Journal of Managerial Economics, 2006, Vol. 4, No. 1 (February), pp. 27-70

DO DISECONOMIES OF SCALE IMPACT FIRM SIZE AND PERFORMANCE? A THEORETICAL AND EMPIRICAL OVERVIEW

Staffan Canbäck[*]
Phillip Samouel[†]
David Price[‡]

Abstract

This article tests Oliver Williamson's proposition that transaction cost economics can explain the limits of firm size. Williamson suggests that diseconomies of scale are manifested through four interrelated factors: atmospheric consequences due to specialisation, bureaucratic insularity, incentive limits of the employment relation and communication distortion due to bounded rationality. Furthermore, Williamson argues that diseconomies of scale are counteracted by economies of scale and can be moderated by adoption of the multidivisional organisation form and by high internal asset specificity. Combined, these influences tend to cancel out and thus there is not a strong, directly observable, relationship between a large firm's size and performance.

A review of the relevant literature, including transaction cost economics, sociological studies of bureaucracy, information-processing perspectives on the firm, agency theory, and studies of incentives and motivation within firms, as well as empirical studies of trends in firm size and industry concentration, corroborates Williamson's theoretical framework. The framework translates into five hypotheses: (1) Bureaucratic failure, in the form of atmospheric consequences, bureaucratic insularity, incentive limits and communication distortion, increases with firm size; (2) Large firms exhibit economies of scale; (3) Diseconomies of scale from bureaucratic failure have a negative impact on firm performance; (4) Economies of scale increase the relative profitability of large firms over smaller firms; and (5) Diseconomies of scale are moderated by two transaction cost-related factors: organisation form and asset specificity.

The hypotheses were tested by applying structural equation models to primary and secondary cross-sectional data from 784 large US manufacturing firms. The statistical analyses confirm the hypotheses. Thus, diseconomies of scale influence the growth and profitability of firms negatively, while economies of scale and the moderating factors have positive influences. This implies that executives and directors of large firms should pay attention to bureaucratic failure.

INTRODUCTION

Why are large firms so small? What stops firms from effortlessly expanding into new businesses? Only fragmentary research exists today as to why the largest business organisations do not have ten, twenty or a hundred million employees rather than a few hundred thousand.

In the early 1920s, 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 no systematic studies of the general issue exist. The basic dilemma is illustrated by the mismatch between theoretical expectations and real-world observations.

[*] Dr Canbäck is Managing Director of Canback Dangel, Boston and Fellow of the World Innovation Foundation
[†] Professor Samouel is Chair in Business Economics at Kingston Business School, Kingston upon Thames, England
[‡] Dr Price is Director of Studies, Doctoral Programmes at Henley Management College, Henley-on-Thames, England

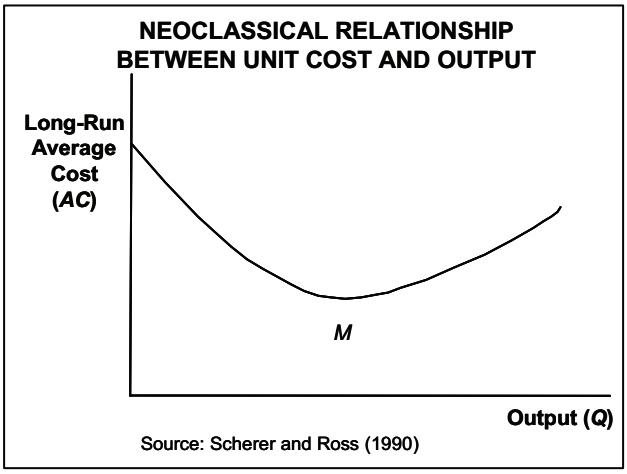
Enquiries: staffan@canback.com

On the one hand, if diseconomies of scale do not exist, then there should be no limits to firm growth and size. We would observe an inexorable concentration of industries and economies until only one global firm was left. The answer to Coase's question (1937, 394): "Why is not all production carried on by one big firm?" would be: it will. Similarly, Stigler (1974, 8) wrote that "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 a given industry has an optimum firm size, then we would expect increased fragmentation as the overall economy grows. This would be in line with Stigler's survivor-principle argument which holds that "the competition between different sizes of firms sifts out the more efficient enterprises" (1958, 55). Again, this is not happening. 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 has been remarkably stable over time for most for the last century, when measured by number of employees or as a share of the total economy.

Cost curves (Figure 1) are used in neoclassical theory to illustrate economies and diseconomies of scale (e.g., Marshall [1920] 1997, 278–292; Scherer and Ross 1990, 101).

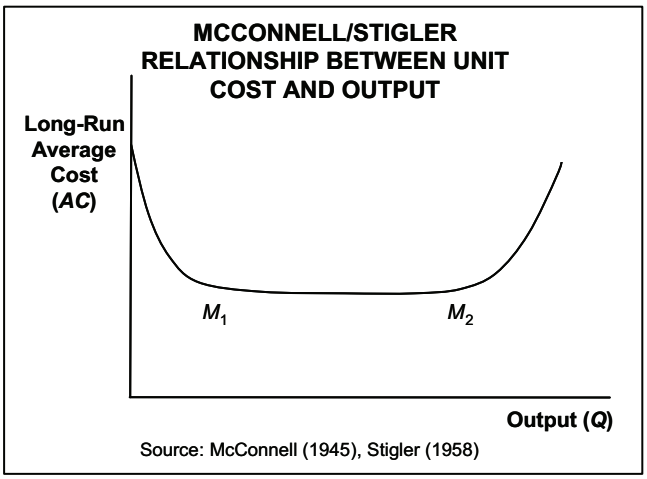
Figure 1. Neoclassical Relationship between Unit Cost and Output



As the output Q increases, the average cost decreases due to economies of scale. At a certain point (M) the economies of scale are exhausted, while diseconomies of scale, presumably driven by diminishing returns to management (e.g., Coase 1937, 395), start to influence the unit cost. As output increases, the unit cost increases. In a competitive market, this implies an equilibrium output M where marginal cost not only equals marginal revenue, but also intersects long-run average cost at its minimum (e.g., Mankiw 1998, 296).

In reality, however, this is not what is observed. Rather, the cost-minimising part of the curve covers a wide range of outputs, and only at high output levels do diseconomies set in, if ever (Panzar 1989, 37–38). 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-world observations. (1) It explains why large and small firms can coexist in the same industry. There is a wide range of outputs, between the points M_1 and M_2, for which the unit cost is more or less constant. (2) It is consistent with Lucas's observation (1978, 509) that, as the economy grows, existing firms tend to expand supply to meet additional demand, because most firms operate with outputs Q below the M_2 inflexion point. (3) It eliminates the supposition that economies of scale are exhausted at approximately the same point that diseconomies of scale start increasing unit cost, which is indicated with M_1 being much to the left of M_2. (4) It demonstrates that there are indeed limits to firm size due to diseconomies of scale, as shown by the increasing unit cost beyond M_2—large firms 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) said: "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".[1] The first part of this statement refers to the negative derivative of the cost curve at outputs smaller than M_1, where economies of scale in production have not yet been exhausted, while the second part applies to the upward slope, where diseconomies of scale due to diminishing returns to management set in beyond M_2.

The article first discusses definitions of firm size and trends over time. Then follows the theoretical framework based on Coase's and Williamson's thinking. The framework is evaluated against the literature and five hypotheses are formulated. Thereafter, the data used to test the hypotheses are discussed and the empirical results are presented. Finally, the findings are discussed and interpreted, ending with concluding remarks. The article is an outgrowth of Canbäck (2002a).[2]

DIMENSIONS OF FIRM SIZE

To describe the essential qualities of diseconomies of scale, it is first necessary to understand the nature of firm size.

---

[1] Simon echoed the writing of Robertson (1923, 25): "It is the economies of large-scale government rather than of large-scale technique which dictate the size of the modern business unit". (Note: *government* here refers to *corporate organisation and governance*, not national government.)

[2] First prize winner in the EDAMBA (*European Doctoral Programmes Association in Management and Business Administration*) competition for best European doctoral thesis in 2002.

Size definitions

There are a number of definitions of what a firm is. The *first*, based on Coase (1937, 389), Penrose ([1959] 1995, 15), and Arrow (1964, 403; 1974, 33) holds that the boundary of the firm is where the internal planning mechanism is superseded by the price mechanism. That is, the firm's border is at the point where transactions are regulated by the market rather than by administration. In most cases this means that the operating firm is equivalent to the legal corporation.

The *second* definition is that ownership sets a firm's boundaries (e.g., Hart 1995, 5–8). With this definition, a firm is the combination of activities for which the bearers of residual risk are one and the same.

A *third* definition sees the firm as a network (Richardson 1972, 884–887). McDonald's Corporation, for example, extends far beyond its corporate ownership, because it also consists of a network of thousands of franchisees over whom McDonald's have a high degree of contractual control (Rubin 1990, 134–144).

The *fourth* definition is based on the firm's sphere of influence. This includes distributors, alliance partners, first- and second-tier suppliers, and so on (Williamson 1985, 120–122). Toyota Motor Corporation, for example, directly employed 215,000 people in 2000, but its sphere of influence probably extended over more than one million people.

For the purposes of this article, the firm is defined as having commonly owned assets—the ownership definition—but employees are also treated as part of the firm. This definition relates closely to Hart's definition (1995, 7), and publicly available data builds on it. It is also commonly used in research (Kumar, Rajan and Zingales 1999, 11). Thus, a firm is an incorporated company (the legal entity) henceforth.

Size measures

Further, there are various ways to measure the size of a firm. Size is most often defined as annual revenue, especially by the business press. However, this measure is basically meaningless because it tells nothing about the depth of the underlying activity. Based on this measure, the world's four largest companies were Japanese trading houses in 1994 (Fortune 1995b). They had between 7,000 and 80,000 employees, but almost no vertical integration.

A better measure of size is value added, which is more or less equivalent to revenue less externally purchased products and services. This metric gives a precise measure of activity, but it is usually not publicly available for individual firms.

Number of employees is the most widely used measure of size. A review by Kimberley claims that more than 80 per cent of academic studies use this measure (1976, 587). In line with Child's observation (1973, 170) that "it is people who are organized", it is not surprising that the number of employees is the most used metric for measuring firm size.

Finally, assets can define size (e.g., as described by Grossman and Hart 1986, 693–694). As with revenue, this measure may not reflect underlying activity; but for manufacturing firms, asset–to–value-added ratios are fairly homogeneous. Asset data for individual firms are usually available back to the 1890s and are therefore a practical measure in longitudinal studies.

In sum, the best measures of size are value added and number of employees, although assets can be used in certain types of studies. This article uses number of employees as the size metric because the data is available and diseconomies of scale should be associated with human frailties. Moreover, this research deals with bureaucratic failure, which in the end is the result of coordination costs. Such costs are best measured in relation to number of employees (Kumar, Rajan and Zingales 1999, 12).

Size trends

The US economy is the basis for the analysis in the current research because it is large, fairly homogenous and transparent, and it has a high level of competition between firms. Within this economy, the research focuses on the manufacturing sector.

Large manufacturing firms play a major role in the US economy. Contrary to popular belief, however, the importance of large firms is not increasing and has not done so for many years. Studies show that large manufacturing firms are holding steady as a share of value added since circa 1965 (Scherer and Ross 1990, 62). Their share of employment in the manufacturing sector has declined from around 60 per cent (1979) to around 50 per cent (1994). Moreover, as a share of the total US economy, they are in sharp decline. Large manufacturing firms employed 16 million people in 1979 versus 11 million in 1994 (Fortune 1995a, 185), while private sector employment grew from 99 to 123 million people (Council of Economic Advisers 1998, 322) over the same time period.

Further evidence that large firms do not increasingly dominate the economy is available from a number of historical studies. Aggregate industry concentration has changed little since the early part of the last century.[3] Nutter (1951) studied the concentration trend between 1899 and 1939 and found no signs of increased aggregate concentration during this period, mainly because new, fragmented industries emerged, while older ones consolidated (pp. 21, 33). Bain (1968) found the same trend between 1931 and 1963, but with less variability between industries. Scherer and Ross (1990, 84) used Nutter's methodology and showed that aggregate concentration increased slightly, from 35 per cent in 1947 to 37 per cent in 1982. Similarly, Mueller and Hamm (1974, 512) found an increase in four-firm concentration from 40.5 per cent to 42.6 per cent between 1947 and 1970, with most (70 per cent) of the increase between 1947 and 1963.

Bock (1978, 83) studied the share of value added contributed by the largest manufacturing firms between 1947 and 1972. There was a large increase between 1947 and 1954, and a further slight increase until 1963. Between 1963 and 1972, there was no increase. Scherer and Ross (1990, 62) confirmed the lack of increase through the end of the 1980s. Sutton (1997, 54–55) reached a similar conclusion in a comparison of concentration in the US manufacturing sector between 1967 and 1987.

The above evidence shows that concentration in the manufacturing sector—defined as the share of value added, employment, or assets held by large firms—has changed little or has declined over much of the last century. The size of large manufacturing firms has kept pace with the overall growth of the manufacturing 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 global corporate sector). This indicates that there is a limit to firm size and that this limit may be decreasing in absolute terms, all of which supports the research findings of this article.

