Organizational capital and the existence of a diversification and size 'discount'

Juan Santalo

University of Chicago

November 27, 2001

Abstract

I analyze the implications of a standard model of the firm where I allow heterogeneity in one unobserved component, which I label organizational capital. Under some reasonable assumptions, firms with better organizational capital are larger and have a lower rate of profit both per unit of output and per unit of input. The model, thus, generates a negative relation between size of firms and Tobin's Q, which I find in the data. When this relation is accounted for, the well-known diversification discount disappears, suggesting that this might simply be a feature of large firms.

1 Introduction[1]

When we use a production function as a simple representation of what a firm is, we know that we are trading a great deal of realism for also a great deal of analytical convenience. Also for simplicity, we tend to understate the degree of heterogeneity among firms. Concepts usually forgotten with these abstractions are the so-called organizational capital and business culture of the firm, concepts that for the purpose of this paper will be almost synonymous. Yet, firms are indeed different. '*Actual organizations seem to have personalities, which are rather stable over time and independent of their actual members*'[2]; and business cultures of corporations like General Motors, Goldman Sachs or IBM are undoubtedly a real world phenomena that seems very far away from the simplified and homogeneous representation of the firm we commonly use.

I try to capture one aspect of this heterogeneity. Firms differ in the quality of their organizational capital. I will show how, using Rosen's (1982) model, we can incorporate the main functions and characteristics of the organizational capital in a simple production function, with interesting implications regarding a whole empirical literature that deals with the 'diversification discount puzzle'. In particular I will show how heterogeneity in the quality of the organizational capital can generate a negative relation between size and profit margins, even in a pure neoclassical model with no agency costs. More importantly, I show that the 'size discount' this model predicts can explain the diversification discount.

What is organizational capital? Nelson -Winter (1982) note that even firms that compete in the same industry operate with very different production processes. *Firms differ in the*

[1]This paper has greatly benefited from the comments and suggestions of Luis Garicano, Kevin Murphy, Boyan Jovanovic, Lester Telser and Luigi Zingales. It also has taken advantage of useful discussions with Bo Becker, Ricard Gil Gilles Hillary, and Claudio Irigoyen. Ernesto Revilla and Esteban Rossi-Hansberg have given me outstanding advice in how to better express the argument of the paper.

[2]Cremer (1993).

procedures for hiring and firing, ordering new inventory, or for stepping out production of items in high demand; they also differ in policies regarding investment, research and development or advertisement. Nelson-Winter also mention that firms have *different business strategies about product diversification and overseas investment*

Nelson-Winter generically call all the above procedures 'routines'. In their evolutionary theory, routines play the same role that genes play in biological evolutionary theory. '*They are a persistent feature of the organism and determine its possible behavior; they are heritable in the sense that tomorrow's organisms generated from today's (for example, by building a new plant) have many of the same characteristics, and they are selectable in the sense that organisms with certain routines may do better than others, and if so, their relative importance in the population (industry) is augmented over time*'[3].

In this paper, I am not interested in an evolutionary theory of the firm; on the contrary, below I will try to introduce the idea of 'routine' from Nelson-Winter in a very simple standard neoclassical production function, but their description of routines represents very well the concept I have in mind when I talk about organizational capital. Instead of the word 'routines' I use the pair of words 'organizational capital', but the concept is the same. Furthermore, I keep the analogy with genes in the sense that the organizational capital will be a persistent feature of the firm, too costly to change. Nelson-Winter exposition serves the purpose of illustrating the concept of organizational capital, but their discussion is too broad and abstract if we want to translate the effect of this concept into properties of the production function. To achieve this goal, I will use in Section 2 the elements and definitions of Cremer (1993) regarding *organizational culture*[4].

Section 3 presents a neoclassical model of diversification with no agency costs. Based in Rosen (1982), the model captures the characteristics of the organizational capital in a simple

[3] Nelson- Winter (1982) page 14.

[4] In fact the organizational culture concept of Cremer considers a subset of the broad set of routines that Nelson-Winter mention.

production function. This specification proves successful in providing a set of implications consistent with the findings of a broad literature about the diversification discount and additional implications regarding the existence of a size discount and the profitability of larger and more diversified firms.

In particular, the model implies that diversified firms and large firms share a common characteristic: a better quality of their organizational capital. This common characteristic translates in a similar behavior of their market to book and market to sales ratios. Firms with better organizational capital have higher total profits, and higher volume of production but less profit both per unit of output and per unit of input. Since market value is the discounted value of the stream of profits, lower profit per unit of input translates in a lower market to book ratio while lower profit per unit of output implies a lower market to sales ratio.

The negative relationship between diversification and market to book and market to sales ratio has been widely reported in the literature. This negative relation has received a name: the diversification discount. Lang-Stulz (1994), Berger-Ofek (1995) and Servaes (1996) among others have documented the existence of a diversification discount for such a long period as 1960-2000. Section 4 replicates and discusses these results.

Maksimovic-Phillips (2002) present a neoclassical model in which an observed diversification discount arises endogeneously due to differences in organizational capabilities. They show how if individual firms have different capabilities in different sectors and if individual firms optimaly choose in which sectors to enter we can observe a discount in the data. In this paper, individual firms have the same capabilities in all sectors. Firms differ among each other in the total quality of their organizational capital. With this different assumption, we also predict a diversification (and size) discount in the data without postulating any inefficient behavior from the part of the firm.

Section 5 documents the inverse relation between size and market to book and market to sales. I follow the exact methodology that has been used to document the diversification discount. Section 5 also shows how this size discount can explain the diversification discount. That is, diversified firms do not have a lower market to book and market to sales ratio when compared with non-diversified firms of similar size.

Section 6 studies the relationship between profitability and both size and diversification. According to the model of section 3, more diversified firms and larger firms should have lower profit both per unit of input and per unit of output. I show evidence that this is indeed the case once we use a measure of earnings free of accounting biases.

Sections 7 and 8 discuss the relation of this paper with other literature related to the diversification discount. Campa-Kedia (1999) and Villalonga (2000) argue that the discount causes the move to diversification and not the reverse. When they control by the endogeneity of the diversification decision, the estimates of the diversification discount become either a premium or non-significant. Section 8 reinterprets these findings. Firms with better organizational capital enter in a larger number of sectors. Also, firms with better organizational capital have less profit per unit of input/output that translates in lower market to book/sales ratio. Therefore, the unobserved firm characteristic that causes the self-selection problem can be the quality of the organizational capital.

Schoar (2001) reports that plants belonging to diversified firms have higher total factor productivity. She also documents the dynamics of plant productivity after a diversifying move. Section 8 discusses how the model of section 3 can help to understand these findings. Section 9 concludes.

2 Characteristics of the Organizational Capital

Cremer (1993) defines organizational culture as part of the stock of knowledge shared by a substantial portion of the employees of the firm, but not by the general population. He specifies that the organizational culture is the portion of the firm-specific human capital that is shared by all the employees. He decomposes the organizational culture in three elements:

- A common language or coding. It is efficient to generate a specialized code to transmit ideas or facts that recur repeatedly. As a consequence, members of an organization share a specialized language or code that increases the speed of communication.

- A shared knowledge of certain facts. A certain number of facts are known by most members of the organization, this knowledge increases the effectiveness of communication allowing certain things to be left unsaid.

- A knowledge of certain established rules of behavior. There are three types: rules of social behavior like choice of clothing, rules of internal behavior like the allowed patterns of communication and rules of action for the firm as a whole like 'we only provide high quality products'.

At this point, we are ready to try to map the characteristics of these elements into properties of a simple production function. Note the following:

1. Some elements of the organizational culture increase the productivity of the factors of production independently of the scale, their effect is the same in a very large enterprise than in a very small one. Examples of these elements are rules to choose which goods to produce or what varieties. If a firm has better rules to choose which goods to produce, the increase in the productivity of a particular employee is independent of the total number of employees of the organization. These type of elements represent an increasing returns property of the organizational capital.

2. There are other elements of organizational culture that can be considered simply as one more input in the production function, and as that, subject to the usual property of decreasing returns to scale when keeping the amount of the rest of the factors of production fixed. Let me explain this very carefully with an example: a better/more efficient internal code will increase the speed of communication among the members of the organization, and

therefore it will increase the productivity of the factors of production. The key point here is to note that a better internal code has an effect on the productivity of a particular member that depends on how many people inside the organization this particular member communicates with. If employee **A** communicates with another ten employees the effect on her productivity of communicating faster is larger than if employee **A** only communicates with a single other employee. This is a source of decreasing returns to scale to the organizational capital. If we increase the quality of the code without increasing the number of people that uses this internal code to communicate, the increase in marginal productivity will decrease with the increase in the quality of the code[5].

3. Finally, note that some elements of the organizational capital seem quite industry specific. There are decreasing returns to the organizational capital when increasing the number of industries in which the firm is operating. For an illustration of this point I will use an idea extracted from Cremer-Garicano-Pratts (2001). Take the common language element of the organizational culture, all members of the organization share the same coding since I assume that there has to be a unique language inside an organization[6].If the organization operates in just one industry, the code in place will be the one that maximizes the speed of communication between the members of the organization when doing the tasks particular to that sector. Now imagine two different industries, A and B. The tasks in A are different than the tasks in B and/or the frequency that a particular task is performed is different in A than in B. Therefore, the optimal code, the one that maximizes the speed of communication in sector A**,** code 1, will be different to the optimal code in sector B**,** code 2**.** If the firm operates in both A and B then whatever code it actually uses, it will not be as efficient as code 1 for sector A and code 2 for sector B. This justifies the decreasing returns to the

[5]A complicated way to say that the second derivative of production with respect to the quality of the organizational capital will be negative.

[6]The argument follows exactly the same if we assume that it is costly to have two languages coexisting in the interior of a organization, for example if we have an interpreter that translates both languages.

organizational capital when increasing the number of sectors. In the appendix I illustrate this intuition with a very simple example.

The discussion of points 1, 2 and 3 above, suggests that a technology that realistically hopes to incorporate the functions of organizational capital in production needs to satisfy three different requirements: first it has to introduce the organizational capital with the usual properties we usually assume in other inputs of production, second it has to introduce an element that captures the increase in the productivity independent of the size of the firm, and third it has to introduce an element that takes into account the decreasing returns of organizational capital when firms diversify in more than one sector. Rosen (1982) specification satisfies these three requirements. I describe a version of his model in the next section.

3 A simple model with Organizational Capital in the production function

In this section, following Rosen (1982) I represent in a simple production function the properties and characteristics that the literature has attributed to the organizational capital. I derive implications regarding the span of control of resources of the organizational capital and the behavior of the market to book and market to sales ratios. These implications will give a guidance for the next sections of a more empirical nature.

3.1 Model and preliminary results

In the model I propose the source of variation among firms is organizational capital. Firms are endowed with organizational capital of a particular quality, $r \in R^+$, that differs among firms, some firms have better organizational capital than others. Organizational capital is not tradable and it is not separable from the property of the firm. The production function of sector i is:

$$y_i = g(r)f(rt_i, q_i(k_i))$$

with

$$\begin{aligned} \sum_{i=1}^{n} t_i &\leq T \\ q_i &= k_i^{a_i} \\ with\ 0 &< a_i < 1 \end{aligned}$$

where $y_i =$ Output in sector i, $t_i =$ 'Time' that the organizational capital devotes to sector i, $n =$Total number of sectors in which the firm is present, $T =$ Fixed endowment of time of the organizational capital, $k_i =$Capital employed in sector i. Furthermore, $f(rt_i, q_i)$ is assumed to be concave in each argument, with constant returns to scale between rt_i and q_i, and with $f(0, x) = f(x, 0) = 0\ \forall x$.

The function $g(r)$, with $g'(.) > 0$, represents a total factor productivity improvement. It captures the increasing returns characteristic of the organizational capital discussed in point 1 of the previous section. The first argument of $f(.)$ reflects the usual decreasing returns characteristics as discussed in point 2 above. The fixed T to share among all sectors captures -in an extreme form- the diminishing returns to organizational capital when entering in different sectors, as also was discussed in the last section. The fiction of a limited amount of time of such an abstract concept as organizational capital serves in a very convenient way the purpose of representing the idea of a extreme form of decreasing returns. Intuitively, if the organizational capital has to fit the production characteristics of n sectors, it will fit worse the production characteristics of each individual sector than if the firm was focused in just one industry. In other words, the more sectors the firm is operating the worse job does the organizational capital in each one of them.

