LIMITS OF FIRM SIZE

AN INQUIRY INTO DISECONOMIES OF SCALE

Doctoral thesis
by Staffan Canbäck
Henley Management College
11 September 2000

TABLE OF CONTENTS

LIMITS OF FIRM SIZE

AN INQUIRY INTO DISECONOMIES OF SCALE

Doctoral thesis
by Staffan Canbäck
Henley Management College
11 September 2000

This thesis tests the proposition by Oliver Williamson (1975, 117) *that transaction cost economics (TCE) can help explain the limits of firm size: "The distinctive powers of internal organization are impaired and transactional diseconomies are incurred as firm size [is] progressively extended." The research suggests, based on a survey of the literature on scale diseconomies and a statistical analysis of the relative performance of large manufacturing companies in the United States, that diseconomies of scale exist and significantly influence profitability and growth.*

1. INTRODUCTION

Why are large companies so small? Why cannot firms effortlessly expand into new businesses? No coherently articulated reason exists today for why the largest business organisations do not have ten, twenty, or a hundred million employees rather than a few hundred thousand.

Limits of firm size have significant strategic implications. A large company operating on the edge where marginal diseconomies start to reduce profitability presumably has to either make an informed choice between geographic reach, product breadth, and vertical depth; and / or it can try to minimise the diseconomies.

This thesis builds on the original research made in the subject area. Specifically, it tests if Oliver Williamson's "limits of firm size" discussions in *Markets and hierarchies: Analysis and antitrust implications* (1975, 117-131) and in *The economic institutions of capitalism* (1985, 131-162) are valid. The findings include a perspective on the nature of the diseconomies of scale and factors that mitigate their impact, and a quantification of the impact of the diseconomies of scale on company results.

The theoretical foundation for the research is found exclusively in transaction cost economics. There are other partial explanations of diseconomies of scale, such as those found in neo-classical economics (e.g. Mas-Colell, Whinston, and Green 1995; Scherer and Ross 1990), agency theory (e.g. Pratt and Zeckhauser 1985; Jensen and Meckling 1976), growth theory (e.g. Penrose [1959] 1995), evolutionary theory (e.g. Nelson and Winter 1982), sociology (e.g. Blau and Meyer 1987), and Marxist theory (e.g. Marglin 1974). These explanations are not of concern here, although they will be used to illuminate particular aspects of the TCE argument.

The purpose of the thesis is to create a model that can be used by executives to delineate strategic and organisational choices for large companies.

2. RESEARCH OBJECTIVES

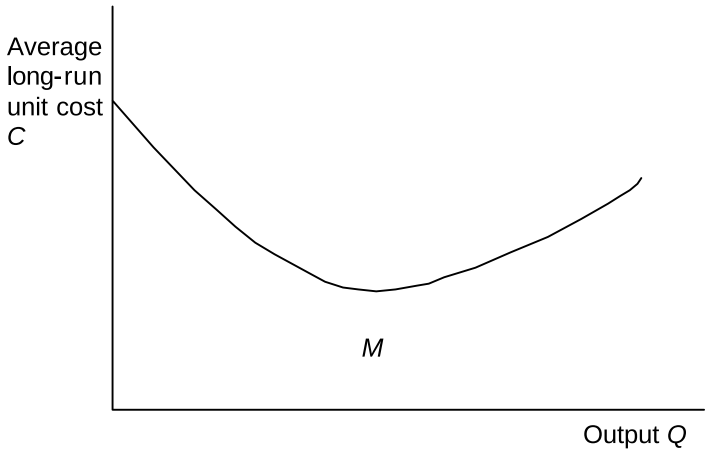
This chapter gives an initial problem definition and discusses the importance of the research.

2.1 PROBLEM DEFINITION

Knight ([1921] 1971, 286-287) observed that the “diminishing returns to management is a subject often referred to in economic literature, but in regard to which there is a dearth of scientific discussion.” Since then, many authorities have referred to the existence of scale diseconomies, but there appear to be no systematic studies of the general issue. The basic dilemma is, on the one hand, that if there are no diseconomies of scale, then there are no limits to firm growth. We would observe an inexorable concentration of industries and economies until there is only one global firm left. As George Stigler (1974, 8) put it: “If size were a great advantage, the smaller companies would soon lose the unequal race and disappear.” This is not happening. On the other hand, if there is an optimum size in an industry, then we would expect increased fragmentation as the overall economy grows, in line with Stigler's survivor principle argument (1958) which holds that “the competition between different sizes of firms sifts out the more efficient enterprises” (p. 55). This is not happening either. Robert Lucas (1982,

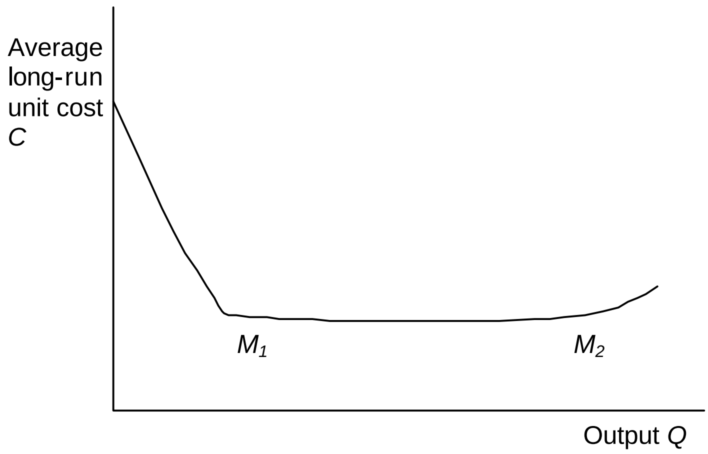
509) observed that “most changes in product demand are met by changes in firm size, not by entry or exit of firms.” The size distribution of firms is remarkably stable over time when measured by number of employees or as a share of the total economy for most of this century (although not lately), as is discussed in chapter 3.

The neo-classical way to illustrate economies and diseconomies of scale is through a cost curve (e.g. Scherer and Ross 1990, 101):



As production increases, the average unit cost *C* decreases due to scale economies. At a certain point, *M*, the scale economies are exhausted while diseconomies of scale start to impact the unit cost. As output increases, the unit cost increases. Thus, a profit-maximising firm should strive for an output at the optimum *M*.

In reality, this is not what is observed. Rather, the cost-minimising (profit-maximising) part of the curve appears to cover a wide range of outputs, and only at high output levels do diseconomies materialise, if ever. McConnell's quantification (1945, 6) and Stigler's illustration (1958, 59) are typical:



This shape of the cost curve reconciles several real life observations. 1) it explains why large and small companies can coexist in the same industry; 2) it is consistent with Lucas's observation that as the economy grows, existing companies tend to expand supply to meet additional demand; 3) it eliminates the supposition that scale economies are exhausted at approximately the same point as scale diseconomies start influencing total cost; and 4) it demonstrates that there are indeed limits to firm size—large companies can not, and have not, expanded indefinitely.

However, if the reasoning above is correct, it is still unclear why the cost curve bends upwards at M_2. Neo-classical theory does not provide a satisfactory answer. As Herbert Simon ([1947] 1976, 292) put it[1]: “the central problem is not how to organize to produce efficiently (although this will always remain an important consideration), but how to organize to make decisions.” The first part of this statement essentially refers to the negative derivative of the cost curve, while the second part refers to the upward slope as diseconomies of scale set in.

The proposed research aims to investigate if transaction cost theory can explain diseconomies of scale, as exhibited by lower relative performance of large firms versus smaller firms, and what drives these diseconomies. It picks up on a debate that has lasted at least since the early 1930s when Florence (1933) and Robinson (1934) argued the case against and for limits of firm size.

---

1 Simon echoed the writing of Robertson (1923, 25): ”It is the economies of large-scale government rather than of large-scale technique which dictate the size of the modern business unit.”

2.2 IMPORTANCE OF RESEARCH

There appears to exist a genuine gap in the understanding of the firm, and that helping fill this gap may have some impact on the way we think about strategy and structure.

Limits of firm size is not a major field of study (Coase 1993a, 228; Holmström and Tirole 1989, 126). There are around 40 articles or books that deal with the topic in a meaningful way (see later in chapters 4 and 5). Williamson (1985, 153), for example, stated that our understanding of bureaucratic failure is low compared to that of market failure. The slowdown in the growth of large companies over the last 30 years (see later in section 3.2) makes it all the more interesting to understand why market based transactions are slowly winning over internally based transactions.

The second reason why the research is academically important is that it uses TCE in a somewhat new fashion. The early 1970s were the defining years of TCE. At that time, large companies still appeared set to become ever more dominant, and the theory is still very much a reflection of this *zeitgeist*. Thus, many of the theory's applications are in antitrust cases. Further, TCE has arguably evolved over time from being a general theory for understanding industrial organisation, to primarily being a tool for analysing vertical integration. This research breaks with this tradition.

Limits of firm size is also a real and difficult problem for business executives. The cost of sub-optimal (i.e. too large) size is probably significant. For example, it has been estimated that 25 percent of the operating budget of a large company is slack due to some of the diseconomies of scale discussed later in chapters 4 and 5. Moreover, large firms have a tendency to slowly decline and disappear (Hannah 1996, 1). Shedding light on why this is the case may be socially and privately beneficial (p. 24), while so far “we have made great strides in storytelling, but a clearer, surer recipe for sustained success for large corporations has remained elusive.”

3. DIMENSIONS OF FIRM SIZE

This chapter defines size and shows the trends in size of firms in the U.S. manufacturing sector.

3.1 DEFINITION OF SIZE

First, there are a number of definitions of what a firm is. The first, based on Coase (1937, 388), Penrose ([1959] 1995, 15), and Arrow (1964, 33), holds that the boundary of the firm is where the internal planning mechanism is superseded by the price mechanism. This means that the operating firm will be equivalent to the legal corporation in most cases. The most important, but rare, exception is a corporation where divisions are totally self-contained profit centres. In this case the parent company is not a firm since the company's divisions by definition trade between themselves through market-based transfer prices. The second definition is that ownership sets a firm's boundaries (e.g. Hart 1995, 7). A firm is the collection of activities for which the bearer of residual risk is the same. A problem is that employees hardly can be part of the firm with this definition. A more serious problem is that a holding company with no control over the operating units is considered a firm. Still, this definition is usually equivalent to Coase's definition since there are few, if any, companies where the divisions are totally self-contained. A third definition is the firm as a network, as defined by Richardson (1972, 884-887). McDonald's, for example, is considerably larger than the ownership definition indicates since it also consists of a network of thousands[2] of franchisees (Rubin 1990, 134-144). The fourth definition is the firm's sphere of influence. This includes distributors, alliance partners, first and second tier suppliers, etc. Toyota is an example (Williamson 1985, 120-122). Toyota directly employed 159,000 people in 1997, but its sphere of influence probably extended over more than 1 million people.

This thesis uses the ownership definition. It relates closely to Coase's definition and all publicly available data build on it. Thus, a firm is a corporation henceforth.

Second, there are various ways to measure the size of a firm. Most business press rankings of size are by revenue. However, this measure is fairly meaningless since it tells nothing about the scope of the underlying activity. With this definition, four of the world's five largest companies are Japanese trading houses (Fortune 1995b) which have almost no vertical integration. A better measure of size is value added, i.e. the sum of factor inputs (basically equivalent to revenue less purchased products and services). This metric gives a precise measure of

[2] 18,000 at the end of 1997.

activity, but is usually not available by individual company. Number of employees is the most widely used measure of size, with more than 80 percent of studies using it according to a review by Kimberley (1976, 587). In line with Child's observation (1973, 170), “it is people who are organized,” it can perhaps be expected that the number of employees is the most important constraint on firm size. Finally, assets can define size (e.g. as used by Grossman and Hart 1986, 693-694). As with revenue, this may not reflect underlying activity, but for manufacturing companies this should not be a major issue since asset to value added ratios are fairly homogeneous outside the financial sector. Assets by firm are usually available back to the 1890s and are therefore a practical measure in longitudinal studies. In sum, the best measure of size are value added and number of employees, while assets can be used. Revenue is not suitable. This research uses number of employees as the size metric.

The definitions are summarised in the table below.

DEFINITION OF FIRM SIZE				
	Internal planning (Coase)	Ownership	Network	Sphere of influence
Revenue				
Value-added				
Employees				
Assets				

3.2 TRENDS IN FIRM SIZE

The U.S. economy is used as the basis for analysis since it is the largest and most competitive economy in the world. Within this economy, the research focuses on the manufacturing sector[3].

Large manufacturing firms play a major role in the U.S. economy. The Fortune industrial 500 companies control more than 50 percent of corporate manufacturing assets and employ more than eleven million people (Fortune 1995a). Their sphere of influence is perhaps 40 million employees out of a total private sector workforce of 123 million. Contrary to popular belief, however, the large companies' importance is not increasing, and has not done so for many years. Studies show that large manufacturing firms are holding steady as a share of value added since circa 1965 (Scherer and Ross 1990, 62). Their share of employment in the manufacturing sector has declined from around 60% to around 50% 1979 and 1994. Moreover, as a share of the total U.S. economy they are in sharp decline. An example is that large manufacturing companies employed 16 million people in 1979 versus 11 million in 1994 (Fortune 1995a, 185), while private sector employment grew from 99 to 123 million people (Council of Economic Advisers 1998, 322) over the same time period.

Further evidence is available from a number of historical studies. Aggregate industry concentration[4] has changed little since the early part of this century. Nutter (1951) studied the concentration trend between 1899 and 1939 and found no signs of increased aggregate concentration during this time (pp. 21, 33), mainly because new, fragmented, industries emerged while older ones consolidated. Bain (1968) found the same trend between 1931 and 1963, but with less variability between industries. Scherer and Ross (1990, 84) used a modified Nutter methodology and showed that aggregate concentration has increased slightly from 35 percent in 1947 to 37 percent in 1982. Similarly, Mueller and Hamm (1974, 512) found an increase in four-firm concentration from 40.5 percent to 42.6 percent between 1947 and 1970, with most (70%) of the increase between 1947 and 1963.

Bain (1968, 87) calculated that the assets controlled by the largest 200 non-financial companies was around 57 percent of total non-financial assets[5] in 1933. He also estimated that the 300 largest non-financial companies accounted for 55 percent of non-financial assets in 1962. Based on this, the top 200 companies

---

3 Alternative approaches would be to look at the global manufacturing sector or the total U.S. private sector, or both. Statistics on the global manufacturing sector are not reliable, and the non-manufacturing sectors are often highly regulated.

4 Although there have been significant changes within industries.

5 A similar study by Berle and Means ([1932] 1991) has been discredited. For example, Scherer and Ross (1990, 60) found that Berle and Means, based on “meager data then available…overestimated the relative growth of the largest enterprises.”

accounted for approximately 50 percent of non-financial assets in 1962, using this researcher's estimate of the assets controlled by the 100 firms smallest companies in the sample. Data from 1994 show the same ratio to be around 40 percent. Adelman (1978) observed a similar pattern when he studied the 117 largest manufacturing firms between 1931 and 1960. He found that concentration was the same at the beginning and at the end of the period (45 percent). He concluded that "overall concentration in the largest manufacturing firms has remained quite stable over a period of 30 years, from 1931 to 1960." This researcher replicated the analysis for 1994 and found the equivalent number to be 45 percent in 1994.

Finally, Bock (1978, 83) studied the share of value added to total value added of the largest manufacturing firms between 1947 and 1972. There was a large increase between 1947 and 1954 and a slight increase up till 1963. Between 1963 and 1972 there was no increase. Scherer and Ross (1990, 62) confirmed the lack of increase up till the end of the 1980s.

The stock market does not expect the largest companies to outperform smaller companies in the future. The stock market valuation of the largest companies relative to smaller companies has declined sharply over the last 34 years (Farrell 1998). In 1964 the largest 20 companies made up 44 percent of total stock market capitalisation, in 1998 they made up 19.5 percent. The market value primarily reflects future growth and profit expectations and thus the market is increasingly sceptical of the large companies' ability to compete with smaller firms. A study of companies on the New York stock exchange (Ibbotson Associates 1999) similarly show that small companies have outperformed large companies between 1925 and 1997. The total annual shareholder return over the period was 11.9 percent for the largest size (1) decile, 13.7 percent in the second size (2) decile, and increasing steadily to 21.8 percent for the smallest size (10) decile (p. 131). The real return after adjustment for risk (using the capital asset pricing model) was –0.30 percent for decile 1, 0.42 percent for decile 2, and rising steadily to 5.36 percent for decile 10 (p. 142). It should be noted though, that the definition of size here is market capitalization; not revenues, value added, employees, or assets.

In sum, the above evidence shows that industry concentration has changed little since the early part of the century. The size of large firms has kept pace with the overall growth of the industrial part of the economy since the 1960s in value added terms, but has declined in employment terms since 1979 (and has declined relative to the total U.S. corporate sector and the global corporate sector). This indicates that there is a limit to firm size and that this limit may be decreasing in absolute terms.

4. THEORETICAL FRAMEWORK AND HYPOTHESES

Transaction cost economics aim to explain the boundary of the firm, what is made internally and what is bought and sold in the marketplace. As firms internalise transactions, bureaucratic diseconomies of scale appear. Thus, a firm will reach a size where the benefit from the last internalised transaction is offset by the bureaucratic diseconomies. Two factors mitigate these diseconomies. First, under conditions of high asset specificity, high uncertainty, or high frequency of transactions, it will be advantageous to internalise transactions. Second, firms can mitigate the diseconomies by organising appropriately.

4.1 TCE AND THE LIMITS OF FIRM SIZE

Four pieces of work within TCE are relevant to the argument. Ronald Coase's original article *The nature of the firm* (1937) established the basic framework. Chapter *7 (Limits of vertical integration and firm size*) in Oliver Williamson's book *Markets and hierarchies* (1975) suggested the nature of limits of size. Chapter 6[6] (*The limits of firms: Incentive and bureaucratic features*) in Williamson's book *The economic institutions of capitalism* (1985) expands on this theme and explains why the limits exist. Riordan and Williamson's article *Asset specificity and economic* *organization* (1985) augments the model by combining transaction costs with neo-classical production costs.

4.1.1 Reason for limits

Coase's paper on transaction costs (1937) is the foundation of the New Institutional Economics branch of industrial organisation. Coase asked the fundamental questions "Why is there any organisation?" (p. 388) and "Why is not all production carried on by one big firm?" (p. 394). His answer was that there are transaction costs that determine what is done in the market, with price as the regulating mechanism, and what is done inside the firm, with bureaucracy as the regulator. Coase pointed out that "the distinguishing mark of the firm is the supersession of the price mechanism" (p. 389). Within this framework, all transactions carry a cost, either an external market transaction cost or an internal bureaucratic transaction cost. "The limit to the size of the firm would be set when the scope of its operations had expanded to a point at which the costs of organising additional transactions within the firm exceeded the costs of carrying out the same transactions through the market or within another firm." (Coase 1993b, 48).

---

[6] Published earlier in a less developed form (Williamson 1984).

According to Coase the most important market transaction costs are the cost of determining the price of a product or service, the cost of negotiating and creating the contract, and the cost of information failure. The most important internal transaction costs are associated with the administrative cost of determining what, when, and how to produce, the cost of resource misallocation, since planning will never be perfect, and the cost of de-motivation, since motivation is lower in large organisations. In any given industry the relative magnitude of market and internal transaction costs will determine what is done where.

Coase thus created a theoretical framework which potentially explains why firms have size limits. However, this is only true if there are decreasing returns to the entrepreneur function (Penrose [1959] 1995, 19). Later, work by Williamson (1975, 130) argued that this is the case. “Why can’t a large firm do everything that a collection of small firms can do and more?” (Williamson 1984, 736). Williamson pointed out that the incentive structure of a firm has to be different from the market. Even if a firm tries to emulate the high-powered incentives of the market there will be unavoidable side effects, and the cost for setting up the incentives is non-trivial. Thus, the combination of small firms into a large firm will never have the same operating characteristics as when they operate independently in the market.

4.1.2 Nature of limits

Williamson (1975) found that the limits of firm size are bureaucratic in origin and can be explained by transaction cost economics. He identified four main categories of scale diseconomies: communication distortion due to bounded rationality, bureaucratic insularity, atmospheric consequences due to specialisation (p. 126), and incentive limits of the employment relation (p. 129).

Communication distortion due to bounded rationality. Since a manager is boundedly rational it is impossible to expand a firm without adding hierarchical layers. As information is passed between layers it is necessarily distorted. This reduces the ability of high level executives to make decisions based on facts and leads to declining return to the entrepreneurial function. In an earlier article (1967), Williamson found that even under static conditions (without uncertainty) there would be a control-loss phenomenon. He developed a mathematical model to demonstrate that control-loss is of critical importance to limitations of firm size and that there is no need to assume rising factor costs to explain the limits (p. 127-130):

$$\ln N^{*} \sim \ln(1/(s-1)) + \ln s\{1+(1/\ln\alpha)[\ln(w_0/(P-r)) + \ln((s/(s-\beta)) + \ln(\ln s/\ln(\alpha s))]\}$$

and

$$n^{*} \sim \ln(N^{*}(s-1))/\ln s$$

Where:

N^* = optimal number of employees

n^* = optimal number of hierarchical levels

s = span of control

α = fraction of work done by a subordinate that contributes to objectives of his/her superior

w_0 = wage of employee

P = price of output

r = non-wage variable cost per unit of output

β = wage multiple between superior and subordinate

Williamson applied data from the 500 largest companies in the United States to the model and showed that the optimal number of hierarchical levels is between 4 and 7. Beyond this, control loss leads to “a static limit on firm size” (p. 135).

Bureaucratic insularity. Williamson (1975) argued that as firms increase in size the senior managers are less accountable to the lower ranks of the organisation (p. 127) and to the shareholders (p. 142). They thus become insulated and will, given opportunism, strive to maximise their personal benefits rather than the corporate goal function (profits). This problem should be most acute in organisations with well-established procedures and rules and where management is well entrenched. This argument is similar to agency theory (Jensen 1989; Jensen and Meckling 1976) which holds that corporate management will tend to overemphasise size over profitability and will keep excess cash flow within the firm rather than distribute it to a more efficient capital market (a lengthier comparison of agency theory and transaction cost theory is found in section 5.1.1). The consequence is that large firms tend to more easily accept organisational slack and that resources are misallocated. If this is correct we will, for example, expect to see wider diversification of large firms, as well as lower profits.

Atmospheric consequences. As firms expand there will be increased specialisation, but also less moral involvement of the employees, according to Williamson (1975, 128-129). The decline in moral involvement is due to the difficulty for the employee to understand the purpose of activities as well as the small contribution each employee makes to the totality. Thus, alienation is more likely to occur in large firms.

Incentive limits of the employment relation. Firms can not compensate their employees perfectly due to a number of limitations, according to Williamson (1975, 129-130). First, large bonus payments may threaten senior managers. Second, performance-related bonuses might affect the employment contract so that less than optimal behaviour is encouraged. The outcome is that large firms tend to pay based on tenure and position rather than on merit. This is especially important in product and process development where the large firms are at a disadvantage to smaller enterprises.

Williamson's four categories are similar to those Coase described in 1937. Coase talked about the determination (or planning) cost, the resource misallocation cost, and the demotivation cost. Williamson's first and second category corresponds broadly to the determination cost, the third category to the demotivation cost, and the fourth category to the resource misallocation cost. Williamson's categories are, however, more specific and allow for easier operationalisation as is shown in chapter 6.

There are a number of consequences of these four diseconomies of scale according to Williamson[7].

- Large companies tend to procure internally when facing a make or buy decision (1975, 119-120).

[7] Williamson's descriptions are confusing. They are found throughout the chapters referenced, in-between theory and examples, and at various levels of the section hierarchies. The outcomes discussed here are this author's attempt to make Williamson's descriptions more explicit.

- They have excessive compliance procedures and compliance related jobs will proliferate. Thus, policing costs such as audits will be excessive (Williamson 1975, 120-121).

- There is a tendency for projects to persist even though they are clear failures (1975, 121-122).

- There is conscious manipulation of information to further individual or sub-unit goals (1975, 122-124).

- Asset utilisation is lower since high-powered market incentives do not exist (1985, 137-138).

- Transfer prices do not reflect reality and cost determination suffers (1985, 138-140).

- Research and development productivity is lower (1985, 141-144).

