# LIMITS OF FIRM SIZE

## Proposal for a Test of the Transaction-Cost Explanation of Diseconomies of Scale

**A research proposal submitted in partial fulfilment of the requirements for the degree of Doctor of Business Administration**

**by**

**Staffan Canbäck**

**Henley Management College**
**Brunel University**

**August 1998**

# CONTENTS

# 1. INTRODUCTION

This research proposal suggests a test of a proposition by Williamson (1975, 117–131): 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" (p. 117). The research aims to test this proposition through a survey of the literature on diseconomies of scale and an empirical analysis of the relative profitability of differently sized manufacturing companies in the United States.

Why are large companies so small? Why doesn't General Motors make diapers? No coherently articulated reason exists today for why the largest business organisations do not have ten or twenty million employees rather than a few hundred thousand.

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

---

This research proposal is an intermediate work product. It contains a number of theoretical, factual and grammatical errors later corrected in the finished DBA thesis "Bureaucratic Limits of Firm Size: Empirical Analysis Using Transaction Cost Economics" (Canbäck 2002).

This effort builds on the original research made in the subject area. Specifically, it will test whether 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 expected outcome is a quantification of the impact of the claimed diseconomies on relative shareholder value, a perspective on the nature of the diseconomies of scale, and a set of hypotheses about how these diseconomies may be minimised by executives.

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 neoclassical 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.

The purpose of the research is to create an application that can be used by academics, consultants, and managers to help delineate strategic and organisational choices and to derive their implications.

# 2. RESEARCH OBJECTIVES

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

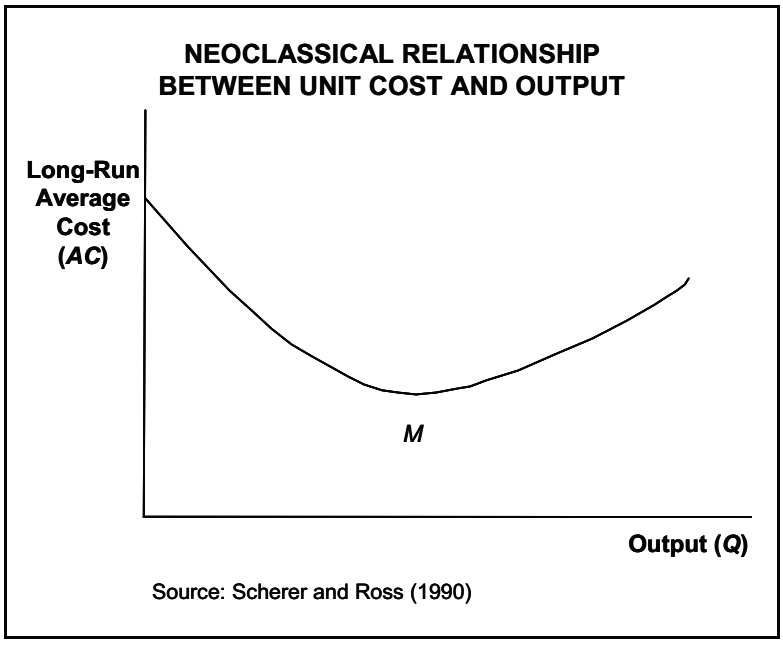
## 2.1 PROBLEM DEFINITION

Knight ([1921] 1964, 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 diseconomies of scale, but there appear to be no systematic studies of the 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 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 (1978, 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 Section 3.2.

The neoclassical way to illustrate economies and diseconomies of scale is with a cost curve such as the one in Figure 1 (e.g., Scherer and Ross 1990, 101).

*Figure 1. Neoclassical Relationship between Unit Cost and Output*

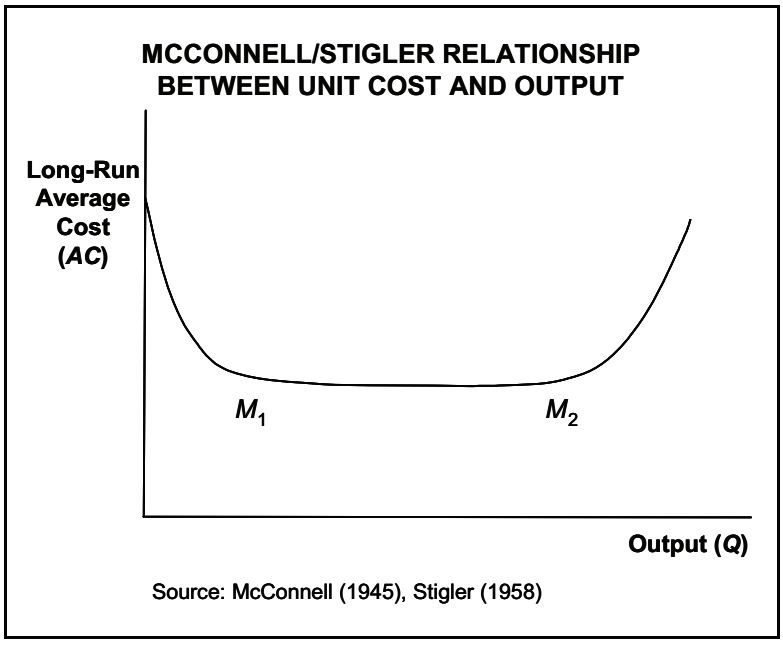



As production increases, the average unit cost *AC* decreases due to scale economies. At a certain point, *M*, the scale economies are exhausted while diseconomies of scale start to impact the unit cost. As output increases, the

unit cost increases. Thus, a profit maximising firm should strive for an output at the optimum output *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), reproduced in Figure 2, are typical.

*Figure 2. McConnell/Stigler Relationship between Unit Cost and Output*




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 $M_2$. Neoclassical theory does not provide a satisfactory answer. As Simon ([1947] 1976, 292) put it: "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 whether transaction cost economics can explain diseconomies of scale, as exhibited in lower relative performance of large firms versus smaller firms in the same industry, and what drives these diseconomies. The description of hypotheses is found in Section 4.2.

## 2.2 IMPORTANCE OF RESEARCH

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

Limits of firm size are often referred to in the literature, but seldom studied (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 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 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 evolved over time from being a general theory for

understanding industrial organisation, to mainly being a tool for analysing vertical integration. The suggested research breaks with this tradition.

Limits of firm size are also a real and difficult problem for business managers. The cost of suboptimal (i.e., too large) size is probably significant. For example, it has been estimated that 25 per cent of the operating budget of a large company is slack (Riahi-Belkaoui 1994, 35–64) due to some of the diseconomies of scale discussed in Chapters 4 and 5.

# 3. DIMENSIONS OF FIRM SIZE

This chapter defines size and shows the trends in size of firms in the US 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), holds that the boundary of the firm is where the internal planning mechanism is superseded by the price mechanism. In most cases the firm will be equivalent to the corporation with this definition. The most important exception is a corporation where divisions are totally self-contained profit centres. In this case the parent company is not a firm because 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 will be considered a firm. Still, this definition is usually equivalent to Coase's definition

because 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). McDonald's, for example, is considerably larger than the ownership definition indicates because it also consists of a network of thousands of franchisees (Rubin 1990, 134–144).

The fourth definition is the firm's sphere of influence. This includes alliance partners, first and second tier suppliers, etc. Toyota is an example (Williamson 1985, 120–122). Toyota employs around 200,000 people directly, but its sphere of influence is perhaps over more than 1 million people.

This paper uses the ownership definition. It relates closely to Coase's definition and most statistics use it. Thus, a firm is a corporation.