Let m be the total number of sectors of the economy, and p_k the price per unit of capital. Assume there is a fixed cost F for entering in a particular sector. I assume that firms are price-takers and normalize all prices of products to one. Firms maximize profits:

$$\begin{aligned} Max_{\{k_i,t_i\}}\pi(r) &= Max_{\{k_i,,t_i\}}\{\sum_{i=1}^{m}[g(r)f(rt_i,q(k_i)) - p_k k_i - nF]\} \\ s.t. \sum_{i=1}^{m} t_i &\leq T \\ s.t.\ t_i &\geqslant 0\ \forall i = 1..m \end{aligned}$$

Assume that all sectors are identical: $a_i = a\ \forall i$. Let n denote the total number of sectors in which the firm is active. Using the assumption of constant returns to scale of f, $f(rt_i,q_i) = q_i\theta(rt_i/q_i)$ we can write the firm's problem as:

$$Max_{\{n,k\}}\pi = Max_{\{n,k\}}\{nk^a g(r)\theta(\frac{rT}{nk^a}) - p_k nk - nF]\} \tag{1}$$

Define $\bar{\beta} = \frac{rT}{nk^a}$, $\bar{\beta}$ is the inverse of the span of control of the organizational capital, where span of control is the amount of resources controlled by unit of quality of organizational capital. Note that

$$n = \frac{rT}{\bar{\beta}k^a}$$

Rewrite (1):

$$Max_{\{\bar{\beta},k\}}\{\frac{r}{\bar{\beta}}[g(r)\theta(\bar{\beta}) - p_k k^{1-a} - \frac{T}{k^a}F]\}$$

Normalize the total time of the organizational capital T to one. Abstracting from integer constraints, the FOCs of the problem determine the optimal number of sectors and the optimal size of each sector:

$$-(1-a)p_k k^{-a} + ak^{-a-1}F = 0 \tag{2}$$

$$[g(r)\theta(\bar{\beta}) - p_k k^{1-a} - \frac{1}{k^a}F]\frac{-r}{\bar{\beta}^2} + \frac{r}{\bar{\beta}}g(r)\theta'(\bar{\beta}) = 0 \tag{3}$$

From (2) we can get:

$$k = \frac{aF}{(1-a)p_k} \tag{4}$$

Note that k is independent both of the quality of the organizational capital, r, and the number of sectors n. There is an optimal size per sector independent of firm characteristics. Although unrealistic, this is a convenient simplification that facilitates considerably the discussion. The next Proposition will be extensively used when finding the empirical implications of the model.

Proposition 1 ***Firms with higher quality of their organizational capital,*** r***, have a higher span of control,*** $\frac{nk^a}{r}$***, that is, they control more resources per unit of quality***

Proof: See Appendix.

This result is already in Rosen (1982). If firm A has an organizational capital 10% better than firm B, the size of A will be more than 10% larger than the size of B. Consequently, the size distribution will be skewed to the right in comparison with the distribution of r.

Proposition 1 comes directly from the increasing returns element $g(r)$. If this was the only function of organizational capital, all resources would be controlled by the highest quality firm. The decreasing returns elements of the organizational capital avoid this extreme result, but the effect is strong enough to get the result stated in Proposition 1.

Next, I consider how the market to book and market to sales ratios change with changes in r. I abstract from dynamic considerations and I assume that the world repeats itself an infinity amount of times. Hence, Market value is equal to the infinite sum of discounted profits, let R be the rate of discount:

$$MV_t = \frac{\pi}{R} = \frac{g(r)\frac{r}{\beta}\theta(\bar{\beta}) - \frac{r}{\beta}p_k k^{1-a} - \frac{rT}{\beta k^a}F}{R} \tag{5}$$

and then the market to sales ratio is given by

$$\frac{MV}{Sales} = \frac{g(r)\frac{r}{\beta}\theta(\bar{\beta}) - \frac{r}{\beta}p_k k^{1-a} - \frac{r}{\beta k^a}F}{g(r)\frac{r}{\beta}\theta(\bar{\beta})R} = \frac{1}{R} - \frac{p_k k^{1-a} - k^{-a}F}{g(r)\theta(\bar{\beta})R} \tag{6}$$

an the market to book ratio by

$$\frac{MV}{BV} = \frac{MV}{np_k k} = \frac{g(r)\frac{r}{\beta}\theta(\bar{\beta}) - \frac{r}{\beta}p_k k^{1-a} - \frac{r}{\beta k^a}F}{\frac{r}{\beta}p_k k^{1-a}R} = \frac{g(r)\theta(\bar{\beta})}{p_k k^{1-a}R} - \frac{p_k k^{1-a} + k^{-a}F}{p_k k^{1-a}R} \tag{7}$$

The next proposition relates r with the market to book and market to sales ratios. Let σ be the elasticity of substitution of f between rt_i and q_i:

Proposition 2 ***If*** $\sigma > 1$ ***firms with higher*** r ***have lower market to book and market to sales ratios.***

Proof: See Appendix.

Proposition 2 provides the condition under which the Marginal and Average Cost curves will shift with an increase in r as they do in Figure 1. Or in other words, the condition under which more efficient firms will find profitable to compensate a lower profit per unit by a large increase in production. A graph clarifies the intuition. Figure 1 depicts the Marginal and Average cost curves of two firms, A and B. B has a better quality of organizational capital than A. I assume a world of perfect competition, therefore both A and B are price takers and produce the quantity that equalizes market price with marginal cost: Q_A and Q_B

respectively. Note that firm B has larger total profits than firm A since the shadow area is smaller than the underlined area. Note also that at the optimum, the profit per unit of output for firm B is lower that the profit per unit of output for firm A since P-AVC$_A$(Q$_A$) > P-AVC$_B$(Q$_B$) That is, we have a situation in which the more efficient firm has higher total profits but lower profits per unit of output. Of course this result totally depends on the way I have drawn the marginal and average cost curves in Figure 1. I could easily draw similar figures in which the more efficient firm has higher or equal profit per unit of output. Proposition 1 states that if $\sigma > 1$ Figure 1 will be an accurate representation of reality.

The simple model I have just described provides a rich set of propositions with important empirical implications. The first set of propositions makes reference to some empirical regularities that have been documented in the literature.

3.2 Propositions related with empirical regularities observed in the data

Proposition 3 ***If*** $\sigma > 1$***, firms with larger*** n ***have lower market to book and market to sales ratios.***

Proof: From Proposition 1 we know that the span of control, $\frac{nk^a}{r}$, increases with r, since by (4) k does not depend on r, it has to be the case that n increases with r. By proposition 2 the market to book and market to sales ratio decrease with r; the only source of variation in the model is r, therefore n and the market to book and market to sales are negatively related.

The intuition is the same we explain in proposition 2. More diversified firms have larger profits when producing more total quantity at a lower rate of profit per unit of input and per unit of output.

A number of papers have documented the robust empirical regularity that diversified firms have lower market to book and market to sales ratios. This regularity has received a name: 'Diversification discount'.

Proposition 4 ***If*** $\sigma > 1$***, the size of a segment of a diversified firm decreases with the number of segments the firm is operating,*** n.

Proof: See Appendix

Intuitively, when a firm diversifies, the span of control increases and then the effective time of the organizational capital available to each sector decreases. Less effective time of the organizational capital translates in a smaller size. Berger-Ofek (1995) gives evidence that the segments of diversified firms are smaller on average than single segment firms.

3.3 Propositions related to 'new' testable empirical implications

Proposition 5 ***If*** $\sigma > 1$***, firms with larger sales have lower market to book and market to sales ratios.***

Proof: See Appendix.

More diversified firms have less sales per sector (Proposition 4) but more total amount of sales, hence diversified firms are also larger and this gives us Proposition 6.

Proposition 6 ***Diversified firms have the same market to book and market to sales ratio as non-diversified firms of equivalent size***

Note that strictly speaking, in the model we would never have 'non-diversified firms of equivalent size'. Since the only possible way of growing is through an increase in the number of sectors, more diversified firms and larger firms are the same and Proposition 7 is trivially true. Note that we could observe 'non-diversified firms of equivalent size' assuming $F = 0$, and $a = 1$. Under this specification firms would be indifferent between growing intensively or extensively and in both cases size would determine their market to book and market to sales ratio. Apart from this particular case, the intuition I want to capture with Proposition 7 is that diversified firms and larger firms share a common characteristic: a better quality of their organizational capital. This common characteristic translates in a similar behavior of their market to book and market to sales ratios.

Proposition 7 ***If*** $\sigma > 1$***, firms with a larger*** n ***-more diversified a firms- have a lower profit to book and a lower profit to sales ratio.***

Proposition 8 ***If*** $\sigma > 1$***, larger firms have a lower profit to book and profit to sales ratio.***

Proof: Since in the model market value is equal to profits discounted, and the rate of discount is fixed, market value and profits are proportional. Therefore Proposition 3 and 6 imply respectively Proposition 7 and 8:

4 Evidence on the diversification discount

This section replicates previous literature that documents the existence of a diversification discount. The sample and variable construction are explained in detail in the Appendix. I follow very closely the procedure used by Berger-Ofek (1995) although I depart from them in some details that I will stress below.

The literature has focused on three ratios to test the existence of a diversification discount: market value to book value, market value to sales and market value to earnings ratios. The results are that these three ratios are systematically lower for diversified firms. I will discuss the behavior of market to book and market to sales ratio in my sample while I postpone the discussion about the market to earnings ratio for the appendix.

In Figure 2 we can look at the relationship between both market to book and market to sales according to different levels of diversification. There is a clear and strong negative relationship between any of the ratios and the number of segments in which the firm is operating. In both cases the most important differences take place when we go from one to two segments as Lang-Stulz (1994) already reported. The median for diversified firms differ from the medians of non-diversified around 15% in the market to sales ratio and 18% in the market to book ratio[7].

Since these ratios vary across industries, the literature has been concerned with the possibility that these differences in means and medians may arise purely due to industry effects. For example, if for whatever reasons diversified firms are concentrated in sectors with lower ratios, then differences in ratios between diversified and non-diversified may just capture industry-specific differences. To control for this type of bias, the usual procedure is to industry-adjust the market to book and market to sales ratio. The idea is to find the 'representative ratio' of each industry and adjust the ratio of each firm by the 'representative ratio' of the industry in which the firm is operating. Since by definition diversified firms operate in more than one industry, the ratio of a diversified firm is adjusted by a weighted average of the 'representative ratios' of all the industries in which the diversified firm is operating. More specifically the adjusted market to book ratio of any firm in year t, $adj(mv/bv)_t$ is:

$$adj(mv/bv)_t = (mv_t/bv_t)/\sum_{i=1}^{n} aweight_t^i * Indmtb_t^i$$

where

- n= Number of segments in which the firm is operating.

[7]Due to the skewness of the distribution of these ratios, the literature pays more attention to medians than to means.

- $Indmtb_t^i$ = Representative market to book ratio of the industry in which segment i is operating in year t.
- $aweight_t^i = bv_t^i/bv_t$, where bv is total book value and bv_t^i is total assets assigned to segment i in year t.

and

$$adj(mv/sales)_t = (mv_t/sales_t)/\sum_{i=1}^{n} sweight_t^i * Indmts_t^i$$

where

- $Indmts_t^i$ = Representative market to sales ratio of the industry in which segment i is operating
- $sweight_t^i = sales_t^i/sales_t$, where *sales* is total sales and sales_t^i is total sales assigned to segment i.

Berger-Ofek find the representative industry ratios of segment i, $Indmtb_t^i$ and $Indmts_t^i$, first by taking all the single segment firms that operate in the same four digit SIC code as segment i in year t; if there are more than five of them, then they take the median of the ratio of interest in these set of observations as the representative of the industry. If there are less than five single segment firms in the same four digit SIC code, then they consider the set of all single segment firms that operate in the same three digit SIC code. If there are more than five observations, they take the median of this last group as the representative ratio, if this group contains less than five observations then they take the median all single segment firms that operate in the same two digit SIC code.

Instead, in order to construct the representative industry ratios I follow Chevalier (1999). I take the median of all the single segment firms that belong to the same three digit SIC code.

The median of this set of observations is the representative ratio of the industry. Finally, again following Berger-Ofek, I eliminate from the sample all observations from which the sum of segment assets is less than 75% of total assets. I lose in this way 4,097 observations resulting in a sample of 60,799 observations.

In Figure 3 we can see the evolution of the median of the adjusted ratios according to different levels of diversification. Although less clear than before, there is a negative correlation between the number of segments and the adjusted ratios. Note that again the main difference is between firms that report one segment and firms that report two segments.

Finally, the literature has checked the robustness of the diversification discount by controlling for size, investment policies and measures of profitability. These authors regress the adjusted ratios on a dummy variable equal to one if a firm operates in more than one segment and zero otherwise. I report in tables 3 and 4 the results of these regressions using our sample[8]. I find a 'discount' of 0.076 for the adjusted market to book and 0.152 for the adjusted market to sales.

Up to now I have been following very closely the procedures used in Berger-Ofek about the Diversification Discount, but their regressions have a potential problem. As we have just seen, they regress the industry adjusted ratios upon non industry adjusted controls, this can potentially bias the estimation of the coefficients, not only of the controls but also of the coefficient on the dummy of interest. In the appendix I explain in detail the magnitude of the bias.

To avoid this possibility, I industry adjust the controls with the same procedure I follow to industry-adjust the market to book and market to sales ratios. First I find the representative industry value at the three digit SIC code for each control: the representative industry value

[8] In results not reported we run the same regressions using the adjusted ratios calculated exactly following exactly the same methodology as in Berger-Ofek. I find a discount of 0.143 using market to book and 0.225 using market to sales. In their study Berger-Ofek find a discount of 0.127 and 0.144 respectively.

of ‘x’ is calculated as the median of the value of ‘x’ in the set of non diversified firms that operate in the same three digit SIC code industry. Having done that, for each firm I construct a weighted average of the representative values of all the industries in which the firm is operating. The weights I use are segment sales over total sales for both earnings to sales and capital expenditures to sales ratios, and segment asset over total assets for the book value.

In Tables 3 and 4, I replicate the results of the literature but now also industry-adjusting the right hand side of the equation. Industry adjusting both sides of the equation has an important effect on the estimation of the diversification discount. Both for market to book and for market to sales the coefficient on the dummy of diversification sky rockets. For the market to book regression, the coefficient increases from -0.07 to -0.26 while in the regression of market to sales it changes from -0.15 to -0.56. Berger-Ofek may have underestimated the true magnitude of the diversification discount!

5 The size discount

5.1 Relationship between size and market to book and market to sales ratios

This section gives evidence of the existence of a ‘size discount’, meaning with this expression that large firms have lower market to book and market to sales ratios than small ones. This fact gives empirical content to Proposition 6.