- The organisation suboptimises by trying to manage the unmanageable, by forgiving mistakes, and by politicising decisions (Williamson 1985, 148-152)

The following linkages seem reasonable between the limiting factors and the outcomes.

LINKAGE BETWEEN LIMITING FACTORS AND OUTCOMES				
Outcomes	**Sources**			
	Communication distortion	**Bureaucratic insularity**	**Atmospheric consequences**	**Incentive limits**
Internal procurement		Strong	Moderate	Strong
Excessive compliance procedures	Strong	Strong	Strong	Strong
Project persistence		Strong	Strong	Moderate
Conscious manipulation of information	Strong	Strong		
Low asset utilisation	Strong		Strong	
Poor internal costing	Strong			Strong
Low R&D productivity	Strong	Moderate	Strong	Strong

These outcomes make it plausible that a large firm exhibits lower relative profitability and growth than a smaller firm with the same product and market mix.

4.1.3 Mitigating influences on the limits of firm size

While the categories discussed in the previous section theoretically impose limits of firm size, there are two mitigating influences that tend to offset the diseconomies of scale. Both of these influences is central to transaction cost economics and thus the argument continues to be confined to this theory. To test the validity of the diseconomies of scale, it is necessary to take these mitigating influences into account.

Asset specificity. There is a vast literature on vertical and lateral integration applications for transaction cost economics and the purpose here is not to review this at length. The theoretical argument is summarised in Williamson (1975, 43-67). Mahoney (1989; 1992) provides overviews of theoretical and empirical work on vertical integration problems. Grossman and Hart (1986) and Teece (1982; 1976; 1980) illustrates the use in lateral relationships. Williamson showed that three factors play a fundamental role in determining the degree of integration: *asset specificity*, *uncertainty*, and *frequency of transactions* under the conditions of bounded rationality (Simon [1947] 1976, xxvi-xxxi) and opportunism (Williamson 1993).

With high asset specificity, market transactions become expensive. By asset specificity is meant physical assets, human assets, site, or dedicated assets (Williamson 1985, 55) which have a specific use and cannot easily be transferred[8]. Opportunistic behaviour can be expected if the asset is part of a market

---

[8] Williamson (1996, 59-60) added brand name capital and temporal specificity.

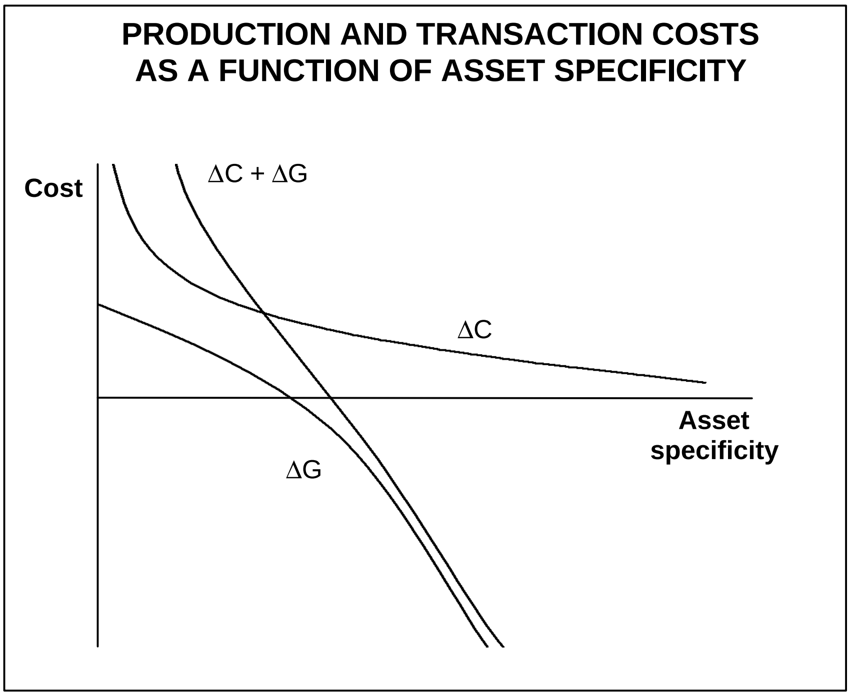
transaction under this condition. An example is when a supplier invests in specific tooling equipment dedicated to one customer. Over time, the customer will be able to put pressure on the vendor since the vendor has no alternative use for the investment. The vendor is ultimately willing to accept a price down to the variable cost of production to cover the fixed cost. By owning the asset the incentive to cheat disappears and the cost of creating contractual safeguards is reduced (Williamson 1985, 32-35).

High uncertainty such as business cycle volatility or technological uncertainty leads to more bureaucratic transactions since it will be difficult, and prohibitively expensive, to create contracts which cover all possible outcomes. Thus, with higher uncertainty firms tend to internalise activities. Finally, if the transactions are frequent there is again a tendency to manage the transaction through bureaucracy since the repeated market contracting cost will be higher than the internal bureaucratic cost. While uncertainty and frequency play some role in creating transaction costs, Williamson considered asset specificity the most important driver (e.g. Riordan and Williamson 1985, 366). Asset specificity is furthermore relatively independent of the drivers of limits of firm size (p. 368).

Neo-classical production costs also exhibit diseconomies as a function of asset specificity (Riordan and Williamson 1985, 369):

> The diseconomies are arguably great where asset specificity is slight, since the outside supplier here can produce to the needs of a wide variety of buyers using the same (large scale) production technology. As asset specificity increases, however, the outside supplier specializes his investment relative to the buyer. This is the meaning of redeployability. As these assets become highly unique, moreover, the firm can essentially replicate the investments of an outside supplier without penalty. The firm and market production technology thus become indistinguishable at this stage.

This is illustrated in the graph below where the differential production cost (ΔC) and transaction cost (ΔG) are shown as a function of asset specificity.



The implication of the asset specificity argument, from both a transaction cost and a production cost perspective, is that firms with high asset specificity will not reach the limits of size as quickly as those with low specificity. Or, alternatively, “larger firms are more integrated than smaller rivals” (p. 376).

Organisational form. Williamson (1975, 117) also recognised that the diseconomies of scale can be reduced by organising appropriately (p. 117). Based on Chandler’s (1962; 1980) pioneering work on the evolution of the American corporation, Williamson argued that the multidivisional (M) form of organisation lowers the internal transaction cost compared to the unitary[9] (U) form. The two key reasons are that the M-form allows the most senior executives to prioritise high-level issues rather than day-to-day operational details, while the whole is greater than the sum of the parts (p. 137). Thus, large firms organised according to the M-form should, *ceteris paribus*, be more profitable than U-form firms.

4.2 TCE-BASED HYPOTHESES OF FIRM SIZE LIMITS

It is now possible to formulate five hypotheses based on the size-distribution of companies discussed in section 3.2 and on the TCE-based model of scale diseconomies discussed in the previous section (4.1). The hypotheses will guide the qualitative evaluation of the model in the literature survey (chapter 5)and the quantitative tests performed in the statistical analyses (chapter 8).

There are no obvious relationships between a company's results (as measured by profitability, growth, or other variables) and its size. If the relationships were easy to see then this would have been done a long time ago. The only obvious observable fact is that no company has more than 600,000 employees, which suggests that diseconomies of scale become severe beyond this size. These effects are not possibly to quantify though. Within the size range where companies exist the following hypothesis can be defined:

H_1: A company's size is not a direct predictor of its results.

Transaction cost economics holds that underlying this are hidden diseconomies of scale which increase as the size of the organisation increases. These diseconomies are bureaucratic in nature and not easily observed. They exist because there are decreasing returns to the entrepreneurial function and because large companies cannot fully replicate the high-powered incentives that exist in the market—leading to bureaucratic failure (the opposite of market failure). The theory predicts that there are four types of diseconomies of scale: communication distortion due to bounded rationality, bureaucratic insularity, atmospheric consequences due to specialisation, and incentive limits of the employment relation. The hypothesis is:

---

[9] Often referred to as functional organisation by other authorities, including Chandler.

H_2: Large companies embed diseconomies of scale of four types: communication distortion, bureaucratic insularity, atmospheric consequences, and incentive limits.

As was shown in chapter 3, the average size of large[10] manufacturing companies in the United States has declined since the 1960s relative to the total economy. Thus, as large companies have become more productive they have on average not been able to fully compensate for the per unit decline in value added by expanding into new geographic markets (reach), product areas (breadth), or by increasing vertical integration (depth). In line with Stigler's survivor principle (1958) this indicates that there are diseconomies of scale (p. 71). These diseconomies are exhibited through lower relative profitability and / or slower growth of the largest firms relative to smaller firms, *ceteris paribus* (such as risk and financial leverage). The hypothesis becomes:

H_3: Scale diseconomies impact company profitability and growth negatively.

It is possible for firms to mitigate the negative impact of scale diseconomies. Transaction cost economics show that large companies benefit from multidivisional structures, while unitary (functional) structures impede

[10] Large is defined as the largest 100 corporations.

performance. Moreover, conscious choices around the degree of integration affects performance. In particular, companies with a high degree of internal asset specificity will outperform companies with low internal asset specificity. The hypothesis can be formulated as:

H_4: Diseconomies of scale can be mitigated by two transaction cost-related factors: organisational form and asset specificity.

In summary, the performance of a firm depends on two counteracting forces. On the one hand, four size-related factors contribute to diseconomies of scale and determine the firm’s limit. If these factors are important, then (all other things equal) the larger firm will show poorer results than the smaller firm. On the other hand, there are two mitigating factors. First, firms can organise to mitigate the effects of the diseconomies. In general, large firms will benefit from using the M-form organisation. M-form companies will on average show better results than U-form companies, or they can be larger with the same results. Second, appropriate levels of integration (primarily dependent on asset specificity, secondarily on uncertainty and frequency of transaction) can improve company results to further mitigate the effects of the diseconomies. Finally, economies of scale in the production function may, or may not, offset the effects of the economies of scale.

TCE-BASED "LIMITS OF FIRM SIZE" MODEL						
Relative value of large firm	**Sources of limits of firm size**				**Offsets**	
	Communication distortion	**Bureau-cratic insula-rity**	**Atmos-pheric conse-quences**	**Incentive limits**	**Asset speci-ficity**	**Organi-sation form**
High	Low	Low	Low	Low	High	M-form
Low	High	High	High	High	Low	U-form

5. LITERATURE REVIEW

This chapter aims to validate the "limits of firm size" model developed above and to modify or complement it if other factors are found. In general, no one has done substantial research on the diseconomies of scale. This is somewhat surprising since many authorities mention the analysis of limits of firm size as critical to our understanding of the modern economy. Fortunately though, there are fragments of evidence in much of the relevant literature. The composite picture of these fragments broadly supports the model developed in the previous chapter.

5.1 DISECONOMIES OF SCALE

The literature relating to the limits of firm size does not, for obvious reasons, follow Williamson's categorisation. Thus, this section will review the evidence by general topic and by author. At the end of the chapter the arguments are summarised and related back to the sources of diseconomies in the "limits of firm size" model.

5.1.1 Previous research

A number of sociological studies describe negative consequences of size which correlate well with Williamson's propositions in the previous chapter. Pugh et al. (1969) and Child (1973), among others, showed that size leads to bureaucracy. Thus, large firms are usually highly bureaucratised through formalisation, and to the extent that there are diseconomies of bureaucracy, these apply to the "limits of firm size" model. Williamson (1996, 266) made a similar point, "almost surely, the added costs of bureaucracy are responsible for limitations in firm size." The diseconomies of bureaucracy fall into three major categories (Blau and Meyer 1987, 139-161): 1) excessive rigidity, 2) conservatism and resistance to change, and 3) perpetuation of social-class differences. Of these, the first one is relevant here (conservatism is essentially a subcategory of rigidity). Excessive rigidity appears as organisations formalise work practices through bureaucratic procedures (Merton 1957, 197-200). Problems are solved by adding structure and the firm reaches a point where the added structure costs more than the problem solved: the "problem—organisation—problem—more organisation" spiral of bureaucratic growth (Blau and Meyer 1987, 147). They showed that external factors, such as increased volume of tasks, have little to do with increased bureaucracy. In the end, the added policies and procedures stifle flexibility. Crozier (1964) also emphasised rigidity as the most important dysfunction of bureaucracy. In fact, he viewed the bureaucratic organisational model as inherently inefficient, especially under conditions of uncertainty. A key problem is that management will be increasingly insulated from reality while lower levels of the organisation will experience alienation. Stinchcombe (1965) demonstrated that a consequence of this rigidity is that companies tend to maintain the organisation form they had when they were created.

Pondy (1969) studied the administrative intensity in different industries and the causes for variations. He found a positive correlation between size of administration and firm size when he included a measure of ownership-management separation. This is in line with Williamson's notion of bureaucratic insularity which argues that management will be more shielded from reality the larger the organisation is and the more distant the owners are.

A similar logic from institutional economics can be found in Olson (1982). His theory holds that as the institutional structure of a country ages, growth-retarding organisations will become more and more abundant. Thus, "states that have been "organized the longest ought, all other things being equal, to have the lowest rates of growth…" The theory specifically predicts that such countries exhibit lower economic growth. If the parallel holds for corporations then older companies should experience lower growth.

A few studies within the “firm as information processor” school of thought relate to diseconomies of scale. Arrow (1974) found that employees in large organisations tend to be highly specialised. Thus, there is an increasing need for co-ordination through communication. Since information flows carry a cost, organisations will code (through formal or informal rules) the information available. The coding brings the benefit of economising on cost, but it also leads to information loss and rigidity (p. 55). The implications are 1) that the longer the hierarchy, the more information loss or distortion; and 2) the older the firm is, the higher the rigidity. Simon ([1947] 1976) made a similar point. Based on his concept of bounded rationality—”human behavior is *intendedly* rational, but only *limited* so” (p. xxviii)—Simon found that information degrades as communications lines are extended. Geanakoplos and Milgrom (1991) added to this perspective by noting that there are inevitable delays of signals in an organisation. The longer the hierarchy, the longer and more frequent the delays. As a practitioner. Barnard ([1938] 1968) found that the size of unit organisations is “restricted very narrowly by the necessities of communication” (p. 110) and that “the size of executive organizations is limited generally by the same conditions that govern the size of unit organizations” (p. 112).

Control-loss problems may contribute to diseconomies of scale. McAfee and McMillan (1995) argued that people in organisations exploit information asymmetries to their advantage (in Williamson's (1993) words: opportunism). Dispersion of knowledge within the organisation combined with individualised incentives make conflict of interest and subgoal pursuit inevitable. They find, among other things, that efficiency will fall as the hierarchy lengthens, and that long hierarchies are not viable in competitive industries (p. 401). Qian (1994), with a logic similar to McAfee and McMillan's, found that large hierarchies will result in low effort levels among the employees. The employees will not have complete information about their role in the enterprise and thus suffer from demotivation. Moreover, there will be a need to monitor effort, leading to higher costs and further demotivation.

An early version of agency theory argues that very large firms will not strive for profit maximisation (Monsen and Downs 1965). They found that such firms need to build "bureaucratic management structures to cope with their administrative problems. But such structures inevitably introduce certain conflicts of interest between men in different positions within them. These conflicts arise because the goals of middle and lower management are different from those of top management. The introduction of these additional goals into the firm's decision-making process also leads to systematic deviations from profit-maximizing behavior." (p. 222). They furthermore found that the motives of managers are different from the motives of owners. Managers tend to maximise personal income while owners maximise profits. It is impossible for owners of large companies to control the behaviour of managers and consequently, profit maximisation does not obtain. The outcome is akin to what Williamson labels bureaucratic insularity.

Silver and Auster (1969) argued that a result of the “divergences of interests within the firm and the costs of dealing with them” (p. 277) is that “the entrepreneur's time is a limitational factor” (p. 280). The reason for this is that employees typically “will shirk their duties unless the employer takes steps to prevent this” (p. 278). This leads to diseconomies in the entrepreneurial function, all other things equal. Silver and Auster furthermore made two predictions based on this argument: 1) the higher the labour content is of an industry's value added, the sooner the total cost curve will turn up. Thus, such industries will be more fragmented; and 2) the higher the need for supervision of employees, the lower the concentration ratio.

Michael Jensen has deepened and extended these arguments over the last 25 years (e.g. Jensen and Meckling 1976; Jensen 1986, 1988, 1989, 1990). He defines agency cost as the sum of the monitoring expenditures by the principal, the bonding expenditures by the agent, and the residual loss. The magnitude of agency costs depends on a number of factors, including the transparency of the firm's activities and the market for managerial talent. Jensen does not, contrary to Monsen and Downs, and Silver and Auster, explicitly state that agency costs increase with the size of the firm. Jensen does demonstrate, however, that managers will emphasise size over profitability: "Managers have incentives to cause their firms to grow beyond optimal size. Growth increases managers' power by increasing the resources under their control. It is also associated with increases in managers' compensation." (Jensen 1986, 323). He demonstrates the point by looking at the profitability of diversified companies and notes that they are less profitable than focused companies.

Agency theory and TCE have many similarities and it is thus not surprising that the two theories lead to the same conclusions. However, it has been argued that agency theory is a special case of TCE, and thus does not capture all the costs associated with transactions. Specifically, Williamson and Mahoney (1992, 566) argue that agency costs correspond to the *ex ante* costs in TCE[11], and that TCE works with both *ex ante* and *ex post* costs while emphasising the latter.

[11] Contrarily, Williamson (1988, 572) argues that agency costs correspond to TCE's *ex post* costs.

COMPARISON OF AGENCY COSTS AND TRANSACTION COSTS		
Transaction costs		**Agency costs**
Ex ante	***Ex post***	
Drafting costs	Maladaptation costs when transactions drift out of alignment	Monitoring expenditures of the principal
Negotiation costs	Haggling costs to correct misalignments	Bonding expenditures by the agent
Safeguard costs	Set-up and running costs associated with the governance structure	Residual losses
	Bonding cost of effecting secure commitments	

Further, it has been argued that agency theory explains the boundaries of the firm poorly (Hart 1995, 20): “the principal-agent view is consistent with there being one huge firm in the world, consisting of a large number of divisions linked by optimal incentive contracts; but it is also consistent with there being many small, independent firms linked by optimal arm's-length contracts.”

A number of authorities argue that job satisfaction is lower in large organisations and large work establishments. Evidence of this is that employees in large companies are paid significantly more than are employees in small companies. This difference is argued to be compensation for a less satisfying work environment. Three studies warrant mention here.

Scherer (1976) is representative of the extensive work done at the establishment level. In a review of the literature, and his own original research, he concluded that worker satisfaction is 30 percent lower in large establishments[12] than in small establishments (p. 109) while compensation is more than 15 percent higher for equivalent job descriptions (p. 119). He concluded that since establishment size is correlated to firm size the effect of alienation is possibly significant. Later work sponsored by the Federal Trade Commission in the United States (Kwoka 1980) confirmed these findings (p. 378).

Brown, Hamilton, and Medoff (1990) found that large firms pay a wage premium of 10-15 percent over small firms when adjustments have been made for other effects such as unionisation and skill levels (p. 42). However, they did not conclude that this differential is necessarily related to alienation. Regardless of the cause though, it appears that large firms pay a substantial wage premium over smaller firms.

Span-of-control problems make it increasingly costly to extend incentive contracts to employees as firms grow (Rasmusen and Zenger 1990, 69). Thus, large firms favour fixed-wage contracts more related to tenure than performance and make extensive use of monitoring to control productivity. Smaller firms link pay and performance closely (p. 80). As a result, the larger firms have a fairly narrow spread of salaries and do not attract top talent, while smaller firms

---

[12] More than 500 employees.

employ both superior talent and low-quality individuals and reward them correspondingly. Rasmusen and Zenger's data strongly support these conclusions, especially in functions with indivisibilities in work (e.g. R&D). The closer match between performance and pay in the small firm puts the large firm at a disadvantage, in line with Williamson's incentive limits as a source of diseconomies of scale. Olson (1982, 31) made a similar observation: "in the absence of selective incentives, the incentives for group action diminishes as group size increases."

It has often been noted that R&D productivity is significantly lower in large firms than in smaller firms. Originally, Cooper surprised many business leaders and academics in 1964 with his article *R&D is more efficient in small companies*. He argued, based on 25 interviews, that small companies have three to ten times higher productivity in development than large companies. The key reasons were: 1) Small companies are able to hire better people since they can offer better (more tailored) incentives; 2) Engineers in small companies have a better attitude towards cost; and 3) The internal communication and co-ordination is more effective in small companies. These reasons match three of Williamson's four sources of diseconomies: communication distortions, atmospheric consequences, and incentive limits.

Later work has confirmed Cooper's anecdotal evidence both theoretically and empirically. Arrow (1983) demonstrated that large firms will invest suboptimally in development because of information loss, and that small firms will have a particular advantage in novel areas of research. Schmookler (1972) found that large firms (more than 5000 employees) trail small firms in the number of patented inventions, the percentage of patented inventions used commercially, and the number of significant inventions (p. 39). Yet, they spend more than twice the resources per patent (p. 37). Schmookler found four reasons for the higher effectiveness and efficiency of small firms in R&D: a better understanding of the problem to be solved, greater cost consciousness, a more hospitable atmosphere for creative contributions, and superior quality of technical personnel (p. 45). Thus, Schmookler confirmed and quantified Cooper's initial evidence. Zenger (1989; 1994) studied employment contracts in R&D in high technology. He found that organisational diseconomies of scale overwhelm technological economies of scale in R&D. His statistical analysis of Silicon Valley companies showed that small firms attract better talent than large firms, that they induce more effort from the employees, and that compensation is more tied to performance (p. 725).

Finally, the leading anti-bigness ideologues make similar observations based on anecdotes. Peters (1992) supported the notion that R&D is less effective in large organisations. He argues that large companies are massively overstaffed in development and that there is little correlation between size of R&D budget and output. He offers several case examples as evidence. Brock (1987) argued that bigness retards technological advance since large companies are overly risk averse.

Tom Peters, who since the early 80s has crusaded against big business, has put forward his own, experience-based, view on the diseconomies of scale in several books and articles. His views were summarised in *Rethinking scale* (1992). Peters believes that decentralisation is necessary in large companies, and that they are far from as decentralised as they can be. Without decentralisation they will not be adaptable enough to respond to changes in the marketplace: “If big is so damn good, then why is almost everyone big working overtime to emulate small?” (p. 13). Moreover, Peters argued that any company is well advised to reduce vertical integration although he does not offer evidence for why this is true. Overall, Peters found that successful firms need to mimic the market as much as possible, while the classical firm creates bureaucratic distortions that will lead to lower profitability and growth. These ideas are in line with Williamson’s description of firm limits, except the notion that companies should always reduce vertical integration.

Schumacher (1989, 245) identified the lack of motivation in large organisations as the key disadvantage of size: “for a large organisation, with its bureaucracies, its remote and impersonal controls, its many abstract rules and regulations, and above all the relative incomprehensibility that stems from its very size, motivation is the central problem.”

5.1.2 Reconciliation with the “limits of firm size” model

The above observations on diseconomies of scale do not map perfectly to Williamson’s four sources of diseconomies. Some are akin to his sources, others to his outcomes. By using a methodology similar to the one used in section 4.1.2, where sources and outcomes are linked, it is possible to match the observations to Williamson’s sources of diseconomies of scale. A question is if rigidity (and / or organisational age) should be introduced as a fifth source of scale diseconomies. Here it is classified as most closely associated with atmospheric consequences and communication distortions.

SOURCES OF LIMITS OF FIRM SIZE				
Communication distortion	**Bureaucratic insularity**	**Atmospheric consequences**	**Incentive limits**	**Other**
Arrow (1974): Specialisation leads to poor communication Arrow (1983a): Information loss in R&D Cooper (1964): R&D co-ordination Geanakoplos and Milgrom (1991): Information signal delays McAfee and McMillan (1995): Lower efficiency Simon ([1947] 1976): Processing bottlenecks	Blau and Meyer (1987): Excessive rigidity Brock (1987): Risk aversion Crozier (1964): Rigidity Jensen (1986): Firms larger than optimum Monsen and Downs (1965): Different owner/manager objectives Olson (1982): Rigidity Pondy (1969): Increase in administration Schmookler (1972): Understanding market needs in R&D Stinchcombe (1965): Perpetuation of organisation form	Arrow (1974): Rigidity to change Blau and Meyer (1987): Excessive rigidity Cooper (1964): R&D cost control Crozier (1964): Alienation Kwoka (1980): Low job satisfaction in large firms Qian (1994): Monitoring costs/inadequate effort levels Scherer (1976): Low job satisfaction in large firms Schmookler (1972): R&D cost consciousness; Climate for innovation Schumacher (1989): Motivation	Blau and Meyer (1987): Excessive rigidity Cooper (1964): R&D incentives Crozier (1964): Rigidity Peters (1992): Low productivity in R&D Rasmusen and Zenger (1990): Employment contracts Schmookler (1972): Quality of R&D employees Silver and Auster (1969): Limits to entrepreneurship Zenger (1989, 1994): Employment contract disincentives in R&D	Brown, Hamilton, and Medoff (1990): Unexplained wage differential

5.2 MITIGATING INFLUENCES

The review of literature relating to Williamson's offsetting mechanisms show that they exist, and that their influence varies by type of industry.