THEORETICAL FRAMEWORK

Transaction cost economics focuses on the boundary of the firm (Holmström and Roberts 1998, 73; Williamson 1981, 548)—that is, the distinction between what is made internally in the firm and what is bought and sold in the marketplace. The boundary can shift over time and for a number of reasons, and the current research looks at one aspect of these shifts. As firms internalise transactions, growing larger, bureaucratic diseconomies of scale appear. Thus, a firm will reach a size at which the benefit from the last internalised transaction is offset by bureaucratic failure.

---

[3] Note that there have been significant changes within individual industries.

Two factors moderate these diseconomies of scale. First, firms can lessen the negative impact of diseconomies of scale by organising activities appropriately and by adopting good governance practices. Second, the optimal degree of integration depends on the level of asset specificity, uncertainty and transaction frequency.

Coase's article "The Nature of the Firm" (1937) establishes the basic framework. "Limits of Vertical Integration and Firm Size" in Williamson's book *Markets and Hierarchies* (1975) suggests the nature of size limits. "The Limits of Firms: Incentive and Bureaucratic Features" in Williamson's book *The Economic Institutions of Capitalism* (1985) expands on this theme and explains why the limits exist. Riordan and Williamson's article "Asset Specificity and Economic Organization" (1985) augments the theoretical framework presented here by combining transaction costs with neoclassical production costs. The remainder of the section discusses the details of the argument.

Diseconomies of Scale

Williamson (1975, 126–130) found that the limits of firm size are bureaucratic in origin and can be explained by transaction cost economics. He identified four main categories of diseconomies of scale: atmospheric consequences due to specialisation, bureaucratic insularity, incentive limits of the employment relation and communication distortion due to bounded rationality.

Williamson's categories are similar to those Coase described in 1937. Coase talked about the determination (or planning) cost, the resource misallocation cost and the cost of lack of motivation. Williamson's first and second categories correspond 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. The four categories are detailed below:

Atmospheric consequences. According to Williamson (1975, 128–129), as firms expand there will be increased specialisation, but also less commitment on the part of employees. In such firms, the employees often have a hard time understanding the purpose of corporate activities, as well as the small contribution each of them makes to the whole. Thus, alienation is more likely to occur in large firms.

Bureaucratic insularity. Williamson (1975) argued that as firms increase in size, senior managers are less accountable to the lower ranks of the organisation (p. 127) and to shareholders (p. 142). They thus become insulated from reality and will, given opportunism, strive to maximise their personal benefits rather than overall corporate performance. According to Williamson, this problem is most acute in organisations with well-established procedures and rules and in which management is well-entrenched.

Incentive limits of the employment relation. Williamson (1975, 129–130) argued that the structure of incentives large firms offer employees is limited by a number of factors. First, large bonus payments may threaten senior managers. Second, performance-related bonuses may encourage less-than-optimal employee behaviour in large firms. Therefore, large firms tend to base incentives on tenure and position rather than on merit. Such limitations may especially affect executive positions and product development functions, putting large firms at a disadvantage when compared with smaller enterprises in which employees are often given a direct stake in the success of the firm through bonuses, share participation, and stock options.

Communication distortion due to bounded rationality. Because a single manager has cognitive limits and cannot understand every aspect of a complex organisation, it is impossible to expand a firm without adding hierarchical layers. Information passed between layers inevitably becomes distorted. This reduces the ability of high-level executives to make decisions based on facts and negatively impacts their ability to strategise and respond directly

to the market. In an earlier article (1967), Williamson found that even under static conditions (no uncertainty) there is a loss of control.

Economies of Scale

Transaction cost economics does not usually deal with economies of scale, which are more often associated with neoclassical production costs. However, Riordan and Williamson (1985) made an explicit attempt to reconcile neoclassical theory and transaction cost economics and showed, among other things, that economies of scale are evident in both production costs (p. 371) and transaction costs (p. 373), and that both can be kept internal to a firm if the asset specificity is positive. That is, the economies of scale can be reaped by the individual firm and are not necessarily available to all participants in a market (pp. 367–369).

Moderating Influences on Firm-Size Limits

While the four categories relating to diseconomies of scale theoretically impose size limits on firms, two moderating factors tend to offset diseconomies of scale: organisation form and degree of integration. Both are central to transaction cost economics.

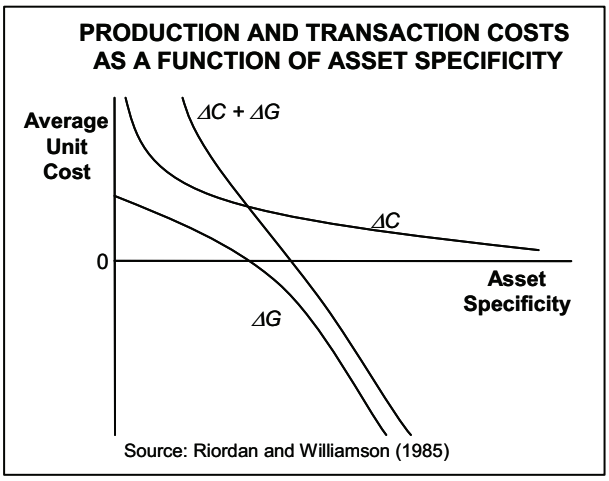
Organisation form. Williamson (1975, 117) recognised that diseconomies of scale can be reduced by organising appropriately. Based on Chandler's pioneering work (e.g., 1962) on the evolution of the American corporation, Williamson argued that the M-form organisation lowers internal transaction costs compared to the U-form organisation. It does so for a key reason: The M-form allows most senior executives to focus on high-level issues rather than day-to-day operational details, making the whole greater than the sum of its parts (p. 137). Thus, large firms organised according to the M-form should perform better than similar U-form firms.

Degree of integration. Williamson showed that three factors play a fundamental role in determining the degree of integration: *uncertainty*, *frequency of transactions* and *asset specificity,* under conditions of bounded rationality (Simon [1947] 1976, xxvi–xxxi) and opportunism (Williamson 1993). While uncertainty and frequency play some role in creating transaction costs, Williamson considered asset specificity the most important driver of integration (e.g., Riordan and Williamson 1985, 366):

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

This is illustrated in Figure 3, in which the differential production cost (ΔC) and transaction cost (ΔG) for markets and hierarchies are shown as a function of asset specificity. The curves show that markets have a large production cost advantage when asset specificity is low, but it approaches zero for high asset specificity (ΔC). For transaction costs, the market has an advantage for low asset specificity and a disadvantage for high asset specificity (ΔG).

Figure 3. Production and Transaction Costs as a Function of Asset Specificity



EVIDENCE IN THE LITERATURE

In general, there exists only limited research on diseconomies of scale. This is somewhat surprising, because many authorities point out that analysing the limits of firm size is critical to our understanding of the modern economy. Fortunately, the relevant literature yields fragments of evidence that not only confirm the existence of diseconomies of scale, but also explicate various features of bureaucratic failure. The composite picture derived from a review of this literature supports the theoretical framework developed in the previous section, and the hypotheses articulated later in the article.

Diseconomies of scale

The literature relating to firm-size limits does not follow Williamson's categorisation. Thus, the relevant studies are reviewed by general topic and author, covering bureaucracy and its negative effect on size, information loss, agency theory, and employee incentive problems.

Bureaucracy: Negative Consequences of Size

A number of sociological studies describe negative consequences of size which correlate well with Williamson's propositions. Pugh et al. (1969) and Child (1973), among others, showed that size leads to bureaucracy. Large firms are usually highly bureaucratised through formalisation, and to the extent that bureaucracies breed diseconomies, this limits the growth of such firms.

Williamson made a similar point: "almost surely, the added costs of bureaucracy are responsible for limitations in firm size" (1996, 266). According to Blau and Meyer the diseconomies of bureaucracy fall into three major categories: (1) excessive rigidity, (2) conservatism/resistance to change, and (3) perpetuation of social-class differences (1987, 139–161).

Of these, the first one is relevant here because conservatism is essentially a subcategory of rigidity, and social-class differences fall outside the scope of this research. Excessive rigidity appears as organisations formalise work practices through bureaucratic procedures (Merton 1957, 197–200). Problems are solved by adding structure and the firm reaches a point at which the added structure costs more than the problem solved; Blau and Meyer referred to this as the "problem—organisation—problem—more organisation" spiral of bureaucratic growth (1987, 147). These researchers showed that factors external to the firm, such as increased number of customers or number of tasks to be performed, have little to do with increased bureaucracy. In the end, the added policies and procedures of bureaucracy 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. Managers become increasingly insulated from reality, while lower levels of the organisation experience alienation. As Stinchcombe (1965) demonstrated, one consequence of such rigidity is that firms tend to maintain the organisation form they had when they were created.

Pondy (1969) studied administrative intensity in different industries and what causes variations in intensity. 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: the larger the organisation is, the more managers are shielded from reality, and the more distant the owners are from daily operations.

Using a demographical research approach, Carroll and Hannan (2000, 289–290) argued that older firms exhibit organisational inertia and find it increasingly difficult to adapt to external changes: "…old organizations are disadvantaged compared to younger ones in changing environments. Alternatively, accumulating rules, routines, and structures might simply impose an overhead cost that reduces the efficiency of organizations even in stable environments".

A similar logic based on institutional economics can be found in Olson (1982). His theory holds that as the institutional structure of a country ages, growth-retarding organisations such as an increasingly complex legal system, special-interest groups and nongovernmental watchdog organisations will become increasingly abundant. The theory and empiry specifically predict that older countries with stable institutions will exhibit lower economic growth (p. 77). If this logic holds for corporations as well, then older firms will experience less growth.

Information Loss and Rigidity

A few studies from the firm-as-information-processor school of thought relate to diseconomies of scale. (Several studies within this school relate to the size distribution of firms, but do not discuss the nature of the diseconomies of scale at length. See Sutton (1997, 43–48) and Axtell (1999, 4–5) for summaries).

Arrow (1974) found that employees in large organisations tend to be highly specialised. Thus, coordination through communication becomes increasingly important. Because information flows carry a cost, organisations code (through formal or informal rules) the information available. Coding economises on resources, but it also leads to information loss and rigidity (p. 55). This means (1) that the more hierarchical levels there are, the more information loss or distortion results; 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)—he found that information degrades as communication lines are extended. Geanakoplos and Milgrom (1991) added to this perspective by noting that there are inevitable signal delays in an organisation. The more hierarchical levels to be traversed, the longer and more frequent the delays are.

Summarising the lessons learnt during a career as a corporate executive, Barnard ([1938] 1968) argued that the size of unit organisations is "restricted very narrowly by the necessities of communication" (p. 110) and that "the size of executive organizations is limited generally by the same conditions that govern the size of unit organizations" (p. 112).

Control-loss problems may contribute to diseconomies of scale as well. McAfee and McMillan (1995) argued that people in organisations exploit information asymmetries to their advantage (or in Williamson's words (1993), they are opportunistic). Dispersion of

knowledge within the organisation combined with individual self-interest make conflict of interest and sub-goal pursuit inevitable.

McAfee and McMillan noted, among other things, that efficiency falls as the hierarchy expands, and that "long" hierarchies are not viable in competitive industries (p. 401). Qian (1994), similarly found that in long hierarchies, employees do not contribute with a high level of effort. Employees have incomplete information about their role in the enterprise and thus suffer from a lack of motivation. Moreover, managers will need to monitor employee effort, leading to higher costs and further resistance or lack of commitment.

However, Mookherjee and Reichelstein (2001) made the case that long hierarchies, under certain restrictive conditions (p. 4), do not lead to control loss: "provided the required conditions on contracting sequence, verifiability of subcontracts and unlimited liability of intermediate agents hold, our model questions the common notion that larger, more complex hierarchies are less efficient owing to 'control losses' with respect to incentives or coordination" (p. 4). It is unclear, however, whether these conditions are met by real-world firms.

Agency Theory

An early version of agency theory argued that very large firms do not strive for profit maximisation. According to Monsen and Downs, 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" (1965, 222). Monsen and Downs furthermore found that the motives of managers differ from those of owners. Managers tend to maximise personal income, while owners maximise profits. It is impossible for owners of large firms to control the behaviour of managers. Consequently, profit maximisation does not occur. The outcome is akin to what Williamson labelled bureaucratic insularity.

Silver and Auster (1969) argued that the "divergences of interests within the firm and the costs of dealing with them" (p. 277) mean that "the entrepreneur's time is a limitational factor" (p. 280). Employees typically "shirk their duties unless the employer takes steps to prevent this" (p. 278). As a result, senior executives will have less time for strategising and entrepreneurialism, all other things being 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, meaning such industries will be more fragmented; and (2) the more supervision employees require, the lower the industry concentration ratio.

More recently, Jensen has deepened and extended these arguments (e.g., Jensen and Meckling 1976; Jensen 1986, 1988, 1989, 2000). He defined agency costs 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 did not, contrary to Monsen and Downs or Silver and Auster, explicitly state that agency costs increase with the size of the firm. Jensen demonstrated, however, that managers emphasise firm 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"

(1986, 323). He looked at the profitability of diversified firms, noting that they are less profitable than focused firms.