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 because 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,

that is, the sum of factor inputs (or revenue less purchased goods). This metric gives a precise measure of activity, but is usually not available by company. Number of employees is the most widely used measure of size, with more than 80 per cent 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 because 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 is value added, but for practical reasons number of employees and assets can be used. The definitions are summarised in Table 1.

*Table 1. Definitions of Firm Size*

| DEFINITIONS OF FIRM SIZE | | | | |
|---|---|---|---|---|
| **Size Metric** | **Internal Planning (Coase)** | **Ownership** | **Network** | **Sphere of Influence** |
| Revenue | | | | |
| Value added | | | | |
| Employees | | | | |
| Assets | | | | |

## 3.2 TRENDS IN FIRM SIZE

The US economy is used as the basis for analysis because it is the largest and most competitive economy in the world. Within this economy, the research focuses on the manufacturing sector.[1]

Large manufacturing firms play a major role in the US economy. The Fortune industrial 500 companies control more than 50 per cent 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). They have, however, reduced their share of employment from around 60 to around 50 per cent in the manufacturing sector between 1979 and 1994. Moreover, as a share of the total US 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

[1] Alternative approaches would be to look at the global manufacturing sector or the total US private sector, or both. Statistics on the global manufacturing sector are not yet reliable, and the non-manufacturing sectors are still often highly regulated.

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[2] 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 per cent in 1947 to 37 per cent in 1982.

Bain (1968, 87) calculated that the assets controlled by the largest 200 nonfinancial companies was around 57 per cent of total nonfinancial assets[3] in 1933. He also estimated that the 300 largest nonfinancial companies accounted for 55 per cent of nonfinancial assets in 1962. Based on this, the top 200 companies accounted for approximately 50 per cent of nonfinancial assets in 1962, using this author's estimate of the assets

---

[2] Although there have been significant changes within industries.

[3] 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."

controlled by the 100 smallest companies in the sample. Data from 1994 show the same ratio to be around 40 per cent. 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 per cent). 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 author replicated the analysis for 1994 and found the equivalent number to be 45 per cent 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 per cent of total stock market capitalisation, in 1998 they make up 19.5 per cent. The 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.

In sum, evidence given so far in the cited statistics and studies 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 US corporate sector and the world corporate sector). This indicates that there is a limit to firm size and that this limit may be decreasing in relative terms.

# 4. THEORETICAL FRAMEWORK AND HYPOTHESES

Transaction cost economics aims 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 such as communications failure, managerial isolation from reality, employee alienation, and misalignment of incentives appear. Thus, a firm will reach a size where the benefit from the last internalised transaction is offset by the bureaucratic diseconomies. Two conditions offset 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. Coase's original article "The Nature of the Firm" (1937) established the basic framework. Chapter 7 ("Limits of Vertical Integration and Firm Size") in Williamson's book *Markets and Hierarchies* (1975) identifies the nature of limits of size. Chapter 6 ("The Limits of Firms: Incentive and Bureaucratic

Features")[4] 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 neoclassical 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 . . . [is reached] when the costs of organizing additional transactions within the firm [exceed] the

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

costs of carrying out the same transactions through the market" (Coase 1993b, 48).

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, because planning will never be perfect, and the cost of demotivation, because 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 that potentially explains why firms have size limits. However, this is only true if there are decreasing returns to the entrepreneur function (Penrose [1959] 1995, 98). 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 incentive structure is non-trivial. Thus, the combination of small firms into a large firm will

never have the same operating characteristics as if they are independent 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 TCE. He identified four main categories of diseconomies of scale: communications distortion due to bounded rationality, bureaucratic insularity, atmospheric consequences due to specialisation (p. 126), and incentive limits of the employment relation (p. 129).

**Communications 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 managers 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 (pp. 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

$\alpha$ = 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

$\beta$ = 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 argument is similar to agency theory (Jensen and Meckling 1976; Jensen 1989) which holds that corporate management will tend to overemphasise size over profitability and will keep excess cashflow within the firm rather than distribute it to a more efficient capital market (a lengthier comparison of agency theory and transaction cost economics is found in Section 5.1.1). The consequences are that large firms tend to more easily accept organisational slack and 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.[5]

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

[5] 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 will 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 will be conscious manipulation of information to further individual or sub-unit goals (1975, 122–124).

- Asset utilisation will be lower because high-powered market incentives do not exist (1985, 137–138).

- Transfer prices will not reflect reality and cost determination will suffer (1985, 138–140).

- Research and development productivity will be lower (1985, 141–144).

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

The links in Table 2 seem reasonable between the limiting factors and the outcomes.

*Table 2. Link between Limits of Firm Size Sources and Outcomes*

| LINK BETWEEN LIMITS OF FIRM SIZE SOURCES AND OUTCOMES | | | | |
|---|---|---|---|---|
| | Sources | | | |
| Outcomes | Communications 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 will exhibit lower relative profitability than a smaller firm with the same product and market mix will.

### 4.1.3 Offsetting Influences on the Limits of Firm Size

While the categories discussed in the previous section theoretically impose limits of firm size, there are two offsetting influences that tend to mitigate the diseconomies of scale. Each of these influences is central to TCE 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 offsetting influences into account.

**Asset specificity.** There is a vast literature on vertical and lateral integration applications of TCE and the purpose here is not to review this at length. The theoretical argument is summarised in Williamson (1975, 43–67). Mahoney (1989; 1992) provided overviews of theoretical and empirical work on vertical integration problems. Grossman and Hart (1986) and Teece (1976; 1980; 1982) illustrated 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.[6] Opportunistic behaviour can be expected if the asset is part of a market transaction under this condition. An example is if a supplier invests in specific tooling equipment dedicated to one customer. Over time, the customer will be able to put pressure on the vendor because the

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

vendor has no alternative use for its investment. The vendor will be willing to accept a price down to the variable cost of production to cover some 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 will lead to more bureaucratic transactions because 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 once again a tendency to manage the transaction through bureaucracy because the repetitive contracting cost will be higher than the bureaucratic cost. While uncertainty and frequency play some role in creating transaction costs, Williamson considered asset specificity as 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).

Neoclassical 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.

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. Based on Chandler's (1962; 1977) 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[7] (U) form. Thus, large firms organised according to the M-form should, *ceteris paribus*, be more profitable than U-form firms should.

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

## 4.2 TCE-BASED HYPOTHESES OF FIRM SIZE LIMITS

It is now possible to formulate five testable hypotheses based on the TCE-based model developed above.

As was shown in Chapter 3, the average size of large[8] 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 beyond a certain point (p. 71). These diseconomies are exhibited through lower future relative profitability and/or slower relative growth of the largest firms relative to smaller competitors, *ceteris paribus* (such as risk and financial leverage). The combination of these two factors is captured in the relative market value of a firm relative to its invested capital (e.g., Rappaport 1998).

**$H_1$: The relative value of large firms is lower than that of small firms.**

**$H_2$: Profitability and growth have a positive influence on firm value.**

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

**$H_3$: The profitability and growth of a firm is negatively correlated with the firm's size.**

A company's costs are usefully divided into two categories: production costs and transaction costs. Production costs are defined as the costs of combining inputs to produce output through a transformation of resources. Thus, they are all the costs that are associated directly with productive activities (Masten 1982) such as manufacturing, logistics, and product development. We usually associate these costs with economies of scale and scope and they arguably[9] decline with size. More importantly, "It can be argued and has been argued that firm and market are identical in production cost respects" (Riordan and Williamson 1985, 369). Transaction costs, on the other hand, are those costs associated with organising economic activity.[10] They are the costs of negotiating, monitoring, and enforcing contracts between and within firms (Alston and Gillespie 1989, 193). They thus vary with organisational form (Masten 1982, 47). Or as Arrow (1983b) put it, "The distinction between transaction costs and production costs is that the former can be varied by a change in the mode of resource allocation, while the latter depend only on the technology and tastes, and would be the same in all economic systems."

