AN EXPLORATION OF THE STOCHASTIC APPROACH TO MODELLING INDUSTRIAL DYNAMICS AND SIZE DISTRIBUTIONS

by

N. M. Paull

A Thesis Submitted in Fulfilment of the Requirements for the Degree
of Doctor of Philosophy of Cardiff University

Economics, Cardiff Business School, Cardiff University

November 2007

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

THESIS DECLARATION AND STATEMENTS

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree

Signed ....N. PAULL......................... (candidate)

Date ....26/11/07...........................

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of Ph.D.

Signed ....N. PAULL......................... (candidate)

Date ....26/11/07...........................

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references.

Signed ....N. PAULL......................... (candidate)

Date ....26/11/07...........................

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ....N. PAULL......................... (candidate)

Date ....26/11/07...........................

CONTENTS

TABLES

FIGURES

SUMMARY

Under the stochastic process known as Gibrat's Law firm growth follows a random walk with no persistence in growth rates and both firm growth rates and the variance of growth rates independent of firm size. Over time, for a fixed population of firms, such a stochastic process gives rise to a lognormal limiting distribution of firm sizes.

Utilising an official U.K. regional database of manufacturing plants, the empirical validity of the propositions of Gibrat's Law and the lognormal distribution are examined across three sequential U.K. economic cycles, making use of a range of econometric and statistical methods to assess the sensitivity of the results.

Formal statistical tests and non-parametric kernel density estimation methods provide evidence of generally stable aggregate plant size distributions, though frequently exhibiting deviations in the form of positive skewness and leptokurtosis compared to the lognormal distribution. Examination of age and selection effects indicates some right-shift of the plant size distributions with time, though not generally in the form of a simple or consistent evolution towards the lognormal distribution. Plant growth rate distributions are found to often exhibit highly peaked distributions with "fatter tails" than the normal or Laplace distributions. Considerable heterogeneity is observed in both plant size and growth rate distributions across industries.

Using a number of econometric model specifications the proposition of equi-proportional plant growth is statistically rejected for aggregate samples with mean reversion observed, a finding robust to corrections for heteroscedasticity and sample selection bias. However, equi-proportional plant growth cannot be rejected for a number of sub-groups. Consistent with previous research, plant growth is found to decline with age. Statistical tests reject the proposition of the variance of plant growth rates being independent of plant size, with a consistent negative relationship identified. Regression approaches suggest the absence of significant plant growth rate persistence effects.

ACKNOWLEDGEMENTS

It is a pleasure to express my sincere gratitude to those individuals who have contributed to the development of this research. In particular, I am especially grateful to my supervisor, Professor Robert McNabb, who has provided considerable encouragement, expertise and practical advice throughout the course of this work, and without whom this research would not have been possible. In addition, the considerable support of colleagues at the Welsh Assembly Government has been much appreciated. This study would again not have commenced without their generous permission to make use of the WRME microdata. However, the analysis and interpretation of the provided microdata contained within this thesis remain my own responsibility, and should not be considered as reflecting the views of the Welsh Assembly Government.

INTRODUCTION

The dynamics of firms, industries and markets continues to provide one of the most fascinating aspects of economic research. In each period new firms emerge from the activities of entrepreneurs and through the formation of new business ventures from existing firms, both domestic and foreign. Incumbent firms also evolve as a consequence of internal forces, external market developments and the introduction of modern technologies in the production process. Simultaneously, some other firms cease to exist in their present form, a consequence of mergers and acquisitions, management buy-outs and buy-ins or exit from the market through the closure of less productive firms and of those ventures unable to adjust effectively to changing market demands.

Whilst many economic models have been proposed as explanations for the observed empirical dynamics of firms, industries and markets, one of the most engaging themes has been the formulation of economic models grounded within the stochastic tradition. These models have provided a theoretical and empirical framework to examine both the growth and survival dynamics of firms and to simultaneously investigate the consistent empirical observations of highly skewed firm size distributions. Of these stochastic models, Gibrat's (1931) Law of Proportionate Effect has provided one of the primary benchmarks within the field of industrial economics.

Since the original formulation of Gibrat's Law and the statistical examination of the skewed distributions of manufacturing plant sizes conducted by Gibrat a voluminous literature has developed examining stochastic models, with periodic reassessments of both the theoretical propositions and the empirical evidence. These reassessments have often involved either the addition of further complexity to the underlying stochastic models through the incorporation of features aiming to replicate particular features of markets and industries or the re-examination of the propositions of the basic stochastic model using more detailed business micro-datasets and employing more sophisticated econometric and statistical approaches.

Given the existence of such an extensive economic and statistical literature examining both methodological issues relating to the assessment of Gibrat's Law and modified Gibrat-type processes, and the associated empirical evidence spanning a wide range of industries, time periods and geographies, a primary objective of this thesis is to provide a detailed survey of this research drawing together some of the main conclusions on the acceptability of Gibrat's Law as a broad description of the patterns of firm dynamics. In this undertaking a significant aim is to consider the effects of these methodological developments, including the role of model specifications, estimation procedures, the coverage of business micro-datasets and the time periods examined on the reported empirical validity of this particular stochastic model.

Building on this survey of previous studies, this thesis further aims to generate new empirical evidence regarding the applicability of Gibrat's Law, providing a detailed examination of manufacturing plant dynamics and plant size distributions. In addressing this objective, the analysis also sets out to provide a broad empirical overview of the historical patterns of plant and employment dynamics within the manufacturing sector. In this research use is made of a relatively under-exploited official business micro-dataset, the Welsh Register of Manufacturing Employment (WRME).

In providing this new analysis the research extends the existing literature in a number of important respects. First, although many previous empirical studies have tested Gibrat's Law for selected (generally short) time periods the long time dimension of the WRME dataset, providing electronic records back to 1966, permits a more comprehensive analysis across a wider range of time dimensions and phases of economic cycles than considered in almost all other previous studies. Specifically, the analysis examines patterns of plant growth over three complete U.K. economic cycles, across both defined periods of macroeconomic expansions and contractions, and additionally for the shortest period of one-year growth, permitting a rigorous assessment of the validity of Gibrat's Law over alternative time series regimes. Given the long time dimension of this dataset this research adds further novelty to the current research base through the investigation of the empirical dynamics of long-surviving plants, a relatively under-researched theme, and in particular by testing the empirical propositions of Gibrat's Law for this particular sub-group of plants.