Agency theory and transaction cost economics are similar in many respects and it is not surprising that the two theories lead to the same conclusions. However, some authorities contend 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.

Other critics have pointed out that agency theory poorly explains the boundaries of the firm (Kumar, Rajan, and Zingales 1999, 5). Hart (1995, 20), for example, noted that "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". For that reason, TCE provides a more nuanced foundation for the current research.

Employee Incentives and Lack of Motivation

A number of authorities have argued that job satisfaction is lower in large organisations and at large work establishments. Employees in large firms are paid significantly more than those in small firms. The reason often given for this disparity is that higher compensation makes up for a less-satisfying work environment (Brown, Hamilton and Medoff 1990, 29).

Scherer's work (1976) is representative of the extensive research conducted at the establishment level. In a review of the literature, including his own original research, he concluded that worker satisfaction was 30 per cent lower in large establishments[4] compared to small establishments (p. 109). Meanwhile, compensation was more than 15 per cent higher for equivalent job descriptions (p. 119). He argued that because establishment size is correlated to firm size, the effects of alienation in large firms appear to be significant. Later work, sponsored by the Federal Trade Commission in the United States, confirmed these findings (Kwoka 1980).

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). They did not conclude that this difference is necessarily related to alienation, but regardless of the cause, large firms seem to pay substantially higher wages than smaller ones.

In addition, span-of-control problems make it increasingly costly to extend incentive contracts to employees as firms grow (Rasmusen and Zenger 1990, 69). Thus, large firms favour fixed-wage contracts based on tenure rather than performance and make extensive use of monitoring to control productivity. In contrast, 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; smaller firms may employ both superior and inferior talent, but they reward individuals accordingly. Rasmusen and Zenger's data strongly supported these conclusions, especially in functions with indivisible work, where success is dependent on joint contributions by several individuals (e.g., in research and development).

The closer match between performance and pay in small firms puts large firms at a disadvantage, in line with Williamson's incentive limits as a source of diseconomies of scale. Olson (1982, 31) noted that: "in the absence of selective incentives, the incentive for group action diminishes as group size increases". A similar argument was made by Axtell (1999), who, based on agent-based computational modelling, found that the number of free riders in a firm grows with firm size and that the limits of firm size are set at the point where the

[4] More than 500 employees.

advantages of joint production (i.e., economies of scale) are smaller than the disadvantages of having many free riders in the firms whose work effort cannot be effectively monitored (p. 54): "We have interpreted firm growth and demise as a process in which agents are attracted to high-income firms, these firms grow, and once they become large get over-run with free-riders."

Many authorities point out that R&D productivity is significantly lower in large firms. Cooper (1964) surprised business leaders and academics with his article "R&D Is More Efficient in Small Companies".

Based on 25 interviews with managers at large and small firms, he argued that small firms have three to ten times higher productivity in development than large firms. The key reasons: (1) small firms are able to hire better people because they can offer more tailored incentives; (2) engineers in small firms are more cost-conscious; and (3) internal communication and coordination is more effective in small firms. These reasons match three of Williamson's four sources of diseconomies: incentive limits, atmospheric consequences and communication distortion.

Later work has confirmed Cooper's anecdotal evidence both theoretically and empirically. Arrow (1983) demonstrated that large firms will invest suboptimally in development because of information loss, and that small firms 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 listed 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 quantified and confirmed Cooper's initial evidence, noting that "big firms tend to provide a haven for the mediocre in search of anonymity" (p. 43).

In addition, 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 firms showed that small firms attract better talent than large firms, motivate employees to try harder and tend to better tie compensation to performance (1994, 725).

Finally, leading anti-bigness ideologues have provided plenty of anecdotal evidence for such arguments, although they are lacking in formal findings. Peters (1992) supported the notion that R&D is less effective in large organisations. He argued that large firms are massively overstaffed in development and that there is little correlation between size of R&D budget and output, offering several case examples as proof. Brock (1987) argued that bigness retards technological advance because large firms are overly risk averse.

Peters, who since the early 1980s has crusaded against large firms, has discussed diseconomies of scale in several books and articles. His views were summarised in "Rethinking scale" (1992). Peters contended there that decentralisation is necessary for large firms, but very few are as decentralised as they can and should be. Without decentralisation, they are not 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 firm would be well advised to reduce vertical integration, although he does not offer evidence for why this is true. Overall, he found that the bureaucratic distortions of traditional firms lead to lower profitability and growth. In contrast, successful firms mimic the market as much as possible. These ideas are in line with

Williamson's description of firm limits, except for the notion that firms should always reduce vertical integration.

Schumacher (1989, 245) identified the lack of motivation in large organisations as the key disadvantage of size, providing a useful summary: "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".

Reconciliation with the "Limits of Firm Size" Framework

The above observations on diseconomies of scale do not map perfectly to Williamson's four sources of diseconomies of scale. Some are similar to his sources, others to his outcomes. Table 1 shows that Williamson's framework is strongly supported. The most important contrary evidence is Mookherjee and Reichelstein's finding (2001) that long hierarchies do not necessarily lead to control loss, and Brown, Hamilton and Medoff's discussion (1990) of the reason for labour cost differentials between large and small firms. They noticed the differential, but found no link to motivation.

Table 1. Sources of Limits of Firm Size

SOURCES OF LIMITS OF FIRM SIZE			
Atmospheric Consequences	**Bureaucratic Insularity**	**Incentive Limits**	**Communication Distortion**
Arrow (1974): Rigidity to change	Blau and Meyer (1987): Excessive rigidity	Blau and Meyer (1987): Excessive rigidity	Arrow (1974): Specialisation leads to poor communication
Blau and Meyer (1987): Excessive rigidity	Brock (1987): Risk aversion	Cooper (1964): R&D incentives	Arrow (1983): Information loss in R&D
Brown, Hamilton and Medoff (1990): Unexplained wage differential	Carroll and Hannan (2000): Firm age leads to insularity	Crozier (1964): Rigidity	Barnard ([1938] 1968): Communication losses
Child (1973): Insularity	Child (1973): Insularity	Peters (1992): Low productivity in R&D	Cooper (1964): R&D coordination
Cooper (1964): R&D cost control	Crozier (1964): Rigidity	Rasmusen and Zenger (1990): Employment contracts	Geanakoplos and Milgrom (1991): Information signal delays
Crozier (1964): Alienation	Jensen (1986): Firms larger than optimum	Schmookler (1972): Quality of R&D employees	McAfee and McMillan (1995): Lower efficiency
Kwoka (1980): Low job satisfaction in large firms	Merton (1957): Rigidity	Silver and Auster (1969): Limits to entrepreneurship	Mookherjee and Reichelstein (2001): No control loss under certain restrictive conditions
Merton (1957): Rigidity	Monsen and Downs (1965): Different owner/manager objectives	Williamson (1996): Weaker incentives in bureaucracies	Simon ([1947] 1976): Processing bottlenecks
Pugh et al. (1969): Insularity from reality	Olson (1982): Rigidity	Zenger (1989, 1994): Employment contract disincentives in R&D	
Qian (1994): Monitoring costs/inadequate effort levels	Pondy (1969): Increase in administration		
Scherer (1976): Low job satisfaction in large firms	Pugh et al. (1969): Insularity from reality		
Schmookler (1972): R&D cost consciousness; Climate for innovation	Schmookler (1972): Understanding market needs in R&D		
Schumacher (1989): Low motivation	Stinchcombe (1965): Perpetuation of organisation form		
	Williamson (1996): Bureaucratic rigidity		

Economies of Scale

This brings us to economies of scale. According to some TCE-authorities (Masten 1982; North and Wallis 1994), these should not be incorporated into the framework because they are independent of the choice of market or hierarchy, once technological indivisibilities are captured within the firm. That is, economies of scale will be reaped regardless of whether all production is carried out in one firm or in many firms. Thus, the intuitively appealing notion that the existence of economies of scale offsets size disadvantages is, according to these authorities, incorrect. This is at odds with Riordan and Williamson's argument (1985) discussed earlier.

The argument has never been tested directly. However, since the 1950s, extensive research has covered the nature and magnitude of economies of scale 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 the findings will only be briefly summarised here. In general, the research shows that economies of scale do not play a major role in explaining firm size.

Bain pioneered this line of research in the 1950s and subsequently revolutionised the study of industry and firm behaviour with his book *Industrial Organization* (1968). "The Rationale of Concentration—Efficiency and Other Considerations" from that book reviews the scale-economies argument. Bain divided the analysis into plant- and firm-level analyses.

At the plant level, economies of scale are exploited by specialising the work force and management, and by using dedicated machinery. Each plant has a minimum optimal scale and beyond this scale few additional economies of scale can be exploited. Bain found that in a study of twenty industries (all within the manufacturing sector), only two (automobiles and typewriters) showed significant economies of scale: "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). In other words, there is scant evidence at the plant level for benefits of size.

At the firm level, Bain's study showed that economies of scale derive from benefits of large-scale management, a large distribution system and purchasing power.[5] He then noted that these firm-level economies of scale are elusive, if they exist at all. His research indicated that "where economies of the multi-plant firm are encountered, they are ordinarily quite slight in magnitude...the unit costs...are typically only 1 or 2 per cent below those of a firm with one plant of minimum optimal scale". Of the twenty industries studied, Bain was able to quantify firm-level economies of scale for twelve industries. Of these twelve industries, none exhibited even moderate scale effects (p. 195).

Bain (1978) later summarised his argument as follows: "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).

Scherer and Ross provided an overview of the economies of scale debate in "The Determinants of Market Structure: Economies of Scale" (1990). They underscored that it is difficult to draw simple conclusions about the relationship between size and returns. In general, they found that economies of scale are exhausted at a surprisingly small firm size.

In a study of twelve industries, they found that market concentration could not be explained by minimally efficient scale considerations. The largest firms in the twelve industries were between two and ten times larger than economies of scale necessitated.

---

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

Scherer and Ross argued that to the extent that economies of scale accrue for large firms in those industries, they derive from savings in overhead costs (including R&D and marketing) and fixed costs in tangible assets. The economies of scale in overhead are similar to the governance-cost scale economies discussed by Riordan and Williamson (1985, 373), indicating some support for their proposition.

A number of theoretical studies (Ijiri and Simon 1964; Lucas 1978; Nelson and Winter 1982; Simon and Bonini 1958) have demonstrated that large firms evolve stochastically, regardless of economies of scale, for the simple reason that they beat the competition over time. Losers disappear, and winners grow at differential rates depending on how many times they won and how much time this took. Given this logic, firms are large because they are winners, not because they realise economies of scale.

Based on realistic assumptions about industry growth rates, variance in firm profitability and so on, simulations have yielded firm-size distributions similar to those observed in real life. As Ijiri and Simon 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" (p. 78).

An empirical test of the stochastic evolution model was carried out by Rumelt and Wensley (1981), who looked at whether high market share led to high profitability, or whether successful firms with high profitability, also achieve high market share. They concluded that "scale economies and/or market power are much less important than stochastic growth processes" (p. 2). Note that the stochastic-growth-process argument also implies that older firms will be more profitable than younger firms. Again, the older firms which still exist are survivors, while younger firms include both winners and losers.

Finally, Peters argued that economies of scale do not exist any more—if they ever existed. In his words: "technology and brainware's dominance is taking the scale out of everything" (1992, 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). In sum, these studies found only slight scale effects. The evidence in the literature review is therefore inconclusive with regard to the argument made by Riordan and Williamson (1985), that economies of scale offset diseconomies of scale.

Moderating Factors

This section reviews the literature to validate Williamson's two moderating factors: organisation form and degree of integration.[6] The literature review lends strong support to Williamson's framework.

Organisation Form

Chandler has argued, in a series of well-known studies (Chandler 1962, 1977, 1982, 1990, 1992; Chandler and Daems 1980), that large firms evolve from functional structures to multidivisional structures as they grow in size and scope of activities. In Chandler's view, the functional (unitary) form is not able to achieve the necessary coordination to be successful in the marketplace; functional economies of scale are too small to make up for this deficiency.

Thus, as firms became more diverse in the early twentieth century they adapted the multidivisional form pioneered by E. I. du Pont de Nemours & Company and General Motors

[6] A third moderating factor, financial synergies, was dismissed because it does not appear to moderate diseconomies when capital markets are efficient (Canbäck 2002a, 71).

Corporation. This line of reasoning is supported by most authorities, including Peters (1992), who found that decentralisation brings major benefits to large firms. Three important quantitative studies illustrate Chandler's argument:

Fligstein (1985, 385–386) showed that between 1919 and 1979, the number of large firms[7] with the multidivisional form went from none to 84 per cent. He estimated that the spread of the multidivisional form is mainly due to the increase of multi-product strategies, in line with Chandler's argument.

Armour and Teece (1978) quantified the difference in profits between functional- and multidivisional-form firms in the petrochemical sector, and summarised as follows: "We find strong support for the M-form hypothesis. In the 1955–1968 period the multidivisional 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).

Later, Teece (1981) studied eighteen manufacturing industries and two retail industries. He found that the multidivisional form outperformed the functional form by an average of 2.37 percentage points (p. 188). He concluded: "the M-form innovation has been shown to display a statistically significant impact on firm performance" (p. 190). These authorities are typical of the strong support for Williamson's view that organisational structure matters and that correct organisational choices can alleviate the effects of diseconomies of scale.