[9] This author has not been able to find any evidence in the literature of rising production costs as size increases, except for transportation costs (Scherer and Ross 1990, 106-108), and scarce resources (Eatwell, Milgate and Newman 1987, 995)

[10] It has been estimated that at least 45 per cent of the gross national product in a developed economy are transaction costs (Wallis and North 1986).

Transaction costs are elusive and can only be observed indirectly: "Empirical research on transaction cost matters almost never attempts to measure such costs directly. Instead, the question is whether organisational relations (contracting practices; governance structures) line up with the attributes of transactions as predicted by transaction cost reasoning or not" (Williamson 1985, 22). It is therefore important to identify the underlying sources of diseconomies and offsetting mechanisms, as was done in Section 4.1.

**$H_4$: The size of firms is determined by diseconomies arising from communications distortion due to bounded rationality, bureaucratic insularity, atmospheric consequences due to specialisation, and incentive limits of the employment relation.**

**$H_5$: Diseconomies of scale are offset by two factors: asset specificity and M-form organisation.**

In summary, the value of a firm ultimately depends on two counteracting forces. On the one hand, four size-related factors determine the firm's size limit. If these factors are important, then (all other things equal) the larger firm will have lower relative value than the smaller firm. On the other hand, there are offsetting factors. First, when vertical integration (asset specificity, uncertainty, or frequency of transaction) is beneficial, then the

firm will tend to internalise more transactions and be larger than otherwise. The more integrated firm will carry a higher relative valuation. Second, a firm that uses the M-form will be more profitable than a U-form company will, or it can be larger with the same profitability. Table 3 summarises the model.

*Table 3. TCE-Based "Limits of Firm Size" Model*

| TCE-BASED "LIMITS OF FIRM SIZE" MODEL | | | | | | |
|---|---|---|---|---|---|---|
| Relative Value of Large Firm | Sources of Limits of Firm Size | | | | Offsets | |
| | Communi-cations 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 because 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. Child (1973) and Pugh et al. (1969), 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. 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 (p. 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 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 coordination 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: "The central problem is not how to organise to produce efficiently, but how to organise to make decisions" (p. 292). 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.

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.

Jensen has deepened and extended these arguments over the last 25 years (e.g., Jensen and Meckling 1976; Jensen 1986, 1988, 1989; Jensen and Murphy 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 or 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 (1985, 20–21) and Mahoney (1992, 566) argued that agency costs correspond to the *ex post* costs of TCE. Meanwhile, TCE works with both *ex ante* and *ex post* costs.[11] Table 4 compares the two theories.

*Table 4. Comparison of Agency Costs and Transaction Costs*

| COMPARISON OF AGENCY COSTS AND TRANSACTION COSTS | | |
|---|---|---|
| **Transaction Costs** | | **Agency Costs** |
| **Ex ante** | **Ex post** | |
| Search and information costs<br><br>Drafting, bargaining and decision costs<br><br>Safeguarding costs | Monitoring and enforcement costs<br><br>Adaptation and haggling costs<br><br>Bonding costs<br>Maladaptation costs | Monitoring expenditures of the principal<br><br>Bonding expenditures by the agent<br><br>Residual losses |

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

[11] In contrast, Williamson (1988, 570) argued that agency costs correspond to TCE's *ex ante* costs.

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 per cent lower in large establishments[12] than in small establishments (p. 109) while compensation is more than 15 per cent 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.

Brown, Hamilton and Medoff (1990) found that large firms pay a wage premium of 10–15 per cent 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).

[12] More than 500 employees.

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 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.

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 because they can offer better (more tailored) incentives. 2) Engineers in small companies have a better attitude towards cost. 3) The internal communication and coordination is more effective in small companies. These reasons match three of

Williamson's four sources of diseconomies: communications distortions, atmospheric consequences, and incentive limits.

Later work has confirmed Cooper's anecdotal evidence both theoretically and empirically. Arrow (1983a) 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, they induce more effort from the employees, and their 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.

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 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 (Table 5). A question is if rigidity (and/or organisational age) should be introduced as a fifth source of diseconomies of scale. Here it is classified as most closely associated with atmospheric consequences and communications distortions.

*Table 5. Sources of Limits of Firm Size*

| SOURCES OF LIMITS OF FIRM SIZE | | | |
|---|---|---|---|
| **Communications Distortion** | **Bureaucratic Insularity** | **Atmospheric Consequences** | **Incentive Limits** |
| Arrow (1974): Specialisation leads to poor communication<br><br>Arrow (1983a): Information loss in R&D<br><br>Blau and Meyer (1987): Excessive rigidity<br><br>Cooper (1964): R&D coordination<br><br>Crozier (1964): Rigidity<br><br>Geanakoplos and Milgrom (1991): Information signal delays<br><br>McAfee and McMillan (1995): Lower efficiency<br><br>Simon ([1947] 1976): Processing bottlenecks | Brock (1987): Risk aversion<br><br>Child (1973): Insularity<br><br>Jensen (e.g., 1986): Firms larger than optimum<br><br>Monsen and Downs (1965): Different owner/manager objectives<br><br>Pondy (1969): Increase in administration<br><br>Pugh et al. (1969): Insularity from reality<br><br>Schmookler (1972): Understanding market needs in R&D<br><br>Stinchcombe (1965): Perpetuation of organisation form<br><br>Williamson (1996): Bureaucratic rigidity | Arrow (1974): Rigidity to change<br><br>Blau and Meyer (1987): Excessive rigidity<br><br>Brown, Hamilton and Medoff (1990): Unexplained wage differential<br><br>Child (1973): Insularity<br><br>Cooper (1964): R&D cost control<br><br>Crozier (1964): Alienation<br><br>Pugh et al. (1969): Insularity from reality<br><br>Qian (1994): Monitoring costs/inadequate effort levels<br><br>Scherer (1976): Low job satisfaction in large firms<br><br>Schmookler (1972): R&D cost consciousness; Climate for innovation<br><br>Schumacher (1989): Motivation | Blau and Meyer (1987): Excessive rigidity<br><br>Cooper (1964): R&D incentives<br><br>Crozier (1964): Rigidity<br><br>Peters (1992): Low productivity in R&D<br><br>Rasmusen and Zenger (1990): Employment contracts<br><br>Schmookler (1972): Quality of R&D employees<br><br>Silver and Auster (1969): Limits to entrepreneurship<br><br>Zenger (1989, 1994): Employment contract disincentives in R&D |

## 5.2 OFFSETTING 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 (Williamson 1989, 150). 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.

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, for example, 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 (Ward 1976; Bane and Neubauer 1981).

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 whether 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 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

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 coordination 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 per cent of large companies,[13] to 75 per cent between 1919 and 1979. He estimated that the spread of the multidivisional form is mainly due to the increase of multiproduct strategies, in line with

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

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-per cent 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 diseconomies of scale.

### 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

Table 6 contains a summary of the support found in the literature for the offsetting factors.