5.2.1 Asset specificity

There is an extensive literature on vertical and lateral integration based on TCE and other theories. Indeed, vertical integration has been called the paradigm problem of TCE. Mahoney (1989; 1992) and Shelanski and Klein (1995) provide summaries. Two issues are relevant here:

- Do asset specificity, uncertainty, and frequency explain vertical integration?
- Does Williamson's model extend to integration in general?

Asset specificity has repeatedly been found to be the most important determinant of vertical integration. A number of empirical studies confirm this (e.g. Masten 1984; Masten, Meehan, and Snyder 1989, 1991; Monteverde and Teece 1982; Joskow, 1993; Klier, 1993; Krickx 1988). Uncertainty and frequency are less important. First, they only contribute to vertical integration in conjunction with asset specificity. Second, the empirical evidence shows only weak explanatory power in regression analyses. Walker and Weber's (1984; 1987) results are typical. They found that volume uncertainty has some impact and that technological uncertainty has no impact on vertical integration. Frequency of transaction has unfortunately not been studied explicitly, perhaps because it is not independent from the various types of asset specificity. Piecemeal evidence from other studies suggests that it is less important than uncertainty when asset specificity is included in the analysis (e.g. Mahoney, 1992, 571).

The answer to the second question appears to be yes. Asset specificity influences integration from a reach, breadth, and depth point of view. Teece (1976) showed that the multinational company would not exist if it were not for the moral hazard resulting from the combination of asset specificity and opportunism. Without, e.g., human asset specificity a firm can just as easily license its technology to a firm in another country and reap the benefits of development. Tsokhas (1986) illustrated this in a case study of the Australian mining industry. Other studies have shown that market diversity (just as product diversity below) reduce profitability (e.g. Bane and Neubauer 1981). Thus, there is support for Coase's view from 1932[13] that the distinction between vertical and lateral integration is without value (1993c, 40).

A number of studies of product breadth show that asset specificity plays a major role in explaining the success and failure of diversification. Rumelt (1974) found a strong correlation between profitability and if a company draws on common core skills or resources (i.e. human asset specificity). In two studies of the Fortune 500 he showed that focused companies will have three to four

[13] Letter to Ronald Fowler 24 March 1932.

percentage points higher return on capital than highly diversified firms. Subsequent studies “have merely extended or marginally modified Rumelt’s (1974) original findings.” (Ramanujam and Varadarajan 1989).

The conclusion is that asset specificity plays a major role in explaining integration in general, not only vertical integration.

5.2.2 Organisational form

Alfred Chandler has, in a series of studies (Chandler 1962; 1977; 1982; 1990; 1992; Chandler and Daems 1980), shown that large corporations have evolved from functional structures to multidivisional structures as they grow in size and scope of activities. He argues that the functional form is not able to achieve the co-ordination necessary to be successful in the marketplace, while functional scale economies are too small to make up for this deficiency. Thus, as companies became more diverse they adapt the multidivisional form pioneered by du Pont and General Motors.

Fligstein (1985) showed that the multidivisional form’s penetration increased from 2 percent of large companies[14], to 75 percent between 1919 and 1979. He

[14] The 131 (120) largest manufacturing companies by assets in 1919 (1979).

estimated that the spread of the multidivisional form is mainly due to the increase of multi-product strategies, in line with Chandler's argument. Armour and Teece (1978) quantified the difference in profits between functional and multidivisional form companies in the petrochemical sector and summarised: "We find strong support for the M-form hypothesis. In the 1955-1968 period the multidivisional (M-form) structure significantly influenced (at better than the 99-percent level) the rate of return on stockholders' equity, raising it on average by about two percentage points...realized by the average functional form firm" (pp. 116-117).

Teece (1981) studied 18 manufacturing industries and two retail industries. He found that the multidivisional form outperformed the functional form by 2.37 percentage points on average (p. 188). He concluded: "the M-form innovation has been shown to display a statistically significant impact on firm performance" (p. 190). He thus supports Williamson's view that organisational structure matters and can alleviate scale diseconomies.

5.2.3 Financial synergies

A potential third offset discussed by Williamson (1986) is that large companies have efficient internal capital markets and thus they realise financial synergies. Bhidé (1990) refuted this line of reasoning and showed

that the improvement in efficiency of external capital markets since the 1960s help explain the trend away from diversification: "Investor power, which goes along with capital market sophistication, has reduced the ability of managers to preserve an inefficient organizational form." Comment and Jarrell (1995) reached the same conclusion based on an exhaustive statistical analysis.

5.2.4 Reconciliation with the "limits of firm size" model

Below is a summary of the support found in the literature for the offsetting factors.

OFFSETS TO LIMITS OF FIRM SIZE		
Asset specificity	**Organisation form**	**Other**
Monteverde and Teece (1982) et al.: Asset specificity strong influence Peters (1992): Vertical integration is bad Rumelt (1974): Product diversity Teece (1976): Asset specificity influences geographic reach Walker and Weber (1984, 1987): Volume uncertainty weak factor Ward (1976): Market diversity	Armour and Teece (1978): M-form increases ROE Chandler (e.g. 1962): M-form alleviates co-ordination and control problems Fligstein (1985): multi-product co-ordination favours M-form Peters (1992): Decentralisation is critical	Bide (1990): Internal capital markets not efficient Comment and Jarrell (1995): Financial synergies not relevant

5.3 NEO-CLASSICAL SCALE ECONOMIES IN THE PRODUCTION FUNCTION

Neo-classical scale (or scope) economies should not be incorporated into the model since they are independent of the form of organisation beyond the point where technological indivisibilities are captured within the firm, according to transaction cost theory. That is, the scale economies will be reaped regardless of if all production is carried out in one firm or in many firms (Masten 1982; North and Wallis 1994; Riordan and Williamson 1985). Thus, the intuitively appealing notion that the existence of scale economies offsets size disadvantages is, according to theory, incorrect.

This proposition has not been tested directly. However, since the 1950s there has been extensive research into the nature and magnitude of scale economies in production costs, much of it emanating from the “structure-conduct-performance” school of thought. This work has been explicated in a number of books and there is no reason to repeat the arguments here, except as a brief summary. In general, the research shows that scale economies do not play a major role in explaining firm size, but they are not insignificant either.

Joe Bain pioneered the research in the 1950s and subsequently revolutionised the study of industry and company behaviour with his book *Industrial organization* (1968). Relevant to this discussion is chapter 6 (*The rationale of concentration—Efficiency and other considerations*) which reviews the scale economies argument. Bain divided the analysis into plant and firm level analyses. At the plant level, scale economies are exploited by specialising the work force and management, and by using dedicated machinery. For each plant there is a minimum optimal scale. Beyond this scale there are few additional scale economies to be exploited. Bain found that in a study of 20 industries, only two showed significant scale economies: “in a preponderance of cases, plant scale curves tend to be at least moderately flat (and sometimes very flat)...in the bulk of cases, then, the relative flatness of plant scale curves virtually diminishes the importance of plant scale economies" (pp. 192-193). He found scant evidence at the plant level for benefits of firm size.

At the firm level, Bain's study showed that scale economies are derived from large-scale management, large-scale distribution, and purchasing power[15]. He then noted that these firm level scale economies are elusive, if they exist at all. His research indicated that "where economies of the multi-plant firm are encountered, they are ordinarily quite slight in magnitude...the unit costs...are typically only 1 or 2 per cent below those of a firm with one plant of minimum optimal scale" (p. 195). Of the 20 industries studied, Bain was able to quantify firm level scale economies for twelve. Of these twelve industries, none exhibited even moderate scale effects (p. 195).

Bain (1978) later summarised his argument that scale economies do not explain firm size: "It is not true that existing degrees of concentration are adequately explained simply as the result of adjustments to attain maximum efficiency in production and distribution...Industries probably tend to be "more than

---

[15] Bain does not mention R&D and marketing, possibly because these factors were less important in the early 1950s.

concentrated than necessary" for efficiency—and the larger firms bigger than necessary" (p. 94).

Scherer and Ross (1990) gave a modern overview of the scale economies debate in chapter 4 of their book. They made the point that it is difficult to draw simple conclusions about the relation between size and returns. In general they found that firm scale economies in production costs are exhausted at a surprisingly small firm[16] size. In a study of twelve industries they found that market concentration could not be explained by minimum efficient scale considerations. The largest companies in the twelve industries were between two and ten times larger than scale economies necessitated. Scherer and Ross argued that to the extent there are scale economies for large companies in an industry, they derive from economies in overhead costs, fixed costs in tangible assets, R&D, and marketing.

A number of theoretical studies (Ijiri and Simon 1964; Lucas 1982; Nelson and Winter 1982; Simon and Bonini 1958) have demonstrated that large firms will evolve, regardless of scale economies, for the simple reason that there will be winners and losers over time. The losers will disappear and the winners will grow at differential rates depending on the length of win periods. Based on this

[16] They make the same argument at the product and plant level.

logic, firms are large because they are winners, not because they realise scale economies. With realistic assumptions about industry growth rates, variance in firm profitability, etc., simulations have created firm size distributions similar to observed real life distributions. As Ijiri and Simon (p. 78) put it: “the observed distributions are radically different from those we would expect from explanations based on static cost curves...there appear to be no existing models other than the stochastic ones that make specific predictions of the shapes of the distribution.”

An empirical test of the stochastic evolution model was done by Rumelt and Wensley (1981) who tested if high market share led to high profitability, or if successful companies, with high profitability, in turn achieve high market share. They concluded that “scale economies and / or market power are much less important than stochastic growth processes” (p. 2).

It should also be noted that the stochastic growth process argument also leads to the prediction that older companies should be more profitable than younger companies. The reason is that older companies that still exist are survivors (cf. p. 7) while younger companies still include both winners and losers.

Finally, Peters (1992) argued that scale economies do not exist any more (if they ever existed): “technology and brainware’s dominance is taking the scale out of everything” (p. 14). Adams and Brock (1986), in case studies of the steel industry, automotive industry and conglomerates, found no evidence that size leads to production scale economies at the firm level. They claimed that it is “the quintessential myth of America’s corporate culture that industrial giantism is the handmaiden of economic efficiency” (p. xiii).

These studies assumed that scale has to be achieved within the firm, but did not find significant scale effects under this assumption. While they do not confirm the transaction cost argument that scale economies are independent of governance, they lend credence to the idea.

At this point two additional hypotheses are introduced:

H_5: Scale economies are weakly related to size.

H_6: Scale economies have no impact on individual company profitability.

5.4 COMMENT ON INDUSTRY INFLUENCE

A number of studies have shown that there is weak correlation between profitability and industry within the manufacturing sector. Schmalensee (1985) suggested methods for disaggregating business unit performance into industry, corporate parent, and market share effects. Rumelt (1981) applied the methodology to manufacturing firms and found that industry effects accounted for 8 percent of explained profitability (63 percent of total profits). McGrahan and Porter (1997) found a 19 percent industry effect for all sectors of the economy and a similar effect as Rumelt (9 percent of explained profitability) for firms in the manufacturing sector (p. 25). Thus, industry appears to influence profitability significantly in the non-manufacturing sector, but only slightly in the manufacturing sector.

5.5 CONCLUSION

The literature review indicates that the TCE model of limits of firm size is fairly robust. All the sources reviewed fit within Williamson's implicit model, and there does not seem to be any reason to change or complement it. The offsets are also validated and asset specificity emerges as the most important driver of both vertical and lateral integration. It may be argued that the "winner" condition should be included among the offsets. The argument is that large firms, especially the ones that are growing, are better managed and will thus generate returns despite the diseconomies of scale. The treatment here though is to leave it as an exogenous category since it does not fit into the TCE logic, except possibly as an illustration of the lack of production cost scale economies at the firm level. Moreover, it can be expected that the winners succeed precisely because they have offset the diseconomies of scale.

The literature did show that the sources of diseconomies are more important in certain contexts. Atmospheric consequences and incentive limits are especially severe in R&D intensive industries. Also, communication distortions are most common in diverse companies and in volatile industries.

The verification also allowed a first cut assessment of the importance of effects and at what size of company the effects have an impact. The importance of effects is a qualitative assessment of the survey samples' collective judgement on each source of diseconomies. The "size impact" parameter roughly indicates at what size (number of employees) the effect sets in. For example, the incentive advantage in R&D for small firms appears to be strong for firms with less than 500 employees according to the literature. Large and medium sized companies do not seem to differ.

The table below extends, but does not change, the summary in section 4.2 by adding estimates of the importance of each factor, the firm size at which the factor impacts profitability, and in which context the factors are more important. The observations are the researcher's interpretation of the literature review.

EXTENDED TCE-BASED "LIMITS OF FIRM SIZE" MODEL						
Value	**Sources of limits of firm size**				**Mitigators**	
	Communication distortion	**Bureaucratic insularity**	**Atmospheric consequences**	**Incentive limits**	**Integration conditions**	**Organisational form**
High	Low	Low	Low	Low	High	M-form
Low	High	High	High	High	Low	U-form
Importance	Strong	Moderate	Moderate	Moderate in general; Strong in R&D	Asset specificity strong; Uncertainty weak; Frequency negligible	Strong
Impact size: **Small (<1000)** **Medium** **Large (>10,000)**	 Strong Strong Strong	 Weak Moderate Strong	 Weak Moderate Strong	 Strong Weak Weak	 Strong Strong Strong	 Weak Moderate Strong
Context	Diverse firms; Unpredictability	Management/ board relation	R&D intensive	R&D intensive		Diverse firms

The table guides the description of sub-hypotheses for hypothesis 3: Scale diseconomies impact company profitability and growth negatively and significantly.

H_{3a}: Communication distortions have a negative impact on large company profitability and growth.

H_{3b}: Bureaucratic insularity has a negative impact on the results of large companies.

H_{3c}: Atmospheric consequences lead to relatively poorer results for large companies.

H_{3d}: Incentive limits exhibit a negative influence on the results of large companies.

Sub-hypotheses for hypotheses 4 (diseconomies of scale can be mitigated by two transaction cost-related factors: organisational form and asset specificity) are:

H_{4a}: Large M-form firms show better results than U-form firms, but this has been fully implemented in today's large companies.

H_{4b}: High asset specificity is positively correlated with results.

The model and sub-hypotheses can be used to test if the TCE explanation of limits of firm size is valid. The literature survey shows that the sources of diseconomies and the offsets are relevant. The key question is if the effects are large enough to make a difference. Only an empirical analysis where the model is operationalised can answer this.

6. APPROACH TO QUANTITATIVE ANALYSIS

This chapter gives a general impression of the analytical approach and the method chosen for the statistical analysis.

6.1 RESEARCH METHODOLOGY

The approach is positivist in nature and aims for universal understanding in the sense of Runkel and McGrath (1972). The expectation is to find general conclusions, while precision and realism are somewhat reduced.

It is possible to approach the issue with a phenomenological approach akin to what Cooper (1964) did in his often quoted study of R&D productivity in large and small firms. Such an approach would most likely be based on case studies. However, the positivist approach was preferred for a number of reasons (Easterby-Smith, Thorpe, and Lowe 1991, 23): A positivist approach allows for more independence from the observations and since individual or group behaviour are not the concern of this research, little additional insight can be gained from action research. Value-freedom is important since the limits of firm size studied are themselves value-laden. Causality can be deduced from the proposed data set and manipulation, and concepts can be operationalised to suit a positivist approach. The problem lends itself to reductionism since the factors are possible to disaggregated. Moreover, as said earlier, it should be possible to draw generalisable conclusions based on the fairly large sample suggested later in this chapter. Finally, cross-industry comparisons will be important and it is easier to do these with a positivist approach. In short, the positivist approach appears to fit the research objective well.

This choice leads to a few success factors (p. 27): The work should focus on the facts and thus it will be important to be careful with the data set. The emphasis should be on looking for causality rather than meaning. The hypotheses should be formulated before the quantitative research rather deduced from the data. The sample should be large and concepts should be operationalised so that they can be measured.

There are no studies of the general type on the particular issue of scale diseconomies. However, generalised studies on for example the profit impact of M-form organisation, and the linkage between size, structure, and complexity are widely quoted in the literature. This indicates that the generalised approach may add substantial value to the study of limits of firm size. An added benefit is that data is widely available to support a generalised study.

On the other hand, there are already several studies aimed at precisely describing particular aspects of limits of firm size, as was shown in chapter 5. Zenger's (1989) study of incentive limits in Silicon Valley is a good example. There are also several case studies that achieve realism, but in the end these studies have had only limited impact on academic thinking. The notable exceptions are in the work on institutions in society based on TCE, where North (1985; 1987; 1992; 1973) was able to merge insights from case studies with a framework for institutional change. Chandler's (1962; 1977; 1990) work on the evolution of large companies has also had major impact.

6.2 STATISTICAL TECHNIQUE

In this section the statistical technique is chosen based on empirical precedents, the nature of the task, and the specific statistical analysis software available.

6.2.1 Empirical precedents

As discussed earlier, general statistical analyses of diseconomies of scale have not been attempted before, except for simple direct comparisons between company size and performance. There are however, a number of empirical studies of particular aspects of the limits of firm size or of general TCE problems. These help inform the choice of statistical methods and the operationalisation of variables. Generally, there are two groups of empirical, quantitative analysis: older, non-regression analyses, and newer, regression-based analyses. The two sections below discuss the statistical approaches taken in a number of these studies. The intent is not to discuss findings but rather to inform the choice of methodology and indicators for this particular research.

Early inquiries (McConnell 1945; Stigler 1958) into relationships between profitability and size made simple, histogram-type, comparisons between the dependent (profitability) and independent variable histogram. These analyses at best indicated relationships but since they did not include modern tests of statistical significance they left many questions unanswered.

Later, researchers including Rumelt (1974), Teece (1981), and Palepu (1985) used comparisons of samples to demonstrate statistical significance. Their methods ranged from simple comparisons of average profitability for two samples, to sophisticated tests of the statistical significance differences. However, these techniques probably did not extract the full information content of the samples. For example, Rumelt later used a regression technique when he updated his analysis of diversification and structure (Rumelt 1982).

Yet another approach to the statistical analysis is to study the amount of variance extracted. Schmalensee (1985) used analysis of variance (ANOVA) and argued that: ”This study employs a simple analysis of variance framework that allows us to focus on the existence and importance of firm, market, and market share effects without having to deal simultaneously with specific hypotheses and measurement issues related to their determinants.” Put differently, analysis of variance is an excellent tool for exploratory analysis. Other examples of ANOVA include Rumelt (1991) and McGahan and Porter (1997).

Finally, most statistical analyses over the last 35 years relating to aspects of diseconomies of scale used multivariate regression techniques (Armour and Teece 1978; Aw and Batra 1998; Child 1973; Comment and Jarrell 1995; Fligstein 1985; Kwoka 1980; Levy 1981; Lucas 1982; Mahoney 1989; Masten 1984; Masten, Meehan, and Snyder 1989; Masten, Meehan, and Snyder 1991; Olson 1982; Pondy 1969; Pugh et al. 1969; Rasmusen and Zenger 1990; Rumelt 1982; Rumelt and Wensley 1981; Walker and Weber 1984; Walker and Weber 1987; Zenger 1994). They ranged from simple regressions to complex structural equation models and were used for both exploratory and confirmatory analysis.

This quick overview indicates a trend over time towards more sophisticated statistical techniques, which usually leads to the use of multivariate techniques.

Stigler’s survivor principle (p. 7) indicates that the added complexity of using these techniques is more than compensated by the added insights they bring, otherwise they would not survive. On the other hand, the simpler techniques have more or less disappeared from the literature.

6.2.2 Selection of statistical technique

The decision was made to use multivariate analysis. With today’s computer technology, no other choice is relevant. The review in the previous section shows that other techniques were used earlier primarily because of lack of computing power. Already in 1966 it was argued that “for the purposes of…any…applied field, most of our tools are, or should be, multivariate. One is pushed to a conclusion that unless a…problem is treated as a multivariate problem, it is treated superficially” (Gatty 1966, 158).Or as Hardyck and Petrinovich put it ten years later: “Multivariate analysis methods will predominate in the future…” (p. 7).

Among different multivariate techniques, structural equation modeling (SEM) was picked based on Hair’s classification scheme for choosing among techniques (1998, 20-21) and a review of the pertinent literature on SEM (Bollen 1989, 1-9; Kelloway 1998, 2-3; Maruyama 1998, 20-24). SEM is the most appropriate technique when multiple relationships between dependent and independent variables are studied, which is the case here. Moreover, SEM is well suited for confirmatory analysis and allows for efficient hypothesis testing, especially of complex models. Finally, SEM allows for the use of latent, unobserved, variables. As is seen in chapters 7 and 8, these three conditions are true in this research.

SEM must be used judiciously however. There are a number of criticisms raised against SEM, as summarised by Maruyama (pp. 272-278). SEM can not be used to confirm a model, only fail to disconfirm it. This makes replication important and this is one of the main reasons why publicly available data was preferred in the statistical analysis in chapter 8. Related to this is the risk of inferring causality where none exists. Strong correlation does not imply causality. For this reason, the path diagrams used in the analysis chapter have causalities based on theory, not on a study of correlation. A further critique is that that incorrect operationalisation and naming of variables (especially latent variables) can lead to erroneous conclusions: "giving something a name does not necessarily make it what we call it or ensure that we understand the thing we have named" (p. 273). This is certainly a valid point for the present work. For example, incentive limits were operationalised, based on the theory, with R&D expense. Does this mean that the relationships found in the SEM models truly represent incentive limits, or do they merely represent R&D expense? Finally, SEM has often been used for model development rather than model confirmation. This has hopefully been avoided here since the path diagrams have been drawn directly from the theory and no model modifications have been made by studying modification indices or critical ratios for differences.

6.2.3 Amos versus LISREL

A choice was made to use Amos[17] instead of the more recognised LISREL SEM software package. LISREL certainly has a much larger installed base and thousands of references in the literature. Amos is less well known and counts 55 references up till 2000. Amos is, however, much easier to use for the occasional statistician since it is based on interactive, graphical path diagrams, and it is gaining ground in research. More importantly, Amos has one attribute that serves this particular research effort well.

As is discussed in chapter 7, the data collected has many missing values. LISREL uses pairwise or listwise deletion to handle missing data. Amos, on the other hand, uses full-information maximum likelihood (FIML) estimation, arguably a leading edge technique. When the data are missing completely at random, then the listwise or pairwise deletion methods used by LISREL are asymptotically

[17] Information on Amos is available at smallwaters.com

equivalent to FIML, but the standard errors of their parameter estimates can be considerably larger. This means that Amos makes more efficient use of the observations, a critical consideration here since there is only a limited number of large companies and thus the sample size cannot be expanded indefinitely. Moreover, if data is not missing at random then the estimates should be less biased than when using pairwise or listwise deletion. It should be noted that FIML does not impute missing values. Instead, Amos calculates the likelihood of the parameters based on each observed case.

The major advantage of LISREL is that it handles ordinal data correctly with polychoric and polyserial correlations. As is discussed in chapter 7, two variables are ordinal in this research. Amos does not have any way of handling ordinals except for treating them as continuous variables., leading to biased estimates. This problem can be reduced by using bootstrapping. On the other hand, it has been argued that effective use of polychoric and polyserial correlations requires 2000-5000 observations, well beyond what is available here. Thus, there are two less than optimal approaches: either use Amos and treat ordinal data as continuous data, or use LISREL with too small sample size. A choice was made to use Amos since it has other positive attributes, but no attempt was made to compare which software package would handle the ordinal data best or if this makes a difference at all.