Degree of Integration

There is an extensive literature on vertical and lateral integration based on transaction cost economics and other theories. Mahoney (1989, 1992) and Shelanski and Klein (1995) provide summaries. Two issues are relevant here:

- Do asset specificity, uncertainty and transaction frequency explain the degree of vertical integration?
- Does Williamson's framework extend to integration in general?

Asset specificity has repeatedly been shown to be the primary 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 does not hold up well in statistical analyses.

Walker and Weber's (1984, 1987) results are typical. They found that volume uncertainty had some impact on the decision to vertically integrate and that technological uncertainty had no impact on vertical integration. Transaction frequency has, unfortunately, not been studied explicitly, perhaps because it is not independent from various types of asset specificity. Piecemeal evidence from other studies suggests that it is even less important than uncertainty when asset specificity is part of the analysis (e.g., Mahoney 1992, 571). Finally, Holmström and Roberts (1998, 79) found that both uncertainty and transaction frequency are less important factors than asset specificity.

As for the second issue, Williamson's framework appears to extend to integration in general. Grossman and Hart (1986) and Teece (1976, 1980, 1982) illustrate the use of TCE in lateral relationships. Asset specificity influences integration from a geographic reach, product breadth, and vertical depth point of view. Teece (1976) showed that multinational firms only

[7] The 131 (120) largest manufacturing firms by assets in 1919 (1979).

exist because the combination of asset specificity and opportunism leads to moral hazard, which is difficult to contain in market transactions.

Without, for example, human asset specificity, a firm could just as easily license its technology to a firm in another country, reaping the benefits of development. Tsokhas (1986) illustrated this in a case study of the Australian mining industry. Other studies have shown that market diversity reduces profitability (e.g., Bane and Neubauer 1981). Thus, there is support for Coase's 1932 view[8] that the distinction between vertical and lateral integration is without value (1993, 40).

A number of studies of product breadth show that asset specificity plays a major role in explaining the success and failure of diversification. Rumelt (1974) found a strong correlation between profitability and human asset specificity—in this case the degree to which a firm draws on common core skills or resources (pp. 121–127).

In two studies of the Fortune 500 list of American firms, he demonstrated that focused firms derive 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, 539). In sum, asset specificity seems to explain integration in general, not only vertical integration.

Reconciliation with the "Limits of Firm Size" Framework

Table 2 summarises the moderating influences on diseconomies of scale. There is again strong support for Williamson's framework. The choice of M-form organisation was found to influence firm performance positively. The determinant of degree of integration has been narrowed down to asset specificity, while uncertainty and transaction frequency were found to be less important.

Table 2. Potential Moderators of Diseconomies of Scale

POTENTIAL MODERATORS OF DISECONOMIES OF SCALE		
M-Form Organisation	**Asset Specificity**	
Armour and Teece (1978): M-form increases ROE	Bane and Neubauer (1981): Market diversity reduces profitability	Masten (1984), Masten et al. (1989, 1991), Monteverde and Teece (1982), Joskow (1993), Klier (1993), Krickx (1988): Asset specificity more important than uncertainty and frequency
Chandler (e.g., 1962), Chandler and Daems (1980): M-form alleviates coordination and control problems	Coase (1993): No distinction between vertical and lateral integration	Rumelt (1974): Product diversity reduces asset specificity
Fligstein (1985): Multi-product coordination favours M-form	Grossman and Hart (1986), Teece (e.g., 1976): TCE applies to lateral integration	Teece (1976), Tsokhas (1986): Asset specificity influences geographic reach
Peters (1992): Decentralisation is critical to firm performance	Mahoney (1992), Holmström and Roberts (1998): Uncertainty and frequency not important	Walker and Weber (1984, 1987): Volume uncertainty is weak factor
Teece (1981): M-form firms are significantly better performers than U-form firms		

HYPOTHESES

The literature review discussed the theoretical and empirical studies that inform the current research. The findings are now translated into five hypotheses:

H_1: Bureaucratic failure, in the form of atmospheric consequences, bureaucratic insularity, incentive limits and communication distortion, increases with firm size

H_2: Large firms exhibit economies of scale

[8] Letter to Ronald Fowler, 24 March 1932.

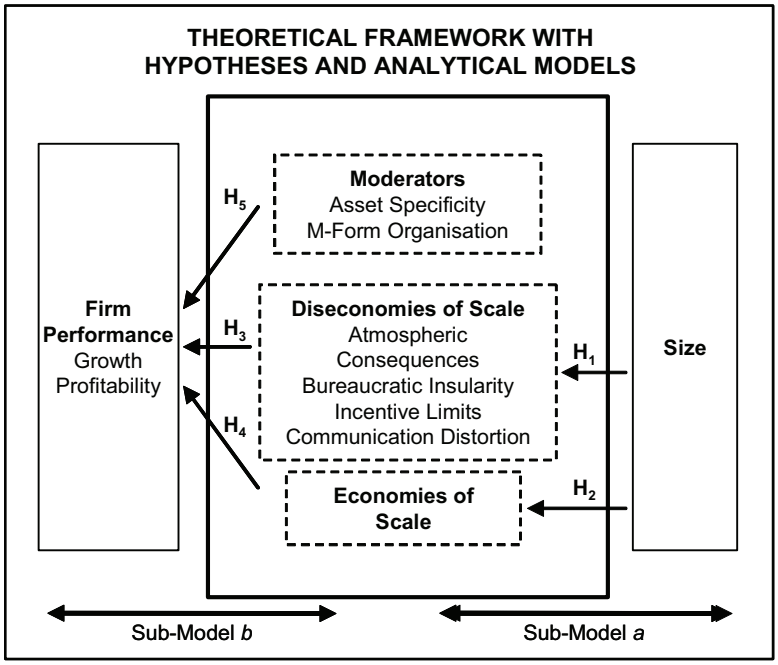
H_3: Diseconomies of scale from bureaucratic failure have a negative impact on firm performance
- H_{3a}: Atmospheric consequences have a negative impact on the performance of large firms
- H_{3b}: Bureaucratic insularity has a negative impact on the performance of large firms
- H_{3c}: Incentive limits have a negative impact on the performance of large firms
- H_{3d}: Communication distortion has a negative impact on the performance of large firms

H_4: Economies of scale increase the relative profitability of large firms over smaller firms

H_5: Diseconomies of scale are moderated by two transaction cost-related factors: organisation form and asset specificity
- H_{5a}: Large M-form firms perform better than large U-form firms
- H_{5b}: High internal asset specificity affects a firm's performance positively

Figure 4 summarises the hypotheses graphically in a theoretical framework. As the figure shows, the hypotheses are broken down into two sub-models, *a* and *b*. Sub-model *a* tests whether corporate size leads to diseconomies and economies of scale (H_1 and H_2). Sub-model *b* then tests whether the diseconomies of scale, economies of scale, and moderating factors (which are not linked to size) in turn affect firm performance (H_3 to H_5)

Figure 4. Theoretical Framework with Hypotheses and Analytical Models



The question remains: are the hypothesised effects large enough materially to influence the performance of a large firm? Only an empirical analysis, in which the framework and hypotheses are operationalised, will answer this. The next two sections focus on this operationalisation and analysis.

DATA OVERVIEW

The positivist approach taken here emphasises universal understanding in Runkel and McGrath's terms (1972, 81–89). There are no studies of this general type on the particular issue of diseconomies of scale. However, generalised studies on, for example, the profit impact of an M-form organisation or the link between size, structure and complexity are widely quoted in the literature (e.g., Rumelt 1974; Ramanujam and Varadarajan 1989). This

indicates that the generalised approach may add substantial value to the study of limits of firm size. An added benefit is that data are available to support a generalised study.

The conducted analyses were cross-sectional. Data were collected for publicly traded manufacturing firms (SIC codes 10–39) with headquarters in the US and with sales of more than $500 million. 1998 was the benchmark year. Canbäck (2002b) contains all data. 1998 was a year of high economic growth, but it was not a peak in the business cycle. Table 3 shows key indicators for the time period surrounding the year and the rank of the indicator for the time period 1961–2000.

Table 3. Select Economic Indicators for the United States

SELECT ECONOMIC INDICATORS FOR THE UNITED STATES						
Indicator	**1996**	**1997**	**1998**	**1999**	**2000**	**'61–'00 Rank**
GDP Growth (%)	3.6	4.4	4.3	4.1	4.1	Medium: 15 of 40
Mfg. GDP Growth (%)	2.4	5.4	4.1	6.1	4.1	Medium: 14 of 40
Mfg. Return on Equity (%)	16.7	16.7	15.8	16.4	15.1	High: 7 of 40
Mfg. Capacity Utilisation (%)	81.6	82.7	81.4	80.6	80.7	Medium: 20 of 40
Inflation (%)	3.0	2.3	1.6	2.2	3.4	Low: 34 of 40
Source: Council of Economic Advisers (2002)						

Primary and secondary data were collected from several sources, including company organisation charts, official filings such as 10-Ks and proxy statements, annual reports, biographies of executives, historical company documents, corporate web sites, articles in *Business Week* and *Fortune*, corporate watchdogs such as the Investor Responsibility Research Center (IRRC), Compustat and academic research. Table 4 depicts the 14 primary variables used in sub-models *a* and *b* (17 supporting variables were used to calculate some of the primary variables or for supporting stand-alone analyses).

Table 4. Overview of Primary Variables Used in the Analysis

OVERVIEW OF PRIMARY VARIABLES USED IN THE ANALYSES				
Use[a]	**Name**[b]	**Label**[c]	**Description**	**Sources**
Size (*a*)	empl	Employees	No. of employees	Compustat
Diseconomies of Scale (*a, b*)	ulabour	Atmospheric Consequences	Unit labour cost	Compustat, annual reports, 10-Ks
Diseconomies of Scale (*a, b*)	tenure	Leadership Tenure	Average years of employment with firm for officers	10-Ks, proxy statements. annual reports, corporate web sites, executive biographies
Diseconomies of Scale (*a, b*)	age	Company Age	Years since founding of company	10-Ks, proxy statements. annual reports, corporate web sites, historical sources
Diseconomies of Scale (*a, b*)	rd	Incentive Limits	Research and development expense/Sales	Compustat, annual reports, 10-Ks
Diseconomies of Scale (*a, b*)	levels	Communication Distortion	No. of hierarchical levels	Annual reports, corporate web sites, 10-Ks, company organisation charts

Economies of Scale (*a, b*)	fixhigh	Economies of Scale	Defined as (fixed cost)^2 ⁄ sales [fixed cost from supporting variables]	Compustat, annual reports, 10-Ks, supporting variables
Moderators (*b*)	foreign	Geographic Reach	% of sales derived outside the United States	Compustat, annual reports, 10-Ks
Moderators (*b*)	dr	Product Breadth	Defined as the diversification ratio (1 − Rumelt's specialisation ratio) [specialisation ratio from supporting variables]	Compustat, annual reports, 10-Ks, supporting variables
Moderators (*b*)	vert	Vertical Depth	2 = Very high; 1 = High; 0 = Average or low	10-Ks, annual reports, corporate web sites, Compustat
Moderators (*b*)	govern	Governance	Qualitative rankings	*Business Week*, IRRC, and *Fortune*
Moderators (*b*)	div	Divisionalisation	2 = Divisionalised; 1 = Hybrid; 0 = Unitary	10-Ks, proxy statements, annual reports, corporate web sites
Performance (*b*)	growth	Growth	5-year compound annual growth rate	Compustat
Performance (*b*)	eva	Profitability	Economic value added defined as return on equity (ROE) less cost of equity (COE) [ROE and COE from supporting variables]	Compustat, supporting variables

[a] *a* and *b* in the "Use" column indicate whether the variable is used in sub-models *a* or *b*
[b] The "Name" column shows the name given to the variable in SPSS and in Amos
[c] The "Label" column shows the label given to the variable in SPSS and in Amos

The original sample contained 901 firm records. After eliminating nonsensical or duplicate observations, 784 remained. The distributional properties of the data were assessed to meet the requirements for covariance-based structural equation modelling. The sample size was deemed appropriate even though two variables had a high share missing data; outliers did not pose a problem; the level of non-normality, heteroscedasticity and non-linearity was not significant after basic data transformations.

EMPIRICAL RESULTS

This section describes the structural equation models used to test the hypotheses. The philosophy of the approach has been to use as simple models and definitions as possible and to use the theoretical framework without alterations. The focus is on practical significance, rather than statistical significance. The analyses are used in a confirmatory sense. That is, the model is derived from the literature review and there is no attempt to explore new relations between variables based on the outcome of the analyses. This means that the correlations and conclusions probably are weaker than they need be in a statistical sense.

In the remainder of the section, sub-model *a* is first discussed, then sub-model *b*. Finally the statistical findings are summarised.

Sub-Model *a*: Relationship between Firm Size and Diseconomies of Scale and Economies of Scale

Sub-model *a* tests the first and second hypotheses: H_1: Bureaucratic failure, in the form of atmospheric consequences, bureaucratic insularity, incentive limits and communication distortion, increases with firm size; and H_2: Large firms exhibit economies of scale.