*Table 6. Offsets to Limits of Firm Size*

| OFFSETS TO LIMITS OF FIRM SIZE | | |
|---|---|---|
| **Asset Specificity** | **Organisation Form** | **Other** |
| Bane and Neubauer (1981): Market diversity reduces profitability<br><br>Masten (1984), Masten et al. (1989, 1991), Monte-verde and Teece (1982), Joskow (1993), Klier (1993), Krickx (1988): Asset specificity more important than uncertainty and frequency<br><br>Peters (1992): Vertical integration is bad<br><br>Rumelt (1974): Product diversity<br><br>Teece (1976), Tsokhas (1986): Asset specificity influences geographic reach<br><br>Walker and Weber (1984, 1987): Volume uncertainty weak factor<br><br>Ward (1976): Market diversity | Armour and Teece (1978): M-form increases ROE<br><br>Chandler (e.g., 1962): M-form alleviates coordination and control problems<br><br>Fligstein (1985): Multiproduct coordination favours M-form<br><br>Peters (1992): Decentralisation is critical<br><br>Teece (1981): M-form firms are significantly better performers than U-form firms | Bhide (1990): Internal capital markets not efficient<br><br>Comment and Jarrell (1995): Financial synergies not relevant |

## 5.3 COMMENT ON NEOCLASSICAL SCALE ECONOMIES IN THE PRODUCTION FUNCTION

Neoclassical scale (or scope) economies should not be incorporated into the model because they are independent of the form of organisation beyond the point where technological indivisibilities are captured within the firm, according to transaction cost economics. 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 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.

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.[14] He then noted that these firm level scale economies are elusive, if they exist at all. His research indicated that "where economies of the multiplant 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 concentrated than necessary' for efficiency—and the larger firms bigger than necessary" (p. 94).

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

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[15] 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 1978; 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 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

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

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).

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.

## 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 and Wensley (1981) applied the methodology to manufacturing firms and found that industry effects accounted for 8 per cent of explained profitability (63 per cent of total profits). McGahan and Porter (1997) found a 19 per cent industry effect for all sectors of the economy and a similar effect as Rumelt (9 per cent 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. The studies do not however, say anything about firm size and its relationship with industry.

## 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 because 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 literature authorities' 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.

Table 7 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 author's interpretation of the literature review.

*Table 7. Extended TCE-Based "Limits of Firm Size" Model*

| EXTENDED TCE-BASED "LIMITS OF FIRM SIZE" MODEL | | | | | | |
|---|---|---|---|---|---|---|
| | **Sources of Limits of Firm Size** | | | | **Offsets** | |
| **Value** | **Communi-cations Distortion** | **Bureau-cratic Insularity** | **Atmos-pheric Conse-quences** | **Incentive Limits** | **Integration Conditions** | **Organi-sational 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:<br>Small (<1000)<br>Medium<br>Large (>10,000) | <br>Strong<br>Strong<br>Strong | <br>Weak<br>Moderate<br>Strong | <br>Weak<br>Moderate<br>Strong | <br>Strong<br>Weak<br>Weak | <br>Strong<br>Strong<br>Strong | <br>Weak<br>Moderate<br>Strong |
| Context | Diverse firms; Unpredicta-bility | Manage-ment/ board relation | R&D intensive | R&D intensive | | Diverse firms |

This model 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. ANALYTICAL APPROACH

This chapter gives a general impression of the analytical approach but is not intended as a full-blown description of all the details. As the research progresses both the operationalisation of the model and the details of the statistical model will evolve.

## 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 because 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 because the factors are easily 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 diseconomies of scale. However, generalised studies on for example the profit impact of M-form organisation, and the link 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 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 for example North and Thomas (1973) and North (1985, 1987, 1992) merged 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 OPERATIONALISATION OF MODEL

The five hypotheses

**$H_1$: The relative value of large firms is lower than that of small firms.**

**$H_2$: Profitability and growth have a positive influence on firm value.**

**$H_3$: The profitability and growth of a firm is negatively correlated with the firm's size.**

**$H_4$: The size of firms is determined by diseconomies arising from communications distortion due to bounded rationality, bureaucratic insularity, atmospheric consequences due to specialisation, and incentive limits of the employment relation.**

**$H_5$: Diseconomies of scale are offset by two factors: asset specificity and M-form organisation.**

can be tested in the following manner.

**$H_1$**, **$H_2$** and **$H_3$** are true if the largest company in each industry exhibits lower profitability and grows at the same pace as the smaller companies in the same industry over a time period, or if the largest company has similar profitability but grows slower than the smaller companies in the same industry. Another way of expressing this is that total stock market return (share appreciation plus dividends) should be lower for the largest company in each industry. An additional test is to check the two largest companies, three largest, etc., against the rest of the industry. If industry is not important in the manufacturing sector (as discussed in Section 5.4) then the same should be true for the whole manufacturing sector.

**$H_4$** and **$H_5$** are true if the value of companies can be significantly explained by the relative magnitude of the four sources of diseconomies and the two

offsetting mechanisms, with consideration given to industry (as discussed in Section 5.4).

Thus, the following needs to be operationalised:

**Value.** The market value-added concept developed by Rappaport (1998) and Stern Stewart & Co (Stewart 1991) is the best measure. This may have to be calibrated against industry returns, although this is unlikely as was discussed in Section 5.4.

**Profits.** Accounting measures of profitability such as return of equity are not ideal because they do not take risk into account. Instead, the best profit measure is to calculate return of equity less cost of equity for each company.

**Growth.** Growth is calculated as annual compounded growth of size.

**Size.** As indicated in Section 3.1, size is best measured as value-added, while number of employees or assets are reasonable substitutes in the manufacturing sector.

**Communications distortion.** The literature (e.g., Child 1973; Williamson 1967) indicates that the best way to operationalise communications distortion is to use number of vertical levels in the organisation.

**Bureaucratic insularity.** The only measure available is the *Business Week* annual ranking of board/management relations, that is, a measure of the quality of governance. Their methodology is basically sound and can be replicated for the full sample. The alternative is to create a new, qualitative, index and measure each company based on analyst reports, press clippings and telephone interviews. Such an index should reflect how entrenched management is.

**Atmospheric consequences.** The best measure should, given Scherer's (1976) research, be to divide the average pay of the largest company in the industry sector with the same sector's average pay.

**Incentive limits.** Incentive limits apply mainly to employees working on indivisibilities (see discussion in Section 5.1). A good proxy is the R&D spending as a share of revenue. Another possibility is the average industry pay compared to average manufacturing sector pay because higher pay is indicative of more qualified work. Another idea is to use the MBA/total employee ratio, although this may be hard to find. Finally, labour cost as a

share of total value-added is in line with Silver and Auster's (1969) argument from agency theory.

**Asset specificity.** Asset specificity has been operationalised many times in the past. Product breadth is usually proxied by a Herfindahl index of SIC codes. Rumelt (1982) used a more sophisticated definition, but it may be difficult to replicate his definition. Geographic reach should be operationalised as the ratio of international revenue to total revenue. Vertical integration is often measured as value added over sales. There are some objections to this approach but it should suffice here. A question is if uncertainty should be included. If so, the best proxy is to measure the standard deviation of the sales volatility because volume volatility is shown to be the most important contributor to uncertainty.

**Organisation form.** The first choice is to follow the lead of Armour and Teece and use only two possibilities: M-form or U-form. If it is necessary to increase precision then Williamson's (1975, 152–154) classification with seven categories[16] can be used.