Finally, by choosing Amos and having missing data, two more choices have been made automatically. Alternative estimation techniques such as generalised least squares and asymptotically distribution-free (ADF) estimation can not be used since they do not allow missing data. ADF could have been useful since it handles non-normality well. Also, Amos does not allow bootstrapping when data is missing. Bootstrapping could have increased the reliability of the analysis since it reduces the impact of non-normality.

7. DATA

7.1 DEFINITIONS AND SOURCES

Data has been collected for public U.S.-headquartered industrial companies (SIC codes 10-39) with sales of more than $500 million and with 1998 as the benchmark year. For these 1998 companies, most of the data has also been collected for 1996, 1995, and 1991, with 1998 and 1995 mild business cycle peak years and 1996 (mild) and 1991 (severe) business cycle troughs (Zarnowitz 1998). In addition, 1998 data was collected for all public U.S.-headquartered companies (i.e. all SIC codes) with sales of more than $500 million. The 1998 industrial companies information is the core data while the other data was used to test longitudinal and cross-sectional extensions of the core analysis. Non-core data is not discussed below but is available together with core data at canback.com \ henley \ data.

In total, 72 variables were collected for 784 companies from Compustat, annual reports, 10K and proxy statements, Business Week, Fortune, and other sources. Of these variables, 30 were ultimately used in the analysis either directly or as parts of calculated variables, leading to 22 variables which were used as observed variables or indicators in the analyses. The union of these variables (34) are described in the table below. Specific information about the variables is given under each analysis section in chapter 8.

OVERVIEW OF VARIABLES					
Name	**Type**	**Description**	**Metric**	**Source**	**Use**
Sales	Direct	Sales	$M	Compustat	Size
Empl	Direct	No. of employees	‘000	Compustat	Size
NA	Calculated	Net assets defined as total assets less current assets plus current liabilities	$M	Compustat	Size
VA	Calculated	Value added defined Sales less cost of purchased goods and services			Size
FA	Direct	Fixed assets	$M	Compustat	Economies of scale
SG&A%	Calculated	SG&A / Sales	%		Economies of scale
SG&A	Direct	Sales, general and administrative expense	$M	Compustat	
Adv	Direct	Advertising expense	$M	Compustat	
Adv%	Calculated	Adv / Sales	%		Performance
R&D	Direct	Research and development expense	$M	Compustat	
R&D%	Calculated	R&D / Sales	%		Performance Diseconomies of scale Economies of scale
Labour	Direct	Total labour cost	$M	Compustat	
ULabour	Calculated	Labour / Empl	$k		Atmospheric consequences
Inputs	Calculated	Factor inputs defined as Sales – NI + COE	$M		
NI	Direct	Net income	$M	Compustat	
ROE	Calculated	NI / Equity	%		
EVA	Calculated	Economic value added defined as ROE – COE	%		Performance
Growth	Direct	5 year compound annual growth in revenue	%	Compustat	
FCF	Direct	Free cash flow	$M	Compustat	
FCF%	Calculated	FCF / Sales	%		Performance
MEV/BEV	Calculated	Market entity value over book equity value defined as (MV + long-term debt) / (BV + long-term debt)	Ratio		Performance
MV	Direct	Market value of equity	$M	Compustat	Performance
BV	Direct	Book value of equity	$M	Compustat	Performance

Foreign%	Direct	Share of revenue from outside the U.S.	%	Compustat	Asset specificity
SR%	Calculated	Specialisation ratio defined as SC% = (Sales of largest business unit) / Sales	%	Compustat	Asset specificity
Return	Direct	5 year total return on investment for shareholder	%	Compustat	Performance
Levels	Direct	Communication distortion defined as number of hierarchical levels	#	Annual reports, web sites, 10Ks	Diseconomies of scale
Govern	Direct		Scale 0 – 400	Business Week, IRRC, and Fortune	Organisation
Division	Direct	Degree of divisionalisation (ordinal)	2 = divi-sionalised; 1 = confu-sed; 0 = unitary	Annual reports, web sites, 10Ks	Organisation
Age	Direct	Years since company founding	#	Annual reports, web sites, 10Ks	Diseconomies of scale
Tenure	Direct	Average tenure of officers in years	#	Annual reports, web sites, 10Ks, proxy statements	Diseconomies of scale
Beta	Direct	Beta of company stock	Ratio	Compustat	
COE	Calculated	Cost of equity defined as COE = risk free rate + risk premium (0.371 + 0.635*Beta)	%	Compustat, Ibbotson	
Vertical	Direct	Vertical integration (ordinal)	2 = Very high; 1 = High; 0 = Normal	Annual reports, web sites, 10Ks	Asset specificity

7.2 SCREENING FOR OUTLIERS AND NONSENSICAL DATA

The original Compustat database contained 901 company records. Of these, 90 were eliminated because they were either based outside the United States, they did not contain revenue figures, or they became defunct during or after 1998. 13 records were eliminated since they contained so-called Pre-FASB data. The 13 companies with pre-FASB data also have regular records which are kept in the database. Four companies with revenues less than $500M were eliminated. Six limited partnerships were eliminated since their records contained too few observations to be of interest. Six companies had duplicate records. One company was recently spun out from its corporate parent and did not have sensical data. Thus, the final database contains 784 records.

For these 784 company records, each variable was first screened for nonsensical data, then outliers were eliminated if a rationale for elimination was found. Examples of nonsensical data are: negative equity; zero foreign sales of well-known international companies; negative beta; and negative market value.

Outliers were identified by standardising the variables and then identifying those records which were more than three standard deviations from the mean, in line with Hair et al (Hair et al. 1998, 65). Those outliners for which the data made sense were kept, while nonsensical data was eliminated. All changes described above to the original database described commented in the master spreadsheet MASTER DATA.XLS at canback.com \ henley \ data.

Missing values are summarized below with an excerpt from the output file from PRELIS, the data screening utility included in LISREL. The effective sample sizes are:

Effective Sample Sizes
Univariate (in Diagonal) and Pairwise Bivariate (off Diagonal)

	Empl	VA	NA	Sales	Adv%	R&D%
Empl	784					
VA	92	92				
NA	744	89	744			
Sales	784	92	744	784		
Adv%	177	21	172	177	177	
R&D%	489	65	478	489	118	489
ULabour	95	92	92	95	21	66
VI%	92	92	89	92	21	65
Fix%	753	92	721	753	170	473
EVA	782	92	744	782	177	489
MEV/BEV	678	86	674	678	162	438
Growth	759	92	719	759	169	474
FCF%	764	92	724	764	170	475
Foreign%	482	61	467	482	121	350
Segments	737	89	701	737	173	462
Return	610	85	579	610	138	384
Levels	271	35	254	271	66	183
Tenure	70	16	67	70	23	52
Age	624	86	588	624	158	413
U	265	34	248	265	65	179
Govern	229	44	215	229	78	162

	ULabour	VI%	Fix%	EVA	MEV/BEV	Growth
ULabour	95					
VI%	92	92				
Fix%	92	92	753			
EVA	95	92	753	782		
MEV/BEV	87	86	662	678	678	
Growth	95	92	729	757	659	759
FCF%	95	92	734	762	663	759
Foreign%	62	61	468	482	459	469
Segments	91	89	713	737	673	715
Return	86	85	590	610	580	609
Levels	35	35	262	270	249	265
Tenure	16	16	66	69	65	70
Age	88	86	600	623	560	603
U	34	34	257	264	243	259
Govern	44	44	224	229	209	228

	FCF%	Foreign%	Segments	Return	Levels	Tenure
FCF%	764					
Foreign%	470	482				
Segments	719	482	737			
Return	610	410	607	610		
Levels	266	177	268	237	271	
Tenure	70	49	67	62	52	70
Age	606	400	603	510	232	67
U	260	174	262	232	264	50
Govern	228	169	223	208	108	33

	Age	U	Govern
Age	624		
U	226	265	
Govern	209	108	229

Percentage of Missing Values
Univariate (in Diagonal) and Pairwise Bivariate (off Diagonal)

	Empl	VA	NA	Sales	Adv%	R&D%
Empl	0.00					
VA	88.27	88.27				
NA	5.10	88.65	5.10			
Sales	0.00	88.27	5.10	0.00		
Adv%	77.42	97.32	78.06	77.42	77.42	
R&D%	37.63	91.71	39.03	37.63	84.95	37.63
ULabour	87.88	88.27	88.27	87.88	97.32	91.58
VI%	88.27	88.27	88.65	88.27	97.32	91.71
Fix%	3.95	88.27	8.04	3.95	78.32	39.67
EVA	0.26	88.27	5.10	0.26	77.42	37.63
MEV/BEV	13.52	89.03	14.03	13.52	79.34	44.13
Growth	3.19	88.27	8.29	3.19	78.44	39.54
FCF%	2.55	88.27	7.65	2.55	78.32	39.41
Foreign%	38.52	92.22	40.43	38.52	84.57	55.36
Segments	5.99	88.65	10.59	5.99	77.93	41.07
Return	22.19	89.16	26.15	22.19	82.40	51.02
Levels	65.43	95.54	67.60	65.43	91.58	76.66
Tenure	91.07	97.96	91.45	91.07	97.07	93.37
Age	20.41	89.03	25.00	20.41	79.85	47.32
U	66.20	95.66	68.37	66.20	91.71	77.17
Govern	70.79	94.39	72.58	70.79	90.05	79.34

	ULabour	VI%	Fix%	EVA	MEV/BEV	Growth
ULabour	87.88					
VI%	88.27	88.27				
Fix%	88.27	88.27	3.95			
EVA	87.88	88.27	3.95	0.26		
MEV/BEV	88.90	89.03	15.56	13.52	13.52	
Growth	87.88	88.27	7.02	3.44	15.94	3.19

FCF%	87.88	88.27	6.38	2.81	15.43	3.19
Foreign%	92.09	92.22	40.31	38.52	41.45	40.18
Segments	88.39	88.65	9.06	5.99	14.16	8.80
Return	89.03	89.16	24.74	22.19	26.02	22.32
Levels	95.54	95.54	66.58	65.56	68.24	66.20
Tenure	97.96	97.96	91.58	91.20	91.71	91.07
Age	88.78	89.03	23.47	20.54	28.57	23.09
U	95.66	95.66	67.22	66.33	69.01	66.96
Govern	94.39	94.39	71.43	70.79	73.34	70.92

	FCF%	Foreign%	Segments	Return	Levels	Tenure
FCF%	2.55					
Foreign%	40.05	38.52				
Segments	8.29	38.52	5.99			
Return	22.19	47.70	22.58	22.19		
Levels	66.07	77.42	65.82	69.77	65.43	
Tenure	91.07	93.75	91.45	92.09	93.37	91.07
Age	22.70	48.98	23.09	34.95	70.41	91.45
U	66.84	77.81	66.58	70.41	66.33	93.62
Govern	70.92	78.44	71.56	73.47	86.22	95.79

	Age	U	Govern
Age	20.41		
U	71.17	66.20	
Govern	73.34	86.22	70.79

The table above shows that all variables except for 4 have more than 200 observations and that 13 out of 21 variables have more than 400 observations. The variables have 490 observations on average. The variables with low number of observations are **Adv%** (177), **Tenure** (>100) , and **ULabour** (95), and **VA** (92). Of these, **Adv%** and **VA** are not critical to the analysis. **ULabour** and **Tenure** are important and the fairly low number of observations reduce the integrity of the analysis, but not to a point to render them useless.

According to Hair (1998, 604-605) four criteria determine an appropriate sample size. When maximum likelihood estimation (MLE) is used, as is the case here, a minimum sample size of 100 to 150 is recommended. Sample sizes of more than 400 to 500 often become “too sensitive.” Hair recommends a sample size of 200 as starting point. In addition, the sample size should be at least five times the number of parameters estimated. The main model in this analysis has 14 variables and thus 105 parameters are estimated, leading to a requirement of at least 565 observations, not too far from the 490 observations actually available. However, even more observations are required if misspecification is suspected (which is not the case here) or if data is heavily non-normal (which is the case). In summary, the sample size appears to be reasonable in although a few variables have too few observations.

7.3 DATA TRANSFORMATION

Many continuous variables where transformed to reach more univariate normal distribution and to reduce heteroscedasticity. The method used was to study histograms of each variable and to transform those variables that deviated significantly from normal distribution. This led to logarithmic transformations of 14 variables and no transformation of 8 variables.

Logarithmic transformation	Empl, VA, NA, Sales, Adv%, R&D%, Fix%, MEV/BEV, Growth, Foreign%, ULabour, Age, Govern, FixHigh
No transformation	Return, EVA, FCF%, Levels, Tenure, SR%
Ordinal	Division, Vertical

The resulting skewness and kurtosis is reported below. In general, the variables are fairly or very skewed and exhibit significant kurtosis. This is clearly a problem, and could have been rectified by crating normal scores for the variables (with e.g. LISREL). This would reduce however make it more difficult to interpret the data and in the choice between optimising test statistics and keeping data transparent, the latter was chosen. Fortunately, the Z-scores are high, but not beyond reason, which made the choice easier to make.

Univariate Summary Statistics for Continuous Variables

Variable	Mean	St. Dev.	T-Value	Skewness	Kurtosis
Empl	2.098	1.090	53.887	0.443	0.583
VA	7.412	1.252	56.794	0.720	0.057
NA	7.242	1.232	160.397	0.758	0.578
Sales	7.500	1.043	201.315	1.177	1.257
Adv%	0.048	0.054	11.751	2.085	4.529
R&D%	0.043	0.053	17.744	2.352	6.382
ULabour	3.968	0.284	135.979	-0.439	-0.332
VI%	0.343	0.075	43.629	0.073	0.304
Fix%	0.231	0.107	59.280	0.792	0.804
EVA	-2.461	17.854	-3.855	-0.042	5.079
MEV/BEV	1.162	0.571	52.978	1.733	3.759
Growth	0.119	0.132	24.892	1.070	3.364
FCF%	0.004	0.100	1.201	-2.808	22.023
Foreign%	-1.146	0.970	-25.939	-1.069	1.422
Return	13.188	21.009	15.504	1.102	6.676
Levels	7.818	0.900	142.991	0.316	-0.155
Tenure	13.709	7.990	14.355	0.342	-1.031
Age	4.049	0.786	128.661	-1.199	1.313
Govern	5.525	0.151	555.535	0.454	0.416

Test of Univariate Normality for Continuous Variables

	Skewness		Kurtosis		Skewness and Kurtosis	
Variable	Z-Score	P-Value	Z-Score	P-Value	Chi-Square	P-Value
Empl	3.105	0.002	2.777	0.005	17.350	0.000
VA	2.849	0.004	0.444	0.657	8.317	0.016
NA	3.747	0.000	2.696	0.007	21.307	0.000
Sales	4.368	0.000	4.768	0.000	41.812	0.000
Adv%	3.732	0.000	4.998	0.000	38.902	0.000
R&D%	4.720	0.000	8.565	0.000	95.640	0.000
ULabour	-1.765	0.078	-0.489	0.625	3.354	0.187
VI%	0.287	0.774	0.895	0.371	0.884	0.643
Fix%	3.814	0.000	3.456	0.001	26.487	0.000
EVA	-0.641	0.522	9.723	0.000	94.946	0.000
MEV/BEV	4.702	0.000	8.053	0.000	86.957	0.000
Growth	4.210	0.000	8.067	0.000	82.800	0.000
FCF%	-5.465	0.000	14.541	0.000	241.310	0.000
Foreign%	-3.773	0.000	4.153	0.000	31.481	0.000
Return	4.026	0.000	9.593	0.000	108.224	0.000
Levels	2.037	0.042	-0.365	0.715	4.283	0.117
Tenure	1.186	0.236	-2.948	0.003	10.095	0.006
Age	-4.152	0.000	4.424	0.000	36.811	0.000
Govern	2.324	0.020	1.347	0.178	7.216	0.027

8. STATISTICAL ANALYSIS

This chapter describes the structural equation model analyses used to test the hypotheses. The philosophy of the approach has been to use as simple models and definitions as possible and to use the theoretical foundation without alterations. The focus is on practical significance, rather than statistical significance, as discussed by Hair (1998, 22): “Many researchers become myopic in focusing solely on the achieved significance of the results without understanding their interpretations, good or bad. A researcher must instead look not only at the statistical significance of the results but also at their practical significance.” Here, the analyses are used in a confirmatory sense in a fairly dogmatic fashion. This means that the correlations and conclusions probably are weaker than they need be in a statistical sense. An example is that the data could have been completely normalised to avoid kurtosis and skew. This was done, and transformations have not been used except for logarithmic transformations where suggested by other empirical studies. Another example is that weak relationships were not eliminated to achieve better parsimonious fit, except in the final section. This could have been done throughout the analyses, but the purpose of the statistical analyses is not to optimise test statistics, but rather to gain insights into the nature of diseconomies of scale. These insights were gained, as is hopefully demonstrated.

The chapter is sectioned by *Introductory analyses*, *Performance and results*, *Diseconomies of scale*, *Economies of scale*, *Mitigators*, and *Complete model*. Each of the hypotheses are discussed at various points in the chapter.

H_1: A company's size is not a direct predictor of its results.

H_2: Large companies embed diseconomies of scale of four types: communication distortion, bureaucratic insularity, atmospheric consequences, and incentive limits.

H_3: Scale diseconomies impact company profitability and growth negatively.

H_{3a}: Communication distortions have a negative impact on large company profitability and growth.

H_{3b}: Bureaucratic insularity has a negative impact on the results of large companies.

H_{3c}: Atmospheric consequences lead to relatively poorer results for large companies.

H_{3d}: Incentive limits exhibit a negative influence on the results of large companies.

H_4: Diseconomies of scale can be mitigated by two transaction cost-related factors: organisational form and asset specificity.

H_{4a}: Large M-form firms show better results than U-form firms, but this has been fully implemented in today's large companies.

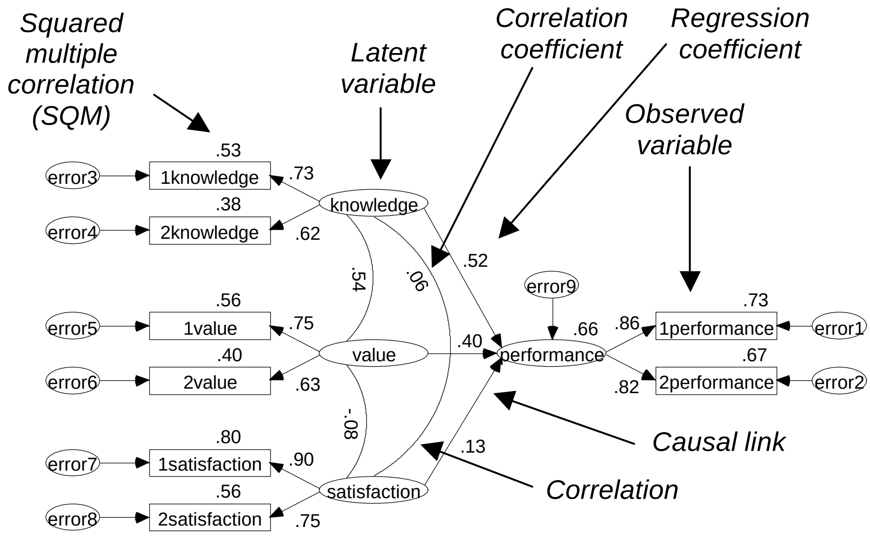
H_{4b}: High asset specificity is positively correlated with results.

H_5: Scale economies are weakly related to size.

H_6: Scale economies have no impact on individual company profitability.

8.1 INTRODUCTORY ANALYSES

Below follow a series of analyses that serve as preamble to the analyses of diseconomies of scale and of mitigating factors. The preamble shows that a superficial analysis does not reveal any material scale diseconomies. The path diagrams follow the standard SEM nomenclature (e.g. Arbuckle 1999, 135)(1999) with rectangles representing observed variables, ovals representing latent variables, curved lines representing correlations, and arrows representing causal links. The example below from Arbuckle shows the standardised nomenclature. Note that the error terms include both errors and influences from variables not included in the model.

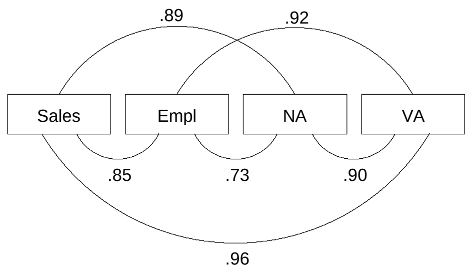


Example 5: Model A
Regression with unobserved variables
Job performance of farm managers
Warren, White and Fuller (1974)

8.1.1 Size metrics

In line with the definitions of size in the chapter 3, *Dimensions of firm size*, size was defined as revenue (**Sales**), number of employees (**Empl**), net assets (**NA**), and value added (**VA**), where correlated against each other. Unsurprisingly, the four definitions are highly correlated.

CORRELATION BETWEEN DIFFERENT DEFINITIONS OF SIZE
Sales, employees, net assets, and value added

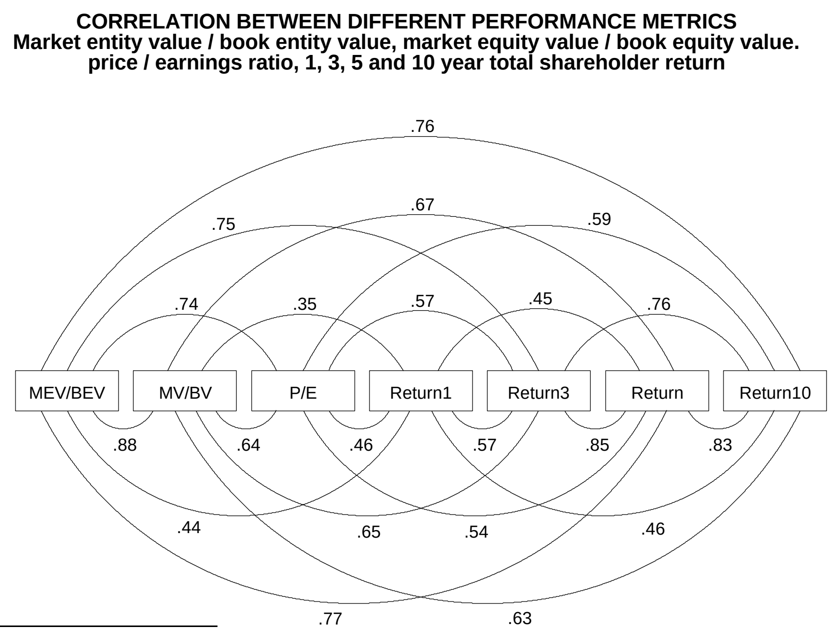


Based on the high level correlation above, it was decided to perform the analysis using one definition of size. The number of employees was chosen since it is available for all companies in the sample, it follows Child's logic in chapter 3, and is highly correlated with the other variables. On the other hand, chapter 3 showed that sales is an unsuitable measure and a separate analysis (not included in the thesis) concluded that value added is too highly correlated with profitability to be a good measure.

8.1.2 Performance and results metrics

The hypotheses argue that there is a link between a company's performance and its size. To establish this link it is convenient to divide the analysis into four linkages. Size leads to diseconomies of scale, diseconomies of scale has an impact on results, and results in turn drives performance in the stock market.

Performance can be defined in many quantitative or qualitative ways. Examples of quantitative performance metrics are annual returns to shareholders, the ratio of market value to book value of the company, bond ratings, and price / earnings ratios. Qualitative measures include analysts' rankings, employee motivation, and customer quality surveys. This analysis uses strictly quantitative metrics based on financial measures: the market entity value to book entity value ratio (**MEV/BEV**), and the 5 year total return on shareholders equity (**Return**).The graph below shows the correlation between these two metrics as well as the market equity value to book equity value ratio (**MV/BV**), the price / earnings ratio (**P/E**), and the 1, 3, 5 and 10 year total return on shareholders equity (**Return1**, **Return3**, **Return**[18], **Return 10**, respectively).



18 Since the 5 year return data was chosen to be used in the remaining analyses the was variable was named Return rather than Return5.

The graph shows strong or very strong correlations between all definitions. The main analyses uses the MEV / BEV ratio and 5 year return for the following reasons. MEV / BEV is strongly correlated with both MV / BV and P / E while either one of these would have slightly weaker correlations with the two other definitions. In addition, there is no theoretical justification in finance theory for using the P / E ratio. MEV / BEV also happens to have the largest numbers of observations (MEV / BEV 677; MV / BV 654; P / E 538).