First, I divide the sample by years. For each year I find the distribution of total sales and then I group the observations according to which decile of the distribution they belong to. The descriptive statistics of market to book and market to sales for each decile are plot in Figure 4 while Tables 5 and 6 show the descriptive statistics. Both ratios are negatively correlated with the size of the firm. Note that the difference in means between the smallest firms, the ones that belong to the first and second deciles, and the largest ones, the ones that belong to the ninth and tenth deciles are around 12% for the market to book and 9% of the market to sales, very similar numbers to the ones I find when I compare diversified and non-diversified firms.

One may suspect that these results might be driven by the presence of diversified firms since they are larger than the non-diversified, so I repeat the above explained procedure taking out from the sample all observations that belong to diversified firms. The results, not reported, are qualitatively the same as the ones shown in tables 5 and 6.

Mimicking the procedure described in the previous section, I industry adjust the ratios. The descriptive statistics of the adjusted ratios are displayed in tables 7 and 8. Although the median of the adjusted ratios does not show evidence of the existence of a 'size discount', the mean of both ratios decrease substantially with size and the regressions displayed in tables 9 and 10 capture this effect. These regressions are the same Berger-Ofek used to determine the diversification discount but instead of estimating the coefficient of a dummy equal to one if the firm is diversified I estimate the coefficient of a dummy, BIG, equal to one if the firm has a level of sales larger than two thirds of the number of firms in the sample. I find evidence of the existence of a size discount in the same way the literature has found evidence of a diversification discount. In the next subsection I explore the relationship between the size discount and the diversification discount.

5.2 The size discount and the diversification discount

Proposition 7 states that size and diversification have the same effect upon market to book and market to sales ratios. In this section we test this hypothesis comparing ratios of diversified firms with non-diversified firms of size in the same range.

In Table 11 we can see some descriptive statistics of three different measures of size for diversified and non-diversified firms. As expected, diversified firms are larger than non-diversified. To match diversified firms with non-diversified of similar size I proceed as follows, for each segment of a diversified firm.

First I get all non-diversified firms that operate in the same three digit SIC code as the segment, I order these observations according to their level of sales and I classify them in five subgroups. The first subgroup has the twenty per cent of observations of smaller size-size measured by sales-, while the last subgroup has the 20 per cent of observations with the largest size. I take the median of the ratio of interest in each subgroup. I look at which subgroup the segment would belong to according to the *total level of sales of the firm*. I use the representative ratio of this subgroup as the ratio the segment would have if standing alone. Once I have done this for all segments, I construct the weighted average of the value of the ratios their segments would have if standing alone. As in Section 4, the weights are segment sales over total sales for the market to sales ratio and segment assets over total assets for the market to book ratio. Finally I divide each ratio by their corresponding weighted average to find the industry-adjusted ratio.

Table 12 shows the statistics of the allocation of segments among subgroups: Not surprisingly, the majority of the segments are in groups corresponding to the largest 20% non-diversified firms, consistent with the fact that diversified firms are larger than single segment firms.

Tables 13 and 14 repeat the regressions of Section 4 but with the new industry-adjusted ratios. The coefficient on the dummy for diversification is either non significative or positive. When I compare diversified firms with non-diversified firms of similar level of size I do not find evidence of a diversification discount.

6 Profitability in more diversified and larger firms

Propositions 8 and 9 have implications regarding the profit per unit of input and per unit of output of more diversified/larger firms. Profits per unit are lower because average costs are higher. In Figure 7 we can see the relationship between the ratio Cost of Goods Sold, from now on COGS, to sales both with the degree of diversification and size where size is measured by sales. COGS includes all accounted costs of the firms directly related to production. Therefore, COGS is a measure of the total variable costs of the firm. As predicted by the theory, the relationship is negative. More diversified/larger firms have higher average costs.

Surprisingly, ratios of profitability that include any measures of earnings behave in opposite way to the predictions of Propositions 8 and 9. For example in Figure 5 we observe how the operating income to sales increase both with the level of diversification and with size, and Campa-Kedia (1999) report that diversified firms have a larger Earnings Before Interest and Taxes to sales ratio than single segment firms.

To reconcile the two apparently contradictory pieces of evidence note the relationship between earnings and COGS:

$$Earnings \backsimeq sales - COGS - SGA \tag{8}$$

where SGA is the abbreviation of 'Selling, General and Administrative expenses', that includes firm's costs that are independent of the levels of production, the fixed costs. Therefore, Figures 5 and 7 jointly with (8) imply that the ratio $SGA/sales$ decreases both with size and the degree of diversification. Figure 6 shows that this is indeed the case.

Does the empirical evidence support Propositions 8 and 9? I will defend that it does based on two arguments:

First, Propositions 8 and 9 have predictions regarding economic earnings, not accounting earnings. Economic earnings differ from accounting earnings because accountants compute as current expenses what in fact are investments. This is very clear in the case of advertisement and R&D. I do not have a way to redo the accounting figures to differentiate investments from costs [9] but I do know where all these misestimated expenses will be captured. Since all of them will be independent of the amount of production they will be accounted in Selling, General and administrative expenses (SGA). As we have just seen this is the component that causes the earnings to sales ratio to decrease both with size and diversification. This suggests that a measure of earnings free of accounting biases may behave according to the implications of Propositions 8 and 9. But, why is this accounting bias more important for conglomerates/larger firms? Figure 8[10] suggests a reason; more conglomerates/larger firms are older. If younger firms have a relatively larger investment, as a life cycle view of the firm would suggest, then the accounting bias would be larger for younger and therefore more diversified/larger firms.

The second argument is of empirical nature. In Tables 15 and 16, I repeat the regressions of the diversification discount but controlling for differences in profitability using as proxy the ratio COGS/sales. We can see that the 'diversification discount' disappears in three of the four regressions, it only remains significant in the regression of market to sales with all the controls, however note that in this case the reduction in the discount when I introduce the ratio COGS/sales is enormous, from -0.56 to -0.09. This is evidence that differences in profitability explain the 'diversification discount'. Note that the coefficient on COGS/sales is negative and as expected strongly significant, correlated with the adjusted market to book and market to sales ratios. On the contrary, compare with the effect of including earnings to sales in the regressions of Section 4, where earnings to sales ratio had a very small effect on the adjusted market to book and market to sales ratios. COGS/sales is a better predictor of the excess value. I interpret it as evidence that COGS/sales represent much better the economic earnings than the actual figure of accounting earnings.

[9] The number of firms that report figures for these accounts in Compustat is very small, around 10% of my sample.

[10] I thank Boyan Jovanovic for giving me access to the age data.

7 Understanding the previous findings on productivity

Schoar (2001) has interesting results regarding the productivity of diversified and non-diversified firms. In particular she finds that plants belonging to diversified firms have higher Total Factor Productivity than plants belonging to non-diversified firms and that after a diversifying move, the productivity of the incumbent plant decreases while the productivity of the new segment that enters the firm increases.

In what follows, I try to reconcile her findings with the model and evidence presented in the rest of this paper.

7.1 Diversified firms have higher TFP than non-diversified

Looking at Figure 1, this Schoar's finding seems to contradict the main argument of the paper. Diversified firms should have lower TFP since they have a lower profit per unit of input. I argue that if the model of Section 3 is a good representation of reality, the estimated TFP will be upward biased for diversified firms and downward biased for non-diversified:

Schoar (2001) assumes a Cobb-Douglas production function at the industry level, therefore she is implicitly assuming σ=1 and all profit maximizing firms should use the same proportion of inputs equal to the parameter of the production function, α. For simplicity assume just one input of production, then:

$$y = \alpha k + TFP \tag{9}$$

where both y and k are in logarithms and TFP is the 'real' total factor productivity. But if $\sigma > 1$, then the model of Section 3 tells us that α will be higher for diversified firms[11], take

[11] To see why the lower π/y implies higher α note that $\pi/y = (y - p_k k)/y = 1 - \alpha$. And α is the paramether of the production function in a Cobb-Douglas.

the simplest case and assume that:

$$\begin{aligned}\alpha_{div} &= \bar{\alpha} + u \\ \alpha_{nondiv} &= \bar{\alpha}\end{aligned}$$

where u is assumed constant. Then we can write (9) as

$$y = (\bar{\alpha} + Du)k + TFP \tag{10}$$

where $D = 1$ if the firm is diversified and zero otherwise. In this specification, I show in the Appendix that if we run an OLS regression on (10), the estimated coefficient, $\hat{\alpha}$, will be

$$\hat{\alpha}_{OLS} = \bar{\alpha} + \Pr\text{ob}(D = 1)uh(k)$$

where $h(k) = \frac{\Pr\text{ob}(D=1)u[Var[k|D=1]+[E[k|D=1]]^2 - E[k|D=1]E[k]]}{Var(k)}$

Note that under the natural assumption $E[k|D = 1] > E[k]$, that is, diversified firms use more inputs than the average, $h(k) > 0$ and then the estimated Total Factor Productivity is:

$$\begin{aligned}EstTFP &= y - \hat{\alpha}_{ols}k = y - (\bar{\alpha} + \Pr\text{ob}(D = 1)uh(k))k \\ &= (D - h(k)\Pr\text{ob}(D = 1)]uk + TFP\end{aligned}$$

Note that if all firms have the same real TFP, the estimated TFP would be higher for diversified firms. This way of proceeding bias upwards the estimation of TFP of diversified firms.

7.2 Dynamics on TFP after acquisition

Schoar (2001) reports that after a non-diversified firm is acquired by another one that operates in a different sector: Plants of the acquirer decrease their TFP while plants of the target increase their TFP:

From the model of Section 3, acquirer's TFP plant before the acquisition is:

$$TFP_{ac}^{before} = \log[g(r_{ac})f(r_{ac}/n_{ac}, k^a)] - \alpha \log k$$

and after the acquisition is

$$TFP_{ac}^{after} = \log[g(r_{ac})f(r_{ac}/n_f, k^a)] - \alpha \log k$$

where $n_f = n_{ac} + n_t$, $n_{ac} =$ Acquirer's n, $n_t =$ Target's n

since $n_f > n_{ac}$, the effective time of the organizational capital has decreased and this decreases the TFP of the target.

The increase in the TFP of the target is not so obvious, for the target is true that the quality of the organizational capital has increased but it is also true that the effective time of the organizational capital has decreased, the total effect on productivity is unclear. Nevertheless if mergers do not destroy value[12], that is, if the combined profits increase, then it has to be the case that the TFP of the target has to increase.

To see this, call π_{ac} and π_t respectively the profits of the acquirer and the target before the acquisition, and let π_f be the profits of the combined firm.

$$\pi_{ac} = n_{ac}g(r_{ac})f(r_{ac}/n_{ac}, k^a) - n_{ac}kp_k$$

$$\pi_t = n_t g(r_t)f(r_t/n_t, k^a) - n_t kp_k$$

$$\begin{aligned}\pi_f &= n_f g(r_{ac})f(r_{ac}/n_f, k^a) - n_f kp_k = n_{ac}g(r_{ac})f(r_{ac}/n_f, k^a) - n_{ac}kp_k + \\ &\quad + n_t g(r_{ac})f(r_{ac}/n_f, k^a) - n_t kp_k\end{aligned}$$

[12] The evidence in the literature favors this hypothesis: see Maquiera-Megginson-Nail (1998) and Bradley-Desai-Kim (1988).

therefore since

$$\pi_f > \pi_{ac} + \pi_t \tag{11}$$

and the productivity of the acquirer's plant has decreased, it has to be the case that:

$$n_t g(r_{ac}) f(r_{ac}/n_f, k^a) - n_t k p_k > n_t g(r_t) f(r_t/n_t, k^a) - n_t k p_k$$

or

$$g(r_{ac}) f(r_{ac}/n_f, k^a) > g(r_t) f(r_t/n_t, k^a)$$

and hence the TFP increases for the plants of the target.

8 Interpretation of the self-selection literature

Campa-Kedia (1999) and Villalonga (2000) give evidence that selection bias causes the diversification discount. They argue that some unobservable firm characteristics are negatively correlated with the adjusted market to book and market to sales ratios. These characteristics also make the probability of being diversified larger. The endogeneity problem created by these unobserved firm characteristics biases the econometric estimation of the diversification discount. They report empirical evidence that supports this hypothesis.

The model of section 3 identifies organizational capital as a firm characteristic that is both negatively related with market to book and market to sales ratios, and positively related with the probability of being diversified. In this sense, the evidence presented by Campa-Kedia (1999) and Villalonga (2000) is consistent with the main argument of this paper. Moreover, the focus on organizational capital derives further implications regarding to size and profitability as I have discussed above.

Villalonga (2000) argues that in order to find the counterfactual of the value of the ratio of the stand alone segment we have to condition not only by industry but also by other characteristics. Following a methodology of Lalonde (1986) she reports that when conditioning by additional firm characteristics, size being one of them, the diversification discount either disappears or becomes a premium The results of tables 13 and 14 would suggest that just finding the matching group conditioning by size suffices to discard the negative effect of diversification on both market to book and market to sales.

Campa-Kedia (1999) present evidence that diversified firms were already trading at a discount before diversifying. The discount exists before and after diversification. They control by fixed effects at the firm level and the discount disappears, this shows that the discount is a result of unobserved endogenous firm characteristics. My interpretation of this evidence is that the quality of the organizational capital of the firm is the same before and after the diversification move.