Most previous studies have also tended to consider either aggregate samples or disaggregated samples relating to a few selected industries. In many cases these samples have also related only to quoted companies, neglecting the dynamic properties of smaller firms. However, here the growth patterns of plants are examined at the 2-digit industry-level, covering the full range of manufacturing industries, for selected plant ownership types and structures and for different plant size-classes and age-classes in producing a thorough investigation of the validity of Gibrat's Law. The use of the WRME dataset provides a detailed record of plants with more than 10 employees, supplemented with some additional records relating to micro-business units. Full use is made of such micro-business unit information: again an under-researched area given the highly incomplete nature of official business surveys for such plant sizes.

Additionally, this thesis further aims to add to the literature on plant (firm) size and growth rate distributions by conducting a detailed analysis using a range of statistical tests and non-parametric estimation methods, initially examining static size distributions, consistent with most previous studies, before considering plant size distribution dynamics, a line of research only recently beginning to be explored in detail, and for which only limited analysis of such issues has been published relating to U.K. firms and industries. The analysis also provides some new research on the roles of plant age and selection effects in the evolution of the distribution of plant sizes using both statistical methods and non-parametric kernel density estimation. In particular, the long time-span of the data enables a somewhat rare opportunity to examine the long-run evolutionary dynamics relating to such plant size distributions. Again, no studies have been identified which provide such a long time period over which to assess the evolution of firm or plant size distributions.

To complete the analysis, the distributions of plant size and growth rates are examined across selected manufacturing industries and for new entrant plant cohorts. Few studies have been observed providing such detail, so that these results may again provide new information on the process of convergence (or otherwise) to specific statistical size distributions.

The structure of this thesis is as follows. The first chapter provides an introduction to the stochastic model of Gibrat (1931) focussing on the statistical formulation of the model, and the resulting propositions and implications. Some comparisons are then drawn with the properties and implications arising from a number of other models of firm growth with deterministic elements or foundations. Consideration is also given to the literature reviewing the assumptions of the model and the effects of alternative plausible modifications to these assumptions on both the postulated firm growth processes and size distributions. The properties of some of the frequently employed alternative statistical distributions within the context of the literature on firm size and growth rate distributions are also briefly described.

The second chapter provides a detailed review of the literature examining both the hypothesised lognormal distribution of firm (or plant) sizes and the evidence relating to some competing skewed distributions considered as approximate descriptions of such firm size distributions, before discussing recent developments in relation to statistical distributions approximating to firm growth rates. The chapter then further considers the evidence emerging from the application of regression approaches to examine the three primary propositions of Gibrat's Law, focussing particularly on the proposition that firm (or plant) growth is independent of size, covering the empirical effects of alternative data sources, correction procedures and econometric estimators. Empirical evidence relating to the variance of firm growth and firm growth persistence is considered before drawing conclusions regarding the empirical regularities emerging from this previous research.

Chapter 3 provides an introduction to the micro-dataset used in this study. In particular, this chapter presents an overview of the structure of the database, the data collection procedures and of the available variables. The broad empirical properties of these variables and some of the trends in manufacturing plant and employment dynamics are then described. The final sections of this chapter provide a discussion on the selection of time periods over which Gibrat's Law will be examined and the properties of the constructed samples. A number of issues relating to the measurement of business size and the choice of the unit of analysis, firms or plants, are then examined.

A primary feature of Gibrat's Law is the prediction of a lognormal limiting distribution of firm sizes. Chapter 4 commences with a brief discussion of a number of quantitative techniques which are employed in providing an assessment of the appropriateness of the lognormal distribution for this particular dataset. The following sections then set out an analysis of aggregate static plant size distributions, making full use of the time dimension of the dataset, and also for a range of sub-samples relating to industries and plant age-classes. Consistent with recent literature, the effects of plant ageing and selection effects on the evolution of size distributions are then examined across a number of time periods. Similar analytical approaches are then employed to consider plant growth rate distributions. A review of the empirical findings and conclusions are presented in the final section of the chapter.

Chapter 5 sets out a detailed econometric and statistical analysis of the propositions of Gibrat's Law, focussing initially on testing for growth persistence using standard regression methods. Statistical methods are then employed to investigate plant growth rates and the variance of growth rates across constructed size classes. To examine the proposition that firm growth rates are independent of initial firm size, regression methods are used to estimate the basic unconditional model specification using traditional cross-section estimators with the analysis conducted over a range of time periods. To supplement the basic model, the effects of the inclusion of additional explanatory variables, such as plant age, the implications of persistence in plant growth and the effects of potential sample selection bias are considered, both at an aggregate-level and for a large range of disaggregated samples. Further specification issues including potential non-linearities in the growth model are also examined.

The final chapter, Chapter 6, sets out a broad review of the research conducted within this thesis, covering the implications of Gibrat's Law, and then presenting the primary conclusions that can be drawn regarding the consistency of the new results generated in this research with the empirical propositions of firm (or plant) growth postulated by Gibrat's Law, and additionally in respect of the examined plant size and growth rate distributions. The results from this investigation are compared with those from previous studies, assessing the extent of coherence, and identifying significant differences. The final sections conclude with a brief discussion of some further initial research that was undertaken within this study, but which could not be included in the final version, and some thoughts on where future research might be usefully conducted.

1

STOCHASTIC PROCESSES AND FIRM DYNAMICS: STATISTICAL THEORY AND MODELS

1.1. Introduction

This chapter provides an introduction to the theoretical foundations of Gibrat's Law. Section 1.2 presents the statistical theory and formal mathematical representation of the model setting out the empirically testable propositions relating both to the growth patterns and size distributions of firms. Section 1.3 considers the appropriateness of the assumptions of Gibrat's Law, before reviewing the implications of some economically plausible modifications to these assumptions on the postulated patterns of firm growth and the associated size and growth rate distributions. Section 1.4 provides a summary of the discussion and some conclusions.

1.2. Theoretical Foundations of Gibrat's Law

The stochastic model of Robert Gibrat known as Gibrat's Law of Proportionate Effect (henceforth Gibrat's Law) has provided a notable theme within the industrial economics literature both as a direct proposition relating to the hypothesised empirical pattern of firm growth and with regards to the size distribution of firms (for brevity the term "firm" is used here to refer to both firms and plants). Gibrat's Law has also been an important benchmark in the assessment and development of more recent economic models (see, for example, Lucas, 1978; Prescott and Visscher, 1980; and Klette and Griliches, 1999).