**Industry.** Lastly, it is necessary to test the influence of which industry a company operates in. The Fortune 500 is normally divided into 20–30

[16] Unitary, holding company, multidivisional, transitional multidivisional, corrupted multidivisional, mixed, matrix.

industries, and other statistics have a similar number of divisions. It may be possible to reduce this further, to perhaps ten industries, because some industries have similar characteristics. An example is electrical manufacturing and mechanical manufacturing. The definition of an industry will most likely be based on the common sense definitions in Fortune (which reflect a structure–conduct–performance way of defining industry (Bain 1968)), but could also be based on SIC codes.

Table 8 summarises the suggested ways to operationalise the variables.

*Table 8. Operationalisation of Variables*

| **OPERATIONALISATION OF VARIABLES** | | |
|---|---|---|
| **Variable** | **Leading Hypothesis** | **Alternative Hypotheses** |
| Value | Market value added | |
| Growth | Compound annual growth in size | |
| Profits | Company return on equity less cost of equity | Company EVA (calibrated against industry EVA) |
| Size | Value-added | Number of employees<br><br>Assets |
| Communications distortion | # of vertical levels | |
| Bureaucratic insularity | Business Week ranking of quality of governance | New, qualitative, index |
| Atmospheric consequences | Average pay at large company/average industry sector pay | |
| Incentive limits | R&D % of revenue | Average pay in industry; MBA intensity; labour cost share of value-added |
| Asset specificity | Breadth (Herfindahl of SIC codes)<br><br>Reach (% international sales)<br><br>Depth (value added/sales) | Asset specificity index and uncertainty index (standard deviation of sales volatility) |
| Organisation form | Dummy: M-form, U-form | Dummies: Williamson's 7 forms |
| Industry | Dummies: Rationalised form of Fortune definitions (10 industries) | Dummies: SIC code based definition |

## 6.3 DATA SOURCES

The data sources are conventional and readily available. The total cost for purchasing information is less than £7,000 (including telephone interviews). Data entry takes at the most 200 hours. Table 9 summarises the potential sources.

*Table 9. Data Sources*

| DATA SOURCES | | |
|---|---|---|
| **Industry Data** | **Company Statistics** | **Proprietary Company Data** |
| FTC Line of Business statistics | Annual reports | Press clippings |
| Standard & Poor's sector analyses | 10Ks and 10Qs | Telephone interviews (e.g., to determine organisation form) |
| Analyst reports on industrial sectors | Compustat (and Tristat) | Earlier research (e.g., Rumelt 1982) |
| Fortune 500 sector rankings | Moody's | Cole Room (HBS) company files |
| U.S. Bureau of the Census employment and wage statistics | Analyst reports | Web sites |
| | Earlier research (e.g., Teece (1982) | |
| | Business Week's governance ranking | |
| | Stern Stewart MVA and EVA tables | |
| | Fortune 500 rankings | |
| | Web sites | |

A critical question is to define the time period to be used. The starting point should ideally be at least 20 years back (1974) to avoid survivor bias (if today's companies are selected then the sample will skewed towards the current survivors, presumably better managed companies). However, it is difficult to find proprietary data this old and thus only the first

analysis, “value, profits and growth versus size”, can be done on this sample. For the “size versus the four plus two factors” it will be necessary to settle for 1989 as the start (the last publication of the Fortune Industrial 500 was in 1994). This will increase survivor bias, but improve the quality of the data.

The research should cover 300 to 400 companies of the Fortune 500. This gives a safety margin since companies disappear and emerge, while it maintains a large enough sample for statistical significance. In particular, the structural equation modelling analyses described below in Section 6.4.2 should use a sample size of 200 (Hair et al. 1998, 604–605) since this is the optimum suggested by model mis-specification, model size, and departures from normality considerations.

## 6.4 STATISTICAL ANALYSIS METHODS

There are two basic methods for testing the hypotheses. The first builds on using traditional matched pair and linear regression methods, the second uses structural equation modelling (SEM). The advantage of the former is that it has better confirmatory value, but it makes inefficient use of the data. The SEM approach is stronger as an explanatory model and uses data efficiently (Hair et al. 1998). Both approaches are summarily described below.

### 6.4.1 Traditional Analyses

**Test of $H_1$, $H_2$ and $H_3$.** To test if the largest companies in each industry have lower relative values than smaller competitors, a matched-pair analysis is appropriate as a first test. Teece (1981) used a matched pair analysis to study the profit impact of the M-from organisation. The reason for this choice of analysis method was that less data was required while the statistical significance was good. The approach is appealing because it is easy to use and it draws on exactly the same data as the linear regressions, but with fewer observations required. Teece matched companies that were similar in all aspects except for organisation form. This gave him 20 pairs in 20 industries on which he could run a Sign Test and a Wilcoxon Matched Pairs Signed Ranks Test.

In this proposal's almost identical case, matched pairs should be between the largest (two largest, three largest, etc.) companies and the average of the rest of the industry. The following should then be tested:

$$\Delta_V = V_L - V_R$$

$$\Delta_\Pi = \Pi_L - \Pi_R$$

$$\Delta_G = G_L - G_R$$

where $\Delta$ is the differential; **V** is the relative market value as defined in Section 6.2; $\Pi$ is the relative profitability, **G** is the annual growth in size, and ; $_L$ denotes the large (largest) company (ies); and $_R$ denotes the average of the rest of the industry participants. If $\Delta < 0$, then the hypothesis is valid (given usual significance tests). However, if $\Delta > 0$ the hypothesis can still be valid because the analysis does not include the universe of even larger companies that do not exist and which may not exist because their performance would be substandard.

A second test is to use the following models:

$$\mathbf{VALUE} = \beta_0 + \beta_1 * \mathbf{PROFIT} + \beta_2 * \mathbf{GROWTH} + \Sigma(\beta_i * \mathbf{IND}_i) + \varepsilon$$

$$\mathbf{PROFIT} = \beta_0 + \beta_1 * \mathbf{SIZE} + \varepsilon$$

$$\mathbf{GROWTH} = \beta_0 + \beta_1 * \mathbf{SIZE} + \Sigma(\beta_i * \mathbf{IND}_i) + \varepsilon$$

where **VALUE** is the relative market value of the company, **PROFIT** is the relative profitability, **GROWTH** is the company's annual growth rate, **SIZE** is the size of the company, **IND** is a dummy for each industry (not used in the second equation because industry does not explain profitability as explained in Section 5.4), and $\varepsilon$ is the error**.** These

regressions will be run for the total sample as well as for each industry large enough to allow for a separate analysis.

**Test of $H_4$ and $H_5$.** The choice of statistical model is straightforward since the dependent variable, value, is continuous, and the independent variables are continuous or standard-form dummies. Other studies rooted in economics using this method for similar problems are Armour and Teece (1978), who studied the correlation between profits and organisation form, D'Aveni and Ravenscraft (1994), who studied the correlation between profits and vertical integration, and Comment and Jarrell (1995), who studied the correlation between profits and corporate focus. Sociology based studies include Pugh et al. (1968), Child (1973), and Pondy (1969). The first two studied the link between size, structure and complexity, the last studied the link between administrative intensity, size and other factors. All these employed the same basic methodology. Based on this methodology, the statistical models become (assuming that the independent variables are not correlated):

$$\mathbf{SIZE} = \beta_0 + \beta_1 * \mathbf{CD} + \beta_2 * \mathbf{BI} + \beta_3 * \mathbf{AC} + \beta_4 * \mathbf{IL} + \Sigma(\beta_i * \mathbf{IND}_i) + \varepsilon$$

$$\mathbf{OFFSET} = \beta_0 + \beta_1 * \mathbf{AS} + \beta_2 * \mathbf{FORM} + \Sigma(\beta_i * \mathbf{IND}_i) + \varepsilon$$

where **SIZE** is the size of the company, **CD** is communications distortion, **BI** is bureaucratic insularity, **AC** is atmospheric consequences, **IL** is incentive limits, **OFFSET** is the offsetting mechanism, **AS** is degree of asset specificity, **FORM** is a dummy for organisational form (M=1), **IND** is a dummy for each industry, and **ε** is the error**.** These regressions will also be run for the total sample as well as for each industry large enough to allow a separate analysis.