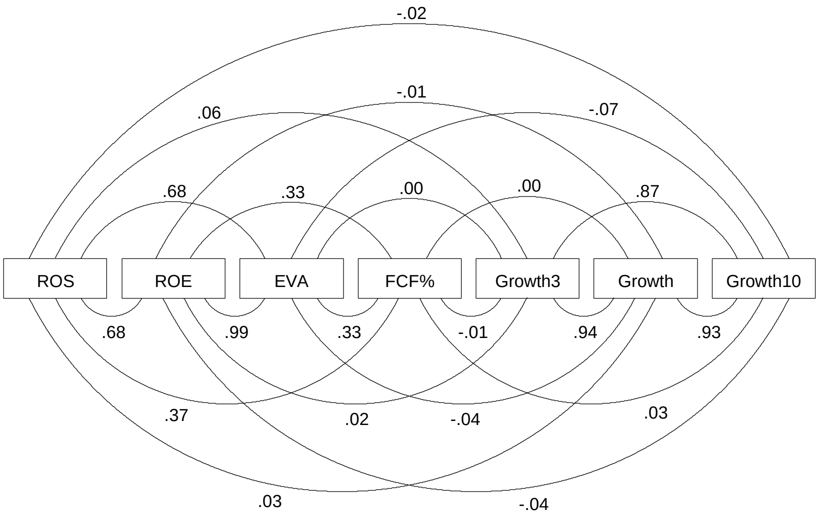
Similarly, The 5 year return is strongly correlated with the 3, 10, and 1 year returns and choosing any other time horizon would reduce the total correlation. Moreover, the 10 year return has fewer observations (496 versus 610), and the 1 year return is more affected by random fluctuations.

Results are distinguished from performance by being the outcome of operating events while performance is the stock market’s valuation of the results. Quantitative measures of results include profitability measures such as gross margin, net margin (i.e. return on sales), return on equity, economic value added (i.e. return on equity less cost of equity), return on capital employed; cash flow measures such as free cash flow over sales, and growth measures such as 1, 3, 5, and 10 year historical revenue growth. Ideally, forward-looking measures such as expected profitability, cash flow, and growth should be used. However, the only way to get such data without relying on analyst estimates is to perform the analyses for a sample older than 5 years. This would significantly reduce the number of variables that can be realistically collected since most nonfinancial data is not available 5 years back. Thus, the decision was made to use historical or current data.

Below are the correlations between seven results metrics. Return on sales (**ROS**), return on equity (**ROE**), and economic value added (**EVA**) are very highly correlated. EVA was chosen as the profitability metric since it most accurately measures profitability and is available for all but two companies. The growth measures also show very correlations. 5 year growth (**Growth**) was chosen as the best measure for similar reasons as 5 year return above. It has 759 observations. Free cash flow (**FCF%**) was not chosen as a measure of results since it is too prone to spurious fluctuations in a single year[19].

[19] An additional benefit of using **EVA** and **Growth** is that it corresponds to Penrose's ([1959] 1995) distinction between static economies (pp. 89-99) and diseconomies (pp. 12-13) of size (**EVA**) and dynamic economies (pp. 99-101) and diseconomies (pp. 212-214) of growth (**Growth**)

CORRELATION BETWEEN DIFFERENT RESULTS METRICS
Return on sales, return on equity, economic value added,
free cashflow / sales, 3, 5, and 10 year sales growth



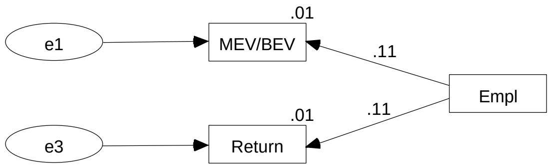
It is worth noting that profitability and growth are uncorrelated (0.01) and that free cash flow is uncorrelated with growth (-0.04) and somewhat correlated with profitability (0.32).

8.1.3 Performance, results and size

As expected, a straight regression of performance or results against size does not indicate any scale diseconomies. If the relationship was obvious then it would have been noted a long time ago. If anything, the regressions indicate the opposite, that scale economies exist although they are small. The regression below shows a positive the relationship between **MEV/BEV** and **Return** when regressed against **Empl** as the independent variable. The sample includes all 784 cases. **Empl** **MEV/BEV** and **Empl**

Return are significant at the 99% level. However, only 1% of the squared multiple correlation (in this case equal to R^2) are explained for **MEV/BEV** and **Return**.

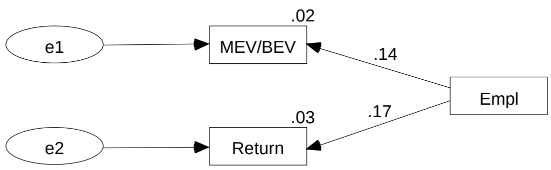
PERFORMANCE VERSUS SIZE
Market equity value / book equity value and shareholder return to employees
Group: All



Regression Weights:	Estimate	S.E.	C.R.
MEV/BEV <------- Empl	0.058	0.020	2.876
Return <-------- Empl	2.148	0.776	2.769

The regression was also run for two subsamples based on company size. The subsamples differed sufficiently to be statistically significant at the 95% level for **Empl MEV/BEV** and at the 90% level for **Empl Return**. Below are the results for the largest 300 companies.

PERFORMANCE VERSUS SIZE
Market equity value / book equity value and shareholder return to employees
Group: Large



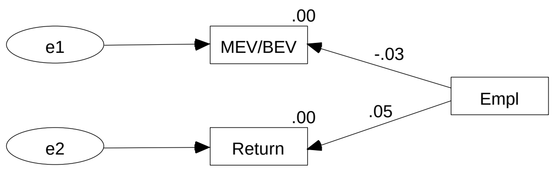
```
Regression Weights:                         Estimate     S.E.      C.R.
-------------------                         --------   -------   -------
               MEV/BEV <------- Empl          0.114     0.048     2.391
               Return <-------- Empl          4.825     1.782     2.708
```

The positive correlations (i.e. economies of scale) are even more pronounced for the large companies but have somewhat lower significance. Again SQM is low:

```
Squared Multiple Correlations:              Estimate
------------------------------              --------
                              Return          0.028
                             MEV/BEV          0.020
```

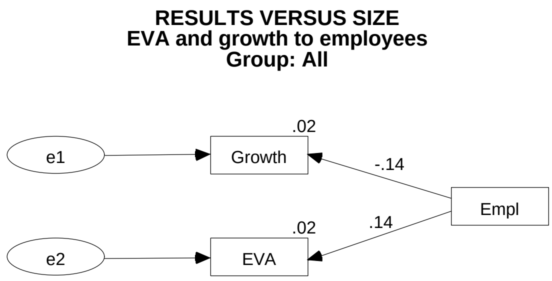
Conversely, the smallest 484 companies show neither scale economies or diseconomies and the correlations are insignificant with critical ratios of – 0.54 and 0.98 respectively. SQM for **MEV/BEV** and **Return** are less than 0.5%.

PERFORMANCE VERSUS SIZE
Market equity value / book equity value and shareholder return to employees
Group: Small




```
Regression Weights:                         Estimate     S.E.      C.R.
-------------------                         --------   -------   -------
               MEV/BEV <------- Empl         -0.024     0.044    -0.544
               Return <-------- Empl          1.718     1.749     0.982
```

Turning to the correlation between results and size, the outcome is similar. The regression is only presented for the total sample of 784 companies since the critical ratios for the differences between the large and the small sample were negligible (0.53 for **Empl Growth** and 0.49 for **Empl EVA**). The regression below show a positive correlation between size and profitability at better than 99% significance and a negative correlation size and growth with 99% significance, and with SQM around 0.02 for both variables.



Regression Weights:	Estimate	S.E.	C.R.
Growth <------ Empl	-0.017	0.004	-3.966
EVA <--------- Empl	2.369	0.580	4.084

Squared Multiple Correlations:	Estimate
EVA	0.021
Growth	0.020

In summary, there is no strong evidence of diseconomies of scale when simple relationships are explored. While the regression coefficients are significant, the explained variance is at the most 3%. As expected, the first hypothesis is not contradicted:

H_1: A company's size is not a direct predictor of its results.

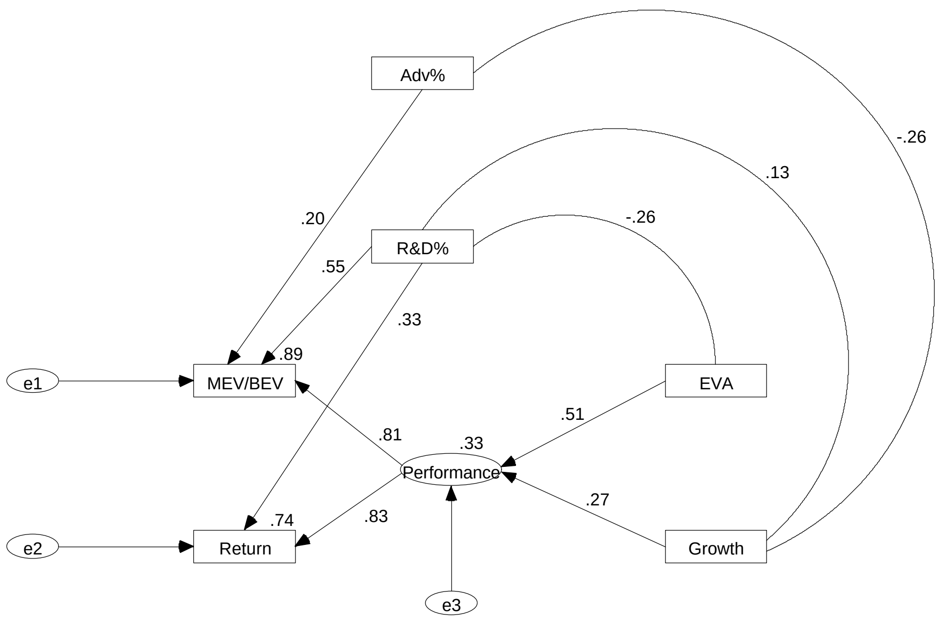
The next step in the analysis is to untangle these relationships using the hypotheses defined in chapters 4 and 5.

8.2 PERFORMANCE AND RESULTS

Before testing the hypotheses it is instructive to study the relationship between corporate stock market performance and results. This relationship is obvious and it is not the purpose of this research to prove what financial research has explained theoretically and empirically over the last 75 years, and which any stock market investor takes for granted. Yet, it is interesting to see what the relationships look like for the particular data set used here.

The path diagram below has a few important characteristics. First, **Performance** has been defined as an unobserved latent variable. Second, **MEV/BEV** and **Return** have two control variables: **R&D%** and **Adv%**. The reason for this is that investments in brand equity and in research and development have value beyond the current year. However, these investments are treated as an expense from an accounting point-of-view. Thus, in line with Fruhan (1979, 35-39), these control variables have been introduced into the structural equation model. The graph below shows that economic value added and growth explain 33% of performance for the whole sample of 784 companies. It should be noted though that the relationships are less than ideal as from a theoretical perspective. **MEV/BEV** is a measure of future, expected results while Return and Growth are historically based, and EVA, R&D% and Adv% are based on the current period. The analysis only has value if there are strong serial correlations, which appears to be the case. If data was available for expected return, EVA, R&D% and Adv%, then the SQM would likely be significantly higher.

RELATIONSHIP BETWEEN PERFORMANCE AND RESULTS
Market entity value / Book entity value and 5 year return versus economic value added and 5 year growth



Returning to the model, the critical ratios for the relationships are all significant beyond the 1‰ level and chi square / degrees of freedom is 3.5, indicating a good fit[20]

Regression Weights:	Estimate	S.E.	C.R.
Performance <------------ EVA	1.000		
Performance <--------- Growth	72.587	11.929	6.085
MEV/BEV <-------- Performance	0.013	0.001	12.669
Return <--------- Performance	0.495	0.041	11.978
MEV/BEV <--------------- R&D%	6.139	0.389	15.801
MEV/BEV <--------------- Adv%	2.165	0.428	5.057
Return <---------------- R&D%	135.040	16.710	8.082

The model was also run with two subsets of the database, large companies (the largest 300 companies by sales) and small companies (the smallest 484 companies by sales). There were no statistically significant differences in regression weights for these subsets (i.e. the critical ratio for the differences were small, in all cases less than 1.105 and not significant at the 10% level) and the SQMs were similar.

Squared Multiple Correlations:	Estimate		
	All	Large	Small
Performance	0.329	0.303	0.355
Return	0.736	0.819	0.810
MEV/BEV	0.889	0.861	0.767

The relationship between and performance and results thus seems to be homogenous within the sample.

---

20 Excellent fit is defined as chi square / degrees of freedom <2, good fit < 5, and acceptable fit < 10. This is broadly in line with (Kelloway 1998, 28) and Hair (1998, 623).

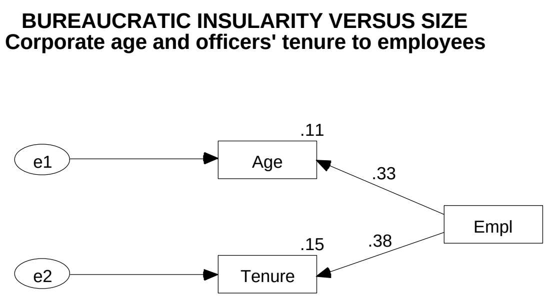
8.3 DISECONOMIES OF SCALE

Chapter 4 showed that there are four types of size-related diseconomies: bureaucratic insularity, communication distortion, atmospheric consequences, and incentive limits. Each of these are analysed individually in this section with aim to determine if they are 1) correlated with size, and 2) if they have a negative impact on results (which in turn impacts performance). Each factor is analysed in isolation and thus the correlations found are not necessarily the same as in an integrated model. Nevertheless, they serve as an initial indicator of dependencies and will guide the design of the complete model. This section also gives more precise definitions of variables and discusses variables collected as part of the research but not included in the final analysis.

8.3.1 Bureaucratic insularity (BI)

Bureaucratic insularity was defined as executives' propensity to become increasingly isolated from market opportunities and the lower levels of the organisation as the firm grows and procedures and processes are added to the organisational fabric. Indicators of bureaucratic insularity could be the compensation of senior executives (since a prediction is that executives will be maximise their own gains rather than shareholder gains), the age of the company (since older companies should have built up more of the insular mechanisms), the tenure of the CEO and officers (since high tenure should lead to higher insularity), and the share of free cash flow generated being reinvested in the business with substandard (below cost of equity) returns.

The choice here was to use corporate age (**Age**) and officers' average tenure (**Tenure**) as indicators. Executive compensation proved impossible to define in a meaningful way, as did cash flow reinvestment, while CEO tenure was uncorrelated with other variables.



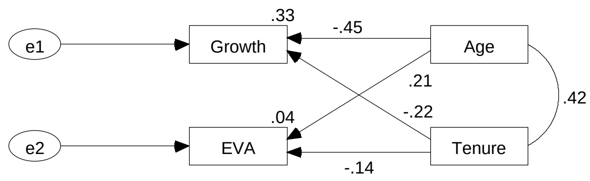
The path diagram shows a strong correlation between corporate age and size and between officers' tenure and size and the regression weights are significant beyond the 1‰ level.

```
Regression Weights:                         Estimate     S.E.      C.R.
-------------------                         --------   -------   -------
                 Tenure <------ Empl           2.503     0.610     4.101
                 Age <--------- Empl           0.238     0.027     8.756
```

The correlation between **Empl** and **Age** is even higher in reality since almost all large companies are old, leading to heteroscedasticity. This is evidenced by running the model with the large company subsample (300 largest company). The critical ratio for **Empl**→**Age** drops from 8.756 for all companies to 2.257 for large companies.

After establishing that the age and tenure indicators are positively correlated with size the analysis now turns to understanding their impact on results, indicated by EVA and growth. The path diagram below shows that age and tenure both have strong negative influences on growth and that tenure also influences EVA negatively. However, age has a positive influence on EVA. It can be hypothesized that this is because of survivor bias. That is, the companies in the sample are by definition survivors and thus can be expected to have higher profitability than non-surviving companies. This implies that companies that have survived over a long time period should be more profitable than young companies since younger companies (usually in younger industries) include a more heterogeneous mix of performers. This hypothesis is tested in a later section where the larger, integrated model is analysed using latent variables.

RESULTS VERSUS BUREAUCRATIC INSULARITY
Economic value added and growth to corporate age and officers' tenure



At this point, the analysis suggests that there is support for Penrose's ([1959] 1995, 261-263) assertion that diseconomies of scale are mainly related to difficulties of growth. Age and tenure explain 33% of the variance of growth and both regression weights have significance better than 1%.

Regression Weights:	Estimate	S.E.	C.R.
Growth <---- Tenure	-0.004	0.002	-2.581
EVA <------- Tenure	-0.354	0.236	-1.499
Growth <------- Age	-0.074	0.008	-8.719
EVA <---------- Age	4.669	1.293	3.611

On the other hand, bureaucratic insularity explains only 4% of profitability (measured as EVA).

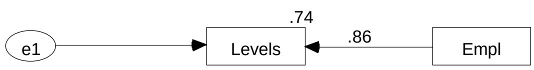
8.3.2 Communication distortion (CD)

Communication distortion can be defined as the length of the communications process through the organisational hierarchy or as the time it takes for decisions to go from issue identification to implementation. The first definition has usually been operationalised with the number of hierarchical levels within the company and this is the definition here as well. Unfortunately, it has not been possible to operationalise the second definition here.

A number of studies, including Child (1973), demonstrate that the number of hierarchical levels increase logarithmically with the number of employees. This research reached the same conclusion by studying the organisational structure of a subset of the 784 companies. The subset included 271 companies with an average of 7.8 levels and a standard deviation of 0.9 levels. No company had more than 10 levels.

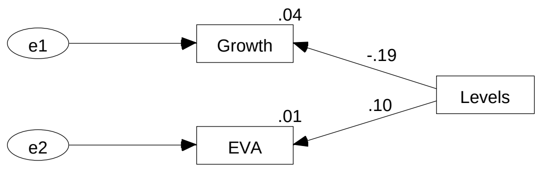
As expected, there is a strong dependency between number of employees and the number of levels, as shown below. The critical ratio is 28.1. Additional regressions were run with control variables such as organisational form (unitary or multi-divisional) but no such control variable proved important.

COMMUNICATIONS DISTORSION VERSUS SIZE
Hierarchical levels to employees



Communication distortion does not have a significant impact on EVA or growth as evidenced by the regression below. Less than 4% of the variance of growth and profitability was accounted for, and only the negative impact on growth is statistically significant (at the 1% level).

RESULTS VERSUS COMMUNICATIONS DISTORTION
Economic value added and growth to hierarchical levels



Regression Weights:	Estimate	S.E.	C.R.
EVA <------- Levels	2.008	1.157	1.735
Growth <---- Levels	-0.028	0.009	-3.250

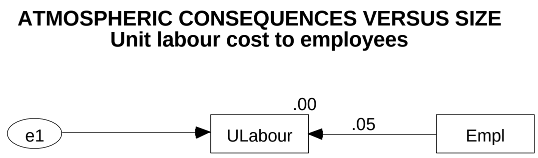
The impact of hierarchical levels will increase in the complete model. Nevertheless, the weak correlations indicate that while there may exist communication distortion within companies, there may exist similar distortions when entrepreneurs operate through the market rather than through hierarchies.

8.3.3 Atmospheric consequences (AC)

Based on the argument on pages 44-45, atmospheric consequences should exhibit themselves as alienation, which in turn requires companies to pay more to keep their employees. Thus, unit labour cost should be a good indicator for atmospheric consequences. Unfortunately, it proved difficult to collect this data for the 784 companies. 1998 labour cost data was only available for 52 companies in Compustat and annual reports, an additional 43 cases where calculated by taking available labour cost data from 1991 till 1997 (most of them from 1997) and extrapolating this data for 1998. This was done by calculating unit labour cost for the observed year and then inflating this by the annual average increase in compensation for the whole sample.

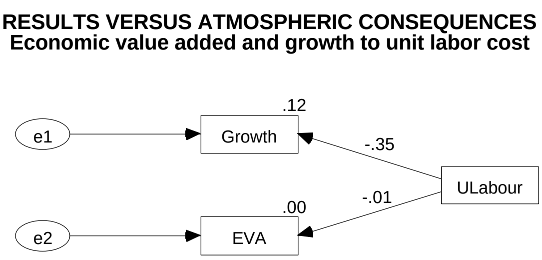
Attempts was made to run a separate regression and to carefully impute additional observations as suggested by e.g. Jöreskog and Sörbom (1996, 78). The regression was run with unit labour cost as the dependent variable and with industry unit labour cost, sales per employee, R&D share, and foreign revenue share as the independent variables. The same variables were used in a LISREL imputation. The results were too erratic to warrant inclusion in the core database even though it would have increased the sample size from 95 to 435 and 399 observations respectively.

The graph below shows that unit labour cost is more or less uncorrelated with number of employees (it is positively correlated with sales). The critical ratio is 0.441.



Thus, to the extent that unit labour cost is a good proxy for atmospheric consequences, very large companies do not seem to suffer negatively from atmospheric consequences. The conclusion was the same when control variables (R&D share, foreign share) were introduced.

Turning to atmospheric consequences impact on results, the graph below shows that unit labour cost is negatively correlated with growth and uncorrelated with profitability. **ULabour→Growth** is significant beyond the 1‰ level. The results were the same for two subsamples: the largest 150 companies and the largest 300 companies.



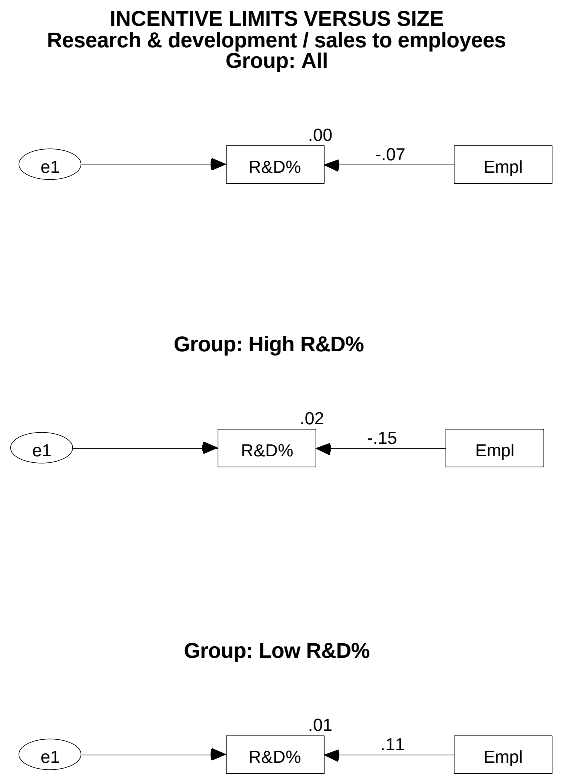
Regression Weights:	Estimate	S.E.	C.R.
Growth <----- ULabour	-0.155	0.039	-3.983
EVA <-------- ULabour	-0.848	4.591	-0.185

The results for atmospheric consequences are inconclusive. Unit labour cost appears to be negatively correlated with growth, but it is not clear that size is the underlying driver.

8.3.4 Incentive limits (IL)

The fourth and final diseconomy factor is incentive limits. As discussed in previous chapters, incentive limits are most serious in companies with indivisible tasks such as product development were the outcome is dependent on the collaboration of many individuals and the contribution of each individual is hard to measure. Two indicators were tested: 1) the research and development intensity measured as R&D expense divided by sales, and 2) general and administrative cost measured as SG&A expense less advertising expense. The assumption behind the latter definition is that the bulk of general and administrative activities are indivisible (e.g. general management tasks). The two indicators have fairly high correlation (52%). However, **R&D%** has 489 observations, while **SG&A% – Adv%** has only 177 observations and thus R&D% was chosen as the better indicator.

In the first analysis, the approach is different from the one taken in the previous three sections. Large companies should try to avoid incentive limits. Consequently, we would expect a negative regression coefficient for **Empl→R&D%**. This should be especially true for companies in high R&D environments. The three graph below show this to be case. The first diagram shows the regression for the whole sample of 784 companies. The second regression includes only those companies with higher than average R&D (245 companies). The third regression includes those companies with lower than average R&D (244 companies)[21].