In the model, Gibrat (1931) assumed that the logarithm of firm sizes would follow a modified Gaussian process under which a large number of small, independent additive components generate a normally distributed variate, the "Law of Laplace" (Sutton, 1997, p.40) so that these components operated multiplicatively. By the Central Limit Theorem[1], for a fixed population of firms, the outcome would be a lognormal limiting distribution of firm sizes[2] (Clarke, 1985, p.34).

Gibrat's Law can be formally stated following, for example, Sutton (1998, pp.242-243). In this representation the size of firm i at time t is denoted by X_{it} with ε_{it} denoting a random variable representing proportionate growth between two periods t and t - 1 (where t - 1 refers more generally to a previous time period). These growth rates (or shocks) are assumed to be drawn in each time period from a random sample of growth opportunities, so that:

$$X_{it} - X_{it-1} = \varepsilon_{it}(X_{it-1}) \tag{1.1}$$

and:

$$(X_{it} - X_{it-1})/X_{it-1} = \varepsilon_{it} \tag{1.2}$$

This may be restated as:

---

[1] The Central Limit Theorem states that for a sample of n independent and random observations drawn from a distribution with mean μ and variance σ^2, as $n \rightarrow \infty$ the sample mean tends towards the normal distribution (Gudjarati, 1988, p.641; Upton and Cook, 2002, p.60).

[2] The probability density function (p.d.f) of the lognormal distribution $LN(\mu, \sigma^2)$ can be represented generally by:

$$f(x) = \frac{1}{x\sqrt{2\pi}\sigma} \exp\left[\frac{1}{2\sigma^2}(\ln x - \mu)^2\right]; x > 0,$$

a distribution with mean $E(x) = \exp[\mu + \sigma^2)/2]$ and variance $var(x) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$. The lognormal distribution has a mean above the median which is in turn larger than the mode (Kleiber and Kotz, 2003, p.113).

$$X_{it} = (1 + \varepsilon_{it})X_{it-1} = X_{i0}(1 + \varepsilon_{i1})(1 + \varepsilon_{i2})\ldots(1 + \varepsilon_{it}) \quad (1.3)$$

where over very short time periods the ε_{it} are approximated by $\ln(1 + \varepsilon_{it}) \approx \varepsilon_{it}$ (Sutton, 1997, p.242) so that:

$$\ln X_{it} \approx \ln X_{i0} + \varepsilon_{i1} + \varepsilon_{i2} + \ldots + \varepsilon_{it} \quad (1.4)$$

Gibrat (1931) assumed that the stochastic term ε_{it} was normally distributed, an assumption justified by the Central Limit Theorem (Davies et al., 1991, p.101), with mean m and variance σ^2, and with the error terms being both mutually independent and independent of initial firm size. Under these assumptions, and making use of the Lindeberg-Levy Central Limit Theorem[3] (Kleiber and Kotz, 2003, p.108-110), for a fixed population of firms, as $t \rightarrow \infty$, the process of multiplicative disturbances (Prais, 1976, p.279) gives rise to the distribution firm sizes corresponding asymptotically to a lognormal limiting distribution (Sutton, 1997, pp.40-41), a distribution skewed to the right (extended right tail), with large firms emerging from multiple periods of above average growth.

Such a skewed distribution with a higher frequency of smaller than larger firms, is consistent with firm size distributions observed across a range of reasonable measures of firm sizes, time periods, industries and countries[4] (Ijiri and Simon, 1977; Hart and

---

[3] As Kleiber and Kotz (2003, p.110) and Hart (1973, 1976) note, more general Central Limit Theorems may also be able to generate such a limiting distribution where the ε_{it} are, for example, heterogeneous or non-normally distributed, though in the latter case giving rise to some slight skewness and kurtosis of the logarithmically transformed variables (Hart, 1976, p.124). Hart and Oulton (1995, 1996) discuss the Weiner and Liapounoff Central Limit Theorems within the context of Gibrat's Law (which make different assumptions regarding the stochastic error term) both theoretically capable of generating a lognormal distribution under appropriate conditions.

[4] In fact, the generality of this stochastic process has given rise to the application and empirical testing of the lognormal distribution in a range of economic contexts including the distributions of incomes, wealth, trade union sizes and city sizes (Aitchison and Brown, 1957; Kleiber and Kotz, 2003).

Oulton, 1996), providing powerful support for the potential role of stochastic influences in shaping the development of firms and market structures (see Clarke, 1985, pp.34-35).

As Geroski (1999) and Kleiber and Kotz (2003, p.108) note, equation (1.1) can also be represented by the non-stationary (random walk) process:

$$\ln X_{it} = \ln X_{i0} + \sum_{t=1}^{n} \varepsilon_{it-n} \qquad (1.5)$$

Equation (1.5) (which can also be specified to include a non-zero drift term, μt) clearly illustrates an important implication of this model: that firm growth will follow an erratic, unpredictable pattern, with firm size at any point in time being the sum of the cumulative shocks to initial firm size, these shocks providing permanent effects on firm size (path dependency) (Geroski, 1999). These changes in firm size occur through fixed adjustments (without phased transitions), which adds a sense of the model implying no meaningful distinction between short-run and long-run dynamics (Geroski, 1999).

Such unpredictability in the patterns of firm growth is further emphasised by the independence of the error term across firms implying the absence of any correlating processes (arising perhaps through inter-dependencies, such as competition, between firms; see, for example, Bottazzi and Secchi, 2002). Consequently, little information can be derived regarding future firm growth based on previous growth performance (Geroski, 1999; Geroski et al. 2003). The time series of firm growth can therefore be described by a stochastic trend rather than a deterministic trend with supplementary noise (Geroski, 1999).

Under the above specified conditions, such a growth process can also be represented by:

$$\ln X_{it} = \alpha + \beta \ln X_{it-1} + \mu_{it} \qquad (1.6)$$

where μ_{it} satisfies the conditions that:

$$E(\mu_{it}|X_{it-1}, x > 0) = 0 \tag{1.7}$$

$$E(\mu_{it}^2 | X_{it-1}, x > 0) = \sigma^2 \tag{1.8}$$

(Cefis et al., 2001). Rearranging equation (1.6) implies that:

$$\ln X_{it} - \beta \ln X_{it-1} = \alpha + \mu_{it} \tag{1.9}$$

with the β coefficient providing the growth effect on the initial size of the firm. It is of note that lagged terms are not included in the adjustment process of firm sizes implying that shocks are not anticipated by firms (see Geroski, 1999, who provides support for this model formulation).