Campa-Kedia (1999) also report regressions estimating the diversification discount with instruments that should affect the probability of being diversified but not the industry-adjusted ratios. Since the industry-adjusted ratios are by construction independent of industry characteristics, they use as instruments industry characteristics presumably correlated with the probability of being diversified, like fraction of firms in the industry that are conglomerates, fraction of sales accounted by diversified firms, number of mergers and acquisitions announcements and announced value of mergers and acquisitions. With these instruments the estimation of the diversification discount drops to almost zero or becomes a premium.

In this identification strategy, the key assumption is the existence of a correlation between industry characteristics and the individual firm probability of diversifying. In the same way, the assumption that makes consistent the model of Section 3 with these results is the existence of a correlation between industry characteristics and firm's quality of organizational capital.

Summarizing the discussion of the above paragraphs, the evidence presented in the self-selection literature and the view presented in this paper are mutually reinforcing. Unobserved heterogeneity in the quality of the organizational capital can 'create' the diversification discount and once we control for it the evidence of a discount vanishes.

9 Conclusion

This papers explores the implications of a very simple specification that captures the effect of heterogeneous quality of organizational capital among firms. These implications are consistent with a set of empirical findings related to diversified firms: namely that diversified firms have lower market to book and market to sales ratio, that self-selection causes the diversification discount, that plants of diversified firms have higher productivity than plants of non-diversified, and that after a diversifying move the productivity of the incumbent plants increases while the productivity of the segment that enters the firm decreases.

I have given evidence of the existence of a size discount, and that the diversification discount disappears when we control properly by size. Additionally, I show that differences in profitability -free from accounting bias- explain the discount. Differences in profitability are not originated by inefficiencies of any kind in the governance of the firm. On the contrary, firms with better organizational capital, optimally choose a level of production with a lower profit per unit of output and per unit of input.

References

Bradley, M., Desai, A., and Kim, E.H., 1988, *Synergistic Gains from Corporate Acquisitions and Their Division between the Stockholders of Target and Acquiring Firms,* Journal of Financial Economics 21, 3-40.

Berger, P.G. and E. Ofek, 1995, *Diversification's effect on firm value*, Journal of Financial Economics 37, 39-65.

Campa J.M. and S. Kedia, 1999, *Explaining the diversification discount*, Working Paper, NYU.

Chevalier, C. A.,1999, *Why do firms undertake diversifying mergers? An examination of the investment policies of merging firms*, Working paper, University of Chicago.

Cremer J. 1993, *Corporate Culture and Shared Knowledge,* Industrial and Corporate Change 2, 351-386.

Cremer, J., Garicano L. and Pratt A., *Codes and Organization, 2001,* Mimeo, University of Chicago.

Gort M. (1962), *Diversification and integration in American industry,* Princeton University Press.

Jovanovic B. 1993, *The diversification of production*, Brooking Papers: Microeconomics, 197-235.

Kaplan, S. and Weisbach, M., 1992, *The success of Acquisitions: Evidence from Divestitures,* Journal of Finance 47, 107-138.

Lang, L.H.P. and R.E. Stulz, 1994, *Tobin's q, corporate diversification and firm performance*, Journal of Political Economy, Vol 102, No. 6 1248-1280.

Lalonde, Robert J., 1986, *Evaluating the econometric evaluation of training programs with experimental data*, American Economic Review 76, 604-620.

Maksimovic, V., and Phillips G. 2002, Do Conglomerate Firms Allocate Resources Inefficiently across Industries?, Journal of Finance, forthcoming.

Maquiera, C. Megginson, W., and Nail, L., 1998, *Wealth Creation versus Wealth Redistributions in Pure Stock for Stock Mergers,* Journal of Financial Economics 48, 3-33.

Matsusaka J.G. 2001, *Corporate diversification, value maximization, and organizational capabilities*, The Journal of Business, Vol 74, No 3 409-431.

Nelson R. and Winter S., 1982, *An Evolutionary Theory of Economic Change,* Cambridge, The Belknap Press of Harvard University Press.

Prescott, E. and Visscher Michael, 1980, *Organization Capital*, Journal of Political Economy, vol 88, No 3, 446-461.

Rajan, R., H. Servaes, and L. Zingales, 2000, *The cost of diversity: The diversification discount and inefficient investment*, Journal of Finance 60, 35-80.

Rajan, R. and L. Zingales, 1998, *Power in a Theory of the Firm,* Quarterly Journal of Economics CXIII: 387-432.

Rosen S., 1982, *Authority, control and the distribution of earnings*, The Bell Journal of Economics, Vol 13, Issue 2, 311-323.

Servaes H., 1996, *The value of diversification during the conglomerate merger wave*, Journal of Finance 51, 4,1201-1225.

Schoar A., 2001, *Effects of Corporate Diversification on Productivity,* Journal of Finance, forthcoming.

Telser L., 1987, *A Theory of efficient cooperation and competition,* Cambridge University Press

Villalonga, B. 2000, *Does diversification cause the diversification discount?*, Working Paper, Harvard University.

A Appendix

A.1 Market to earnings ratio

Berger-Ofek (1995) provided evidence of the existence of the diversification discount using a third ratio different from the two ones discussed in the paper: market value to earnings. Up to my knowledge these authors are the only ones in doing so. The problem with this measure is how to proceed when earnings are negative. First the weights used when constructing the industry-adjusted ratio will be negative, and second the regressions are usually taken in logs. Berger-Ofek, deal with this problem as follows: when earnings are negative they add to earnings the accounting value for depreciation of the assets, if the sum is still negative, then they use segment sales as multiplier. This procedure can bias their results since they are systematically treating differently firms that perform badly.

In my sample, I look at the behavior of the *unadjusted* market to earnings ratio without taking logs. In this way observations with negative earnings are not a problem. Table A1 shows that there is not evidence of a discount using the market to earnings ratio. Table A2 replicates for the unadjusted market to earnings ratio the 'usual' regression that documents the diversification discount. The coefficient on the dummy for diversification is not statistically significative.

The evidence provided both in tables A1 and A2 shows that there are reasons to believe that the results of Berger-Ofek might be driven by the way they adjust the observations with negative value for earnings.

A.2 Example with decreasing returns to organizational capital when increasing the number of sectors

In this appendix I illustrate with a simple example from Cremer-Garicano-Pratt (2001)how there can be decreasing returns to the organizational capital when the number of sectors in which the firm operates increases.

Sector j has two tasks, A and B. A needs to be mentioned α_A^j times per period and B α_B^j times. The firm can choose between creating a word for every task or just using plain English. Creating a word has the advantage of increasing the speed of communication, since the time you need to communicate the new word is $l, l < S$, where S is the time needed to communicate in plain English, but it requires a fixed time investment: Γ. The firm wants to maximize the speed of communication, consequently, a specialized word will only be created if

$$\alpha_i^j l + \Gamma < \alpha_i^j S$$

for $i = A, B$ or

$$\alpha_i^j > \frac{\Gamma}{S - l}$$

that illustrates the intuition that words will be created only for those tasks that are used enough number of times

Assume two sectors, 1 and 2 such that:

$$\alpha_A^1 > \frac{\Gamma}{S - l}$$

$$\alpha_B^1 < \frac{\Gamma}{S - l}$$

$$\alpha_A^2 < \frac{\Gamma}{S - l}$$

$$\alpha_B^2 > \frac{\Gamma}{S - l}$$

If a firm operates in sector 1, the total communication time will be

$$\alpha_A^1 l + \Gamma + \alpha_B^1 S$$

and if a firm operates in sector 2 the total communication time would be

$$\alpha_B^2 l + \Gamma + \alpha_A^2 S$$

What happens if the firm operates in both sectors? I assume that the organization has to share the same internal code. Two words cannot coexist for the same task inside the organization. This means that whenever a firm operates in both sectors, the total communication time in any sector will be higher or equal than the communication time if the firm is specialized[13]

A.3 Proof of propositions

Proof of proposition 1: Reduce (3) to get:

$$g(r)\theta(\bar{\beta}) + C = \bar{\beta} g(r)\theta'(\bar{\beta}) \tag{12}$$

where $C = -(p_k k^{1-a} + k^{-a} F)$

Differentiate (12) to get[14]:

$$\frac{d\bar{\beta}}{dr} = \frac{g'(r)}{g(r)} \frac{\bar{\beta}\theta'(\bar{\beta}) - \theta(\bar{\beta})}{-\theta''(\bar{\beta})\bar{\beta}} < 0 \tag{13}$$

since we have assumed f concave $\theta'' < 0$ $\bar{\beta}\theta'(\bar{\beta}) - \theta(\bar{\beta}) < 0$ since $\theta(\bar{\beta}) - \bar{\beta}\theta'(\bar{\beta})$ is the marginal productivity of q_i, assumed positive.

Since $\bar{\beta}$ is the inverse of the span of control (13) proves proposition 1

Proof of Proposition 2: From (6) and (7) note that both ratios, market to book and market to sales would be constant, -meaning by constant independent of r- if not by the

[13] The decision rule of the integrated firm will be to create a word for code i if

$$(\alpha_{1i} + \alpha_{2i})l + 2\Gamma < (\alpha_{1i} + \alpha_{2i})S$$

whatever the decision of the firm is, the cost of communication will be higher in at least one sector.

[14] Note that $\theta'' < 0$ since we have assumed f concave and that $\bar{\beta}\theta'(\bar{\beta}) - \theta(\bar{\beta}) < 0$ since it is the marginal productivity of q_i

presence of the term: $g(r)\theta(\bar{\beta})$. Note also that both ratios depend positively on $g(r)\theta(\bar{\beta})$: $\frac{dMV/Sales}{dg(r)\theta(\bar{\beta})} > 0$ $\frac{dMV/BV}{dg(r)\theta(\bar{\beta})} > 0$,differentiate totally this last term:

$$d(g(r)\theta(\bar{\beta})) = g'(r)\theta(\bar{\beta})dr + g(r)\theta'(\bar{\beta})d\bar{\beta}$$

substitute (13) to get:

$$\frac{d(g(r)\theta(\bar{\beta}))}{dr} = g(r)[-\frac{-\bar{\beta}\theta''(\bar{\beta})\theta(\bar{\beta}) + \theta'(\bar{\beta})(\bar{\beta}\theta'(\bar{\beta}) - \theta(\bar{\beta}))}{-\bar{\beta}\theta'(\bar{\beta}) + \theta(\bar{\beta})}]d\bar{\beta}/dr \tag{14}$$

since

$$\sigma = -\frac{(-\bar{\beta}\theta'(\bar{\beta}) + \theta(\bar{\beta}))\theta'(\bar{\beta})}{\bar{\beta}\theta''(\bar{\beta})\theta(\bar{\beta})}$$

we can write (14) as:

$$\begin{aligned} \frac{d(g(r)\theta(\bar{\beta}))}{dr} &= g(r)[-\frac{-\bar{\beta}\theta''(\bar{\beta})\theta(\bar{\beta}) + \bar{\beta}\theta''(\bar{\beta})\theta(\bar{\beta})\sigma}{-\bar{\beta}\theta'(\bar{\beta}) + \theta(\bar{\beta})}]d\bar{\beta}/dr = \\ &= g(r)[-\frac{\bar{\beta}\theta''(\bar{\beta})\theta(\bar{\beta})(\sigma - 1)}{-\bar{\beta}\theta'(\bar{\beta}) + \theta(\bar{\beta})}]d\bar{\beta}/dr \end{aligned} \tag{15}$$

hence (15) shows that if $\sigma > 1$ and from proposition 1 (since $\bar{\beta} = 1/\beta$)

$$\frac{d(g(r)\theta(\bar{\beta}))}{dr} < 0$$

Proof of Proposition 4: It is immediate from Proposition 3 and from $\frac{d(g(r)\theta(\bar{\beta}))}{dr} < 0$, since $g(r)\theta(\bar{\beta})$ is output per segment

Proof of Proposition 6: $dy/dr = d(k^a rg(r)\theta(\bar{\beta})/\bar{\beta})/dr = k^a[rg'(r) + g(r)]\theta(\bar{\beta})/\bar{\beta} + g(r)r[\theta'(\bar{\beta}) - \theta(\bar{\beta})](1/\bar{\beta}^2)\frac{d\bar{\beta}}{dr} > 0$

A.4 Sample selection and variable construction

Our sample is composed by all the firms registered in Compustat from 1980 to 1999, both active and inactive. An observation is data for a firm in a particular year. Following the methodology of Berger Ofek (1995), I exclude all observations that have a segment in the financial services industry, SIC codes 6000 to 6999, I eliminate all observations with sales less than $20 million, and all observations with the sum of segment sales different in more than one per cent of the value of total sales. This procedure gives us a sample of 64,896 observations, 43,264 belonging to non-diversified firms and 21,632 belonging to diversified firms. As is custom in the literature, I take a firm as diversified if it reports that operates in more than one segment.

The variables are constructed as follows, market value, MV is:

$$MV_t = P_t * shares_t + shorttermdebt_t + longtermdebt_t$$

where P is the price of shares at the end of the year, data23 in Compustat, shares is the number of common shares outstanding. Long and short term debt are measured according to its book value, data9 and data34 respectively. As book value of the assets I use data6 and for sales data12.