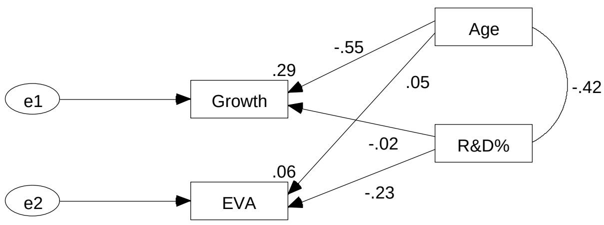


[21] 295 companies did not specify their R&D%.

The critical ratio for the difference between the regression coefficient **Empl→R&D%** is 2.624, i.e. the difference is significant at the 1% level. This lends support to the hypothesised impact of incentive limits.

The second analysis tests if incentive limits have an impact on profitability and growth. Age was introduced as a control variable since older companies tend to be active in mature industries. The regression shows a strong negative impact of R&D% on growth (at better than the 1‰ significance level). The impact on growth is negligible.

RESULTS VERSUS INCENTIVE LIMITS
Economic value added and growth to research & development / sales, controlled for age
Group: All




```
Regression Weights:                          Estimate    S.E.      C.R.
-------------------                          --------   -------   -------
                  Growth <------ R&D%         -0.057     0.107    -0.533
                  EVA <--------- R&D%        -76.278    16.011    -4.764
                  EVA <---------- Age          1.199     0.966     1.240
                  Growth <------- Age         -0.091     0.006   -14.688
```

Once again, there should be a difference between the high R&D% and low R&D% subsamples. The **R&D%→EVA** standardized coefficient for the low R&D% subsample is +0.05 and –0.18 for the high R&D subsample, in support of the theory. However, the critical ratio for the difference is an inconclusive 1.075.

In summary, there is considerable support for the hypothesis that incentive limits increase with corporate size and that they affect performance by having a negative impact on profitability.

* * *

Each of the four factors have now been shown to increase with firm size. Thus the second hypotheses can not be refuted.

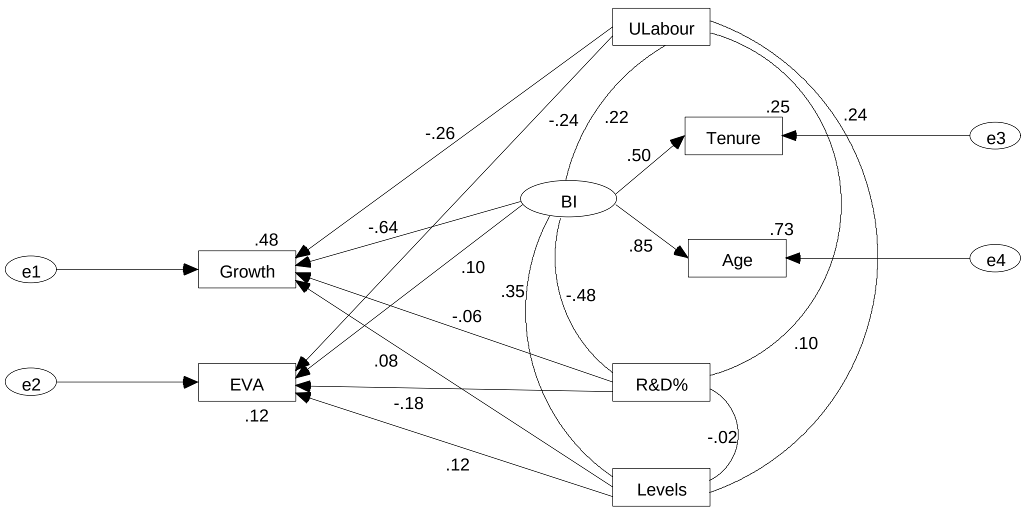
H_2: Large companies embed diseconomies of scale of four types: communication distortion, bureaucratic insularity, atmospheric consequences, and incentive limits.

For three factors the impact of size was significant, while it was marginal for atmospheric consequences (as measured with unit labour cost).

8.3.5 Integrated diseconomies of scale model

It is now possible to put together an integrated diseconomies of scale model, bearing in mind that the mitigating factors (M-form organization and asset specificity) and scale economies have not yet been introduced, nor has survivor bias. The path diagram below shows the design. A latent variable, BI, has been introduced to capture bureaucratic insularity since there are two indicators (corporate age and officers' tenure) for this factor.

DISECONOMIES OF SCALE
Impact of bureaucratic insularity, communications distorsion, atmospheric consequences, and incentive limits on results
(excluding survivor bias, mitigating factors and economies of scale)



The model has a chi square of 9.260 and 5 degrees of freedom, leading to an excellent chi square / df = 1.852 and a commensurate P=0.099. The normed fit index (NFI) is 0.998 and the parsimonious fit index is 0.179—

both well beyond any reasonable requirement. The model thus appears to be robust.

Below are the unstandardised regression weights with standard errors and critical ratios:

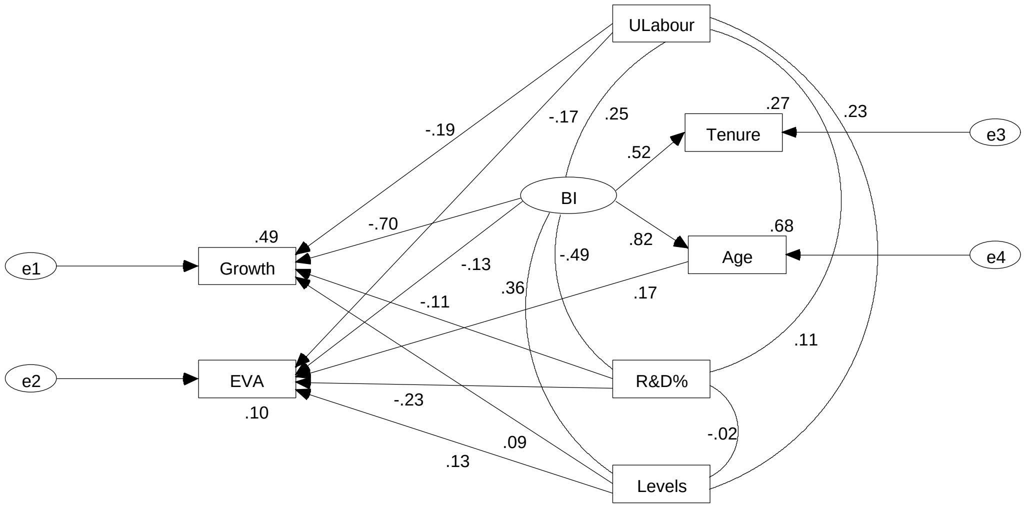
Regression Weights:	Estimate	S.E.	C.R.
EVA <----------- R&D%	-59.701	22.728	-2.627
EVA <--------- Levels	2.272	1.463	1.553
Growth <------ Levels	0.012	0.014	0.848
Growth <---------- BI	-0.124	0.042	-2.984
Growth <-------- R&D%	-0.160	0.289	-0.556
EVA <------------- BI	2.490	2.141	1.163
EVA <-------- ULabour	-13.987	5.393	-2.594
Tenure <---------- BI	5.261	1.304	4.035
Age <------------- BI	1.000		
Growth <----- ULabour	-0.112	0.037	-3.032

Bureaucratic insularity has a strong negative impact on growth (**BI→Growth** = 0.64) and is significant at the 1% level (critical ratio = 2.984), while there is a non-significant positive impact on EVA (BI→EVA = 0.10) before the model adjustment for survivor bias. Communication distortion (**Levels**) has a non-significant (worse than 10% level) positive impact on both growth and EVA (critical ratio = 0,848 and 1.553 respectively), weakly contradicting the hypothesis. Atmospheric consequences has a negative impact on both growth (**ULabour→Growth** = -0.26) and EVA (**ULabour→EVA** = -0.24) with significance better than 1% (critical ratios 3.032 and 2.594 respectively). It should be remembered that unit labour cost has few observations. Finally, incentive limits had a negative impact on EVA with a significance almost at the 1% level (critical ratio = 2.527 is equal to the 1.15% level), while there is a negative, non-significant impact on growth. The structural equation model thus is either directionally in line with the hypotheses for each factor or does not show any correlation. In no case there a statistically significant contradiction of the hypotheses.

Finally, the squared multiple correlation is a high 0.476 for growth and a moderate 0.122 for EVA.

A second path diagram was constructed to take survivor bias into account. This was done by introducing a link between corporate age and profitability as explained earlier in section 5.3. This model is less parsimonious (PFI = 0.143) and has a slightly lower chi square / df ratio (8.259 / 4 = 2.065). Yet it is probably a slightly more realistic representation of the underlying theory and hypotheses.

DISECONOMIES OF SCALE
Impact of bureaucratic insularity, communications distorsion, atmospheric consequences, and incentive limits on results
(Including survivor bias, excluding mitigating factors and economies of scale)



The unstandardised regression coefficients, standard errors, and critical ratios are shown below. Without repeating the discussion from the previous path diagram it should be noted that bureaucratic insularity now has the hypothesised (but non-significant) negative impact on profitability.

Regression Weights:	Estimate	S.E.	C.R.
Age <------------- BI	1.000		
EVA <----------- R&D%	-76.088	26.554	-2.865
EVA <--------- Levels	2.469	1.409	1.753
Growth <------ Levels	0.013	0.015	0.887
Growth <---------- BI	-0.140	0.050	-2.803
Growth <-------- R&D%	-0.266	0.339	-0.785
EVA <------------- BI	-3.594	5.733	-0.627
EVA <-------- ULabour	-9.802	6.719	-1.459
Tenure <---------- BI	5.645	1.329	4.247
Growth <----- ULabour	-0.084	0.048	-1.769
EVA <------------ Age	3.703	3.428	1.080

In some respects, this integrated diseconomies of scale model is the best model since it is fairly parsimonious and has good predictive power. It does not contradict the theory and explains a significant share of the variance in results, especially that of corporate growth.

Thus, a test of the hypotheses related to diseconomies of scale shows that the theory can not be refuted (the sub-hypotheses are discussed in section 8.6:

H_3: Scale diseconomies impact company profitability and growth negatively.

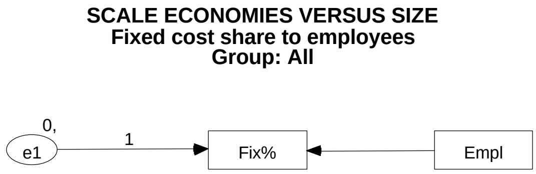
The next steps are to incorporate the potential effects of economies of scale and mitigating factors into the model. While additional insights are gained from this, it should also be noted that the model becomes less robust as variables are added.

8.4 ECONOMIES OF SCALE

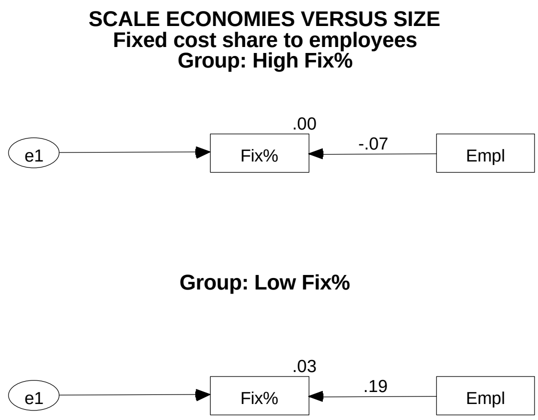
The literature survey showed that the effects of scale economies should be small or non-existent. The reason was that scale economies set in at a fairly small size and that they, according to transaction cost theory, apply to entire industries rather than individual companies since information travels fast between companies. This conclusion is somewhat at odds with neoclassical economics. The objective of this research is not to prove or disprove the existence of scale economics. However, it is instructive to incorporate scale economies in the model to see if there are any effects.

Two indicators were tested, both building on the assumption that scale economies exist when fixed costs are high. The chosen definition was to take all fixed and semi-fixed costs on the income statement and divide these by total inputs (inputs are similar to sales but differ since they are the sum of all factor inputs including cost of equity, regardless of if the sum of these are larger or smaller than sales; this definition is equivalent to sales less net income plus cost of equity). By using inputs rather than sales spurious business cycle effects are reduced. The observed variable **Fix%** was defined as fixed cost share = (interest + depreciation+ SG&A) / inputs. This definition assumes that fixed costs are related to more than fixed assets, specifically that the level of SG&A expense (including R&D) is not easily varied. It is equivalent to Penrose's definition ([1959] 1995, 89-95). The second definition, which was discarded, was to use the classical definition of fixed assets / sales. This definition had no statistical significance, lending some credence to the transaction cost theoretical argument that scale economies do not exist, at least for bricks and mortar assets.

The graph below shows the relationship between number of employees and fixed cost share. The correlation is negligible (with a critical ratio of 0.418), possibly because there are countervailing forces at work, or because scale economies do not exist. The countervailing forces argument is that scale economies leads large companies to be active in fixed cost-intensive industries; on the other hand, these very companies would realise the benefits of scale and thus have a lower share of fixed costs.



With a similar logic as was employed in the case of incentive limits, it is possible to test which of the two arguments is correct. If there are no scale economies then two subsamples consisting of 1) high fix%, and 2) low fix% should not differ, while the opposite is true if scale economies exist. Below are the path diagrams for these subsamples with high fix% defined as those 300 companies with the highest fixed cost ratio, and low fix% as the 453 companies remaining companies (with an additional 31 companies having missing values).



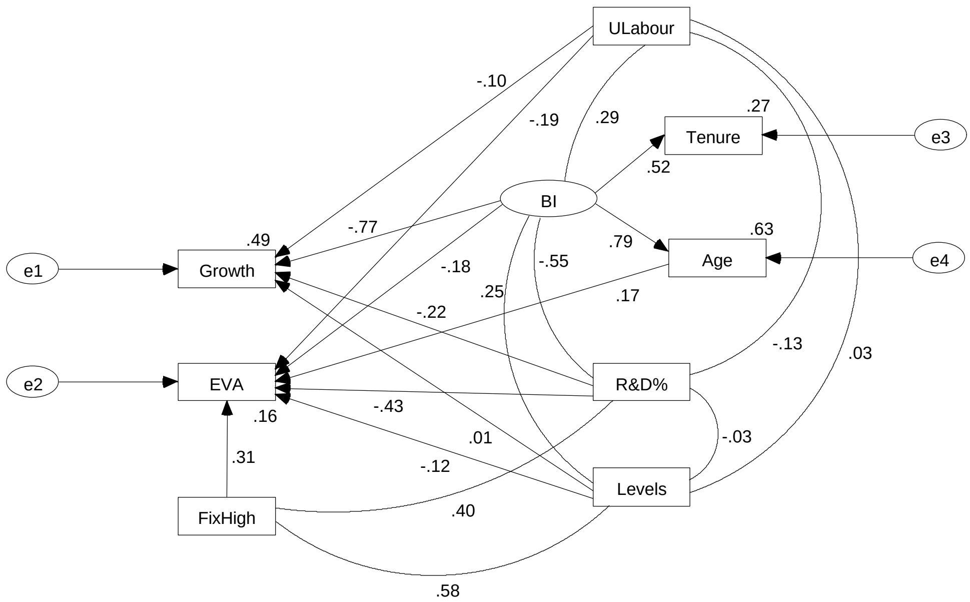
The standardised regression coefficients differ significantly for the two subsets and the critical ratio for the difference is 2.871, implying a significance better than the 1% level. At this point, the argument that scale economies exist among large and very large companies cannot be rejected.

Economies of scale is now introduced directly into the diseconomies of scale path diagram. Rather than performing the regression on the two subsets as above, the variable **Fix%** was replaced with **FixHigh**. The reason was that using two subsets reduces the number of observations too much for the variables **ULabour** and **Tenure**, which already in the total sample have uncomfortably few observations. **FixHigh** was constructed using the following logic: Scale economies should be large for those companies that **simultaneously** are active in high fixed cost environments and have high fixed costs. Thus, the variable **FixHigh** multiplies the fixed cost ratio with the absolute level of fixed cost (**FixHigh** = **Fix%** * **FA** = $\mathbf{FA}^2$

/ **Sales**). The structural equation model was also run with a dummy variable set at 1 for a high fixed cost ratio and 0 otherwise, and with only the fixed cost ratio. This yielded similar, but less significant, results as should be expected since the first definition leads to information loss and the second is atheoretical.

The path diagram below shows the results. Economies of scale has a positive impact on EVA with a significance better than the 1‰ level since the critical ratio is 4.210. Most of the other regression coefficient become more significant, with a particular increase in incentive limits negative impact on profitability (R&D% EVA = -0.43; critical ratio = 5.721)

DISECONOMIES AND ECONOMIES OF SCALE
Excludes mitigating factors (includes survivor bias)



The chi square / df ratio is 4.991 which is good, the normed fit index is an excellent 0.994, and the parsimonious fit index is 0.222 which is reasonable.

Regression Weights:	Estimate	S.E.	C.R.
Age <------------- BI	1.000		
Growth <---------- BI	-0.159	0.049	-3.235
Growth <-------- R&D%	-0.514	0.299	-1.719
EVA <------------- BI	-5.122	4.567	-1.122
EVA <-------- ULabour	-11.436	6.300	-1.815
Tenure <---------- BI	5.910	1.315	4.493
Growth <----- ULabour	-0.044	0.049	-0.892
Growth <------ Levels	0.001	0.013	0.095
EVA <--------- Levels	-2.538	1.816	-1.397
EVA <------------ Age	3.702	2.697	1.373
EVA <-------- FixHigh	3.559	0.845	4.210
EVA <----------- R&D%	-139.069	24.310	-5.721

The introduction of scale economies into the model has further strengthened the argument that the statistical analysis does not contradict the theoretical predictions and the associated hypotheses.

H_5: Scale economies are weakly related to size.

H_6: Scale economies have no impact on individual company profitability.

Next, Williamson's mitigating factors are introduced and the complete model is tested.

8.5 MITIGATORS

At this stage the final two variables are introduced into the model: the mitigating factors organisational form and asset specificity.

8.5.1 Organisational form

Williamson (Williamson 1970, 120-139) argued that the large multidivisional firm (M-form) on average outperforms the large unitary firm (U-firm). Williamson's definition of M-form was (pp. 120-121):

1. The responsibility for operating decisions is assigned to…operating divisions or quasifirms.
2. The elite staff attached to the general office performs both advisory and auditing functions…
3. The general office is principally concerned with strategic decisions…
4. …separation of the general office from operations…
5. The resulting structure displays both rationality and synergy…the whole is greater than the sum of the parts.

This definition can be operationalised by introducing two variables. The first describes the ability of the company to effectively divide the tasks performed by the senior executives and their staffs as well as the division of responsibility between the board of directors and executive management (p. 138-139), what today often is referred to as governance. The second variable measures whether the organisational structure is multidivisional or functional, or somewhere in between.

The governance variable was operationalised as **Govern** based on CalPERS's[22] definitions (1999). CalPERS evaluates corporate governance based on financial results (three year shareholder returns and EVA) and a corporate governance screen. For the purposes here the governance screen is of interest so that co-linearity with the independent variables is avoided. The screen uses 25 criteria divided into four main categories: board

---

[22] CalPERS (California Public Employees' Retirement System) is arguably the world's leading authority on governance issues.

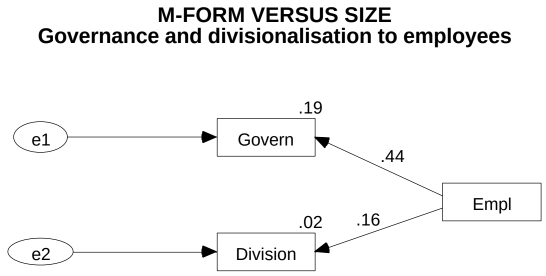
composition / structure, director compensation / stock holdings, management, and anti-takeover devices. The data was taken from three sources: 1) Business Week's annual survey of corporate governance (Byrne 2000). Business Week measures six attributes of governance and of these, four where used in **Govern** since they correspond reasonably well to the CalPERS criteria and two (relating to the quality of the board members) where excluded. 2) The compilation of governance data for 1,500 companies published by the Investor Responsibility Research Center[23] (1999), and 3) Fortune's annual ranking of America's most admired companies (Colvin 2000) in which three variables (of eight) relate to CalPERS governance criteria. The three sources were merged into a single index. No attempt was made to independently validate the definitions and underlying research, except for using this researcher's general understanding of the quality of governance at those companies surveyed. This crude test corroborated the data.

The organisational structure variable was operationalised as **Division**. The variable is ordinal with 2 representing a clean, multidivisional structure, 1 representing a mixture of multidivisional and unitary structures, and 0 representing functional structures. The classification was done by the researcher using annual reports, web sites, and 10Ks. The data collection

[23] CalPERS subcontracts the governance screening to IRRC.

approach was similar to Rumelt's (1974, 43) with three factors influencing the judgments made: the titles of senior executives, the descriptions of large operating units, and the reporting or lack of reporting of operating unit financials. For example, if the senior executives at headquarters have titles such as Senior Vice President of Business Development and similar staff descriptions, while the senior executives of the operating units are called President, Operating Unit, then that would suggest a multidivisional structure. Conversely, if a company does not talk about its operating units as businesses and there is no unit reporting then this would suggest a functional structure. **Govern** has 229 observations, mainly among the largest 400 companies in the sample, and **Division** was randomly collected for 265 companies.

The graph below shows that quality of governance and use of divisionalisation increases with company size, supporting Williamson's prediction and in line with previous empirical research. **Empl→Govern** has a critical ratio of 7.335 (better than 1‰ significance) and **Empl→Division** has a critical ratio of 2.591 (better than 1% significance). It should be remembered though that **Division** is an ordinal variable and Amos treats it as continuous. Results would be more accurate if polyserial correlations were used and the squared multiple correlation would then increase somewhat[24].

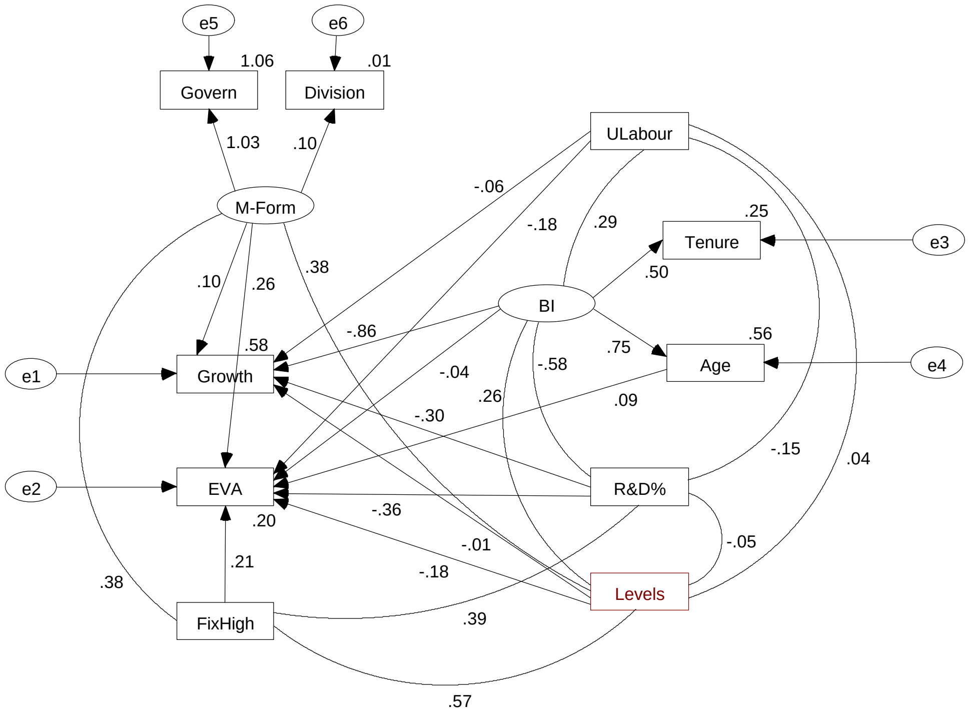


Turning to the second prediction, that M-form organisation improves results, the latent variable **M-Form** was introduced into the master path diagram together with the indicators **Govern** and **Division**. The results are shown below. M-form organisation has a significant, positive impact on both growth and profitability with all the impact emanating from governance. This is not surprising since, as was discussed in the literature survey, the positive impact of divisionalisation was fully exploited more than 25 years ago.