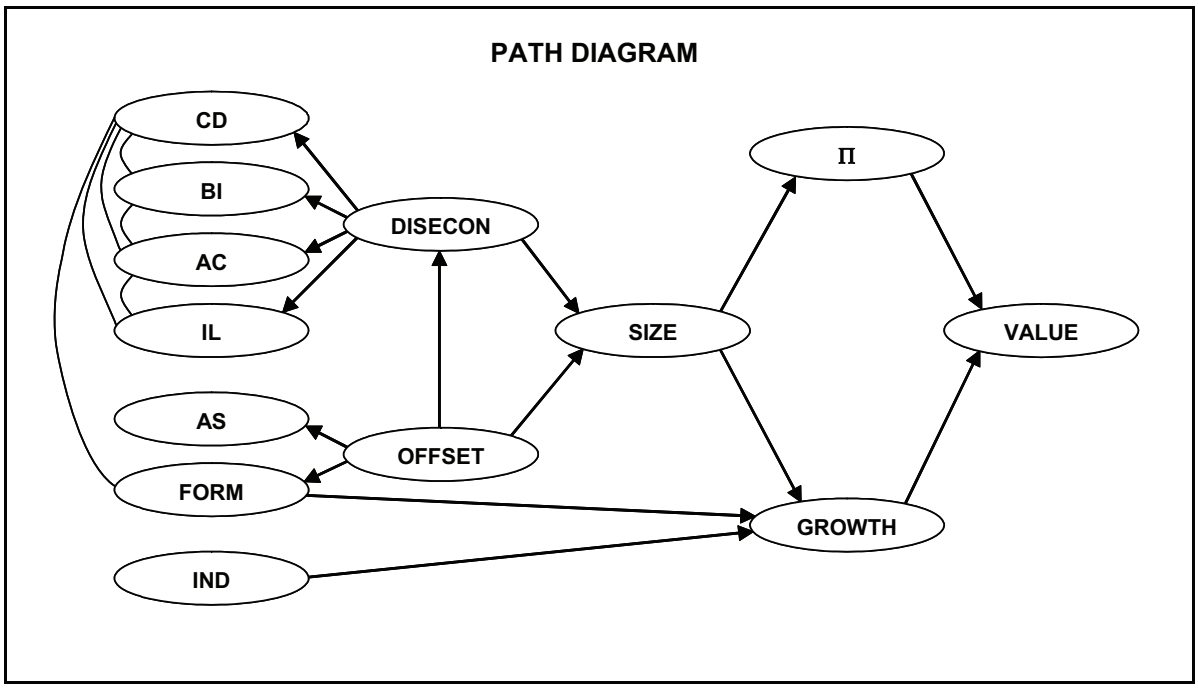
SPSS is the preferred analysis software.

### 6.4.2 Structural Equation Modelling

SEM, an extension of factor analysis, allows for exploring multiple data relationships simultaneously while providing for a high level of confirmatory analysis (Hair et al. 1998, 578). This makes the technique unique compared to multiple regression, multivariate analysis of variance, and discriminant analysis. It appears to be highly suitable for the analysis at hand since the dependencies and causal links in the "limits of firm size" model are unclear. Moreover, it allows for the use of latent variables (hypothesised and unobserved variables) which makes it possible to estimate the "sources of limits of firm size" and the "offsetting mechanisms."

Figure 3 contains a preliminary path diagram. It is only intended to illustrate the approach, and the reasoning behind the causal dependencies and correlations is not explicated. As the work progresses, a nested model approach (p. 591) will be used to test different dependencies.

*Figure 3. Path Diagram*




The seven manifest variables are defined as before: **CD** is communications distortion, **BI** is bureaucratic insularity, **AC** is atmospheric consequences, **IL** is incentive limits, **AS** is degree of asset specificity, **FORM** is a dummy for organisational form (M=1), and **IND** is a dummy for each industry. The latent variables are **DISECON,** which captures Williamson's sources of limits of firm size, and **OFFSET,** which corresponds to Williamson's

offsetting mechanisms. The dependent variables are **VALUE** for market value, **Π** for profits, **GROWTH** for growth, and **SIZE** is size of company.

LISREL or Amos is the preferred analysis software package.

# 7. WORKPLAN

This chapter contains a rough outline of the workplan, including timing of activities, approach, end product and anticipated diffculties.

## 7.1 TIMING OF ACTIVITIES

The research and writing of the thesis should at the most take 18 months. There will be a continuous research effort in parallel with regular work, with 5-day total immersion periods once every four months. At the end one can expect to spend 2–3 weeks full time writing up the thesis. The steps are:

1. Define general analysis needs (1 month)

   - Identify supporting facts required

   - Assess data collection difficulty

   - Identify sources

2. Revise theoretical model (1 week)

3. Plan statistical analysis approach (2 months)

- Learn about statistical methods

- Update definition of dependent, manifest, and latent variables, and create validity test

4a. Collect data (6 months)

- Copy Baker Library information

- Buy commercial information

- Collect individual company data

4b. Carry out general statistical analysis (3 months)

- Run regressions

- Interpret results

5. Write thesis and do complementary research (3 months)

- Identify poor logic and missing analyses

- Write core document

- Review and refine

- Update working papers

- Make intermediate checks with supervisors

6. Adjust, finalise, and defend thesis (1 to 4 months)

During this process various formal reviews (meetings, telephone conversations or written status reports) with supervisors should be scheduled every 3 months.

## 7.2 APPROACH

The approach to getting the work done draws on the researcher's experience as a management consultant. Work will progress in parallel on collecting and structuring the underlying data and creating the statistical model (based on LISREL or Amos, complemented with SPSS). The data collection requires two types of information, general corporate data available through Compustat, government sources, Fortune, etc.,[17] and

[17] Including revenues and profits over time, economic value added, number of employees over time, geographic reach (e.g. per cent of revenue outside the US), product breadth (SIC codes), and value added (cost of goods sold + employee costs + depreciation).

company-specific information[18] available through earlier studies (e.g., Rumelt 1982), published information available in the Cole Room at Harvard Business School, and telephone surveys with companies in the sample. Collecting the company-specific information is a fairly large undertaking and it will be necessary to employ a temp over a six-month period to assist with this. Hopefully the temp can be found on the Harvard Business School campus among student spouses.

Finally, it will be important to draw on the Henley resources by participating in workshops and by selectively engaging faculty. Those workshops focusing on statistical analysis and on giving feedback on the research will be particularly important.