A.5 Industry-adjusting both sides of the equation

In this appendix I illustrate how industry-adjusting just the left hand side of the equation can bias the estimation of the diversification discount. To see this consider the simplest model with just one control:

$$y = \alpha_0 + \alpha_1 div + \alpha_2 x + u$$

where y is the ratio, div the dummy of diversification, x the control and u a random variable such that $E[u|x, div] = 0$

Assume that both y and x have industry variation denoted by e_{ind}:

$$\begin{aligned} y &= Y + e_{ind} \\ x &= X + e_{ind} \end{aligned}$$

with $E[e_{ind}|Y] = E[e_{ind}|X] = 0$

In the literature they industry-adjust y but they do not industry-adjust x so in fact they are running the regression of Y on div and x. Let δ_i be the OLS estimator of α_i since:

$$Y = \alpha_0 + \alpha_1 div + \alpha_2 x + u - e_{ind}$$

then

$$p\lim \begin{pmatrix} \delta_2 \\ \delta_3 \end{pmatrix} = \begin{pmatrix} \alpha_2 \\ \alpha_3 \end{pmatrix} + \begin{pmatrix} Var(div) & cov(div,x) \\ cov(div,x) & Var(x) \end{pmatrix}^{-1} \begin{pmatrix} cov(div,-e_{ind}) \\ cov(x,-e_{ind}) \end{pmatrix}$$

assume that $cov(div,-e_{ind}) = 0$ then

$$\begin{aligned} p\lim \begin{pmatrix} \delta_2 \\ \delta_3 \end{pmatrix} &= \begin{pmatrix} \alpha_2 - \frac{cov(div,x)cov(x,-e_{ind})}{Var(div)Var(x)-cov^2(div,x)} \\ \alpha 3 + \frac{Var(div)cov(x,-e_{ind})}{Var(div)Var(x)-cov^2(div,x)} \end{pmatrix} = \\ &= \begin{pmatrix} \alpha_2 \\ \alpha_3 \end{pmatrix} + \begin{pmatrix} \frac{prob(Div=1)[E[x|div=1]-E[x]]Var(e_{ind})}{Var(div)Var(x)-cov^2(div,x)} \\ -\frac{Var(div)Var(e_{ind)})}{Var(div)Var(x)-cov^2(div,x)} \end{pmatrix} \end{aligned}$$

If $E[x/div = 1] - E[x] > 0$,as it is reasonable to assume for example with the case of book value since book value of diversified firms is larger than book value of non-diversified firms. then we are going to overestimate the coefficient on the dummy, and since in this case $\alpha_2 < 0$, we are going to underestimate the diversification discount since:

$$|\alpha_2 + \frac{prob(Div = 1)[E[x|div = 1] - E[x]]Var(e_{ind})}{Var(div)Var(x) - cov^2(div,x)}| < |\alpha_2|$$

A.6 OLS estimation of α

Schoar (2001) assumed the same coefficient α for diversified and non-diversified firms and then:

$$\begin{aligned} \hat{\alpha}_{OLS} &= cov(y,k)/Var(k) = \frac{cov((\bar{\alpha}+Du)k+TFP,k)}{Var(k)} = \\ &= \bar{\alpha} + cov(Duk,k)/Var(k) = \\ &= \bar{\alpha} + \frac{E[Duk^2]-E[Duk]E[k]}{Var(k)} \end{aligned} \tag{16}$$

If we assume that k and D are statistically independent then the (16) is:

$$\hat{\alpha}_{OLS} = \bar{\alpha} + \Pr\text{ob}(D=1)u$$

If D and k are not independent:

$$\begin{aligned} \hat{\alpha}_{OLS} &= \bar{\alpha} + \frac{\Pr\text{ob}(D=1)u[E[k^2|D=1]-E[k|D=1]E[k]]}{Var(k)} = \\ &= \bar{\alpha} + \Pr\text{ob}(D=1)uh(k) \end{aligned}$$

where $h(k) = \frac{\Pr\text{ob}(D=1)u[Var[k|D=1]+[E[k|D=1]]^2-E[k|D=1]E[k]]}{Var(k)}$

FIGURE 1

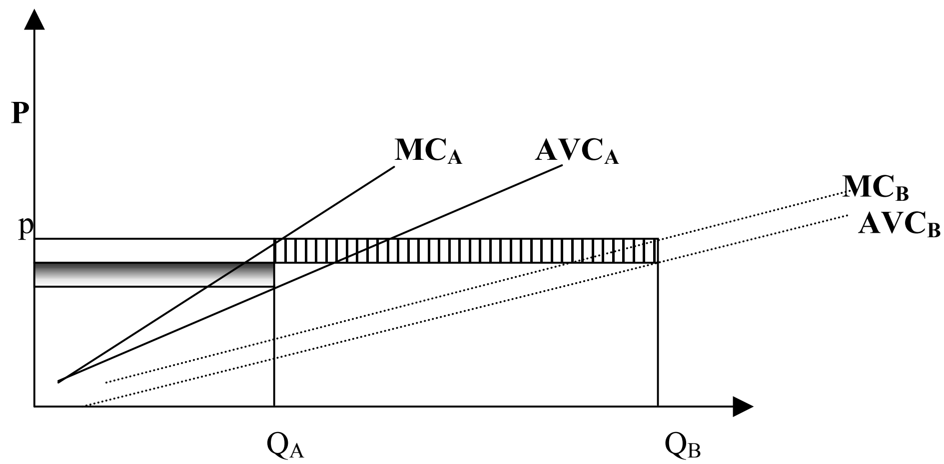
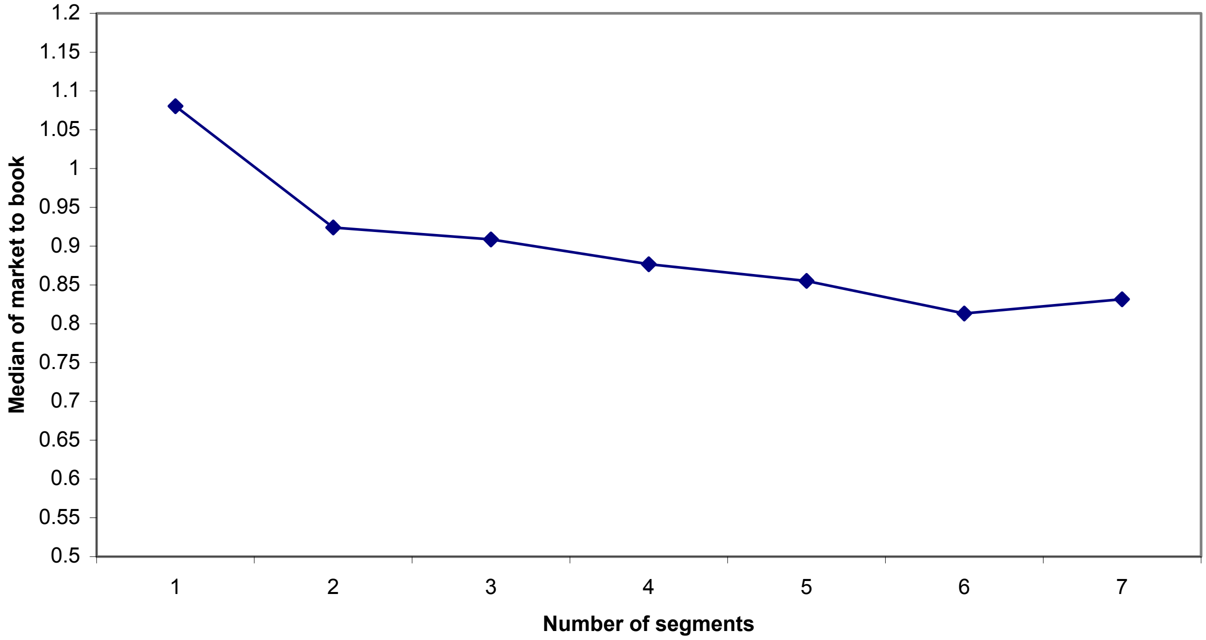


FIGURE 2


1.2
1.15
1.1
1.05
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
Median of market to book
1
2
3
4
5
6
7
Number of segments

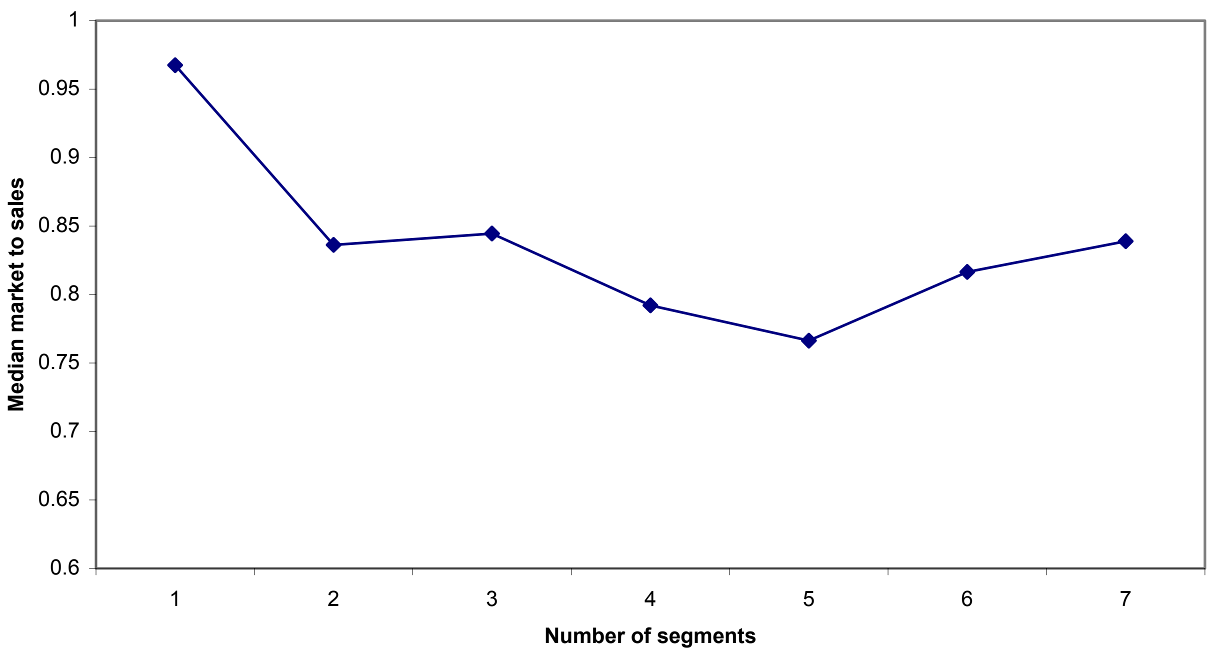
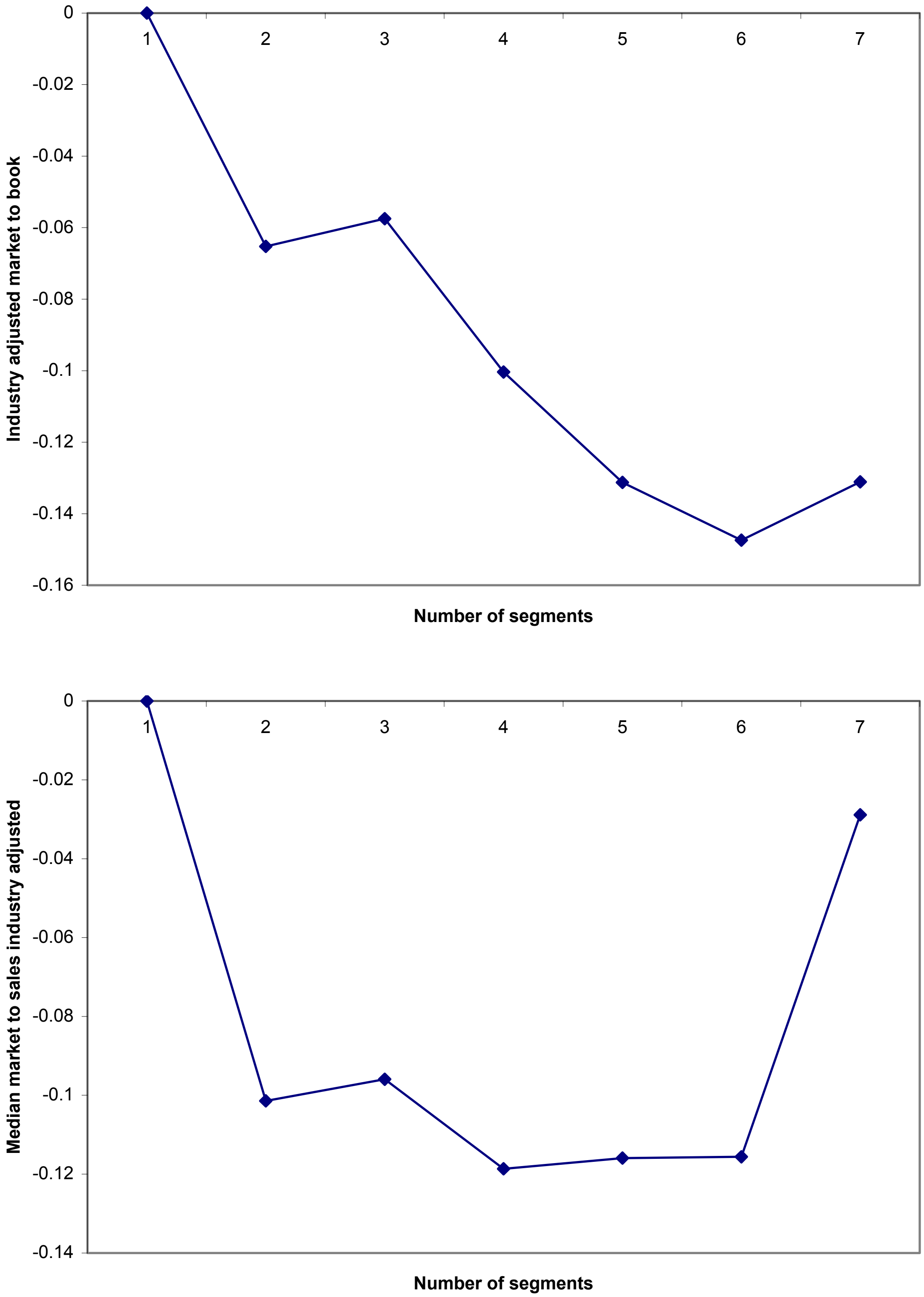