---

[24] This analysis was performed using LISREL. The regression coefficient increased to 0.22 from 0.18 and SQM to 0.049 from 0.025.

DISECONOMIES OF SCALE, ECONOMIES OF SCALE AND ORGANISATIONAL FORM
Excludes asset specificity (includes survivor bias)



The table below shows similar results as in earlier analyses with all coefficients having the predicted sign and critical ratios ranging from highly significant to non-significant.

Regression Weights:	Estimate	S.E.	C.R.
Age <--------------- BI	1.000		
EVA <------------- R&D%	-118.174	43.051	-2.745
Growth <------------ BI	-0.189	0.057	-3.319
Growth <---------- R&D%	-0.712	0.337	-2.113
EVA <--------------- BI	-1.205	6.095	-0.198
EVA <---------- ULabour	-10.904	6.244	-1.746
Tenure <------------ BI	6.048	1.327	4.558
Growth <------- ULabour	-0.027	0.053	-0.515
Govern <-------- M-Form	0.034	0.044	0.767
Division <------ M-Form	0.019	0.018	1.026
Growth <-------- Levels	-0.001	0.017	-0.071
EVA <----------- Levels	-3.745	2.484	-1.507
EVA <-------------- Age	2.014	2.260	0.891
Growth <-------- M-Form	0.003	0.002	1.427
EVA <---------- FixHigh	2.432	1.661	1.464
EVA <----------- M-Form	1.000		

8.5.2 Asset specificity

The last variable to be introduced is asset specificity. As was discussed in the literature survey, companies can, according to transaction cost theory. mitigate diseconomies of scale by increasing asset specificity. Asset specificity is measured in three ways: product breadth, geographic reach, and vertical depth. Each of these were operationalised.

Product breadth was defined in several ways. In the end, Rumelt's (1974, 14-15) specialisation ratio (SR%) was used since it is commonly accepted (Ramanujam and Varadarajan 1989, 539) and it minimises information loss. It also has the benefit of not being based on the researcher's judgement. SR% is defined as (sales from the largest business unit) / (total company sales). Alternative measures (available at canback.com \ henley \ data in the downloadable file ASSET SPECIFICITY.XLS) are 1) number of business segments the company is active in (ordinal), 2) number of SICs the company is active in (ordinal), 3) Rumelt's relatedness ratio (pp. 15-16), RR% = (sales from the largest business unit plus other business units with related activities) / (total company sales), 4) Rumelt's five category classification of companies into single, dominant-unrelated, dominant, related, and unrelated businesses (p. 31), and 5) a Herfindahl index of corporate diversity. Measures 1, 2 and 4 are based on categorical data and as such are less rich on information than continuous data. Measures 3 and 5 are continuous but less accurate than the specialization ratio for the current purpose.

Geographic reach was measured as the percent of foreign sales. No other measures were available from Compustat or other sources and the measure appears to be logical.

Vertical depth is more problematic to define, however. Vertical integration has been studied empirically many times. The best measure (described by D'Aveni and Ravenscraft (1994, 1994)) is arguably to measure the amount of internal forward or backward transfers by line of business. Unfortunately such data is not available for individual companies. A second measure often used is value added (factor inputs other than purchased goods) divided by sales. This approach has been criticised since it is sensitive to current year profitability (spuriously high or low profitability increases or decreases vertical integration) (Levy 1981, 86). A modification to this ratio is to adjust the nominator by using cost of equity rather than net income and the denominator by using cost of equity and total inputs (**VI%**) as described in the "Economies of scale" section (p. 119). Even with this adjustment the definition is open to criticism[25]. For example, the large U.S. oil companies are among the most vertically integrated companies in the world, ranging in activities from exploration, to production, to refining, to retailing. Their vertical integration ratio as described above is among the lowest (ExxonMobil has a vertical integration of 19.1% versus 41.3% for the total sample).

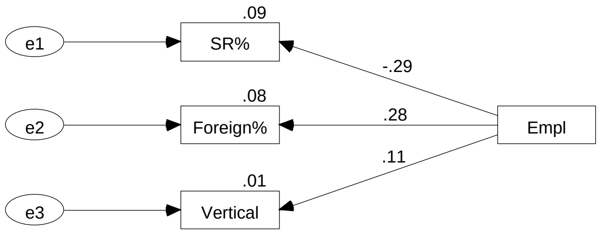
In the end, the vertical integration variable was based on a qualitative assessment by the researcher. The 784 companies were categorised based on their degree of vertical integration with **Vertical** equal to 2 for companies with very high vertical integration (13 companies), 1 for high integration (145 companies), and 0 for companies with normal or low integration (512 companies). No judgement was passed on 114 companies. Vertically integrated companies were mainly found among resource based companies and among aerospace contractors. Again, the data is found in the file ASSET SPECIFICITY.XLS. Finally, the polyserial correlation between **Vertical** and **VI%** was –0.418 which seems to confirm the criticism of the use of the value added / sales ratio.

The graph below shows the relationship between product breadth (**SR%**), foreign reach (**Foreign%**), vertical depth (**Vertical**) and number of employees (**Empl**). Note that high **SR%** implies high asset specificity, while high **Foreign%** and high **Vertical** implies low asset specificity.

---

[25] Furthermore, there were only 92 observations of the vertical integration ratio in the sample.

ASSET SPECIFICITY TO SIZE
Specialization ratio, foreign sales ratio, and vertical integration versus employees



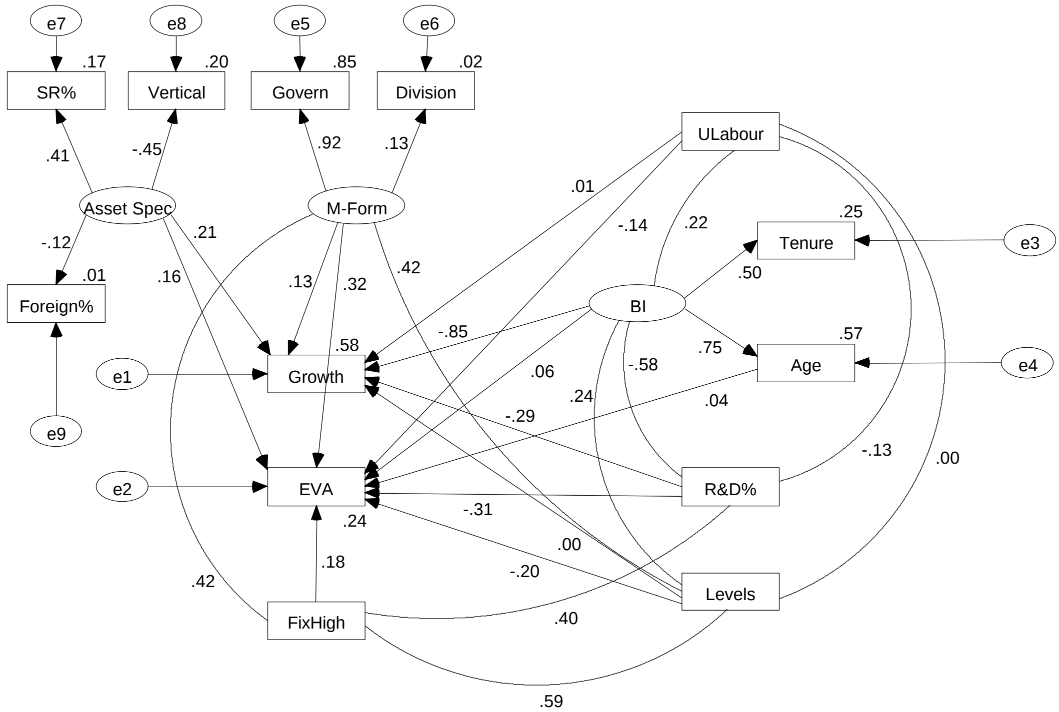
Thus, large size on average leads to less asset specificity along all three dimensions even though only a small part of the variance is explained. The critical ratios below show significance better than 1% for **Vertical**[26] and better than 1‰ for **SR%** and **Foreign%**.

Regression Weights:	Estimate	S.E.	C.R.
Vertical <-------- Empl	0.049	0.017	2.899
SR% <------------- Empl	-0.064	0.008	-7.983
Foreign% <-------- Empl	0.044	0.006	7.580

At this point it is possible to complete the model and do the final test of whether degree of asset specificity affects profitability and growth. Again, a latent variable, **Asset Spec**, is introduced to capture the total impact of asset specificity. The graph below shows the complete path diagram.

26 Using a polyserial correlation between **Vertical** and **Empl** in LISREL increases the significance of **Vertical**→**Empl** beyond the 1‰ level (critical ratio 3.406).

COMPLETE MODEL
Includes diseconomies of scale, economies of scale, and mitigating factors (adjusted for survivor bias)



High asset specificity appears to have a positive impact on profitability and growth and the three indicators for asset specificity have the hypothesised signs. The critical ratios imply a significance at the 5% for all coefficients except **Asset Spec→Foreign,** which is non-significant.

Regression Weights:		Estimate	S.E.	C.R.
SR% <-----------	Asset Spec	1.000		
Vertical <------	Asset Spec	-2.199	0.850	-2.588
Growth <--------	Asset Spec	0.273	0.103	2.651
EVA <-----------	Asset Spec	29.075	13.738	2.116
Foreign% <------	Asset Spec	-0.211	0.129	-1.637

While the hypotheses regarding asset specificity are somewhat confirmed, it is not clear that the inclusion of asset specificity in the model adds to the overall purpose of the research at hand. There are a issues with the definitions of two indicators (product breadth and vertical depth) and the theory is not clear on whether long geographic reach really implies low asset specificity[27].

* * *

The section on mitigators suggests that the fourth hypothesis and its two sub-hypotheses are not disproved:

H_4: Diseconomies of scale can be mitigated by two transaction cost-related factors: organisational form and asset specificity.

H_{4a}: Large M-form firms show better results than U-form firms, but this has been fully implemented in today's large companies.

H_{4b}: High asset specificity is positively correlated with results.

8.6 COMPLETE MODEL

At this point it is possible to analyse the complete model. One additional variable, industry, could have been included but was left out for two

27 A separate analysis was conducted by testing if companies with low fixed cost ratios had a stronger negative correlation between share foreign sales and profitability (the assumption being that such companies would have less common costs that could be spread between geographic entities). The test was inconclusive.

reasons. First, section 5.4 showed that a company's industry does not influence results significantly, at least not in the manufacturing sector. Second, it proved impossible to collect relevant variables to test this proposition and thus industry has been left as an exogenous factor, included in the error terms. Attempts were made to relate each company to its industry's averages but since most companies are active in several industries this proved impossible to do based on the data at hand. In the end, the constructs created did not improve the statistical analysis. This is in line with Rumelt's findings (1974, 98) that "industry corrected results were not only elusive, but essentially unattainable and possibly meaningless."

8.6.1 Basic model

It is now time to interpret the complete model. The structural equation model explains 59% of the variance in growth and 24% of the variance in profitability. However this has been achieved by adding variables. Even though the complete model reflects the underlying theory and the hypotheses fully and can be viewed as a confirmatory model, it is equally true that it is unwieldy. Having said this, the relationships are always in the hypothesised direction and many coefficients are significant at the 5% or better level (the significance has dropped since so many variables are included that each individual coefficient cannot have a high significance). The chi square / df ratio is a reasonable 6.896 (331.012 / 48) and the normed fit index is 0.978

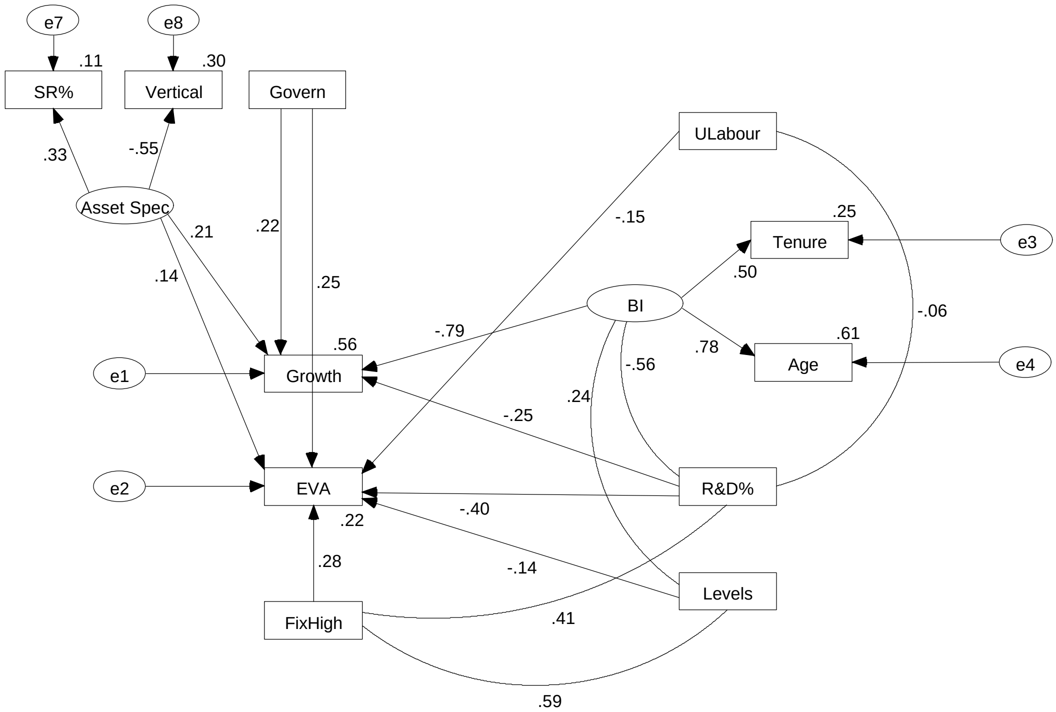
Regression Weights:	Estimate	S.E.	C.R.
Age <------------------- BI	1.000		
EVA <----------------- R&D%	-101.464	49.605	-2.045
Growth <---------------- BI	-0.184	0.055	-3.337
Growth <-------------- R&D%	-0.688	0.338	-2.037
EVA <------------------- BI	1.844	6.566	0.281
EVA <-------------- ULabour	-8.894	6.246	-1.424
Tenure <---------------- BI	5.955	1.347	4.422
Growth <----------- ULabour	0.006	0.052	0.111
Govern <------------ M-Form	0.024	0.026	0.929
Division <---------- M-Form	0.018	0.017	1.086
Growth <------------ Levels	-0.001	0.017	-0.043
EVA <--------------- Levels	-4.184	2.657	-1.574
EVA <------------------ Age	0.897	2.375	0.378
Growth <------------ M-Form	0.003	0.002	1.535
SR% <----------- Asset Spec	1.000		
Vertical <------ Asset Spec	-2.199	0.850	-2.588
Growth <-------- Asset Spec	0.273	0.103	2.651
EVA <----------- Asset Spec	29.075	13.738	2.116
Foreign% <------ Asset Spec	-0.211	0.129	-1.637
EVA <-------------- FixHigh	2.126	1.842	1.154
EVA <--------------- M-Form	1.000		

The statistical analyses has not been able to disprove the hypotheses at any level of analysis—with or without economies, with or without mitigating factors, and with or without survivor bias. The final statistical analysis section tries to modify the model so that these positive attributes remain but the parsimony improves from the current level of 0.527 (parsimonious fit index).

8.6.2 Model adjusted for parsimony

Since the model already reflects the theory and there are no competing models, the task here is to reduce the number of relationships with the assumption that certain relationships exist, but are not significant. The first simplified model below achieves has marginally lower squared multiple correlations for **Growth** (0.56 versus 0.58) and **EVA** (0.22 versus 0.24). The major changes in the model was to: 1) eliminate non-significant correlations for the four diseconomies of scale factors, 2) eliminate the **Division** variable which did not contribute to **M-Form** and to eliminate **M-Form** since it now only has one indicator (thus, **Govern** now is directly linked to **Growth** and **EVA**), and 3) eliminate **Foreign%** since it did not contribute significantly to asset specificity and did not have strong theoretical support.

SIMPLIFIED COMPLETE MODEL
Adjusted for parsimony
Includes asset specificity
Chi square = 210.687; df = 36; Chi square / df = 5.852



The chi square / df ratio improved from 6.896 to 5.852 and is slightly higher at 0.980. The parsimonious fit ratio is 0.545 compared to 0.527. The unstandardised regression coefficients below show that all coefficients have the right sign and all except two are significant at the 5% level, while the two exceptions are significant at the 10% level.

Regression Weights:	Estimate	S.E.	C.R.
EVA <----------------- R&D%	-128.655	16.879	-7.622
Growth <---------------- BI	-0.164	0.025	-6.456
Growth <-------------- R&D%	-0.583	0.174	-3.344
EVA <-------------- ULabour	-9.551	5.845	-1.634
Tenure <---------------- BI	5.721	1.276	4.483
Age <------------------- BI	1.000		
EVA <--------------- Levels	-2.899	1.582	-1.832
SR% <----------- Asset Spec	1.000		
Vertical <------ Asset Spec	-3.390	1.498	-2.264
Growth <-------- Asset Spec	0.358	0.120	2.975
EVA <----------- Asset Spec	33.049	15.300	2.160
EVA <-------------- FixHigh	3.256	0.782	4.163
EVA <--------------- Govern	29.302	6.594	4.444
Growth <------------ Govern	0.185	0.043	4.292

Hypotheses three and its four sub-hypotheses are not contradicted:

H_3: Scale diseconomies impact company profitability and growth negatively.

H_{3a}: Communication distortions have a negative impact on large company profitability and growth.

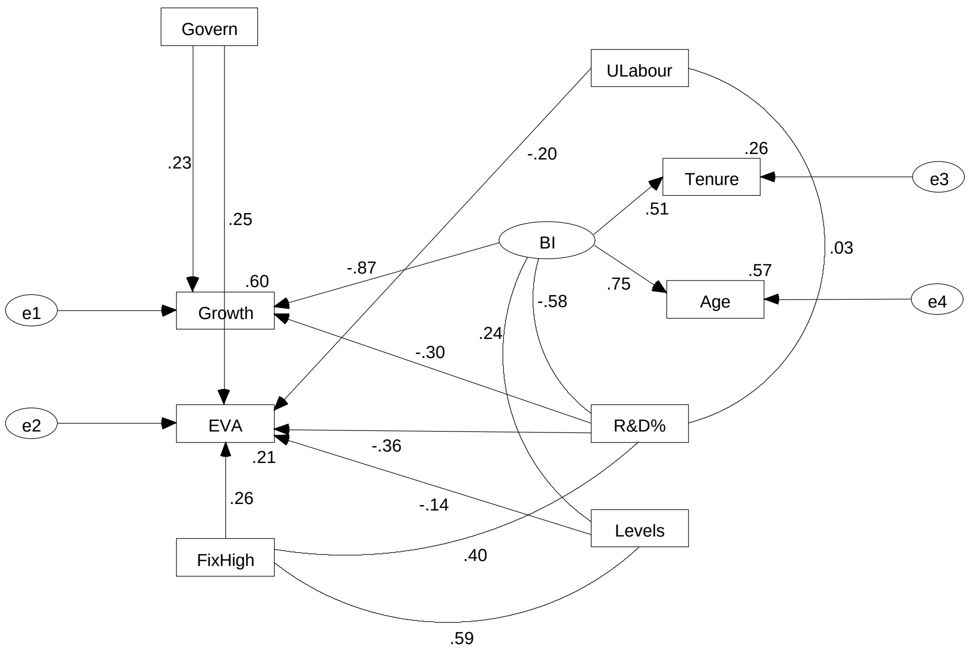
H_{3b}: Bureaucratic insularity has a negative impact on the results of large companies.

H_{3c}: Atmospheric consequences lead to relatively poorer results for large companies.

H_{3d}: Incentive limits exhibit a negative influence on the results of large companies.

As a final simplification, asset specificity was eliminated entirely. The reason for this was that theoretical support is weak and that the indicators (especially **Vertical**) may not be correctly defined or measured. The resulting path diagram is shown below. Squared multiple correlations are 0.60 for Growth (up from 0.58) and 0.21 for EVA (down from (0.24).

SIMPLIFIED COMPLETE MODEL
Adjusted for parsimony
Excludes asset specificity
Chi square = 102.022; df = 21; Chi square / df = 4.858



The chi square / df ratio is down to 4.848 from 6.896, a good fit[28] as discussed in footnote 20. The normed fit index is 0.988 and the parsimonious fit index is 0.467. The regression coefficients all have the right sign and all except one are significant at least the 5% level (the

[28] See e.g. Masten (1991, 15) which has a chi square / df ratio of 4.76.

exception having a 7% significance). Six coefficients are significant at better than the 1‰ level, one at better than the 1% level.

Regression Weights:	Estimate	S.E.	C.R.
EVA <----------- R&D%	-116.963	17.114	-6.834
Growth <---------- BI	-0.190	0.027	-6.941
Growth <-------- R&D%	-0.713	0.184	-3.873
EVA <-------- ULabour	-12.700	5.672	-2.239
Tenure <---------- BI	6.081	1.293	4.702
Age <------------- BI	1.000		
EVA <--------- Levels	-2.944	1.583	-1.859
EVA <-------- FixHigh	2.995	0.782	3.829
EVA <--------- Govern	29.739	6.501	4.575
Growth <------ Govern	0.201	0.043	4.723

In conclusion, the simplified model continuous to lend support to the theory and the hypotheses have not been disproved in any analysis. The diseconomies of scale appear to be both strong and significant.

9. DISCUSSION

9.1 SUMMARY OF FINDINGS

Diseconomies of scale appear to be real and large. The literature overview discussed the theoretical underpinnings and showed that a wide range of empirical research, quantitative and qualitative, supports pieces of the theoretical predictions. The statistical analysis section took a broader and more general approach to test the hypotheses. Nothing in the statistical analysis disproved the hypotheses. The analysis also showed that the diseconomies of scale vary in size and impact and that economies of scale and mitigating factors can be important. The findings regarding the hypotheses are summarised in the table below:

SUMMARY OF FINDINGS[29]		
Hypothesis	**Literature findings**	**Statistical findings**
H1: A company's size is not a direct predictor of its results.	Confirmed	Confirmed since less than 3% of variance explained
H2: Large companies embed diseconomies of scale of four types: communication distortion, bureaucratic insularity, atmospheric consequences, and incentive limits.	All four factors confirmed	Confirmed for CD, BI, and IL (at 1‰, 1‰, and 1% respectively). AC inconclusive
H3: Scale diseconomies impact company profitability and growth negatively.	Confirmed	Confirmed with 60% of the variance in growth, and 22% of the variance in profitability explained
H3a: Communication distortions have a negative impact on large company profitability and growth.	Confirmed	Confirmed for profitability (at 10% significance). Inconclusive for growth
H3b: Bureaucratic insularity has a negative impact on the results of large companies.	Confirmed	Confirmed for growth (at 1‰ significance). Inconclusive for growth
H3c: Atmospheric consequences lead to relatively poorer results for large companies.	Confirmed	Confirmed for profitability (at 10% significance)
H3d: Incentive limits exhibit a negative influence on the results of large companies.	Confirmed	Confirmed for growth (at 1‰ significance) and for profitability (at 1% significance)
H4: Diseconomies of scale can be mitigated by two transaction cost-related factors: organisational form and asset specificity.	Confirmed	Confirmed for organisational form versus growth and profitability (at 1‰ significance. Confirmed for asset specificity versus growth and profitability (at 1% and 5% significance, respectively)
H4a: Large M-form firms show better results than U-form firms, but this has been fully implemented in today's large companies.	Confirmed	Confirmed since organisational structure did not explain growth or profitability
H4b: High asset specificity is positively correlated with results.	Confirmed	Confirmed for both growth (at 1% significance) and profitability (at 5% significance)
H5: Scale economies are weakly related to size.	Inconclusive	Confirmed at better than 1% significance
H6: Scale economies have no impact on individual company profitability.	Inconclusive	Not confirmed. Scale economies have positive impact on profitability (at 1‰ significance)

29 For simplicity., the word "confirmed" is used here, although "not disconfirmed" is more accurate

Given the high explanatory power of the analysis it may now be possible to be prescriptive when analysing underperforming companies, especially when poor performance stems from low growth. Bureaucratic insularity at both the institutional and individual levels appears to be endemic in large companies, leading to low growth. Incentive limits influences both profitability and growth negatively. Atmospheric consequences have a moderately negative impact on growth, while communication distortion has a mild negative impact on profitability.