Under the condition that $\beta = 1$ firm growth follows a random walk[5] (a non-stationary process with unit root) with constant, possibly non-zero, drift (where α is the drift parameter), and with the firm size mobility effect given by the residual variance, σ_μ^2, (Hart, 1980) - larger estimated residual variances indicating increased mobility of firms between size-classes. Under such a parameter the model implies that there are no systematic advantages or disadvantages arising from firm size, at variance with the standard deterministic theory of optimal firm sizes, which permits the possibility of declining unit production costs for larger firms, at least to a minimum efficient scale (MES) of production (Clarke, 1985, p.35).

---

[5] Tests for random walk or mean reversion processes have been conducted across a range of economic issues in addition to Gibrat's Law including, for example, by Hall (1978) and Nelson and Plosser (1982) in relation to macroeconomic aggregates, Cootner (1964), Granger and Morgenstern (1970) and Poterba and Summers on stock market prices, Wu and Zhang (1996) for interest rates and Smith et al. (2003) on real exchange rates.

A further implication of the condition that $\beta = 1$ is that, subject to the existence of some random growth rate variation across firms, industrial concentration will increase (continuously) over time[6], *ceteris paribus* with more dispersed growth rates resulting in concentration rising more quickly (Prais, 1976). In fact, Prais (1976, p.279) notes that under the condition that $\beta = 1$, concentration will rise providing that the disturbances are cumulative, with such a process occurring whether these effects operate either multiplicatively or additively. It is of note that such a prediction of increasing concentration in the absence of constraining factors has led some authors (such as Prais, 1976; see Clarke, 1985) to suggest the need for policy interventions to alleviate such concentration implications.

Where the estimated coefficient $\beta > 1$, large firms, on average, exhibit growth rates above that of smaller firms. In the case of $\beta < 1$, termed mean reversion or regression (see Galton, 1892), smaller firms are subject, again on average, to proportionately faster growth than larger firms (Prais, 1976, p.279; Hart and Oulton, 1996, p.1245) with firm sizes following a stationary process so that the effects of shocks on firm sizes are transitory. Such regression implies a tendency for firm sizes to return to the mean size of the population (Prais, 1976, p.281; Hay and Morris, 1991, p.537).

An alternative firm growth model specification has also been employed in the literature, which may be represented as:

$$\Delta \ln X_{it} = \alpha + \beta_1 \ln X_{it-1} + \nu_{it} \qquad (1.10)$$

---

[6] The development of firms owned by shareholders rather than individuals may also support such continuous rises in concentration since there is no implied (fixed) firm survival duration relating to the retirement of the owner. Prais (1976, p.283) argues that such changes in firm ownership practices may have altered the equilibrium distribution of firms (see also, Marshall's (1890, 1919), analogy of "trees in the forest" described in Prais, 1976). See further Hannah and Kay (1977) for an investigation of the roles of Gibrat-type processes and mergers in observed industrial concentration patterns using simulation methods, finding more influence from the latter.

where $\Delta \ln X_{it} = \ln X_{it} - \ln X_{it-1}$ and $\beta_1 = \beta - 1$. The test of the first proportion therefore becomes for a β_1 coefficient not statistically significantly different from zero. Given that $\alpha > 0$ for all firms, there is convergence to a uniform long-run size given by $-\alpha/\beta_1$, with deviating firms converging at a rate given by the coefficient β_1 (Geroski, 1999; Geroski et al., 2003). Both model specifications imply a linear relationship between firm size and growth. However, extensions to this model through the inclusion of higher-order (usually quadratic) firm size terms have also been examined in the empirical literature.

In summary, Gibrat's Law therefore provides a number of empirically testable propositions. First, that the regression coefficient $\beta = 1$ in equation (1.6) (the data exhibiting a unit root) or that $\beta_1 = 0$ in equation (1.10) so that firm growth rates across size-classes are independent of initial firm size (or equi-proportional across firm sizes); second, that the variance of growth rates is independent of initial firm size; and third, that there is no serial correlation (or persistence) in firm growth rates across time periods. Following Tschoegl (1983) these propositions can be formally stated for the standard model of equation (1.6) as:

Proposition 1: $\beta = 1$ (1.11)

Proposition 2: $E(\mu_t^2) = \sigma_t^2$ (1.12)

Proposition 3: $\mathrm{cov}[\mu_{it}, \mu_{it-1}] = 0$ (1.13)

Although such stochastic models have been subject to some criticism for their lack of economic foundations (Sutton, 1998) (and the explanation of causal mechanisms, and hence direct policy implications, see Davies et al., 1991), whilst Gibrat's Law postulates that stochastic influences dominate the shaping of firm sizes it does not imply the absence of any role for more deterministic forces, a point recognised by, for example, Simon and Bonini (1958), Evans (1987a, 1987b) and Hall (1987). The model simply asserts that given the numerous and varied factors which may be considered as incorporating some element of probability or chance, such as R&D outcomes[7], recruitment, the effects of advertising (Scherer, 1970, p.148) natural phenomena and disasters, financial bubbles and political change (Hart and Oulton, 1996, p.1244), industrial strikes, movements in exchange rates and the outcomes of mergers (Clarke, 1985, p.33), that such stochastic factors dominate the deterministic influences giving rise to a process closely replicating that of a pure stochastic model (Hart and Oulton, 1996, pp.1244-1245).

It is also relatively straightforward to establish some degree of coherence between the first proposition of Gibrat's Law, that firm growth is independent of firm size, with aspects of the standard deterministic models[8]. For example, standard microeconomic theory frequently tends to reflect an industry-specific "U-shaped" long-run average cost curve, average costs declining at first through economies of scale before rising beyond the least-cost point of production due to diseconomies of scale. Within such a framework, firm growth can be interpreted as a transitionary process towards the "optimal" size determined by these economies of scale, and other market factors (Geroski, 1999). Such a model therefore implies a convergence of firm sizes through a gradual adjustment process (with persistence in firm growth rates) to a static, optimal size

---

[7] As Geroski (1999) observes, whilst many firms persistently commit expenditure to research and development activities, innovation outcomes, including patents and commercialised products, do not generally exhibit a smooth and consistent pattern, such irregularity being consistent with the random walk process implied by Gibrat's Law.