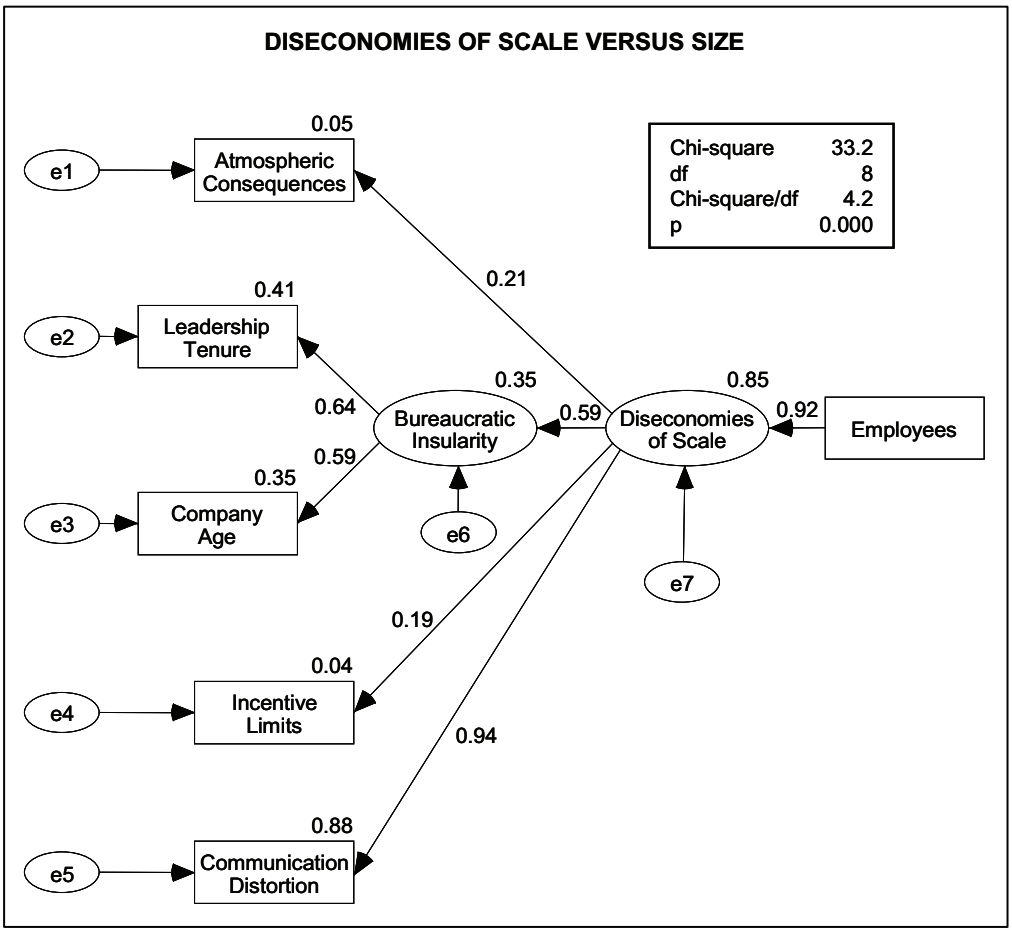
At this point, nothing is said about the importance of the diseconomies of scale and economies of scale. That is, while firm size may lead to diseconomies and economies of scale, this does not necessarily imply that firm performance is influenced. That relationship is explored in sub-model *b*.

Diseconomies of Scale

The "Theoretical Framework" section showed that there are four types of scale-related diseconomies: atmospheric consequences, bureaucratic insularity, incentive limits and communication distortion. Each of these factors is analysed in this section with the aim to determine whether it is driven by firm size.

Figure 5 shows the structural equation model for the diseconomies of scale. **Incentive Limits** has **Leadership Tenure** and **Company Age** as indicators. **Diseconomies of Scale** is a latent variable constrained on the one hand by **Employees**, on the other hand by the four factors driving diseconomies of scale.

Figure 5. Diseconomies of Scale versus Size



The path diagram in Figure 5 shows the standardised regression weights and the squared multiple correlations. All the regression weights are positive. This shows that the relationships have the hypothesised sign: increasing size leads to increasing diseconomies of scale and all four factors contribute to this increase. The squared multiple correlations vary significantly though, from 0.04 to 0.88.

Table 5 shows that the critical ratios are significant at better than the 5% level for all regression coefficients available (two coefficients were set to 1 to constrain the model). Furthermore, the model has a normed chi-square of 4.152, indicating a good fit,[9] and the normed fit index (NFI) is 0.995, well above the threshold of 0.900.[10]

Table 5. Regression Weights for Diseconomies of Scale versus Size

REGRESSION WEIGHTS FOR DISECONOMIES OF SCALE VERSUS SIZE				
	Std. Coeff	**Unstd. Coeff.**	**SE**	**CR**
Employees → Diseconomies of Scale (γ_{11})	0.924	0.333	0.138	2.424*
Diseconomies of Scale → Atmospheric Consequences	0.212	1		
Diseconomies of Scale → Bureaucratic Insularity	0.589	1.981	0.901	2.197*
Diseconomies of Scale → Incentive Limits	0.189	0.072	0.035	2.082*
Diseconomies of Scale → Communication Distortion	0.938	0.534	0.220	2.424*
Bureaucratic Insularity → Leadership Tenure	0.642	1		
Bureaucratic Insularity → Company Age	0.594	2.491	0.526	4.733***
* p<5%, *** p<0.1% (two-tailed)				

The first hypothesis—H_1: Bureaucratic failure, in the form of atmospheric consequences, bureaucratic insularity, incentive limits and communication distortion, increases with firm size—cannot be rejected. The standardised coefficient γ_{11} is 0.92 and the significance is better than the 5% level, supporting the hypothesis. The practical statistical significance is good because of the strength of most of the relationships and the high explanatory power of the analysis.

Economies of Scale

The literature survey was inconclusive regarding the effects of economies of scale. The reasons were that while it is easy to conjecture that average cost per unit of output falls with firm size, the scale effects may be exhausted at fairly small firm sizes and they may apply to entire industries rather than individual firms (because information travels fast and easily between firms). Thus the choice of market or hierarchy may not matter.

Economies of scale were quantified building on the assumption that economies of scale exist when relative fixed costs are high. The chosen definition was to take fixed and semi-fixed costs from the income statement and divide these by total factor costs (including purchased goods and services). Factor costs differ slightly from revenue because they are the sum of all inputs, including cost of equity, regardless of if the sum of these inputs is larger or smaller than revenue.

By using factor costs rather than sales, spurious business cycle effects due to yearly fluctuations in net income are eliminated. The observed variable **Fixed Cost%** was

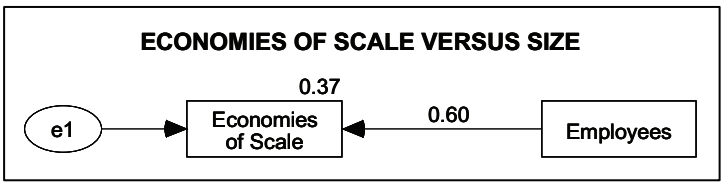
[9] Excellent fit is defined as normed chi-square (chi-square/degrees of freedom) <2, good fit < 5, and acceptable fit < 10. This is in line with Kelloway (1998, 28) and Hair et al. (1998, 623).

[10] Recommended by Hair et al. (1998, 635–636).

consequently defined as (**Interest** + **Depreciation** + **SG&A**)/**Factor Costs**. The definition assumes that fixed costs are composed of more than the contribution from fixed assets. Specifically, the level of SG&A expense (including R&D) is not easily varied and can be considered fixed. The definition is equivalent to Penrose's definition ([1959] 1995, 89–95).

The variable **Economies of Scale** was then constructed using the following logic: economies of scale are large for those firms which simultaneously are active in high fixed cost environments and have high (absolute) fixed costs. Thus, the variable **Economies of Scale** multiplies the fixed cost ratio with the absolute level of fixed cost. The relationship between economies of scale and firm size is shown in Figure 6.

Figure 6. Economies of Scale versus Size



Not surprisingly, the regression coefficient is highly significant at better than the 0.1% level (Table 6).

Table 6. Regression Weight for Economies of Scale versus Size

REGRESSION WEIGHT FOR ECONOMIES OF SCALE VERSUS SIZE				
	Std. Coeff.	**Unstd. Coeff.**	**SE**	**CR**
Employees → Economies of Scale (γ_{61})	0.605	1.797	0.086	20.800***
*** p<0.1% (two-tailed)				

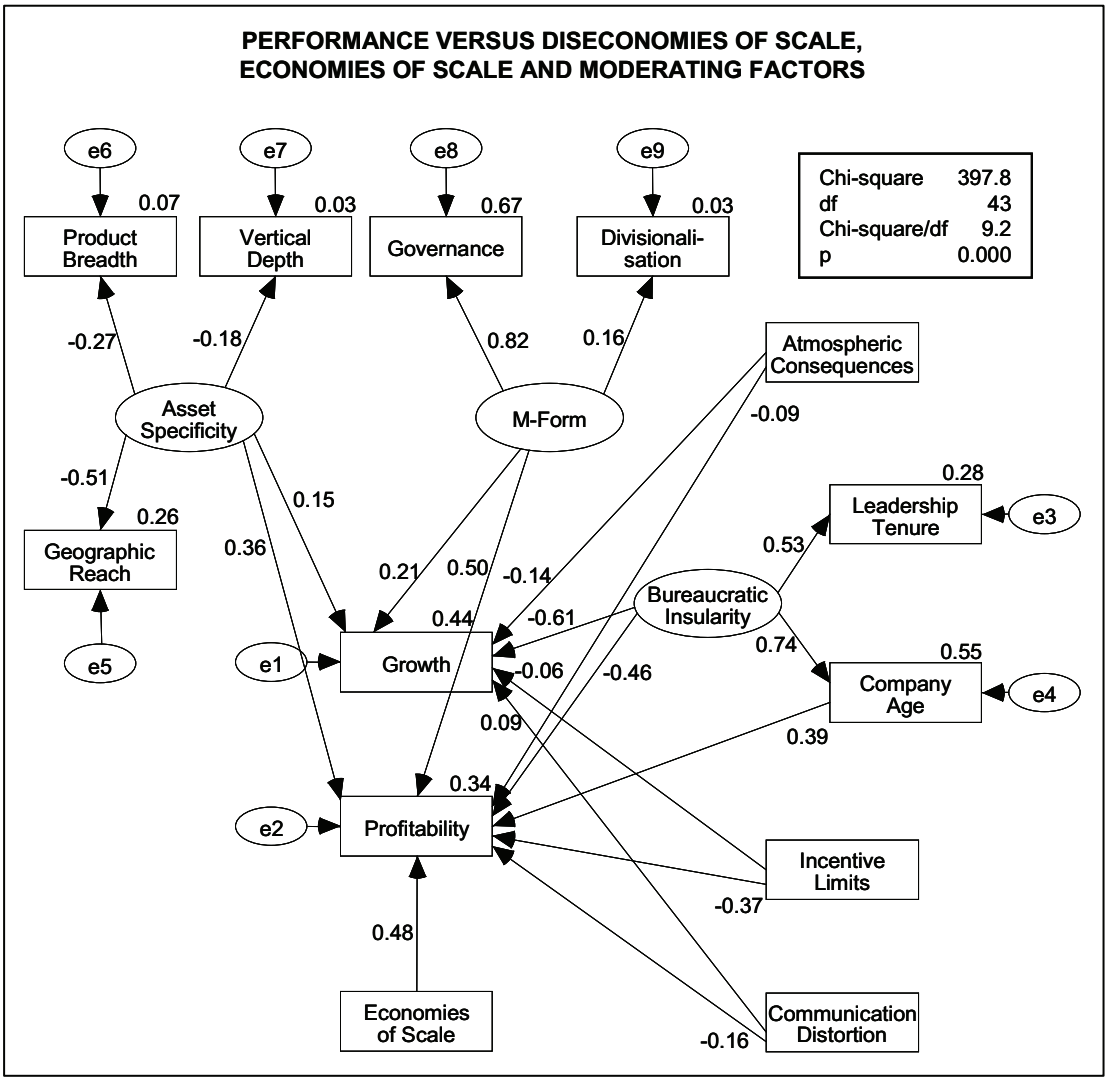
At this point, the argument that economies of scale exist among large firms cannot be rejected because $\gamma_{61}>0$, with better than 0.1% significance and —H_2: Large firms exhibit economies of scale—is confirmed. Thus, both diseconomies of scale and economies of scale tend to increase with firm size.

Sub-Model *b*: Relationship between Firm Performance and Diseconomies of Scale, Economies of Scale and Moderating Factors

Sub-model *b* tests the final three hypotheses: H_3: Diseconomies of scale from bureaucratic failure have a negative impact on firm performance; H_4: Economies of scale increase the relative profitability of large firms over smaller firms; and H_5: Diseconomies of scale are moderated by two transaction cost-related factors: organisation form and asset specificity. H_3 and H_5 have further sub-hypotheses.