## 7.3 END PRODUCT

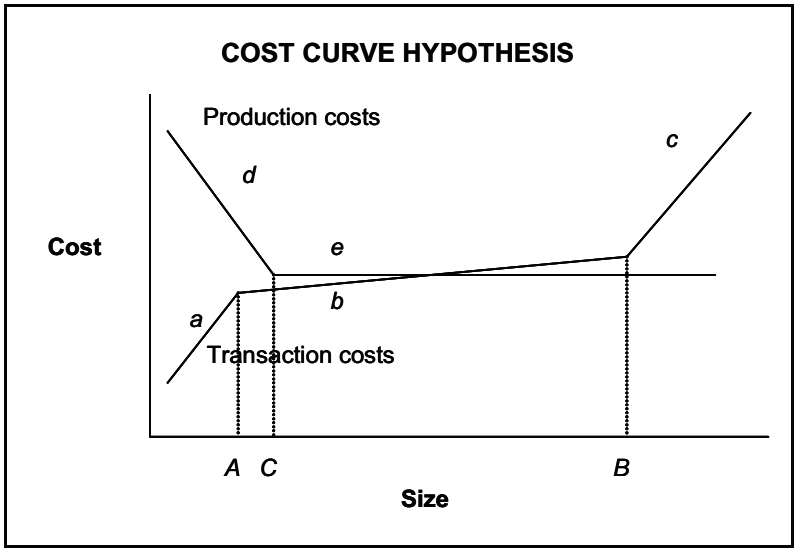
The end products required for the DBA will be delivered—nothing more, nothing less—but with high quality. One of the lessons learnt so far is to stay focused on the task at hand and not to expand beyond the original plan.

[18] Including internal organisation structure, board/management relations, R&D productivity, and labour relations.

### 7.3.1 Expected Results

The purpose of the research is to account empirically for the limits of firm size and test if they can be explained by Williamson's TCE model. The literature survey strongly suggests that size limits exist. The more important question is how strong the limiting forces are. The schematic cost curves in Figure 4 can be expected

*Figure 4. Cost Curve Hypothesis*




What is unclear is the slope of the curves and the breakpoints. For the transaction cost curve, slope *a* is probably fairly steep, especially in R&D intensive industries. Slope *c* is steep in fixed cost type industries, including R&D intensive industries. Breakpoint *A* is somewhere in the region of 500 to 2,000 employees, while breakpoint *B* varies considerably by industry and is somewhere in the range of 10,000 to 200,000 employees. For the

production cost curve, Slope *d* is steeply negative, while *e* is flat since production costs are common to the whole industry and independent of each individual company. Breakpoint *C* varies considerably by industry and is set by technological indivisibilities as described in Section 5.3. These are informed guesses based on the actual size distribution of American industry.[19]

The statistical analysis will confirm or disprove this hypothesis. The expectation is that the statistical significance will be fairly low—but it will exist, and that the results will be questioned mainly on the merits of the operationalisation of the independent variables.

### 7.3.2 Table of Contents

The thesis will follow a traditional structure, similar to this proposal (Table 10).

[19] Note that the production cost curve applies to the whole industry, in line with the argument in Section 5.3.

*Table 10. Preliminary Table of Contents of Thesis*

| PRELIMINARY TABLE OF CONTENTS OF THESIS | |
|---|---|
| 1. | Introduction (5 pages) |
| 1.1 | Purpose |
| 1.2 | Results |
| 2. | Research Objectives (20 pages) |
| 2.1 | Description of the Dilemma |
| 2.2 | Problem Definition |
| 2.3 | Importance and Uniqueness of Research |
| 3. | Theoretical Framework (20 pages) |
| 3.1 | Limits of Size |
| 3.2 | Offsets |
| 4. | Literature Survey (80 pages) |
| 4.1 | Limits of Firm Size |
| 4.2 | Moderators |
| 4.3 | Modified Theoretical Framework |
| 5. | Research Methodology (60 pages) |
| 5.1 | Overview of Relevant Prior Research |
| 5.2 | Operationalisation of Theoretical Model |
| 5.3 | Data Sources |
| 6. | Research Results (20 pages) |
| 6.1 | Data Presentation |
| 6.2 | Findings |
| 6.3 | Sensitivity Analysis |
| 6.4 | Alternative Interpretations |
| 7. | Conclusion (5 pages) |
| 7.1 | Summary of Results |
| 7.2 | Limitations of Research |
| 7.3 | Further Research |

The thesis is expected to be 40,000 to 50,000 words with a distribution as indicated above. In addition, there will a number of appendices with the key raw data.

## 7.4 ANTICIPATED DIFFICULTIES

There are a few foreseeable difficulties:

1. The task to operationalise the theoretical concept will be non-trivial. Attempts have been made before in the analysis of vertical integration, and the results are not totally encouraging.

2. The research requires massive amounts of data and it will be difficult to limit the scope of analysis. The risk is that the focus will be on data collection at the detriment of insightful analysis.

3. Interpreting the results from the analysis and adjusting the analytical approach takes time, mental concentration, and sparring from others. It will be difficult to free up this time given work commitment.

There should not be any problems with the theoretical model, finding the data, and doing the regression and other analyses.

# 8. CONCLUSION

This research proposal 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 scale. 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 proposal is the TCE-based model which combines four distinct aspects of Williamson's thinking: 1) The sources of limits of firm size: communications distortion due to bounded rationality, bureaucratic insularity, atmospheric consequences due to specialisation, and incentive limits of the employment relation. 2) The offsetting influence of asset specificity on both transaction cost and production cost diseconomies. 3) The importance of choice of organisation form to reduce diseconomies. 4) The unimportance of neoclassical scale economies at the firm level.

As far as can be determined, no one has used the TCE paradigm to empirically test the diseconomies of scale at the firm level before. The research therefore complements functional level research to add to our understanding of the limits of firm size, and in the end to our understanding of bureaucratic failure.

There seem to be a number of real life implications of the research, regardless of if Williamson's model can be empirically proven. If his theory is supported then we will add to our understanding that strategy and structure are intimately linked. Executives at large corporations have real trade-offs to make when they think about expansion (as was shown in the literature survey, they always think about expansion). For example, strategically sound acquisitions may lead to declining profitability if the diseconomies of scale are real. If there is scant support for Williamson's model then the strategic degrees of freedom are larger for the chief executive.

# 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.

Alston, L. J. and W. Gillespie. 1989. Resource Coordination and Transaction Costs: A Framework for Analyzing the Firm/Market Boundary. Journal of Economic Behavior and Organization 11: 191–212.

Armour, H. O. and D. J. Teece. 1978. Organizational Structure and Economic Performance: A Test of the Multidivisional Hypothesis. Bell Journal of Economics 9 (1): 106–122.

Arrow, K. J. 1964. Control in Large Organizations. Management Science 10 (3): 397–408.

———. 1974. The Limits of Organization. New York: Norton.

———. 1983a. Innovation in Large and Small Firms. In Entrepreneurship, edited by J. Ronen. Lexington, Mass.: Lexington Books.

———. 1983b. The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Nonmarket Allocation. In Collected Papers of Kenneth J. Arrow. Volume 2. Cambridge, Mass.: Belknap.

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

———. 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 (3): 219–233.

Berle, A. A. and G. C. Means. [1932] 1991. The Modern Corporation and Private Property. Revised ed. Reprint, with a new introduction by M. L. Weidenbaum and M. Jensen. New Brunswick, N.J.: Transaction Publishers. Original edition, New York: Commerce Clearing House.

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. 3d 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.

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.

Canbäck, S. 2002. Bureaucratic Limits of Firm Size: Empirical Analysis Using Transaction Cost Economics. DBA thesis, Henley Management College, Brunel University, Henley-on-Thames.

Chandler, A. D., Jr. 1962. Strategy and Structure: Chapters in the History of the American Industrial Enterprise. Cambridge: MIT Press.

———. 1977. The Visible Hand. Cambridge, Mass.: Belknap.