[8] Geroski (1999) examines a range of deterministic models of firm growth, assessing their consistency with the empirical evidence. Geroski (1999) criticises stage-theory models given their general assumption of firm growth being a consequence of long-run deterministic trends, rather than stochastic trends as supported by the empirical evidence. He also demonstrates that under appropriate assumptions, given that lagged firm size variable terms beyond one year generally are not significant, a model degenerating to Gibrat's Law can be derived from the theories postulated by Penrose (1980), additionally noting that the implied variable adjustment costs of the constructed Penrose model are inconsistent with both the empirical evidence and predictions of Gibrat's Law. Geroski (1999) further observes that in the case of the organisational capabilities models that a random walk model of firm growth can be determined, but based on random capabilities which are not supported by the existing evidence.

for all firms (within an industry), contradicting the predictions of Gibrat's Law[9].

However, frequently the empirical evidence is more supportive of "L-shaped" long-run average cost curves over reasonable output ranges (see, Clarke, 1985, p.30 and Scherer, 1980). Under such a constant cost structure beyond the (industry-specific) minimum efficient scale of production, it is plausible to envisage that growth rates would not be systematically related to size, even though firms may be of differing sizes above the respective minimum efficient scales. Indeed the models of Ijiri and Simon (1977) reflect a Gibrat-type process operating above a threshold size, which can be interpreted as consistent with such a minimum efficient scale hypothesis.

With either "L-shaped" or "U-shaped" long-run average cost curves though, average costs would still decline initially with increasing scale[10], a feature interpreted in some studies as likely to necessitate the more rapid growth of new, smaller (surviving) firms as they strive towards the minimum efficient scale of production, this feature giving rise to the possibility of non-linearities in the firm size-growth relationship. Where economies of scale are particularly significant firm start-up size and immediate post-entry growth are likely to be relatively higher to permit adjustment to the minimum efficient scale of production (with a disproportionately high number of larger firms). The survival of new entrant firms is also likely to be relatively lower in these conditions (Audretsch et al., 2002). Where economies of scale are relatively modest, consistency with Gibrat's Law across a broader range of firms may be observed. Sunk costs may also influence survival probabilities and growth rates, with exit rates expected to be lower in industries with higher sunk costs (see for example, Audretsch et al., 2002).

---

[9] Geroski (1999) illustrates that such "optimal size" models of firm growth can be re-stated to allow optimal size to drift unpredictably, but this gives rise to a model of Gibrat's Law. He further notes a re-interpretation in the form of the "island models" (see Ijiri and Simon, 1977; Sutton, 1998), where firms are subject to independent (finite) growth opportunities, where if firm size does not impact of the probability of obtaining these opportunities, a model corresponding to Gibrat's Law emerges.

[10] Clarke (1985, p.30) observes that if there is some degree of marker power, so that price is above long-run average costs, firms below the minimum efficient scale of production, and hence on the downward sloping portion of the standard long-run average cost curve, may still continue to operate despite being subject to cost disadvantages.

In considering such issues, it is worth noting a point that has frequently been neglected in studies of Gibrat's Law, that there may be reasons for suspecting that whilst plants may be subject to diseconomies of scale, there may be less reason for suspecting that such diseconomies may impinge of the growth of multi-plant firms (unless, for example, these are specifically related to multi-site co-ordination issues), so that there may be no observed regression in the latter case (Prais, 1976, p.283). Conversely, if plants are subject to such diseconomies, this may make analysis at the plant-level more likely to exhibit mean reversion patterns. Such features may well, however, be influenced by factors specific to firms or industries, an area meriting further investigation.

In reviewing the predictions of Gibrat's Law it is also of value to reflect upon a range of models of industrial dynamics that have recently been developed to incorporate industry-specific factors, such as economies of scale, capital intensity and sunk costs into firm growth processes with stochastic elements. The learning and evolutionary models, for example, particularly associated with Jovanovic (1982), Ericson and Pakes (1995) and Audretsch (1995), have utilised stochastic frameworks incorporating, for example, facets of competition and maximisation behaviour, with entry, learning and selection process. In these models, firms in each period make size adjustments and exit decisions, which by implication shape the distribution of firm sizes. Nelson and Winter (1982) have also developed a model of selection with random innovation search and imitation (Scherer, 1980, p.147) based on rationality requirements rather than stricter maximisation behaviour (Sutton, 1998, p.244). Such learning and selection models suggest greater variation in the performance of younger, less informed, firms, with the shake-out of less efficient and non-optimally size firms over time.

As Evans (1987b) reports, under certain specifications of the Jovanovic (1982) model, for mature firms, as firm age increases towards the limit, growth would be unrelated to size, consistent with Gibrat's Law. In contrast, the Nelson and Winter

(1982) model gives rise to the implication of an inverted "U-shaped" relationship between growth and size for firms over 20 years of age (Evans, 1987b)[11]. Given the observations from these models, a key theme in the empirical work is therefore to reflect age related issues in examining Gibrat's Law, a point returned to later in this research.

1.3. Modified Stochastic Processes and Statistical Firm Size Distributions

Gibrat's Law, applied to a fixed population of firms, postulates that the distribution of firm sizes approximates to a (continuous) lognormal distribution, a distribution skewed to the right[12], and for which the tail of the distribution provides a downward (concave) parabolic representation on a double-logarithmic scale (de Wit, 2005, p.16, p.28). However, a broad range of statistical distributions have been shown to also be capable of replicating some of the features of empirical firm size distributions, at least for some industries and over certain ranges of the size distribution of firms (see Kleiber and Kotz, 2003).

Such distributions, presented in the economics literature have frequently been generated from modifications to the basic Gibrat-type process, either addressing particular implications of Gibrat's Law, by weakening the primary assumptions or through the addition of greater complexity into the model. Such extensions have been primarily aimed at more fully reflecting empirical observations relating to, for example, firm entry[13] and exit processes (which in combination with changes in the sizes of

---

[11] Evans (1987b) further reflects that Lucas's (1967) model also predicts consistency with Gibrat's Law. Prescott and Visscher's (1980) adjustment costs model generates a prediction of equal firm growth rates independent of size but with declining growth rate variance with increased firm size.

[12] Although, as Kleiber and Kotz (2003, p.110) note, as the value of σ decreases the lognormal distribution LN(μ, σ^2) tends to approximate towards the normal distribution N(expμ, σ^2).