1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
Median market to sales
1
2
3
4
5
6
7
Number of segments

FIGURE 3



0
-0.02
-0.04
-0.06
-0.08
-0.1
-0.12
-0.14
-0.16
1
2
3
4
5
6
7
Industry adjusted market to book
Number of segments
0
-0.02
-0.04
-0.06
-0.08
-0.1
-0.12
-0.14
1
2
3
4
5
6
7
Median market to sales industry adjusted
Number of segments

FIGURE 4

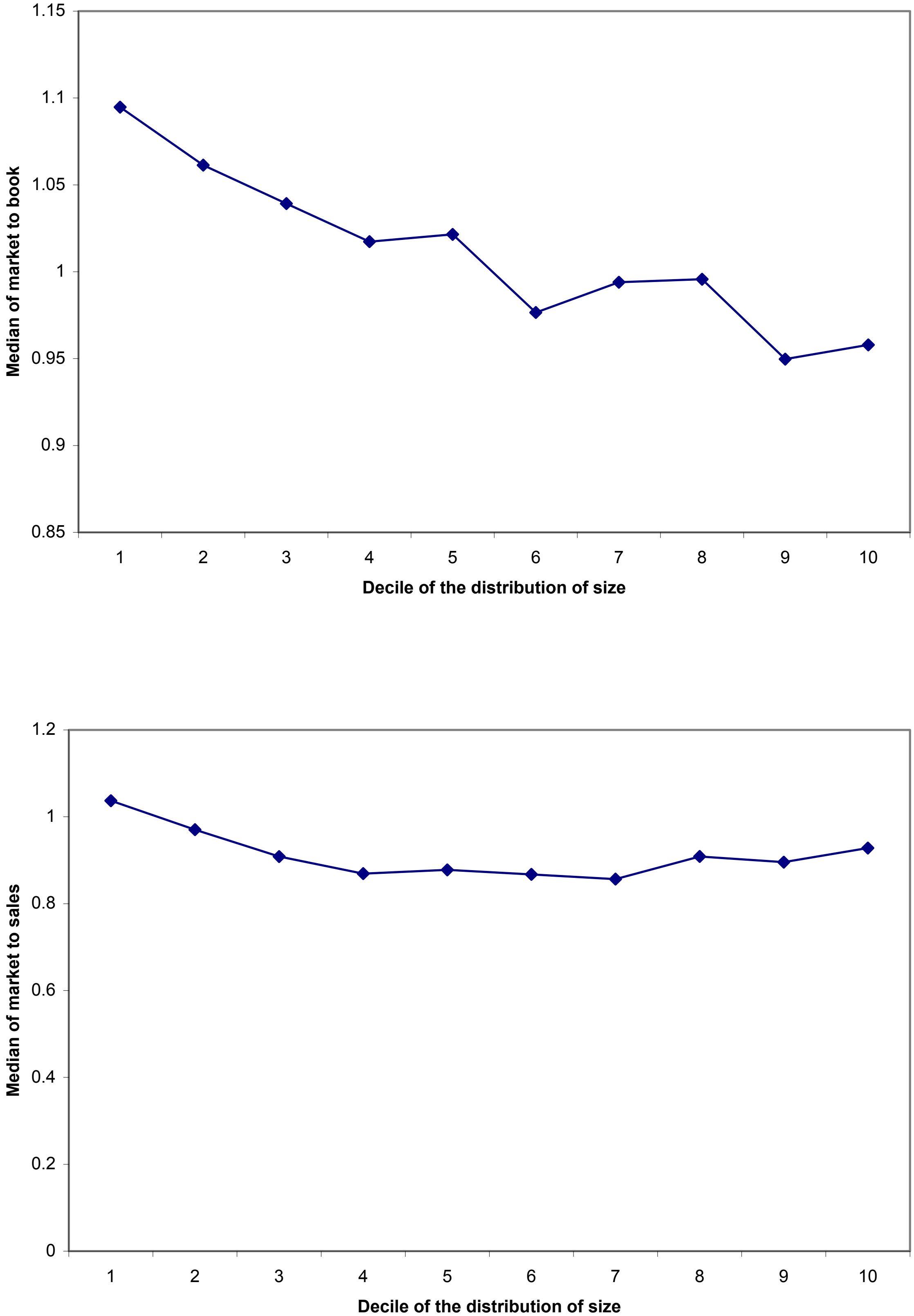
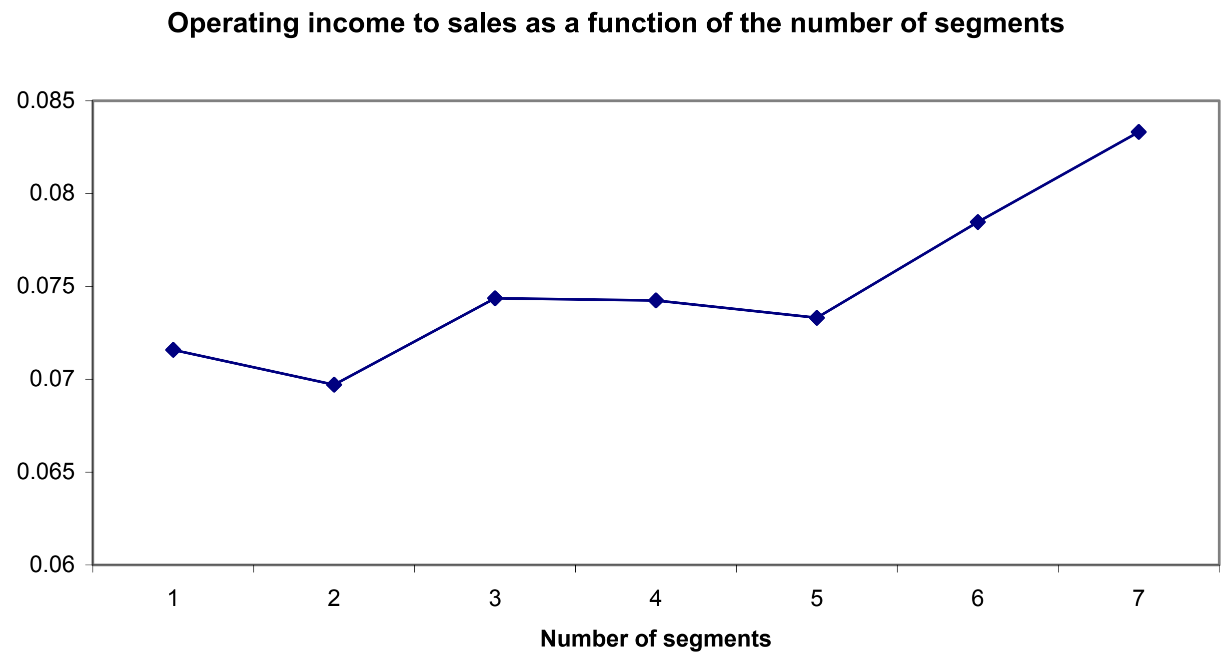
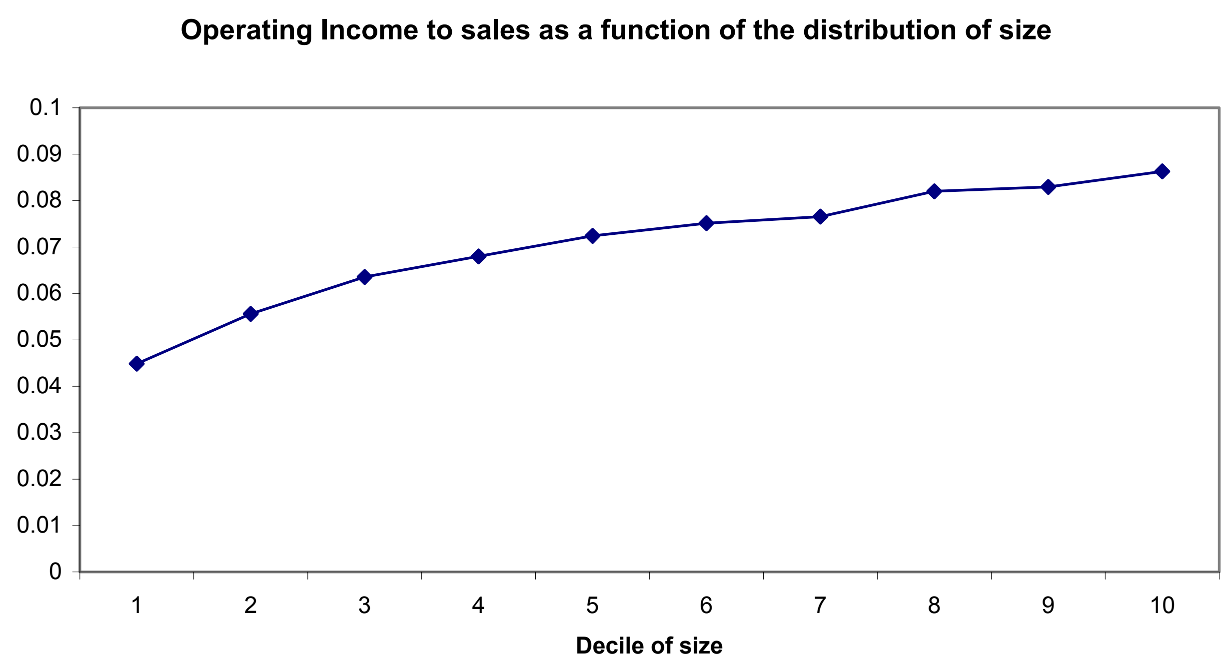

1.15
1.1
1.05
1
0.95
0.9
0.85
Median of market to book
1
2
3
4
5
6
7
8
9
10
Decile of the distribution of size
1.2
1
0.8
0.6
0.4
0.2
0
Median of market to sales
1
2
3
4
5
6
7
8
9
10
Decile of the distribution of size

FIGURE 5

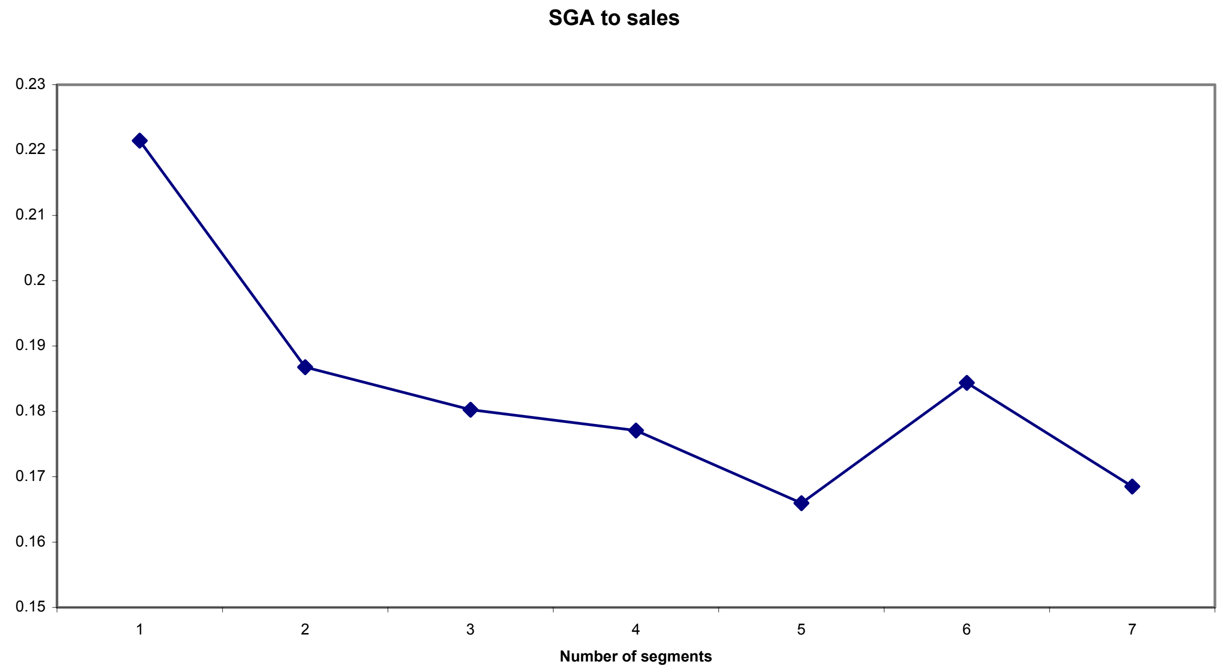


Operating income to sales as a function of the number of segments
0.085
0.08
0.075
0.07
0.065
0.06
1
2
3
4
5
6
7
Number of segments



Operating Income to sales as a function of the distribution of size
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
1
2
3
4
5
6
7
8
9
10
Decile of size

FIGURE 6


SGA to sales
0.23
0.22
0.21
0.2
0.19
0.18
0.17
0.16
0.15
1
2
3
4
5
6
7
Number of segments