In the original research (Canbäck 2002a), sub-model *b* was analysed in three steps, but the details are not reported here. First, the influence of diseconomies of scale on firm performance was analysed. Second, the influence of economies of scale on firm performance was added to the model. Third, the moderating factors were added. In the first step, the model had good explanatory power with a normed chi-square of 4.7, but with mixed significance of the individual variables. The second and third step reduced the overall model fit, but increased the variable significance. The results after step three are reported in Figure 7. The structural equation model explains 44% of the variance in growth and 34% of the variance in profitability.

Figure 7. Performance versus Diseconomies of Scale, Economies of Scale and Moderating Factors



The regression coefficients (Table 7) have the hypothesised sign (except for the non-significant **Communication Distortion → Growth**) and many coefficients are significant at the 5% or better level.

Table 7. Regression Weights for Performance versus Diseconomies of Scale, Economies of Scale and Moderating Factors

REGRESSION WEIGHTS FOR PERFORMANCE VERSUS DISECONOMIES OF SCALE, ECONOMIES OF SCALE AND MODERATING FACTORS				
	Std. Coeff.	**Unstd. Coeff.**	**SE**	**CR**
Atmospheric Consequences → Growth (γ_{11})	−0.142	−0.057	0.041	−1.417
Atmospheric Consequences → Profitability (γ_{21})	−0.087	−0.049	0.066	−0.746
Bureaucratic Insularity → Growth (γ_{12})	−0.609	−0.120	0.036	−3.348***
Bureaucratic Insularity → Profitability (γ_{22})	−0.465	−0.128	0.103	−1.244
Bureaucratic Insularity → Leadership Tenure	0.531	0.263	0.050	5.244***
Bureaucratic Insularity → Company Age	0.740	1.000		
Company Age → Profitability	0.386	0.079	0.047	1.689†
Incentive Limits → Growth (γ_{13})	−0.059	−0.019	0.027	−0.706
Incentive Limits → Profitability (γ_{23})	−0.375	−0.170	0.063	−2.688***
Communication Distortion → Growth (γ_{14})	0.092	0.333	0.312	1.067
Communication Distortion → Profitability (γ_{24})	−0.157	−0.793	0.833	−0.952
Economies of Scale → Profitability (γ_{25})	0.483	0.176	0.079	2.232*
Asset Specificity → Growth (γ_{16})	0.149	1.000		
Asset Specificity → Profitability (γ_{26})	0.365	3.431	2.213	1.550
Asset Specificity → Geographic Reach	−0.507	−1.487	0.675	−2.201*
Asset Specificity → Product Breadth	−0.268	−0.880	0.421	−2.091*
Asset Specificity → Vertical Depth	−0.179	−1.510	0.806	−1.872†
M-Form → Growth (γ_{17})	0.213	0.168	0.117	1.427
M-Form → Profitability (γ_{27})	0.498	0.548	0.409	1.339
M-Form → Governance	0.819	1.000		
M-Form → Divisionalisation	0.163	0.270	0.169	1.596
† p<10%, * p<5%, ** p<1%, *** p<0.1% (two-tailed)				

The hypotheses relating to sub-model *b* have now been tested and the results largely support the theoretical framework:

H_{3a}: 'Atmospheric consequences have a negative impact on the performance of large firms' is neither confirmed nor rejected. The regression coefficients have the hypothesised sign ($\gamma_{11} = -0.14$ and $\gamma_{21} = -0.09$), but are not significant.

H_{3b}: 'Bureaucratic insularity has a negative impact on the performance of large firms' is confirmed. Bureaucratic insularity has a strong negative impact on growth ($\gamma_{12} = -0.61$) with a significance better than 0.1%. It also has a strong negative impact on profitability ($\gamma_{22} = -0.46$), but without meeting threshold levels of significance.

H_{3c}: 'Incentive limits have a negative impact on the performance of large firms' is confirmed. Incentive limits have a strong negative impact on profitability ($\gamma_{23} = -0.37$) with better than 1% significance. The impact on growth is also negative ($\gamma_{13} = -0.06$), but the significance is low.

H_{3d}: 'Communication distortion has a negative impact on the performance of large firms' is neither confirmed nor rejected. Communication distortion has a non-significant positive impact on growth ($\gamma_{14} = 0.09$), contrary to the hypothesis, and a non-significant negative impact on profitability ($\gamma_{24} = -0.16$), in line with the hypothesis. The inconclusive nature of the finding may, however, agree with Mookherjee and Reichelstein (2001).

H_4: 'Economies of scale increase the relative profitability of large firms over smaller firms' is confirmed. The presence of economies of scale have a strong positive influence on firm profitability ($\gamma_{25} = 0.48$) at a significance better than the 5% level.

H_{5a}: 'Large M-form firms perform better than large U-form firms' is possibly confirmed. M-form appears to lead to both higher growth ($\gamma_{17} = 0.21$) and higher profitability

(γ_{27} = 0.50). The significance is low in both cases though, mainly because **Divisionalisation** reduces the significance.

H_{5b}: 'High internal asset specificity affects a firm's performance positively' is confirmed. Asset specificity has the predicted positive impact on both growth (γ_{16} = 0.15) and profitability (γ_{26} = 0.36), but the significance is low. The non-normal nature of the indicators probably leads to a large underestimate of significance. Using the **AS** indicator, the significance is better than the 1% level for growth and 10% level for profitability.

The practical significance is quite high at this point. The fit between the theoretical framework and the statistical analysis for sub-model *b* is in some ways surprisingly good, even though the test statistics vary in strength.

Competing and Parsimonious Models

Four competing models were tested. Alternative 1: a correlation was added between **M-form** and **Atmospheric Consequences**. The logic behind this is that employees in M-form firms presumably are more motivated than employees in U-form firms because they work in smaller organisational units and with better governance. Alternative 2: a correlation was added between **M-form** and **Bureaucratic Insularity** because individual units in an M-form firm should be more exposed to the surrounding market and less isolated from external pressures. Alternative 3: both the above correlations were added. Alternative 4: the correlation between **M-form** and **Communication Distortion** was deleted. The logic for this is that the adoption of M-form organisation may not be driven by communication distortion, but rather by other, exogenous factors such as established practices in a given industry. Note that the added or deleted correlations are theoretically plausible, but not theoretically prescribed.

Table 8. Comparison of Parsimony for Competing Models

COMPARISON OF PARSIMONY FOR COMPETING MODELS		
Description	**Normed Chi-Square**	**Parsimonious Fit Index (PFI)**
Chosen model	9.252	0.473
Alternative 1	9.465	0.462
Alternative 2	9.414	0.462
Alternative 3	9.642	0.451
Alternative 4	8.799	0.484

Table 8 demonstrates that the alternative models are similar to the chosen model. Alternative 4 is the only model with a better fit and parsimony, but only marginally so. It was nevertheless rejected because the exclusion of the correlation between **M-Form** and **Communication Distortion** does not agree as well with the theory as the chosen model.

The second step was to reduce the number of relationships in the model. This builds on the assumption that while the theoretical predictions captured in the hypotheses may be correct, they are not significant for certain relationships and thus the theory should be modified.

The pruned model uses the **AS** indicator, which is a composite of **Geographic Reach**, **Product Breadth** and **Vertical Depth**. The pruned model also eliminates the non-significant regression coefficients and correlations for the four diseconomies of scale factors. The squared multiple correlation for **Profitability** improves dramatically from 0.34 to 0.64 (because **AS** is more well-behaved than the three individual measures of asset specificity), while it remains the same for **Growth** (0.44 versus 0.42).

The normed chi-square improves from 9.252 to 6.999 and the normed fit index is slightly higher at 0.980 versus 0.966. The parsimonious fit ratio is 0.424 compared to 0.473. The regression coefficients in Table 9 show that all coefficients have the hypothesised sign and all, except one, are significant at better than the 10% level.

Table 9. Regression Weights for Pruned Sub-Model b

REGRESSION WEIGHTS FOR PRUNED SUB-MODEL *B*				
	Std. Coeff.	**Unstd. Coeff.**	**SE**	**CR**
Atmospheric Consequences → Growth	−0.131	−0.053	0.028	−1.926†
Bureaucratic Insularity → Growth	−0.508	−0.092	0.013	−6.961***
Bureaucratic Insularity → Profitability	−0.695	−0.179	0.102	−1.754†
Bureaucratic Insularity → Leadership Tenure	0.500	0.231	0.045	5.107***
Bureaucratic Insularity → Company Age	0.797	1.000		
Company Age → Profitability	0.588	0.120	0.064	1.870†
Incentive Limits → Profitability	−0.391	−0.178	0.047	−3.806***
Communication Distortion → Profitability	−0.182	−0.921	0.739	−1.247
Economies of Scale → Profitability	0.459	0.166	0.053	3.111**
AS → Growth	0.318	0.386	0.120	3.231**
AS → Profitability	0.525	0.902	0.433	2.081*
M-Form → Growth	0.355	0.479	0.247	1.940†
M-Form → Profitability	0.786	1.499	0.859	1.746†
M-Form → Governance	0.474	1.000		
M-Form → Divisionalisation	0.270	0.769	0.222	3.468***
† p<10%, * p<5%, ** p<1%, *** p<0.1% (two-tailed)				

The validity of the hypotheses has been strengthened. H_{3a} is now confirmed at the 10% level, H_{3b} is (even more) strongly supported, H_{3c} is (more) strongly confirmed, H_{3d} has increased its significance, but is still not at the 10% level, H_4 is strongly supported, H_{5a} is supported at the 10% level, while H_{5b} is strongly supported.

Finally, the 784 observations were randomly divided in two groups to test whether similar results are achieved for different samples. The procedure was repeated eight times and the critical ratios of the differences were compiled for the main-effects model. Out of 104 possible differences, the analysis indicated sixteen instances of differences significant at better than the 10% level, of which ten where significant at better than the 5% level, of which two were significant at better than the 1% level. This leads to the conclusion that the results are homogenous across samples.

Summary of Statistical Findings

Table 10 shows the hypotheses and their associated findings. As was seen throughout this section, most of the hypotheses were confirmed. The findings seem to be robust for a number of reasons. The data were screened and tested extensively. They were found to be well-behaved in most respects. The path diagrams confirm well with the underlying theory. The indicators appear to reflect the unobserved phenomena fairly well. Finally, the results were similar when random sub-samples were used.

Table 10. Summary of Statistical Findings

SUMMARY OF STATISTICAL FINDINGS[a]				
Hypothesis	**Test**	**Result**	**CR and Sign.**	**Interpretation**
H_1: Bureaucratic failure, in the form of atmospheric consequences, bureaucratic insularity, incentive limits and communication distortion, increases with firm size	$\gamma_{11}(a) > 0$	$\gamma_{11}(a) = +0.92$	+2.424 (p<1%)	Confirmed
H_2: Large firms exhibit economies of scale	$\gamma_{61}(a) > 0$	$\gamma_{61}(a) = +0.60$	+20.800 (p<0.1%)	Confirmed
H_{3a}: Atmospheric consequences have a negative impact on the performance of large firms	$\gamma_{11}(b) < 0$ $\gamma_{21}(b) < 0$	$\gamma_{11}(b) = -0.13$ –	−1.926 (p<10%) –	Confirmed
H_{3b}: Bureaucratic insularity has a negative impact on the performance of large firms	$\gamma_{12}(b) < 0$ $\gamma_{22}(b) < 0$	$\gamma_{12}(b) = -0.51$ $\gamma_{22}(b) = -0.70$	−6.961 (p<0.1%) −1.754 (p<10%)	Confirmed
H_{3c}: Incentive limits have a negative impact on the performance of large firms	$\gamma_{13}(b) < 0$ $\gamma_{23}(b) < 0$	– $\gamma_{23}(b) = -0.39$	– −3.806 (p<0.1%)	Confirmed
H_{3d}: Communication distortion has a negative impact on the performance of large firms	$\gamma_{14}(b) < 0$ $\gamma_{24}(b) < 0$	– $\gamma_{24}(b) = -0.18$	– −1.247 (p=21.2%)	Inconclusive
H_4: Economies of scale increase the relative profitability of large firms over smaller firms	$\gamma_{25}(b) < 0$	$\gamma_{25}(b) = +0.46$	+3.111 (p<1%)	Confirmed
H_{5a}: Large M-form firms perform better than large U-form firms	$\gamma_{17}(b) > 0$ $\gamma_{27}(b) > 0$	$\gamma_{17}(b) = +0.36$ $\gamma_{27}(b) = +0.79$	+1.940 (p<10%) +1.746 (p<10%)	Confirmed
H_{5b}: High internal asset specificity affects a firm's performance positively	$\gamma_{16}(b) > 0$ $\gamma_{26}(b) > 0$	$\gamma_{16}(b) = +0.32$ $\gamma_{26}(b) = +0.52$	+3.231 (p<1%) +2.081 (p<5%)	Confirmed
[a] For simplicity, the word "confirmed" is used, although "not rejected" is more accurate.				

The practical significance of the statistical analyses is that both sub-model *a* and sub-model *b* validate the theoretical framework. Both the main analyses and the supporting analyses that tested particular aspects of the theory are in line with the theoretical predictions.

DISCUSSION AND INTERPRETATION

Diseconomies of scale appear to be real. The literature overview discussed the theoretical underpinnings of this paper, indicating that a wide range of theoretical development and empirical research, quantitative and qualitative, supports pieces of the current theoretical predictions. The statistical analysis section took a broader and more general approach to testing the hypotheses, and nothing uncovered there disproved them. The analyses also showed that diseconomies of scale vary in magnitude and impact, and economies of scale and the moderating factors are important when we try to understand the limits of the firm (Table 11).