[13] Although Chapter 3 presents some broad empirical evidence on aggregate-level patterns of plant entry, no formal analysis of the factors influencing the process of entry or of the size of new entrant plants is conducted in this study. For further research on firm entry see, for example, Cabral (1997) and Geroski (1991, 1995). Dunne, Roberts and Samuelson (1988) and Baldwin and Gorecki (1991) provide examples of empirical evidence on the properties of firm entry and exit in the United States and Canada, respectively.

incumbent firms would contribute to the dynamic properties of observed size distributions), and mergers and acquisitions, which may provide particularly large and immediate changes in firm sizes and which would simultaneously have the effect of reducing the (fixed) firm population. Some of these modifications to Gibrat's Law and alternative, competing statistical distributions are now considered.

As noted in an early paper by Kalecki (1945) the unmodified Gibrat's Law implies a steadily increasing variance of firm sizes over time for a fixed population of firms. Demonstrating this point, following Hay and Morris (1990, pp.537-538), if the initial variance of firm sizes is given by $\mathrm{var}(\ln X_{it})$, after one period:

$$\ln X_{it} = \ln X_{it-1} + \mu_{it} \qquad (1.14)$$

Therefore:

$$\mathrm{var}(\ln X_{it}) = \mathrm{var}(\ln X_{it-1}) + \sigma^2 \qquad (1.15)$$

and after n periods,

$$\ln X_{it+n} = \ln X_{it-1} + \mu_{it} + \mu_{it+1} + \ldots + \mu_{it+n} \qquad (1.16)$$

so that

$$\mathrm{var}(\ln X_{it+n}) = \mathrm{var}(\ln X_{it}) + n\sigma^2 \qquad (1.17)$$

Prais (1976) referred to this feature of increasing variance as "spontaneous drift", a process implying continuously increasing industrial concentration, arising most significantly where $\beta > 1$, and potentially even where $\beta < 1$ where mean reversion provides an insufficiently strong counteracting effect (Hay and Morris, 1990, p.538). Additionally, under the conditions of either positive serial correlation or of a higher probability of exit for smaller firms relative to larger firms this increasing concentration would be further amplified. Such a process of concentration may however be moderated under the conditions of mean reversion and sufficient new firm entry[14].

This feature of an increasing variance of the firm size distribution with time implies that there is no steady-state for the predicted lognormal distribution, at least under the standard Gibrat assumptions (de Wit, 2005, p.5). However, Kalecki (1945) demonstrated that a steady-state lognormal distribution with a constant variance could be derived by replacing Gibrat's Law with the assumption of negatively correlated average firm size and growth rates (de Wit, 2005, p.6). Ijiri and Simon (1977) also confirmed that such an increasing variance could be stabilised through the introduction of firm entry and exit processes into a Gibrat-type stochastic model.

In fact, the lognormal distribution may still emerge under a range of modifications to the three propositions of Gibrat's Law. Taking first the weakening of the first assumption that firm growth rates independent of initial firm size, or equi-proportionate across size classes[15]. In an example of such an approach weakening the first proposition, Hay and Morris (1990, p.539) present a model of a non-random relationship between firm size and growth as:

---

[14] The extent of regression required to alleviate the implication of increasing concentration depends on the variation in firm growth rates (Prais, 1976, p.281). Hart and Prais (1956, p.172) show that whilst equation (1.17) emerges in the absence of regression, under such regression, with a correlation coefficient, ρ, given by $\rho^2 = 1 - \sigma^2/\mathrm{var}(\ln X_{t+1})$, the variance relationship can be represented by $\mathrm{var}(X_{t+1})/\mathrm{var}(\ln X_t) = \beta^2/\rho^2$, so that industrial concentration declines where $\beta < \rho$.

[15] The condition of stochastically distributed growth opportunities giving rise to equi-proportionate, or independent, growth would be violated if firm size is related to the probability of firm survival. A positive relationship between firm size and survival would suggest that small firms with low growth rates would be more likely to exit than larger firms with similar growth rates, with the full (or at least more complete) distribution of growth rates only observed for larger firms (Sutton, 1997).

$$\ln X_{it} = (1 - \beta)\varphi + \beta \ln X_{it-1} + \mu_{it} \qquad (1.18)$$

where φ is a constant term. By induction:

$$\ln X_{it} = \varphi + \mu_{it} + \beta\mu_{it-1} + \beta^2\mu_{it-2} + \ldots + \beta^t\mu_{i0} \qquad (1.19)$$

so that firm size is determined by the sum of the error terms, μ_{it}. However, despite this weakened assumption, in this specification the lognormal distribution still emerges as the limiting distribution. Hart (1975, see also Hart, 1976) and Hart and Oulton (1996, p.1246) also note that under either (modest) regression-to-the-mean or regression-from-the-mean, the limiting distribution may closely approximate to the lognormal.

Arguments have also been advanced for weakening the second proposition of Gibrat's Law, that the variance of firm growth rates is independent of initial firm size[16]. For example, it may plausibly be argued that larger multi-product firms may have an opportunity to alleviate some of the variability in output through the acquisition of a portfolio of products which exhibit some degree of negative correlation in their performance[17]. Such an opportunity for risk diversification would not be available to smaller single-product firms (Prais, 1976) so that their growth outcomes would be more reliant on a single product outcome. A similar argument regarding the greater opportunities for diversifying risk for larger firms can also be made in terms of the outcomes of research and development activities. The implication is that whilst smaller

---

[16] Although a rejection of a firm growth rate variance being independent of firm size would invalidate this second proposition of Gibrat's Law, this need not necessarily invalidate the implication of the lognormal distribution as the (approximate) limiting distribution. For example, as Hart and Oulton (1995; 1996, p.1245) observe, providing that β (in equation 1.6) is not less than unity, the Liapounoff Central Limit Theorem could still generate a lognormal distribution (although the Weiner and Lindeberg-Levy Central Limit Theorems would no longer be appropriate given their assumption of a constant variance).

[17] Recent research, including by Stanley et al. (1996) and Sutton (2001), has attempted to generate models and empirically examine the existence of independence or correlations between business units within firms and the effects on firm growth rate variances.

firms may occasionally experience very high or low growth (and rapid size changes), larger firms would tend to exhibit more stable sizes and growth patterns.

Further, as Dunne and Hughes (1994) note, firm age and particularly learning experience may also provide a role in the observed patterns of firm growth rate variances, perhaps declining on average as firms obtain more information about their markets and efficient production processes (see Jovanovic, 1982; Ericson and Pakes, 1995; and Audretsch, 1999). Consequently, experienced firms might exhibit lower growth rate variances than new entrant firms of a similar size, although differences may emerge here between single-plant firms and those of multi-plant structures where younger plants within the latter may be able to draw on the experiences from other units within the firm.