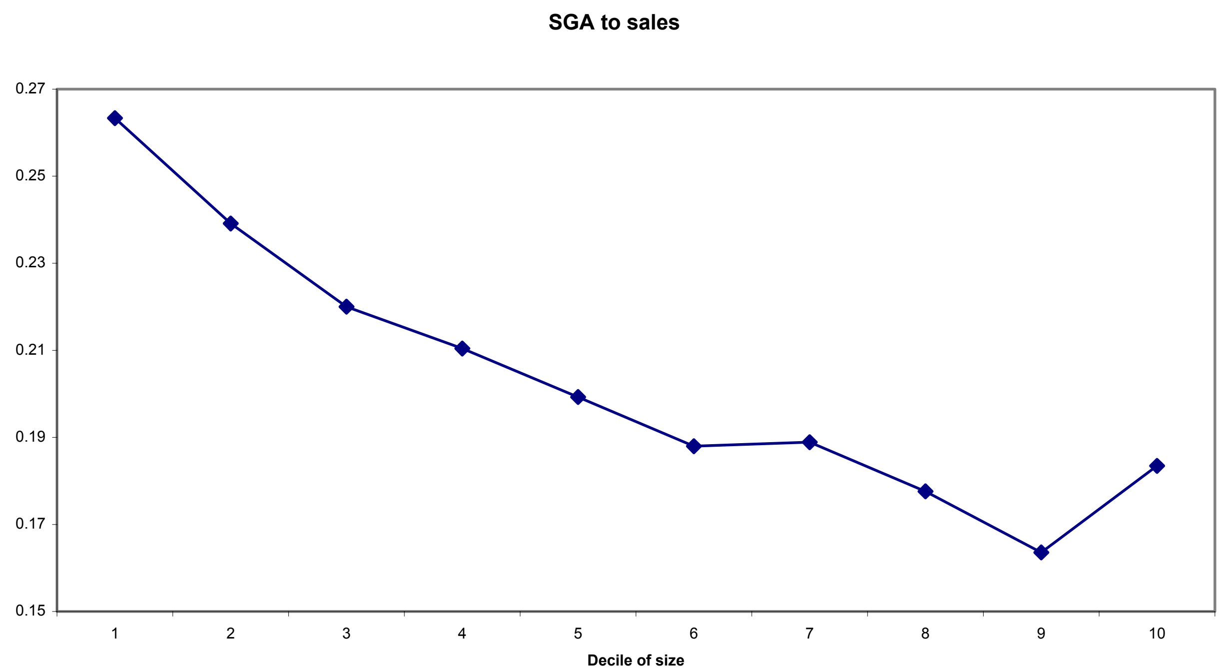
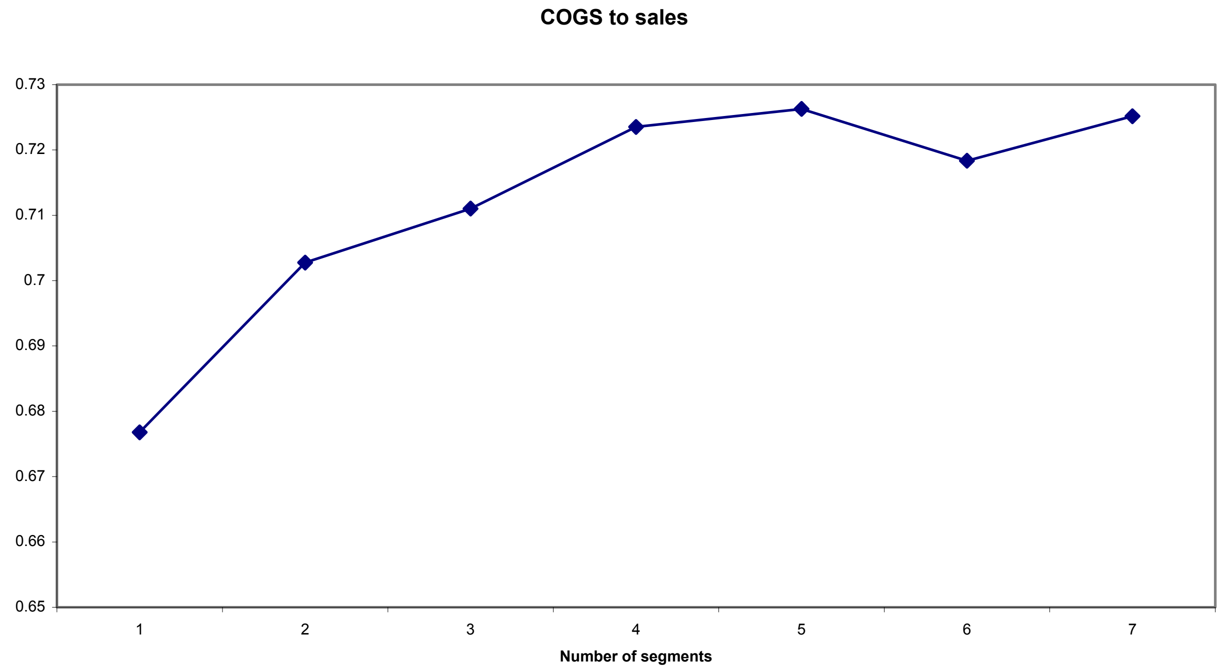
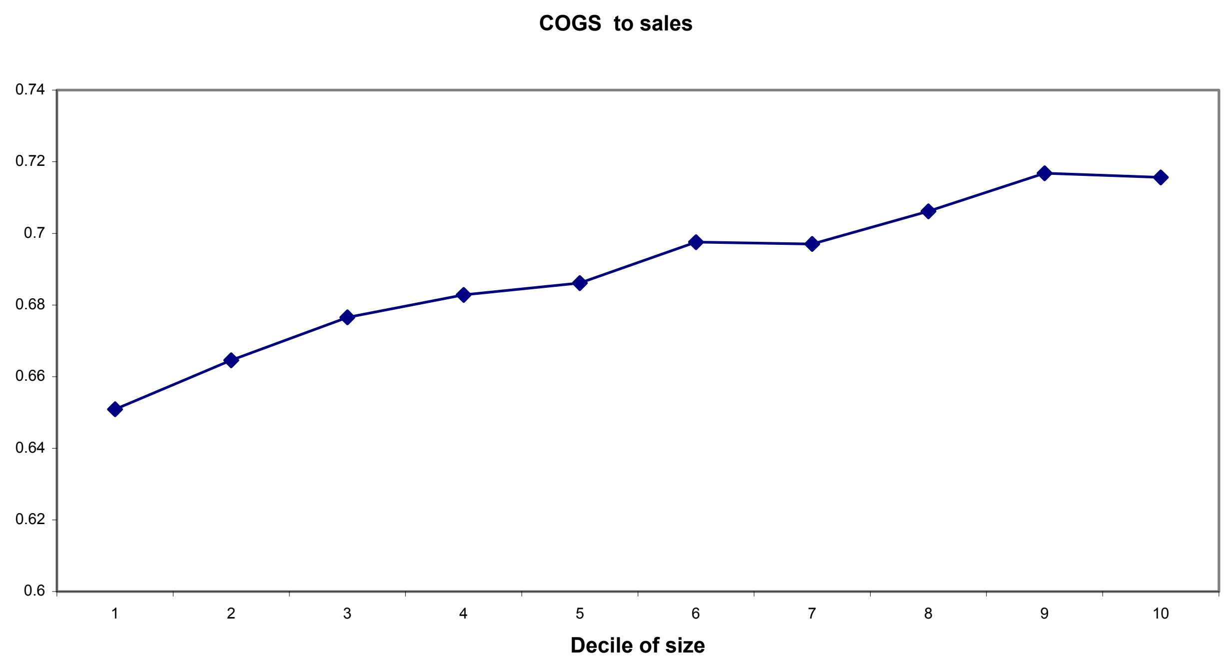

SGA to sales
0.27
0.25
0.23
0.21
0.19
0.17
0.15
1
2
3
4
5
6
7
8
9
10
Decile of size

FIGURE 7

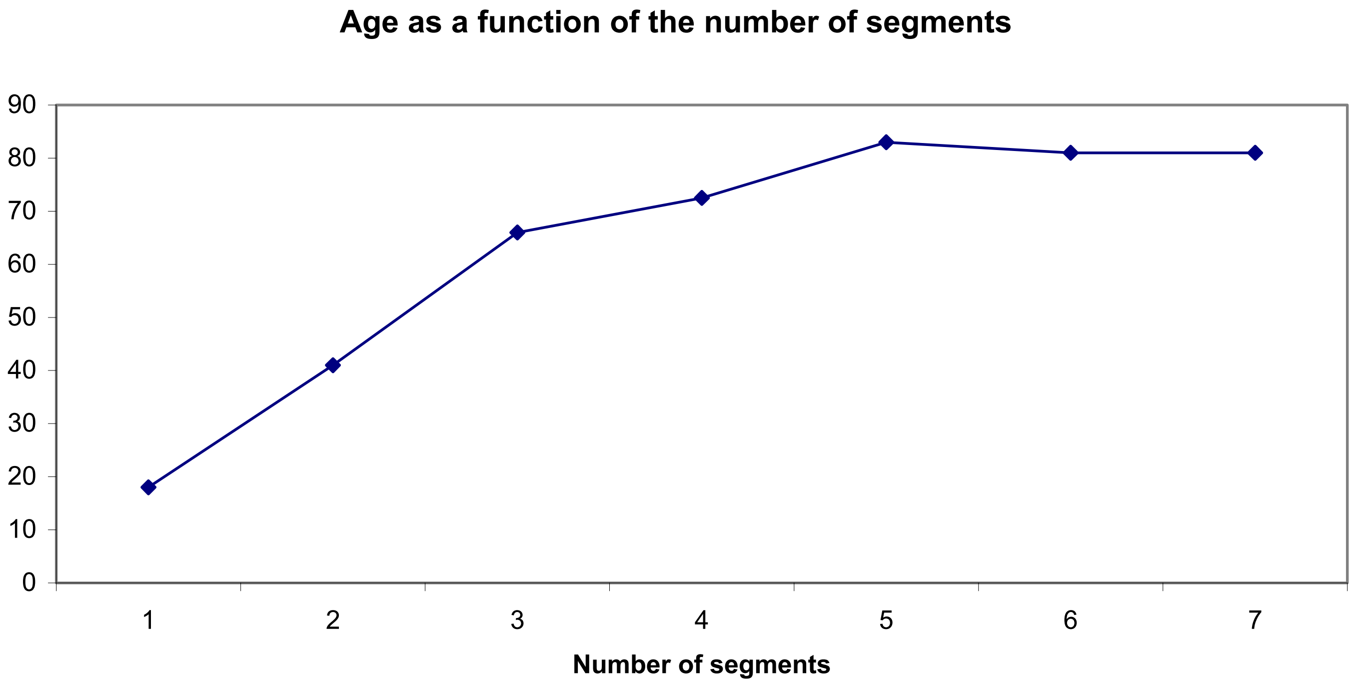


COGS to sales
0.73
0.72
0.71
0.7
0.69
0.68
0.67
0.66
0.65
1
2
3
4
5
6
7
Number of segments



COGS to sales
0.74
0.72
0.7
0.68
0.66
0.64
0.62
0.6
1
2
3
4
5
6
7
8
9
10
Decile of size

Figure 8


Age as a function of the number of segments
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
Number of segments

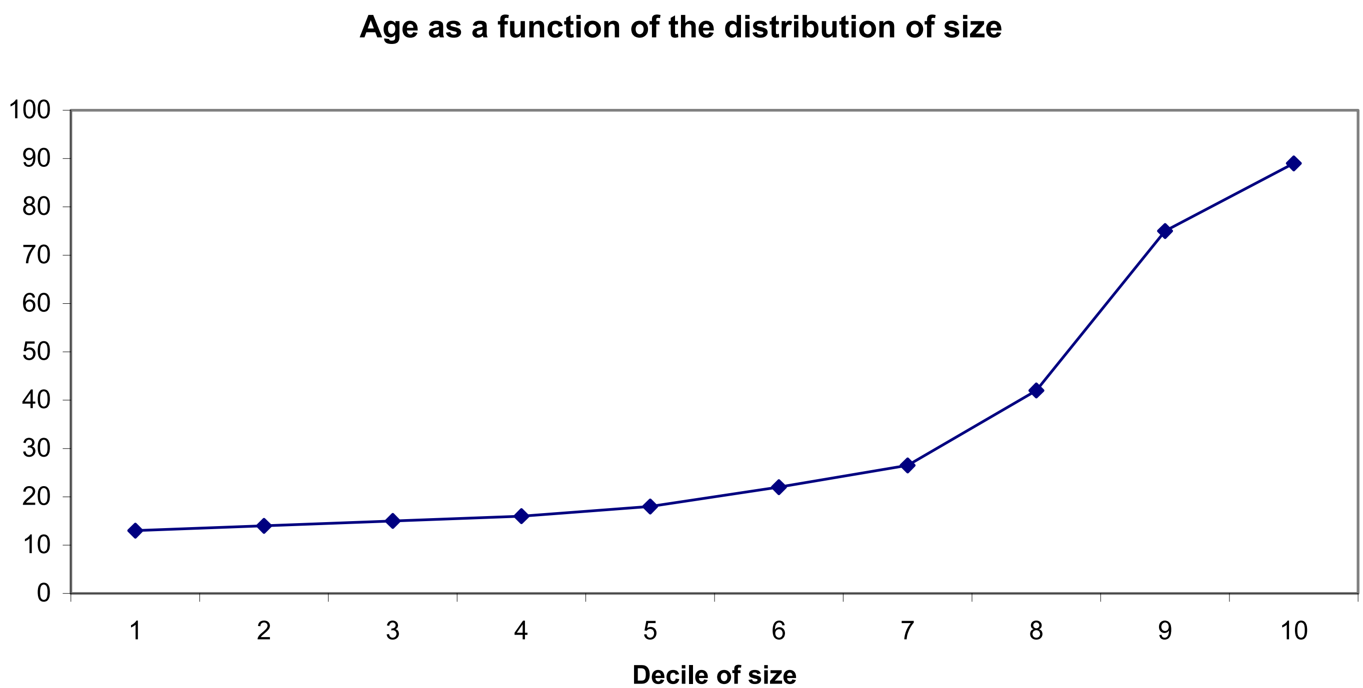

Age as a function of the distribution of size
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
Decile of size