Table 11. Summary of Findings

SUMMARY OF FINDINGS		
Hypothesis	**Literature Finding**	**Statistical Finding**
H_1: Bureaucratic failure, in the form of atmospheric consequences, bureaucratic insularity, incentive limits and communication distortion, increases with firm size	Confirmed	Confirmed
H_2: Large firms exhibit economies of scale	Confirmed	Confirmed
H_3: Diseconomies of scale from bureaucratic failure have a negative impact on firm performance	Confirmed	Confirmed
H_{3a}: Atmospheric consequences have a negative impact on the performance of large firms	Confirmed	Confirmed
H_{3b}: Bureaucratic insularity has a negative impact on the performance of large firms	Confirmed	Confirmed
H_{3c}: Incentive limits have a negative impact on the performance of large firms	Confirmed	Confirmed
H_{3d}: Communication distortion has a negative impact on the performance of large firms	Confirmed	Inconclusive
H_4: Economies of scale increase the relative profitability of large firms over smaller firms	Inconclusive	Confirmed
H_5: Diseconomies of scale are moderated by two transaction cost-related factors: organisation form and asset specificity	Confirmed	Confirmed
H_{5a}: Large M-form firms perform better than large U-form firms	Confirmed	Confirmed
H_{5b}: High internal asset specificity affects a firm's performance positively	Confirmed	Confirmed

It is now possible to interpret the findings by returning to the neoclassical cost curves. First, the cost curve shown in Figure 2 is modified to reflect the characteristics of diseconomies of scale, economies of scale and the moderating factors. Second, a similar curve is constructed for firm growth. Third, these two curves are combined to show the overall impact of these two factors on firm performance.

Average cost. To begin with, the elongated U-shaped average total cost curve used in neoclassical theory can be split into two parts: the average production cost curve and the average transaction cost curve. Not much evidence exists for what the relative magnitude of production and transaction costs is. However, Wallis and North (1986) attempted to quantify the relative contribution each type of cost makes to the overall economy. They found that the transaction-cost part of the economy grew from 25 per cent to 50 per cent of gross national product between 1890 and 1970 (p. 121). This suggests that an even split is a reasonable assumption.

The modified cost curves are depicted in a stylised fashion in Figure 8. The top graph shows a curve for average production cost (AC_{P}) consistent with the findings in the current research. One characteristic of the curve is important: the curve has a negative slope for all levels of firm output (Q). This agrees with the view that economies of scale can be kept proprietary to the firms that reap them. It is also agrees with the statistical finding that economies of scale are not exhausted at small firm sizes.

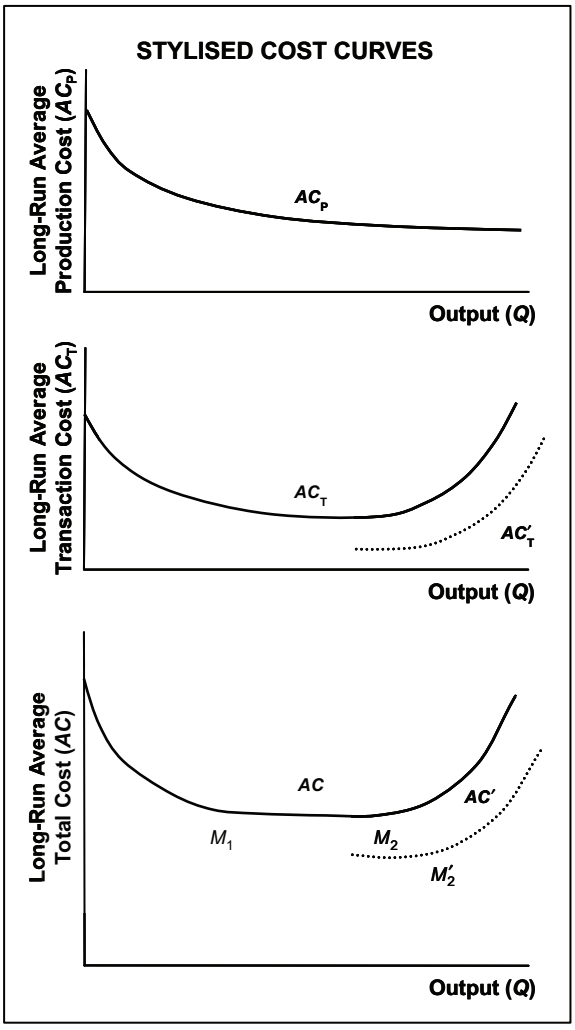
The middle graph in Figure 8 shows the average transaction cost curve (AC_{T}). The negative slope for smaller firms, indicating bureaucratic economies of scale, is supported in the literature review (but was not tested in the statistical analysis). The positive slope for larger firms, indicating diseconomies of scale and bureaucratic failure, is supported by both the literature and by the statistical analysis.

The middle graph also shows a shifted and slightly tilted average transaction cost curve (AC'_{T}). The curve reflects the positive contribution from the moderating factors. AC'_{T}

is supported by the literature and by the statistical analysis. This analysis indicates that the shift can be quite large.

Finally, the bottom graph in Figure 8 shows the average total cost curve (AC), with a shifted curve AC' for the moderators ($AC = AC_P + AC_T$; $AC' = AC_P + AC'_T$). The curve resembles the neoclassical curve in Figure 2. The question now is: where along this curve do firms operate? The statistical analyses suggest that, on average, the largest firms in the sample operate at outputs somewhere close to M'_2 in the upward-sloping region of AC'. That is, they show some diseconomies of scale, but they also benefit from economies of scale and they manage to take advantage of the moderating factors.

Figure 8. Stylised Cost Curves



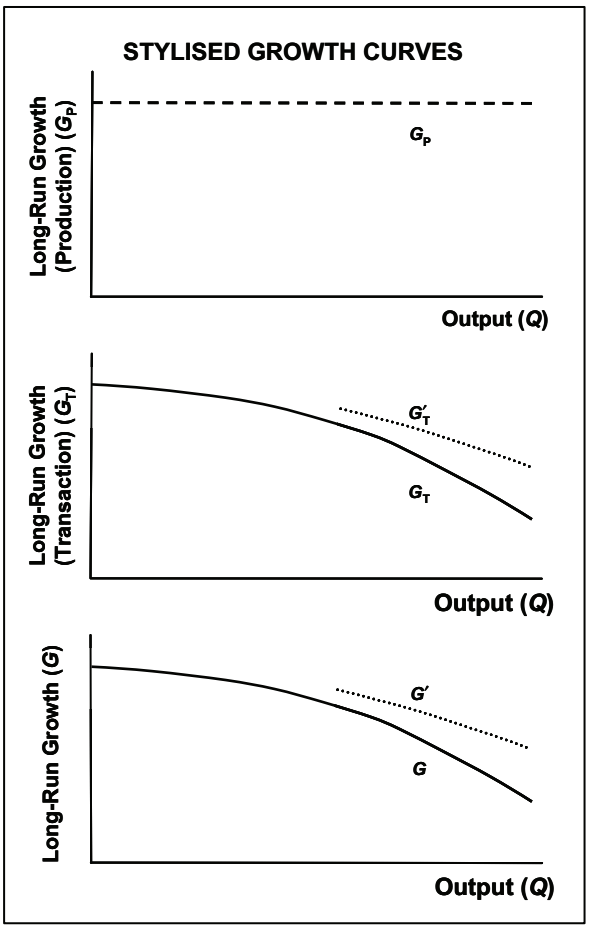
Growth. The underlying logic of the cost curves can also be applied to firm growth. Figure 9 shows the same set of graphs as above for the relationship between firm growth and output. The top graph illustrates the relationship between growth and output, under the hypothetical assumption that firms only have neoclassical production costs (G_P). Neither the literature nor the statistical analysis indicated an influence and thus the graph shows a constant relationship.

The middle graph in Figure 9 portrays the growth curve resulting from bureaucratic, transaction cost-based, failure (G_T). The literature and the statistical analysis make it fair to assume that G_T should be monotonously declining for increasing outputs. Again, the

moderating influences can shift the curve, which is illustrated by G'_T in the graph. The statistical analysis indicates that the shift is smaller than in the case of average costs (AC').

The bottom graph in Figure 9 convolutes the production- and transaction-cost contributions to growth into overall growth (G). The graph shows that the growth capacity of firms is steadily declining as a function of output, but it can be moderated (G'). Interestingly, this interpretation of the research contradicts Gibrat's law of proportional effects (1931, 74–81), which will be discussed later in this section.

Figure 9. Stylised Growth Curves



Performance. Finally, it is instructive to combine the cost and growth curves to see how they jointly contribute to a firm's performance (Figure 10). Other factors also contribute to firm performance and the graph shows the partial contribution to performance.[11] By convoluting the average total cost (AC) and growth (G) curves, the partial performance curve Ψ results.

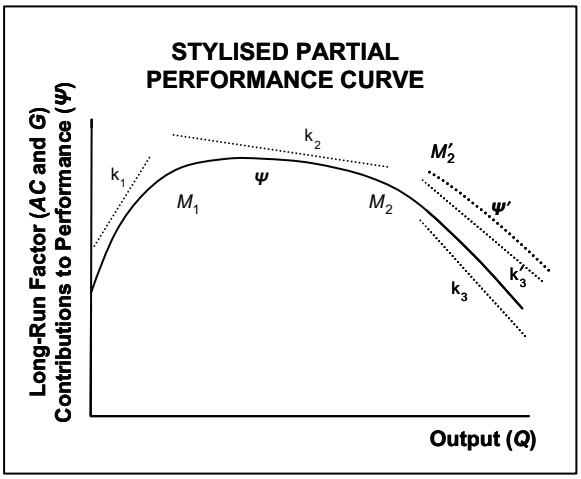
Several, perhaps speculative, interpretations can be derived from the graph: (1) Firms operating at small outputs suffer from a lack of economies of scale and this is most likely not compensated for by the higher relative growth achievable by smaller firms. Thus, the slope $k_1 > 0$. (2) There is an area where performance is fairly independent of firm size. On the one hand, economies of scale should lead to steadily lower costs. On the other hand, diminishing

[11] Total performance (Ψ_{TOT}) is a function of, profitability(π), growth(G), risk(β) and other factors (ε):
$\Psi_{TOT} = f(\pi, G, \beta, \varepsilon) = f(TR - TC, G, \beta, \varepsilon) = f(TR - AC \cdot Q, G, \beta, \varepsilon)$

growth prospects reduce performance. On balance, the analyses show that $k_2<0$, but only slightly so. (3) As diseconomies of scale due to bureaucratic failure set in, the combined negative contribution of increasing transaction costs and lower growth far outweigh economies of scale. Thus, $k_3<0$. (4) The moderating factors shift the performance curve outwards from Ψ to Ψ' and $k_3<k_3'<0$, while $M_2'>M_2$. That is, if firms judiciously apply the moderating factors, then bureaucratic failure will set in at a larger level of output and the impact from the failure will be less severe.

The four interpretations above are supported by the literature review; while the last three are supported by the statistical analysis (the statistical analysis did not explore what happens at small firm sizes).

Figure 10. Stylised Partial Performance Curve



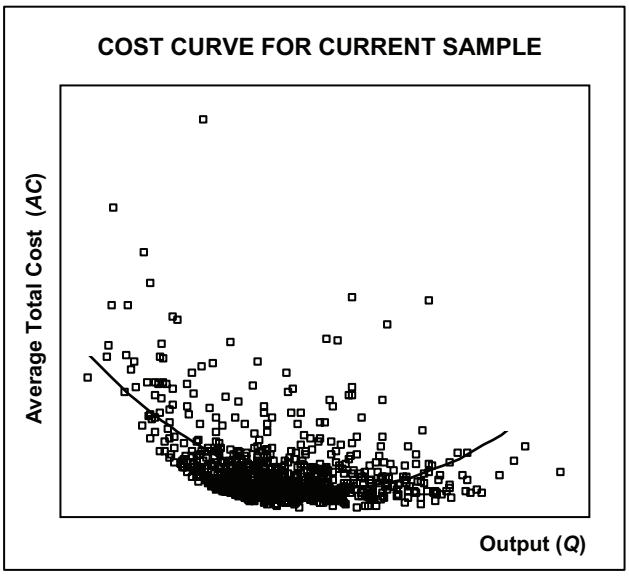
The set of curves discussed above agree well with neoclassical theory (e.g., Panzar 1989) and transaction cost economics (e.g., Williamson 1975), individually. The curves also agree with the joined perspectives on production and transaction costs expressed by, for example, Riordan and Williamson (1985) and Wallis and North (1986). What may make them interesting is the unbundling of the production cost and transaction cost contributions to firm performance, and the attempt to transform the research findings into rough estimates of the shapes of the curves.

The conceptual curves depicted in Figures 8 to 10 can also be used to show the shape of the data in the sample of 784 firms. This was done with three analyses which replicated the cost (*AC*), growth (*G*) and partial performance (Ψ) curves. Figures 11 to 13 show the resulting graphs, which are surprisingly similar to the conceptual curves.