Weakened assumptions relating to the third proposition that there is no persistence in firm growth rates have further been incorporated into models of Gibrat's Law. Such persistence effects can be easily theorised, arising perhaps from the introduction of successful innovations or business practices (see, for example, the evolutionary models of Nelson and Winter, 1982) giving firms several periods of above average growth. Similarly, serially correlated growth rates are also an implication of the life-cycle models of firm growth. Such a modification to this third assumption can be formulated relatively simply by assuming a linear persistence relationship of the form:

$$\ln X_{it} = \alpha + \ln X_{it-1} + \delta(\ln X_{it-1} - \ln X_{it-2}) + \mu_{it} \qquad (1.20)$$

where the parameter δ provides a measure of growth persistence between periods. However, under this weakening of the third proposition of Gibrat's Law, Hay and Morris (1990) demonstrate by induction that the following relationship can still be derived:

$$\ln X_{it} = \mu_{it} + (1 + \delta)\mu_{it-1} + (1 + \delta + \delta^2)\mu_{it-2} + (1 + \delta + \delta^2 + \ldots + \delta^t)\mu_{i0} \qquad (1.21)$$

which provides a lognormal distribution under the condition that the μ_{it} are independent random variables following a normal distribution (Hay and Morris, 1990, p.539).

Although, as discussed and illustrated above, particular modifications to the assumptions of Gibrat's Law may still give rise to the lognormal distribution as the limiting distribution of firm (or plant) sizes, slight variations to the nature of the modifications and assumptions can also produce a number of other potential steady-state (or non-steady-state) distributions, which have also been considered as providing a reasonable approximation to at least part of observed firm size distributions (see, for example, de Wit, 2005, for an overview of a number of statistical distributions that may be generated through modified Gibrat-type processes).

In practice, such modifications have often been to assumptions regarding whether the population of firms is defined to be fixed or allowed to vary through firm entry and exit, assumptions permitting the decline of incumbent firms, the inclusion of a minimum firm size condition (or relative size condition), and whether the nature of the assumed growth process (being either consistent with Gibrat's Law where growth is proportional to firm size, or with growth rates being non-proportional to firm size). The mean reversion and minimum firm size assumptions have frequently been employed to provide stability conditions bounding the operation of the random walk process (de Wit, 2005, p.5). The nature of the units examined, and specific modelling assumptions may also give rise to implied continuous or discrete distributions (Ijiri and Simon, 1977, p.142; de Wit, 2005). Some of these distributions are briefly considered to provide context for the review of literature conducted in Chapter 2.

Within the literature on firm size distributions, of these alternative distributions, power law (or scaling law) distributions such as Pareto-type distributions (Pareto, 1895, 1896, 1897; see also Zipf, 1949), both continuous and discrete, have been the most frequently examined within the literature on firm size distributions. Such distributions are, however, generally reasonable only above a minimum size value (corresponding to the minimum value to which the distribution applies rather than to either the smallest observed value or an economically determined size, such as the as MES) and therefore of primary relevance as an approximation to the upper tails of the size distribution of firms[18] (Hart, 1975, p.428; Hannah and Kay, 1977, p.99), diverging as the value of the examined variate declines below this minimum threshold, as $X \rightarrow 0$.

In his original work Pareto (1895, 1896; see also Kleiber and Kotz, 2003; and Gabaix, 2006) found some evidence of a negative (approximately) linear relationship between the logarithm of the number of income earners with incomes above a minimum value, N_x, and the logarithm of income, X, so that:

$$\ln N_x = A - \alpha \ln X \qquad \text{for } A, \alpha > 0 \qquad (1.22)$$

where A is a coefficient reflecting the size of the largest income (providing the vertical intercept (Ijiri and Simon, 1977, p.185). When normalised for the total number of income earners the relationship:

$$N_x/N = 1 - F(X) = (X/X_0)^{-\alpha} \quad X \geq X_0 > 0 \qquad (1.23)$$

can be formulated (Kleiber and Kotz, 2003, p.61), where X_0 is a scale term reflecting some minimum income size for which the linear representation is valid (Vining Jr., 1974, p.1277) and α is a shape parameter providing an indication of the mass of the distribution in the right tail (Kleiber and Kotz, 2003, pp.59-60). Such a Pareto Law is generally referred to as the Pareto Type I (or Classical) distribution. The Pareto Type II

---

[18] Although as Dosi (2005, p.5) notes this is the area of the distribution where statistical data has tended to be more frequently available.

distribution is asymptotically equivalent to the Type I distribution[19]. Both have been examined in respect of the firm size distribution literature, and are discussed here under the general terminology of the Pareto distribution given their asymptotic equivalence[20].

The Pareto Law (equation 1.23) can easily be translated into a relationship for firm sizes (see, for example, Steindl, 1965; and Ijiri and Simon, 1977). Ijiri and Simon (1977, p.185) present a Pareto Law of the form:

$$Xr^{\beta} = M \tag{1.24}$$

both M and β denoting constant terms, X being firm size and r the rank order of the firm by size within the population of firms (Ijiri and Simon, 1977, p.185). In logarithms such a relationship can be represented by:

$$\ln X = \ln M - \beta \ln r \tag{1.25}$$

where β is the parameter of primary interest, providing the coefficient of the (approximately) linear relationship[21]. Higher estimated values, of the slope coefficient β implies increased relative sizes of large versus small firms (Ijiri and Simon, 1977, p.196),

---

[19] The Pareto Type II distribution is represented by a cumulative distribution function (c.d.f.) given by $F(X) = 1 - [1 + (X/X_0)]^{-\alpha}$, with these variables as denoted in the main text (Kleiber and Kotz, 2003, p.60).

[20] A third variation, known as the Pareto Type III, given by: $F(X) = 1 - [Ce^{-\beta X}/(X - \mu)^{\alpha}]$, where $X \geq \mu$, $\alpha > 0$, $\mu \in \Re$, and C is a function of these parameters (Kleiber and Kotz, 2003, p.60) has received little empirical investigation in the context of firm size distributions, an exception being the recent study of Bottazzi (2007a) who, examining a sample of large firms, found support for the Pareto Type III distribution (which is convex in the double-logarithmic plot of the distribution tail) being a better fit of the data than the Pareto Type I distribution.