TABLE 1: REPLICATING PRIOR STUDIES: ESTIMATION OF THE DIVERSIFICATION DISCOUNT USING ADJUSTED MARKET TO BOOK

Regression of adjusted market to book upon a variable, div, that is equal to 1 if the firm is diversified and zero otherwise. The results are found using exactly the same procedure as in Berger-Ofek (1995) with the exception that for finding the representative ratio of the industry I use all the firms that belong to the same three digit SIC code. All variables are in logs except Earnings to sales. All the P-values are found correcting by heteroskedasticity

	Coefficient	P-Value	Coefficient	P – Value
Intercept	0.027	0.000	0.229	0.000
Div	-0.067	0.000	-0.076	0.000
Book value			0.001	0.271
Cap. Exp./sales			0.065	0.000
Earnings/sales			0.012	0.650
R^2	0.002		0.022	
Observations	60,798		58,415	

TABLE 2: REPLICATING PRIOR STUDIES: ESTIMATION OF THE DIVERSIFICATION DISCOUNT USING ADJUSTED MARKET TO SALES

Regression of adjusted market to sales upon a variable, div, that is equal to 1 if the firm is diversified and zero otherwise. The results are found using exactly the same procedure as in Berger-Ofek(1995) with the exception that for finding the representative ratio of the industry I use all the firms that belong to the same three digit SIC code. All variables are in logs except Earnings to sales. All the P-values are found correcting by heteroskedasticity

	Coefficient	P-Value	Coefficient	P-Value
Intercept	0.011	0.001	0.291	0.000
Div	-0.101	0.000	-0.152	0.000
Book value			0.043	0.000
Capexp/sales			0.161	0.000
Earnings/sales			0.018	0.333
R^2	0.003		0.099	
Observations	64,798		58,413	

TABLE 3: DIVERSIFICATION DISCOUNT ESTIMATION INDUSTRY-ADJUSTING BOTH SIDES OF THE EQUATION. MARKET TO BOOK CASE

Regression of adjusted market to book upon a variable, div, that is equal to 1 if the firm is diversified and zero otherwise. The results are found using exactly the same procedure as in Berger-Ofek(1995)) with the exception that for finding the representative ratio of the industry I use all the firms that belong to the same three digit SIC code. The explanatory variables are also industry adjusted. All the P-values are found correcting by heteroskedasticity

	Coefficient	P-Value
Intercept	0.091	0.0000
Div	-0.269	0.0000
Capexp/sales	0.125	0.0000
Earnings/sales	-0.000	0.7997
Book value	0.010	0.0001
R^2	0.046	
Observations	58,235	

TABLE 4: DIVERSIFICATION DISCOUNT ESTIMATION INDUSTRY-ADJUSTING BOTH SIDES OF THE EQUATION. MARKET TO SALES CASE

Regression of adjusted market to sales upon a variable, div, that is equal to 1 if the firm is diversified and zero otherwise. The results are found using exactly the same procedure as in Berger-Ofek(1995)) with the exception that for finding the representative ratio of the industry I use all the firms that belong to the same three digit SIC code. The explanatory variables are also industry adjusted. All the P-values are corrected by heteroskedasticity

	Coefficient	P value
Intercept	0.024	0.081
Div	-0.560	0.000
Book value	-0.005	0.854
Capexp/sales	0.314	0.000
Earnings/sales	-0.000	0.550
Cogs/sales		
R^2	0.169	
Observations	58,231	

TABLE 5: DESCRIPTIVE STATISTICS OF MARKET TO BOOK AS A FUNCTION OF SIZE.

The total sample is divided in 10 groups according to which decile of the distribution of sales in a particular year the observation belongs to. The First decile is the one with lowest level of sales.

	MEAN	MEDIAN	STD	OB
First decile	1.732	1.091	2.738	6351
Second decile	1.658	1.061	2.568	6469
Third decile	1.572	1.039	2.549	6657
Fourth decile	1.525	1.017	2.334	6695
Fifth decile	1.455	1.021	1.533	6696
Sixth Decile	1.387	0.976	1.461	6555
Seventh Decile	1.368	0.993	1.677	6380
Eigth Decile	1.303	0.995	1.113	6411
Nineth Decile	1.247	0.949	1.107	6242
Tenth decile	1.266	0.957	1.181	6308

TABLE 6: DESCRIPTIVE STATISTICS OF MARKET TO SALES AS A FUNCTION OF SIZE.

The procedure I follow is the same as in table 5

	MEAN	MEDIAN	STD	OB
First decile	2.700	1.037	11.429	6351
Second decile	2.354	0.970	12.931	6469
Third decile	1.995	0.908	6.435	6657
Fourth decile	1.819	0.869	5.735	6695
Fifth decile	1.613	0.877	3.516	6696
Sixth Decile	1.536	0.867	2.816	6555
Seventh Decile	1.471	0.856	3.393	6380
Eigth Decile	1.388	0.908	1.658	6411
Ninth Decile	1.315	0.895	1.494	6242
Tenth decile	1.283	0.928	1.380	6308

TABLE 7: DESCRIPTIVE STATISTICS OF INDUSTRY-ADJUSTED MARKET TO BOOK AS A FUNCTION OF SIZE

The procedure used is the same as in table 5.

	MEAN	MEDIAN	STD	OB
First decile	0.367	0	2.494	4250
Second decile	0.329	0	2.193	4313
Third decile	0.341	0	2.574	4424
Fourth decile	0.316	0	2.848	4447
Fifth decile	0.304	0	2.250	4500
Sixth Decile	0.257	0	1.608	4414
Seventh Decile	0.238	0.000	1.382	4256
Eigth Decile	0.236	0.033	1.909	4192
Nineth Decile	0.208	0.039	1.375	4069
Tenth decile	0.257	0.032	1.488	3798

TABLE 8: DESCRIPTIVE STATISTICS OF INDUSTRY-ADJUSTED MARKET TO SALES AS A FUNCTION OF SIZE

The procedure used is the same as in Table 5.

	MEAN	MEDIAN	STD	OB
First decile	1.147	0	11.519	4250
Second decile	0.795	0	9.690	4313
Third decile	0.751	0	12.845	4424
Fourth decile	0.553	0	7.702	4447
Fifth decile	0.470	0	5.117	4500
Sixth Decile	0.288	0	2.677	4414
Seventh Decile	0.245	0.000	2.007	4256
Eigth Decile	0.290	0.052	3.645	4192
Nineth Decile	0.195	0.068	1.763	4069
Tenth decile	0.213	0.090	1.621	3798

TABLE 9: SIZE DISCOUNT USING MARKET TO BOOK

Regression of adjusted market to book upon a variable, big, that is equal to 1 if the observation belongs to the third quartile of the distribution of size. I use a Sample of only non-diversified firms. All the P-values are found correcting by heteroskedasticity

	Coefficient	P-Value	Coefficient	P-Value
Intercept	0.302	0.000	0.304	0.000
Big	-0.067	0.001	-0.060	0.002
Capexp/sales			0.082	0.016
Earnings/sales			-0.123	0.356
R^2	0.000			0.000
Observations	42,662			41,116

TABLE 10: SIZE DISCOUNT USING MARKET TO SALES

Regression of adjusted market to sales upon a variable, big, that is equal to 1 if the observation belongs to the third quartile of the distribution of size**.** I use a Sample of only non-diversified firms**.** All the P-values are found correcting by heteroskedasticity

	Coefficient	P-Value	Coefficient	P-Value
Intercept	0.582	0.000	0.265	0.000
Big	-0.351	0.000	-0.062	0.204
Capexp/sales			2.632	0.000
Earnings/sales			-4.343	0.000
R^2	0.000		0.049	
Observations	42,662		41,119	

TABLE 11: DESCRIPTIVE STATISTICS OF SIZE
(All Values in millions of dollars)

	Non-Diversified firms		**Diversified Firms**	
	Mean	Median	Mean	Median
Book Value	741.99	106.99	1962.72	298.59
Sales	729.22	124.33	1894.72	344.12
Market Value	1043.44	135.18	2357.72	298.42
Number of seg.	1	1	3.02	3

TABLE 12
Matching of segments with non-diversified firms in the same industry according to the *total firm* size.

	Lowest 20%	**20-40%**	**40-60%**	**60-80%**	**80-100%**
Total	7,253	5,754	7,267	9,323	41,048
Percentage	10.26	8.14	10.28	13.19	58.10

TABLE 13: REGRESSION OF MARKET TO BOOK ADJUSTED BOTH BY SIZE AND INDUSTRY

Regression of adjusted market to book upon a variable, div, that is equal to 1 if the firm is diversified and zero otherwise. The industry-adjustment is made taking into account firms of comparable size. All variables are in logs except Earnings to sales. All the P-values are found correcting by heteroskedasticity

	Coefficient	P-Value	Coefficient	P - Value
Intercept	0.006	0.009	0.264	0.000
Div	0.140	0.000	0.149	0.000
Book value			-0.014	0.000
Cap. Exp./sales			0.058	0.000
Earnings/sales			0.034	0.239
R^2	0.012		0.013	
Observations	63,822		63,893	

TABLE 14: REGRESSION OF MARKET TO SALES ADJUSTED BOTH BY SIZE AND INDUSTRY

Regression of adjusted market to book upon a variable, div, that is equal to 1 if the firm is diversified and zero otherwise. The industry-adjustment is made taking into account firms of comparable size. All variables are in logs except Earnings to sales. All the P-values are found correcting by heteroskedasticity

	Coefficient	P-Value	Coefficient	P-Value
Intercept	-0.011	0.000	0.305	0.000
Div	-0.002	0.765	-0.033	0.000
Book value			0.020	0.000
Capexp/sales			0.134	0.000
Earnings/sales			0.031	0.246
R^2	0.000		0.056	
Observations	64,503		61,955	

TABLE 15: `USUAL' REGRESSION OF THE DIVERSIFICATION DISCOUNT CONTROLLING BY DIFFERENCES IN PROFITABILITY FREE OF ACCOUNTING BIASES. THE ENDOGENOUS VARIABLE USED IS MARKET TO BOOK.

Regression of adjusted market to book upon a variable, div, that is equal to 1 if the firm is diversified and zero otherwise. The results are found using exactly the same procedure as in Berger-Ofek(1995)) with the exception that for finding the representative ratio of the industry I use all the firms that belong to the same three digit SIC code. The explanatory variables are also industry adjusted. All the P-values are found correcting by heteroskedasticity

	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P - Value
Intercept	0.0227	0.000	0.091	0.000	0.138	0.000
Div	0.136	0.000	-0.269	0.000	-0.024	0.023
Capexp/sales			0.125	0.0000	0.140	0.000
Earnings/sales			-0.000	0.799	-0.000	0.703
Book value			0.010	0.000	0.021	0.000
Cogs/sales	-0.139	0.000			-0.193	0.000
R^2	0.016		0.046		0.071	
Observations	60,782		58,235		58,228	

TABLE 16: `USUAL' REGRESSION OF THE DIVERSIFICATION DISCOUNT CONTROLLING BY DIFFERENCES IN PROFITABILITY FREE OF ACCOUNTING BIASES. THE ENDOGENOUS VARIABLE USED IS MARKET TO SALES.

Regression of adjusted market to sales upon a variable, div, that is equal to 1 if the firm is diversified and zero otherwise. The results are found using exactly the same procedure as in Berger-Ofek(1995)) with the exception that for finding the representative ratio of the industry I use all the firms that belong to the same three digit SIC code. The explanatory variables are also industry adjusted. All the P-values are corrected by heteroskedasticity

	Coefficient	P-Value	Coefficient	P value	Coefficient	P Value
Intercept	-0.005	0.091	0.024	0.081	0.113	0.000
Div	0.269	0.000	-0.560	0.000	-0.097	0.000
Book value			-0.005	0.854	0.021	0.000
Capexp/sales			0.314	0.000	0.343	0.000
Earnings/sales			-0.000	0.550	-0.000	0.451
Cogs/sales	-0.247	0.000			-0.365	0.000
R^2	0.030		0.169		0.225	
Observations	60,786		58,231		58,228	

Table A1: Market to Earnings to ratio and Diversification

	MEAN	MEDIAN	STD	OBSERVATIONS
segn>1	24.4205	17.2002	165.8771	20927
segn=1	24.5351	17.4246	170.093	42146
segn=2	25.9477	16.5557	175.158	9013
segn=3	23.8261	17.6491	147.192	6244
segn=4	20.3197	17.5637	173.35	3291
segn=5	26.5661	17.7297	153.903	1490
segn=6	27.0983	18.2796	150.134	555
segn=7	23.5953	19.4466	286.592	205
segn=8	8.5102	17.4868	79.372	71
segn=9	27.7179	21.3527	51.999	35
segn=10	21.5493	18.3587	26.981	23

TABLE A2
Regression of market to earnings upon a variable, div, that is equal to 1 if the firm is diversified and zero otherwise.

	Coefficient	P-Value
Intercept	22.89	0.0000
Div	-0.86	0.5334
Capexp/sales	13.33	0.0182
Earnings/sales	23.40	0.0000
Book value	0.000307	0.2832
R^2	0.0015	
Observations	62,243	