and (17) for e, δ and θ.

The three equations can be re-written as follows:

$$1.13108 = \left((1-e) + \left((e)^{\frac{1}{\delta}} + 0.115385 \right)^{\delta} \right)^{\theta} \tag{15}$$

$$1.08651 = \left((1-e) + \left((e)^{\frac{1}{\delta}} + 0.076923 \right)^{\delta} \right)^{\theta} \tag{16}$$

$$1.0428 = \left((1-e) + \left((e)^{\frac{1}{\delta}} + 0.038462 \right)^{\delta} \right)^{\theta} \tag{17}$$

Divide equation (15) by equation (17) and divide equation (16) by equation (17). A system of two equations results. Take logarithms of both sides of each of the two equations; θ is down from the power position. Substitute for θ using equation (19)

$$\theta = \frac{e^{\frac{1-\delta}{\delta}}}{0.90866\delta} \tag{A.1}$$

Solve the system of two equations for e, and δ. Substitute in equation (A.1) to get θ.

References

Ang, J.S. and Beck, K.L. (2000). A Comparison of Marginal and Average Tobin's Ratios. *International Journal of Business*, 5(1), 27-54.

Banz, R.W. (1981). The Relationship between Return and Market Value of Common Stocks. *Journal of Financial Economics*, 9, 3-18.

Chan, K.C., Chen, N., and Hsieh, D.A. (1985). An Exploratory Investigation of the Firm Size Effect. *Journal of Financial Economics*, 14, 451-471.

Diewert, W.E., Nakajima, T., Nakamura, A., Nakamura, E. and Nakamura, M. (2011). Returns to scale: concept, estimation and analysis of Japan’s turbulent 1964–88 Economy. *Canadian Journal of Economics*, 44 (2), 451–485.

El-Erian, M.A. (2016). The Only Game in Town – Central Banks, Instability, and Avoiding the Next Collapse. (Random House, New York N.Y.)

Fama, E.F. and French, K.R. (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, 33, 3-56.

Handa, P., Kothari, S.P., and Wasley, C. (1989). The Relation between the Return Interval and Betas, Implications for the Size Effect. *Journal of Financial Economics*, 23.

Horowitz, J.L., Loughran, T., and Savin N.E. (2000). Three Analyses of the Firm Size Premium. *Journal of Empirical Finance*, 7, 143-153.

Keim, D.B. (1983). Size-Related Anomalies and Stock Return Seasonality, Further Empirical Evidence. *Journal of Financial Economics*, 12, 13-32.

Reinganum, M.R. (1981). Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earnings’ Yields and Market Values. *Journal of Financial Economics*, 9, 19-46.

Tobin, J. (1969). A General Equilibrium Approach to Monetary Theory. *Journal of Money, Credit and Banking* (February), 15-29.