Scale economies can offset this to some extent and there is a tendency to find large companies in industries where scale economies are important. Furthermore, the negative effect can be mitigated by attention to governance and by reducing asset specificity (product and service breadth, and vertical depth). These factors more or less offset the diseconomies of scale for large companies, resulting in a low overall correlation between results and size.

9.2 LIMITATIONS OF RESEARCH

The research has a number of limitations. Many of the variables have not been properly operationalised. Other, more targeted research has often used better definitions but replicating these definitions here would expand the work too much. As a consequence, simple but somewhat unreliable definitions have been used.

The selection of data also has a number of limitations. Potential industry effects—while hypothesised to be small—were not included. The data was only collected for the manufacturing sector (strictly manufacturing, construction, and mining) which only represents 22% of the economy (26% of the private sector), includes less than half of all large companies, and which is in steady relative decline. No attempt was made to make international comparisons. Longitudinal comparisons proved difficult to make.

Furthermore, No competing theories or models were introduced. It may very well be that while the transaction cost economics approach to studying diseconomies of scale yields some insights, other theoretical approaches may be superior.

Finally, the statistical analyses consciously sacrificed precision at several points. The data had significant skew and kurtosis. Ordinal values were not analysed with polyserial and polychoric correlations. The model was not optimised to extract the maximum explanatory power.

Thus, this effort has only captured a small amount of the insights that may come out of more careful and ambitious studies of diseconomies of scale.

9.3 FURTHER RESEARCH

Four major avenues four further research appear to be open:

1. An attempt to prove or disprove diseconomies of scale by studying a more narrowly defined problem such as within an industry.

2. Expanding the analysis across geography and time.

3. Finding better ways to operationalise the unobserved diseconomies of scale, perhaps by using panel data from primary research.

4. Replicating this research with better statistical approaches and a large sample. In particular, industry effects are unknown (but hypothesised to be small).

10. CONCLUSION

This research demonstrates the need for research on the issue of limits of firm size, creates a model for thinking about the problem, and indicates—based on the literature survey—that there are real and quantifiable diseconomies of size. The proposal also suggests a tentative analytical model and discusses the expected outcome. Finally, it discusses the work plan for delivering the results.

The heart of the research is the TCE-based model which combines four distinct aspects of Williamson's thinking: 1) the sources of limits of firm size: communication distortion due to bounded rationality, bureaucratic insularity, atmospheric consequences due to specialisation, and incentive limits of the employment relation. 2) the importance of organisation form to reduce diseconomies 3) the mitigating influence of asset specificity on both transaction cost and production cost diseconomies, and 4) the unimportance of neo-classical scale economies at the firm level. The qualitative and quantitative analyses confirm the predictive power of the theory, except possibly for the fourth point. As such, the research may contribute to our understanding of the mechanisms behind bureaucratic failure.

There seem to be a number of real life implications of the research. First, it suggests that strategy and structure are intimately linked. Executives at large corporations have real trade-offs to make when they think about expansion. Certain growth strategies are easier to execute than others and the choice of organisation has major implications for which strategies make sense. In a sense, structure does not follow strategy; strategy and structure are intimately linked with both informing each other continuously and forever.

Second, much of the rational for mergers and acquisitions seem to be weak at best. Typical arguments for mergers are that the new, larger entity will realise economies of scale to the benefit of customers and shareholders and that growth will be accelerated with the introduction of new products and services hitherto too expensive to develop. The analysis shows that while there may be some scale economies they are likely to be offset by the scale diseconomies. Furthermore, there is no indication that the new, larger entity will innovate more and grow quicker. Instead the opposite appears to be true: innovation and growth will decline on average. This is particularly true in knowledge-intensive industries such as pharmaceuticals. This is not to imply that mergers and acquisitions do not make sense—often they do. But executives seem to be well advised to think through how to minimise the diseconomies of scale and to maximise the mitigating factors as the post-merger integration is carried out.

Third, the board of directors may want to emphasise the importance of executive renewal and the elimination of rigid processes to stimulate growth. Old, large companies with entrenched management have a fundamental dilemma. There is no indication that they can achieve above average, profitable, growth. The choice thus is to either pay out excess cash flow to shareholders (as is often done) or to try to find ways to break the bureaucratic insularity. Also, maximising the quality of governance, which is part of the board's fiduciary duties, appears to be an important lever for maximising the value of large corporations.

Fourth, successful growth through geographic expansion should be easier to achieve rather than broadening the product and service scope or increasing vertical integration. By and large, anecdotal and empirical evidence suggests that this has happened over the last 20-30 years. "Focus on the core business" and "outsourcing" have been hallmarks of restructuring programs for many years and this research shows that this often is a valid strategy.

Finally, in a world where companies increasingly try to sell solutions rather than basic products and services it is instructive to note that incentive limits are real and problematic. In businesses with indivisibilities such as team selling or large product development efforts attention should be paid on defining well-functioning incentive schemes. The superior productivity of research and development in small companies compared to large companies due to incentives tailored to individual performance, demonstrate the differential returns that will accrue to the companies that invest in structuring solid incentive schemes.

There are numerous other implications. At the end of the day it may be said that on average, large companies have neither advantages or disadvantages compared to small and medium-sized companies. However, the individual company can prosper or fade depending on how it manages the diseconomies of scale.

LIST OF REFERENCES

Adams, W, and J W Brock. 1986. *The bigness complex: Industry, labor, and government in the American economy*. New York: Pantheon.

Adelman, M A. 1978. Changes in industrial concentration. In *Monopoly power and economic performance*, edited by E. Mansfield. New York: Norton.

Arbuckle, J L, and W Wothke. 1999. *Amos 4.0 User's Guide*. Chicago: SPSS.

Armour, H O, and D J Teece. 1978. Organizational structure and economic performance: A test of the multidivisional hypothesis. *Bell Journal of Economics* 9:106-122.

Arrow, K J. 1964. Control in large organizations. *Management Science* 10 (3):397-408.

Arrow, K J. 1974. *The limits of organization*. New York: Norton.

Arrow, K J. 1983. Innovation in large and small firms. In *Entrepreneurship*, edited by J. Ronen. Lexington, Mass.: Lexington Books.

Aw, B, and G Batra. 1998. Firm size and the pattern of diversification. *International Journal of Industrial Organization* 16:313-331.

Bain, J S. 1968. *Industrial organization*. 2d ed. New York: Wiley.

Bain, J S. 1978. The general explanation of the development of concentration. In *Monopoly power and economic performance*, edited by E. Mansfield. New York: Norton.

Bane, W T, and F Neubauer. 1981. Diversification and the failure of new foreign activities. *Strategic Management Journal* 2:219-233.

Barnard, C. [1938] 1968. *The functions of the executive*. Reprint, with an introduction by K. R. Andrews. Cambridge: Harvard University Press.

Berle, A A, and G C Means. [1932] 1991. *The modern corporation and private property*. Reprint, with a new introduction by M. Weidenbaum and M. Jensen. New Brunswick, N.J.: Transaction Publishers. Original edition, New York: Harcourt, Brace & World.

Bhidé, A. 1990. Reversing corporate diversification. *Journal of Applied Corporate Finance* 3 (2):70-81.

Blau, P M, and M W Meyer. 1987. *Bureaucracy in modern society*. 3rd ed. New York: McGraw-Hill.

Bock, B. 1978. New numbers on concentration: Fears and facts. In *Monopoly power and economic performance*, edited by E. Mansfield. New York: Norton.

Bollen, K A. 1989. *Structural equations with latent variables*. New York: Wiley.

Brock, J W. 1987. Bigness is the problem, not the solution. *Challenge* 30 (July-August):11-16.

Brown, C, J Hamilton, and J Medoff. 1990. *Employers large and small*. Cambridge: Harvard University Press.
Byrne, J A. 2000. Special report: Corporate governance. *Business Week*, 24 January, 142-152.
CalPERS. 2000. *CalPERS shareowner forum* CalPERS, 1999 [cited 12 September 2000]. Available from www.calpers-governance.org.
Chandler, A D, Jr. 1962. *Strategy and structure: Chapters in the history of the American industrial enterprise*. Cambridge: MIT Press.
Chandler, A D, Jr. 1977. *The visible hand*. Cambridge, Mass.: Belknap.
Chandler, A D, Jr. 1982. Evolution of the large industrial corporation: An evolution of the transaction cost approach. *Business and Economic History* 2d series, 11:116-134.
Chandler, A D, Jr. 1990. *Scale and scope: The dynamics of industrial capitalism*. Cambridge, Mass.: Belknap.
Chandler, A D, Jr. 1992. The emergence of managerial capitalism. In *The sociology of economic life*, edited by M. Granovetter and R. Swedberg. Boulder, Col.: Westview Press.
Chandler, A D, Jr., and H Daems, eds. 1980. *Managerial hierarchies: Comparative perspectives on the rise of the modern industrial enterprise*. Cambridge: Harvard University Press.
Child, J. 1973. Predicting and understanding organization structure. *Administrative Science Quarterly* 18:168-185.
Coase, R H. 1937. The nature of the firm. *Economica* n.s., 4:386-405.
Coase, R H. 1993a. 1991 Nobel lecture: The institutional structure of production. In *The nature of the firm: Origins, evolution, and development*, edited by O. E. Williamson and S. G. Winter. New York: Oxford University Press.
Coase, R H. 1993b. The nature of the firm: Meaning. In *The nature of the firm: Origins, evolution, and development*, edited by O. E. Williamson and S. G. Winter. New York: Oxford University Press.
Coase, R H. 1993c. The nature of the firm: Origin. In *The nature of the firm: Origins, evolution, and development*, edited by O. E. Williamson and S. G. Winter. New York: Oxford University Press.
Colvin, G. 2000. America's most admired companies. *Fortune*, 21 February, 108-117, F1-F7.
Comment, R, and G A Jarrell. 1995. Corporate focus and stock returns. *Journal of Financial Economics* 37 (1):67-87.
Cooper, A C. 1964. R&D is more efficient in small companies. *Harvard Business Review* 42 (May-June):75-83.
Council of Economic Advisers. 1998. Economic report of the president. Washington, D.C.: U.S. Government Printing Office.
Crozier, M. 1964. *The bureaucratic phenomenon*. Chicago: University of Chicago Press.

D'Aveni, R A, and D J Ravenscraft. 1994. Economies of integration versus bureaucracy costs: Does vertical integration improve performance? *Academy of Management Journal* 37 (5):1167-1206.

Easterby-Smith, M, R Thorpe, and A Lowe. 1991. *Management research: An introduction*. London: Sage.

Farrell, C. 1998. Does Wharton's Jeremy Siegel have Alan Greenspan's ear? *Business Week,* July 13.

Fligstein, N. 1985. The spread of the multidivision form among large firms, 1919-1979. *American Sociological Review* 50 (June):377-391.

Florence, P S. 1933. *The logic of industrial organisation*. London: Kegan Paul.

Fortune. 1995a. The Fortune 500, 1994. *Fortune,* May 15.

Fortune. 1995b. The Fortune Global 500, 1994. *Fortune*, August 7.

Fruhan, W E. 1979. *Financial strategy: Studies in the creation, transfer, and destruction of shareholder value*. Homewood, Ill.: Irwin.

Gatty, R. 1966. Multivariate analysis for marketing research: An evaluation. *Applied Statistics* 15 (November):157-172.

Geanakoplos, J, and P Milgrom. 1991. A theory of hierarchies based on limited managerial attention. *Journal of the Japanese and International Economies* 5 (3):205-225.

Grossman, S J, and O D Hart. 1986. The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy* 94:691-719.

Hair, J F, Jr., R E Anderson, R L Tatham, and W C Black. 1998. *Multivariate data analysis*. 5th ed. Upper Saddle River, N.J.: Prentice Hall.

Hannah, L. 1996. Marshall's 'trees' and the global 'forest': Were 'giant redwoods' different? Paper read at NBER Conference, 19 October, at Cambridge, Mass.

Hardyck, C D, and L F Petrinovich. 1976. *Introduction to statistics for the behavioral sciences*. 2d ed. Philadelphia: Saunders.

Hart, O. 1995. *Firms, contracts, and financial structure.* Oxford: Clarendon Press.

Holmström, B R, and J Tirole. 1989. The theory of the firm. In *Handbook of industrial organization, vol. 1,* edited by R. Schmalensee and R. Willig. New York: North-Holland.

Ibbotson Associates. 1999. *Stocks, bonds, bills, and inflation 1999 yearbook.* Chicago: Ibbotson Associates.

Ijiri, Y, and H A Simon. 1964. Business firm growth and size. *American Economic Review* 54 (2):77-89.

IRRC. 1999. *Board practicies 1999: The structure and compensation of board of directors at S&P super 1,500 companies*. Washington, D.C.: Investor Responsibilitry Research Center.

Jensen, M C. 1986. Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review* 76:323-329.

Jensen, M C. 1988. Takeovers: Their causes and consequences. *Journal of Economic Perspectives* 2:21.

Jensen, M C. 1989. Eclipse of the public corporation. *Harvard Business Review* 67 (September-October):61-74.

Jensen, M C, and W H Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3:305-360.

Jensen, M C, and K J Murphy. 1990. CEO incentives -- It's not how much you pay, but how. *Harvard Business Review* 68 (May-June):138-153.

Jöreskog, K, and D Sörbom. 1996. *LISREL 8: User's reference guide*. Chicago: Scientific Software International.

Joskow, P L. 1993. Asset specificity and the structure of vertical relationships: Empirical evidence. In *The nature of the firm: Origins, evolution, and development*, edited by O. E. Williamson and S. G. Winter. New York: Oxford University Press.

Kelloway, E K. 1998. *Using LISREL for structural equation modeling*. Thousand Oaks, Cal.: Sage.

Kimberley, J R. 1976. Organizational size and the structuralist perspective: A review, critique and proposal. *Administrative Science Quarterly* 21:571-597.

Klier, T H. 1993. Transaction cost theory and just-in-time manufacturing: A new look at vertical integration in the United States automobile market. Ph.D. diss., Department of Economics, Michigan State University.

Knight, F H. [1921] 1971. *Risk, uncertainty, and profit*. Reprint, with a foreword by George J. Stigler. Chicago: University of Chicago Press.

Krickx, G A M. 1988. Historical evidence on the evolution of vertical exchange mechanisms: Examples from the computer systems industry. Ph.D. diss., University of California.

Kwoka, J E. 1980. Establishment size, wages, and job satisfaction: The trade-offs. In *The economics of firm size, market structure, and social performance*, edited by J. J. Siegfried. Washington, D.C.: Federal Trade Commission.

Levy, D T. 1981. The transactions cost approach to vertical integration: An empirical examination. Ph.D. diss., University of California, Los Angeles.

Lucas, R E, Jr. 1982. On the size distribution of business firms. *Bell Journal of Economics* 9:508-523.

Mahoney, J T. 1989. Organizational rationalization and innovation: Vertical integration and multidivisional organization. Ph.D. diss., University of Pennsylvania.

Mahoney, J T. 1992. The choice of organizational form: Vertical financial ownership versus other methods of vertical integration. *Strategic Management Journal* 13:559-584.

Marglin, S A. 1974. What do bosses do? The origins of hierarchy in capitalist production. *Review of Radical Political Economics* 6:33-60.
Maruyama, G M. 1998. *Basics of structural equation modeling*. Thousand Oaks, CA: Sage.
Mas-Colell, A, M D Whinston, and J R Green. 1995. *Microeconomic theory*. New York: Oxford University Press.
Masten, S E. 1982. Transaction costs, institutional choice and the organization of production. Ph.D. diss., University of Pennsylvania.
Masten, S E. 1984. The organization of production: Evidence from the aerospace industry. *Journal of Law and Economics* 27 (October):403-417.
Masten, S E, J W Meehan, Jr., and E A Snyder. 1989. Vertical integration in the US auto industry: A note on the influence of transaction specific assets. *Journal of Economic Behavior and Organization* 12:265-273.
Masten, S E, J W Meehan, Jr, and E A Snyder. 1991. The costs of organization. *Journal of Law, Economics and Organization* 7 (1):1-25.
McAfee, R P, and J McMillan. 1995. Organizational diseconomies of scale. *Journal of Economics & Management Strategy* 4 (3):399-426.
McConnell, J L. 1945. Corporate earnings by size of firm. *Survey of Current Business* 25 (May):6-12.
McGahan, A M, and M E Porter. 1997. How much does industry matter, really? *Strategic Management Journal* 18 (Summer):15-30.
Merton, R K. 1957. *Social theory and social structure*. Revised and enlarged ed. Glencoe, Ill.: Free Press.
Monsen, R J, and A Downs. 1965. A theory of large managerial firms. *Journal of Political Economy* 73:221-236.
Monteverde, K, and D J Teece. 1982. Supplier switching costs and vertical integration in the automobile industry. *Bell Journal of Economics* 13:206-213.
Mueller, W F, and L G Hamm. 1974. Trends in industrial market concentration, 1947 to 1970. *Review of Economics and Statistics* 56:511-520.
Nelson, R, and S G Winter. 1982. *An evolutionary theory of economic change*. Cambridge, Mass.: Belknap.
North, D C. 1985. Transaction costs in history. *Journal of European Economic History* 14 (3):557-576.
North, D C. 1987. Institutions, transaction costs and economic growth. *Economic Inquiry* 25 (3):419-428.
North, D C. 1992. *Transaction costs, institutions, and economic history, Occasional papers*. San Francisco: ICS Press.
North, D C, and R P Thomas. 1973. *The rise of the western world: A new economic history*. Cambridge: Cambridge University Press.
North, D C, and J J Wallis. 1994. Integrating institutional change and technical change in economic history: A transaction cost approach.

Journal of Institutional and Theoretical Economics (Zeitschrift für die gesamte Staatswissenschaft) 150 (4):609-24.

Nutter, G W. 1951. *The extent of enterprise monopoly in the United States, 1899-1939*. Chicago: University of Chicago Press.

Olson, M. 1982. *The rise and decline of nations: Economic growth, stagflation, and social rigidities*. New Haven: Yale University Press.

Palepu, K. 1985. Diversification strategy, profit performance and the entropy measure. *Strategic Management Journal* 6:239-255.

Penrose, E. [1959] 1995. *The theory of the growth of the firm*. 3d ed. New York: Oxford University Press.

Peters, T. 1992. Rethinking scale. *California Management Review* 35 (1):7-29.

Pondy, L R. 1969. Effects of size, complexity, and ownership on administrative intensity. *Administrative Science Quarterly* 14:47-60.

Pratt, J W, and R J Zeckhauser. 1985. *Principals and agents: The structure of business*. Boston: Harvard Business School Press.

Pugh, D S, D J Hickson, C R Hinings, and C Turner. 1969. The context of organization structures. *Administrative Science Quarterly* 14:91-114.

Qian, Y. 1994. Incentives and loss of control in an optimal hierarchy. *Review of Economic Studies* 61:527-544.

Ramanujam, V, and P Varadarajan. 1989. Research on corporate diversification: A synthesis. *Strategic Management Journal* 10 (6):523-551.

Rasmusen, E, and T Zenger. 1990. Diseconomies of scale in employment contracts. *Journal of Law, Economics, and Organization* 6 (1):65-92.

Richardson, G B. 1972. The organisation of industry. *Economic Journal* 82 (September):883-896.

Riordan, M H, and O E Williamson. 1985. Asset specificity and economic organization. *International Journal of Industrial Organization* 3:365-378.

Robertson, D H. 1923. *The control of industry*. London: Nisbet.

Robinson, A. 1934. The problem of management and the size of firms. *Economic Journal* 44 (June):242-257.

Rubin, P H. 1990. *Managing business transactions*. New York: Free Press.

Rumelt, R P. 1974. *Strategy, structure, and economic performance*. Cambridge: Harvard University Press.

Rumelt, R P. 1982. Diversification strategy and profitability. *Strategic Management Journal* 3:359-369.

Rumelt, R P. 1991. How much does industry matter? *Strategic Management Journal* 12 (3):167-185.

Rumelt, R P, and R C Wensley. 1981. In search of the market share effect. *Academy of Management Proceedings* 41:2-6.

Runkel, P J, and J E McGrath. 1972. *Research on human behavior: A systematic guide to method*. New York: Rinehart & Winston.

Scherer, F M. 1976. Industrial structure, scale economies, and worker alienation. In *Essays on industrial organization in honor of Joe S. Bain,* edited by R. T. Masson and P. D. Qualls. Cambridge, Mass.: Ballinger.

Scherer, F M, and D Ross. 1990. *Industrial market structure and economic performance*. 3rd ed. Boston: Houghton Mifflin.

Schmalensee, R. 1985. Do markets differ much? *American Economic Review* 75 (3):341-351.

Schmookler, J. 1972. The size of firm and the growth of knowledge. In *Patents, invention, and economic change,* edited by Z. Griliches and L. Hurwicz. Cambridge: Harvard University Press.

Schumacher, E F. 1989. *Small is beautiful: Economics as if people mattered.* Perennial Library ed. New York: HarperCollins.

Shelanski, H A, and P G Klein. 1995. Empirical research in transaction cost economics: A review and assessment. *Journal of Law, Economics, and Organization* 11 (2):335-361.

Silver, M, and R Auster. 1969. Entrepreneurship, profit, and limits on firm size. *Journal of Business* 42 (April):277-281.

Simon, H A. [1947] 1976. *Administrative behavior*. 3rd ed. New York: Free Press.

Simon, H A, and C P Bonini. 1958. The size distribution of business firms. *American Economic Review* 48 (4):607-617.

Stigler, G J. 1958. The economies of scale. *Journal of Law and Economics* 1:54-71.

Stigler, G J. 1974. The case against big business. In *Monopoly power and economic performance,* edited by E. Mansfield. New York: Norton.

Stinchcombe, A L. 1965. Social structure and organizations. In *Handbook of organizations,* edited by J. G. March. Chicago: Rand McNally.

Teece, D. 1982. Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior and Organization* 3:39-63.

Teece, D J. 1976. *The multinational corporation and the resource cost of international technology transfer*. Cambridge, Mass.: Ballinger.

Teece, D J. 1980. Economies of scope and the scope of the enterprise. *Journal of Economic Behavior and Organization* 1:223-247.

Teece, D J. 1981. Internal organization and economic performance: An empirical analysis of the profitability of principal firms. *Journal of Industrial Economics* 30 (2):173-199.

Tsokhas, K. 1986. *Beyond dependence: Companies, labour processes and Australian mining.* Melbourne: Oxford University Press.

Walker, G, and D Weber. 1984. A transaction cost approach to make-or-buy decisions. *Administrative Science Quarterly* 29:373-391.

Walker, G, and D Weber. 1987. Supplier competition, uncertainty and make-or-buy decisions. *Academy of Management Journal* 30:589-596.

Williamson, O E. 1967. Hierarchical control and optimum firm size. *Journal of Political Economy* 75:123-138.

Williamson, O E. 1970. *Corporate control and business behavior: An inquiry into the effects of organization form on enterprise behavior*. Englewood Cliffs, N.J.: Prentice-Hall.

Williamson, O E. 1975. *Markets and hierarchies: Analysis and antitrust implications*. New York: Free Press.

Williamson, O E. 1984. The incentive limits of firms: A comparative institutional assessment of bureaucracy. *Weltwirtschaftliches Archiv* 120 (4):736-763.

Williamson, O E. 1985. *The economic institutions of capitalism*. New York: Free Press.

Williamson, O E. 1986. *Economic organization: Firms, markets and policy control*. New York: New York University Press.

Williamson, O E. 1988. Corporate finance and corporate governance. *Journal of Finance* 43 (3):567-591.

Williamson, O E. 1993. Opportunism and its critics. *Managerial and Decision Economics* 14:97-107.

Williamson, O E. 1996. *The mechanisms of governance*. New York: Oxford University Press.

Zarnowitz, V. 1999. *How to date business cycles: Real GDP vs. coincident index* [pdf]. Foundation for International Business and Economic Research, 23 July, 1998 [cited 8 December 1999]. Available from http://www.emgmkts.com/fiber/fiber.htm.

Zenger, T R. 1989. Organizational diseconomies of scale: pooling vs. separating labor contracts in Silicon Valley. Ph.D. diss., University of California, Los Angeles.

Zenger, T R. 1994. Explaining organizational diseconomies of scale in R&D: Agency problems and the allocation of engineering talent, ideas, and effort by firm size. *Management Science* 40 (6):708-729.