It should be remembered though, that the scatterplots presented are somewhat simplistic. They use the sample data as is and no attempt was made to include control variables or to make other corrections.

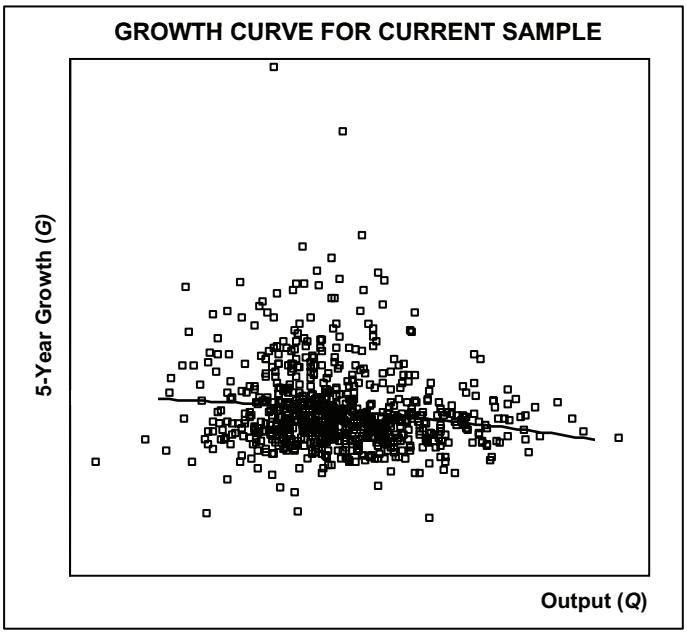
First, Figure 11 reports the results for the cost curve (*AC*), which plots average total cost (average factor costs, defined in, was used as the proxy) against output (firm size was used as the proxy). A quadratic regression line has been added to show the underlying trend in the data. The data conforms well to the conceptual *AC* curve in Figure 8.

Figure 11. Cost Curve for Current Sample



Second, growth data was plotted against output (Figure 12). Again, the curve has the predicted shape and the quadratic regression line is similar to the conceptual *G* curve in Figure 9. The plot points are quite scattered though, and firms seem to have considerable leeway to deviate from the growth rate prescribed by their size.

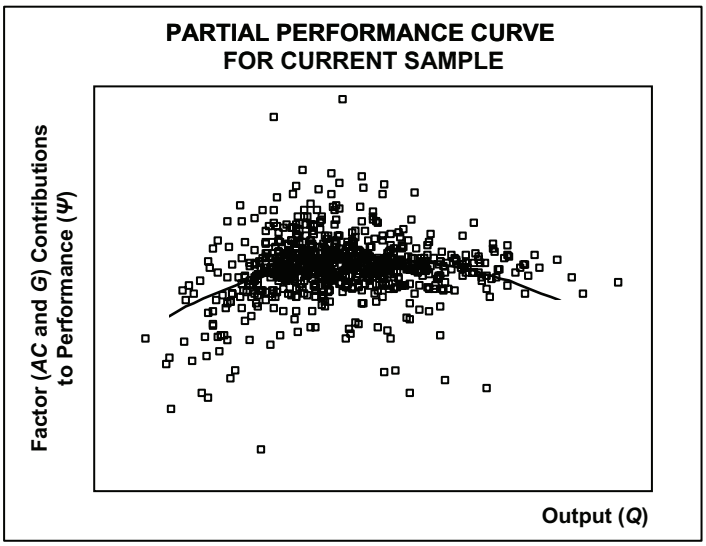
Figure 12. Growth Curve for Current Sample



Third, the joint contribution to firm performance by the two factors is shown in Figure 13. The average cost and growth data have been weighted and added ($\Psi = -0.6AC + 0.4G$, normalised). The weights for the current sample came from an analysis of the relative contribution of *AC* and *G* to Tobin's *Q*, a commonly used composite measure of a firm's performance (e.g., Brainard and Tobin 1968; Lang and Stulz 1993).

The performance curve (Ψ) is not unlike the conceptual curve shown in Figure 10. There is significant variation around the trend line, but overall the data conforms to the theoretical and empirical predictions.

Figure 13. Partial Performance Curve for Current Sample



These results now make it possible to prescribe certain remedies for underperforming large firms, especially when poor performance stems from low growth. Consider the impact of diseconomies of scale: Bureaucratic insularity at both the institutional and individual levels appears to be endemic in large firms, leading to low growth and low profitability. Incentive limits negatively influence both growth and profitability. Atmospheric consequences have a moderately negative impact on growth, while communication distortion does not seem to be an important source of diseconomies of scale.

Economies of scale can offset this to some extent; indeed, large firms tend to exist in industries in which economies of scale are important. Moreover, the negative effects of diseconomies of scale can be moderated by paying attention to governance and organisational issues and by increasing asset specificity. These factors more or less offset the diseconomies of scale for large firms, resulting in a low overall correlation between performance and size.

The diseconomies of scale exhibit a stronger negative influence on growth than on profitability. This may indicate that Penrose's suggestion that the limits of a firm are related to dynamic factors rather than static factors is correct. A large firm will find it relatively easy to maximise profitability, but difficult to spur growth. An extension of this argument is that Gibrat's law of proportional effects (1931, 77) may not be valid for growth and firm size, in line with corporate demography research (Carroll and Hannan 2000, 315–319) and the findings of Sutton (1997).

CONCLUSION

There are a number of real-life implications of the research. First, strategy and structure appear to be intimately linked. Executives at large corporations have to grapple with real trade-offs when they consider expansion. Certain growth strategies are easier to execute than others, and the choice of organisation has major implications for which strategies make sense. Indeed, structure does not necessarily follow strategy; strategy and structure inform each other continuously and forever.

Second, much of the rationale for mergers and acquisitions seems to be weak, at best (see also Canbäck 2005). Proponents of mergers typically argue that the resulting larger entity after a merger will realise economies of scale, benefiting customers and shareholders; in addition, they claim that growth will be accelerated through the introduction of new products and services that were previously too expensive to develop. But the analysis here shows that although some economies of scale may be realised, they are likely to be offset by

diseconomies of scale. Furthermore, there is no evidence that larger, merged entities innovate more and grow faster. Instead, the opposite appears to be true: innovation and growth decline.

Third, boards of directors may want to emphasise the importance of executive renewal and the elimination of rigid processes to stimulate growth. Maximising the quality of governance, which is part of the board's fiduciary duties, appears to be an important lever for addressing these issues.

Fourth, firms that strive for high internal asset specificity appear to be better off than those that expand reach, breadth, or depth. This does not imply that single-product or single-geography strategies are optimal (because this reduces growth in the long run), but it does imply that any expansion strategy should strive for high asset specificity and that some firms are best off reducing their scope of activities.

Finally, in a world in which companies increasingly try to sell solutions rather than basic products and services, incentive limits have become real and problematic. In businesses that involve team selling or large product-development efforts, attention should be paid to creating well-functioning incentive schemes for employees. The superior productivity of research and development in small firms, in which incentives are tailored to individual performance, demonstrates why effective incentive schemes matter.

From a research perspective, the current work indicates a number of opportunities for further study. For example, the statistical analyses indicate yet another way to put Gibrat's law of proportional effects (1931, 74–81) into doubt. The research also suggests four areas for further research: (1) proving the existence of diseconomies of scale by studying a more narrowly defined problem such as focusing on an industry rather than a whole economic sector; (2) expanding the analysis across geography and time; (3) finding better ways to operationalise unobserved diseconomies of scale; and (4) replicating the current research with better statistical approaches and a larger sample, with a particular eye towards industry effects.

REFERENCES

Adams, W. and J. W. Brock. 1986. *The Bigness Complex: Industry, Labor, and Government in the American Economy*. New York: Pantheon.

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.

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

Axtell, R. 1999. The Emergence of Firms in a Population of Agents: Local Increasing Returns, Unstable Nash Equilibria, and Power Law Size Distributions. Working paper, Washington: Center on Social and Economic Dynamics, Brookings Institution.

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.

Barnard, C. [1938] 1968. *The Functions of the Executive*. Revised ed. Reprint, with an introduction by K. R. Andrews. Cambridge: Harvard University Press. Original edition, Cambridge: Harvard University Press.

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.

Brainard, W. C. and J. Tobin. 1968. Pitfalls in Financial Model Building. *American Economic Review* 58 (May): 99–122.

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. 2002a. Bureaucratic Limits of Firm Size: Empirical Analysis Using Transaction Cost Economics. DBA thesis, Henley Management College, Brunel University, Uxbridge, England.

———. 2002b. Supporting Thesis Materials. http://canback.com/henley.htm (accessed 17 January 2002).

———. 2005. Does Corporate Size Matter? *Axess*, 4 (7). http://www.axess.se/english/2005/07/theme_canback.php (accessed 2 February 2006).

Carroll, G. R. and M. T. Hannan. 2000. *The Demography of Corporations and Industries*. Princeton: Princeton University Press.

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.

———. 1993. The Nature of the Firm: Origin. In *The Nature of the Firm: Origins, Evolution, and Development*, edited by O. E. Williamson and S. G. Winter. New York: Oxford University Press.

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.

———. 2002. *Economic Report of the President*. Washington: U.S. Government Printing Office.

Crozier, M. 1964. *The Bureaucratic Phenomenon*. Chicago: University of Chicago Press.

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.

Gibrat, R. 1931. *Les inégalités économiques*. Paris: Librairie du Recueil Sirey.

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. and J. Roberts. 1998. The Boundaries of the Firm Revisited. *Journal of Economic Perspectives* 12 (4): 73–94.

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.

———. 2000. *A Theory of the Firm: Governance, Residual Claims, and Organizational Forms*. Cambridge: Harvard University Press.

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.

Joskow, P. L. 1993. Asset Specificity and the Structure of Vertical Relationships: Empirical Evidence. In *The Nature of the Firm: Origins, Evolution, and Development*, edited by O. E. Williamson and S. G. Winter. New York: Oxford University Press.

Kelloway, E. K. 1998. *Using LISREL for Structural Equation Modeling*. Thousand Oaks, Calif.: Sage.

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.

Kumar, K. B., R. G. Rajan and L. Zingales. 1999. What Determines Firm Size? Working paper, Cambridge, Mass.: NBER.

Kwoka, J. E. 1980. Establishment Size, Wages, and Job Satisfaction: The Trade-Offs. In *The Economics of Firm Size, Market Structure, and Social Performance*, edited by J. J. Siegfried. Washington: Federal Trade Commission.

Lang, L. H. P. and R. M. Stulz. 1993. Tobin's *Q*, Corporate Diversification and Firm Performance. Working paper, Cambridge, Mass.: NBER.

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.

Marshall, A. [1920] 1982. *Principles of Economics: An Introductory Volume*. 8th ed. Philadelphia: Porcupine Press. Original edition, London: Macmillan.

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.
Merton, R. K. 1957. *Social Theory and Social Structure*. Revised and enlarged ed. Glencoe, Ill.: Free Press.
Monsen, R. J. and A. Downs. 1965. A Theory of Large Managerial Firms. *Journal of Political Economy* 73 (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.
Mookherjee, D. and S. Reichelstein. 2001. Incentives and Coordination in Hierarchies. *Advances in Theoretical Economics* 1 (1): Article 4. http://www.bepress.com/bejte/advances/vol1/iss1/art4 (accessed 12 January 2002).
Mueller, W. F. and L. G. Hamm. 1974. Trends in Industrial Market Concentration, 1947 to 1970. *Review of Economics and Statistics* 56 (4): 511–520.
Nelson, R. and S. G. Winter. 1982. *An Evolutionary Theory of Economic Change*. Cambridge, Mass.: Belknap.
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.
Olson, M. 1982. *The Rise and Decline of Nations: Economic Growth, Stagflation, and Social Rigidities*. New Haven: Yale University Press.
Panzar, J. C. 1989. Technological Determinants of Firm and Market Organization. In *Handbook of Industrial Organization, Vol. 1*, edited by R. Schmalensee and R. Willig. New York: North-Holland.
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.
Pugh, D. S., D. J. Hickson, C. R. Hinings and C. Turner. 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.
Rasmusen, E. and T. Zenger. 1990. Diseconomies of Scale in Employment Contracts. *Journal of Law, Economics, and Organization* 6 (1): 65–92.

Richardson, G. B. 1972. The Organisation of Industry. *Economic Journal* 82 (September): 883–896.
Riordan, M. H. and O. E. Williamson. 1985. Asset Specificity and Economic Organization. *International Journal of Industrial Organization* 3 (4): 365–378.
Robertson, D. H. 1923. *The Control of Industry*. London: Nisbet.
Rubin, P. H. 1990. *Managing Business Transactions*. New York: Free Press.
Rumelt, R. P. 1974. *Strategy, Structure, and Economic Performance*. Cambridge: Harvard University Press.
Rumelt, R. P. 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*. 3d ed. Boston: Houghton Mifflin.
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.
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.
Sutton, J. 1997. Gibrat's Legacy. *Journal of Economic Literature* 35 (March): 40–59.
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.

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.

———. 1981. The Economics of Organization: The Transaction Cost Approach. *American Journal of Sociology* 87 (3): 548–577.

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

———. 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.