———. 1982. Evolution of the Large Industrial Corporation: An Evolution of the Transaction-Cost Approach. Business and Economic History 2d series, 11: 116–134.

———. 1990. Scale and Scope: The Dynamics of Industrial Capitalism. Cambridge, Mass.: Belknap.

———. 1992. The Emergence of Managerial Capitalism. In The Sociology of Economic Life, edited by M. Granovetter and R. Swedberg. Boulder, Colo.: 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 (2): 168–185.

Coase, R. H. 1937. The Nature of the Firm. Economica n.s., 4 (16): 386–405.

———. 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.

———. 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.

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: 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.

Eatwell, J., M. Milgate and P. Newman, eds. 1987. The New Palgrave: A Dictionary of Economics. 3 vols. New York: Stockton Press.

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.

Fortune. 1995a. The Fortune 500, 1994. Fortune, 15 May, 165–227, F1–F64.

———. 1995b. The Fortune Global 500, 1994. Fortune, 7 August, 123–136, F1–F42.

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 (4): 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.

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.

Ijiri, Y. and H. A. Simon. 1964. Business Firm Growth and Size. American Economic Review 54 (2): 77–89.

Jensen, M. C. 1986. Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. American Economic Review 76 (2): 323–329.

———. 1988. Takeovers: Their Causes and Consequences. Journal of Economic Perspectives 2 (1): 21–48.

———. 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 (4): 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.

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.

Kimberly, J. R. 1976. Organizational Size and the Structuralist Perspective: A Review, Critique and Proposal. Administrative Science Quarterly 21 (4): 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. PhD thesis, Department of Economics, Michigan State University.

Knight, F. H. [1921] 1964. Risk, Uncertainty and Profit. Reprint. New York: Augustus M. Kelley. Original edition, Boston: Houghton Mifflin.

Krickx, G. A. M. 1988. Historical Evidence on the Evolution of Vertical Exchange Mechanisms: Examples from the Computer Systems Industry. PhD thesis, University of California.

Lucas, R. E., Jr. 1978. On the Size Distribution of Business Firms. Bell Journal of Economics 9 (2): 508–523.

Mahoney, J. T. 1989. Organizational Rationalization and Innovation: Vertical Integration and Multidivisional Organization. PhD thesis, University of Pennsylvania.

———. 1992. The Choice of Organizational Form: Vertical Financial Ownership versus Other Methods of Vertical Integration. Strategic Management Journal 13 (8): 559–584.

Marglin, S. A. 1974. What Do Bosses Do? The Origins of Hierarchy in Capitalist Production. Review of Radical Political Economics 6 (2): 60–112.

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. PhD thesis, University of Pennsylvania.

———. 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 (2): 265–273.

———. 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 and 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.

Monsen, R. J. and A. Downs. 1965. A Theory of Large Managerial Firms. Journal of Political Economy 73 (3): 221–236.

Monteverde, K. and D. J. Teece. 1982. Supplier Switching Costs and Vertical Integration in the Automobile Industry. Bell Journal of Economics 13 (1): 206–213.

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.

———. 1987. Institutions, Transaction Costs and Economic Growth. Economic Inquiry 25 (3): 419–428.

———. 1992. Transaction Costs, Institutions, and Economic Performance. 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): 609–624.

Nutter, G. W. 1951. The Extent of Enterprise Monopoly in the United States, 1899–1939. Chicago: University of Chicago Press.

Penrose, E. [1959] 1995. The Theory of the Growth of the Firm. 3d ed. New York: Oxford University Press. Original edition, New York: Wiley.

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 (1): 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. 1968. Dimensions of Organization Structure. Administrative Science Quarterly 13 (1): 65–105.

———. 1969. The Context of Organization Structures. Administrative Science Quarterly 14 (1): 91–114.

Qian, Y. 1994. Incentives and Loss of Control in an Optimal Hierarchy. Review of Economic Studies 61 (3): 527–544.

Ramanujam, V. and P. Varadarajan. 1989. Research on Corporate Diversification: A Synthesis. Strategic Management Journal 10 (6): 523–551.

Rappaport, A. 1998. Creating Shareholder Value. Revised and updated ed. New York: Free Press.

Rasmusen, E. and T. Zenger. 1990. Diseconomies of Scale in Employment Contracts. Journal of Law, Economics, and Organization 6 (1): 65–92.

Riahi-Belkaoui, A. 1994. Organizational and Budgetary Slack. Westport, Conn.: Quorum.

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 (4): 365–378.

Rubin, P. H. 1990. Managing Business Transactions. New York: Free Press.

Rumelt, R. P. 1974. Strategy, Structure, and Economic Performance. Cambridge: Harvard University Press.

———. 1982. Diversification Strategy and Profitability. Strategic Management Journal 3 (4): 359–369.

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: Holt, Rinehart and 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. 3d ed. New York: Free Press. Original edition, New York: Macmillan.

Simon, H. A. and C. P. Bonini. 1958. The Size Distribution of Business Firms. American Economic Review 48 (4): 607–617.

Stewart, G. B., III. 1991. The Quest for Value. New York: HarperCollins.

Stigler, G. J. 1958. The Economies of Scale. Journal of Law and Economics 1 (October): 54–71.

———. 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. J. 1976. The Multinational Corporation and the Resource Cost of International Technology Transfer. Cambridge, Mass.: Ballinger.

———. 1980. Economies of Scope and the Scope of the Enterprise. Journal of Economic Behavior and Organization 1 (3): 223–247.

———. 1981. Internal Organization and Economic Performance: An Empirical Analysis of the Profitability of Principal Firms. Journal of Industrial Economics 30 (2): 173–199.

———. 1982. Towards an Economic Theory of the Multiproduct Firm. Journal of Economic Behavior and Organization 3 (1): 39–63.

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 (3): 373–391.

———. 1987. Supplier Competition, Uncertainty and Make-or-Buy Decisions. Academy of Management Journal 30 (3): 589–596.

Wallis, J. J. and D. C. North. 1986. Measuring the Transaction Sector in the American Economy, 1870–1970. In Long-Term Factors in American Economic Growth. Studies in Income and Wealth Series, edited by S. L. Engerman and R. E. Gallman. Chicago: University of Chicago Press.

Ward, J. L. 1976. The Opportunity to Measure Strategic Variables: An Attempt to Quantify Product-Market Diversity. Journal of Economics and Business 28 (3): 219–226.

Williamson, O. E. 1967. Hierarchical Control and Optimum Firm Size. Journal of Political Economy 75 (2): 123–138.

———. 1975. Markets and Hierarchies: Analysis and Antitrust Implications. New York: Free Press.

———. 1984. The Incentive Limits of Firms: A Comparative Institutional Assessment of Bureaucracy. Weltwirtschaftliches Archiv 120 (4): 736–763.

———. 1985. The Economic Institutions of Capitalism. New York: Free Press.

———. 1986. Economic Organization: Firms, Markets and Policy Control. New York: New York University Press.

———. 1988. Corporate Finance and Corporate Governance. Journal of Finance 43 (3): 567–591.

———. 1989. Transaction Cost Economics. In Handbook of Industrial Organization, Vol. 1, edited by R. Schmalensee and R. Willig. New York: North-Holland.

———. 1993. Opportunism and Its Critics. Managerial and Decision Economics 14 (2): 97–107.

———. 1996. The Mechanisms of Governance. New York: Oxford University Press.

Zenger, T. R. 1989. Organizational Diseconomies of Scale: Pooling vs. Separating Labor Contracts in Silicon Valley. PhD thesis, University of California, Los Angeles.

———. 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.