[21] However, under a process of mean reversion this linear Pareto Law relationship is no longer exhibited with the Pareto distribution understating the mass of the actual firm size distribution in the middle sections (Vining, 1976, p.369). Similarly, deviations from the linear Pareto relationship may also potentially arise as a result of firms entering the population at sizes other than the minimum size or through higher relative growth rate variances in small firms (de Wit, 2005, p.17). Ijiri and Simon (1974, p.318) modelled such curvature by adding a non-linear term $c(\ln r)^2$ to equation (1.25) with convex curvature in the double-logarithmic plots observed where $c > 0$ and concave curvature where $c < 0$.

higher firm entry rates into the smaller size-classes or an increased minimum firm size, with less steep slopes being observed in the case of smaller relative minimum firm sizes or either higher growth rates or growth rate variances for smaller firms[22] (de Wit, 2005, p.17).

Such steady-state Pareto distributions can, for example, following Simon (1955) be generated through the addition of assumptions relating to both a minimum firm size threshold (defined as $X \geq X_0$, where X_0 is some minimum size) and a constant rate of entry for new small firms[23] (but with no exit so that the population of firms is increasing) to the assumption of Gibrat's Law of the likelihood of obtaining growth opportunities being proportional to firm size (de Wit, 2005, p.6) (see also Levy and Solomon (1996) who appended a minimum firm size condition to Gibrat's Law to produce a steady-state continuous Pareto distribution[24]). As de Wit (2005, p.16) notes, with relatively modest minimum sizes the Pareto parameter will be close to unity[25], with larger minimum sizes giving rise to less mass of the distribution in the tails.

Further, continuous Pareto distributions can be generated by modifying the first proposition of Gibrat's Law (but retaining an assumption that current firm size is the outcome of previous sizes plus a random influence) so that the likelihood of firm growth giving rise to a change from size-class i to size-class j is a function of the difference in the positions of the size-classes (Hannah and Kay, 1977, p.99; Kleiber and Kotz, 2003,

---

[22] In fact, Ijiri and Simon (1977, p.185) formally describe this relationship so that given the ratio of firm sizes X_1/X_2 equals the reciprocal of the ratio of the ranks r_1, r_2, to the power β, a doubling of the rank of r to 2r gives rise to a prediction of a firm size given by $X/2^{\beta}$.

[23] See also Adelman (1958) who examined the size distribution of U.S. iron and steel firms between 1929-1939 and 1945-1956 using a Markov probabilistic model incorporating entry and exit processes to determine the equilibrium firm size distributions.

[24] As discussed in de Wit (2005, p.12), in the absence of such a minimum firm size condition the size distribution would approach the lognormal with, under increasingly large values of σ^2, the distribution also approximating to a Pareto distribution with parameter zero. de Wit (2005, p.17) further recognises, a stretched lognormal distribution can have a similar appearance to the Pareto distribution, with double-logarithmic plots for parts of the lognormal p.d.f. potentially appearing close to a straight line, so that visual inspection may not provide conclusive evidence on any superiority of either statistical distribution.

[25] The parameters of this model have generally been estimated using OLS, although Gabaix and Ibragimov (2006) suggest such a procedure may provide biased estimates in small samples, and require some assumptions regarding the appropriate minimum size of firm to which such a Law applies.

pp.65-66; see also Champerknowne, 1953, in respect of income distributions), again under the assumption of a minimum firm size, X_0. The application of the Kesten process, a variant of Gibrat's Law which provides a lower bound to firm sizes given by $X_{it+1} = \max[X_0, \gamma_t X_{it}]$, where γ is a random growth rate, also generally gives rise to a Pareto distribution (Axtell, 2001, p.1819).

Returning to the representation given by equation (1.23), the Zipf distribution (also known as Zipf's Law), for which the c.d.f. is plotted with the firm size on the horizontal axis and rank/frequency on the vertical axis (opposite to that of the Pareto plots), according to Axtell (2001, p.1818) and Dosi (2005, pp.4-5) emerges as a special case of the (discrete) Pareto distribution when the parameter value of $\alpha = 1$. Under such a parameter value, and more generally, the empirical law of Zipf is frequently stated as the frequency of a given firm size being of inverse proportion to the ranking of the firm, so that the most frequent size is observed twice as often as the next most frequent, and the nth most 1/n times as frequent as the first. As Gabaix (2006, p.3) considers, whether the Pareto or Zipf distributions emerges depends on the extent of frictions (factors that restrict firms from declining in size) operating along with the random growth process: the Pareto being observed where frictions are more significant, the Zipf with less significant frictions (Gabaix, 2006, p.3).

A further distribution related to that of the Pareto and Zipf is the Yule (or Yule-Simon)[26] distribution, examined by Ijiri and Simon (1964, 1977), a distribution which can be derived from a stochastic model of firm growth with a modified Gibrat-type process (Gibrat's Law applying for firms above a minimum size, often considered to be the MES)

---

[26] Formally, the Yule distribution can be represented by $f(i) = \rho B(x, \rho + 1)$, where $f(i)$ is the probability density of firms above size $i = 1$, ρ is a parameter of the distribution, and B is a Beta function (Ijiri and Simon, 1977, p.143). They also note that as $i \rightarrow \infty$, $f(i) \rightarrow \rho\Gamma(\rho + 1)\, i^{-(\rho + 1)}$ providing a Pareto distribution (in fact coinciding when $\rho = 1$) Γ being a Gamma function (Ijiri and Simon, 1977, p.143). Ijiri and Simon (1977, p.157) also represent the Yule distribution in the form $F(i) = \Gamma(\rho + 1)i^{-\rho}$ where F(i) denotes the number of firms at or above some minimum size, ρ reflects the new firm entry rate and Γ is the Gamma function. In logarithms $\ln F(i) = -\rho \ln i + c$, where c is a constant term (Ijiri and Simon, 1977, p.75).

incorporating an assumption of a constant rate of new firm entry into the smallest size-classes. The Yule distribution approximates to the Pareto distribution (and for which the tails asymptotically correspond to Zipf's Law) in the upper tails relating to larger firms (Ijiri and Simon, 1977, p.143; Curry and George, 1983) and also exhibits a straight line (with an approximately unity parameter with relatively modest new firm entry) when plotted on a double-logarithmic scale (de Wit, 2005, p.28). As de Wit (2005, p.16) reflects whether a Pareto or Yule distribution emerges depends on the specific form of the assumed firm entry process[27] (see Ijiri and Simon, 1977 and de Wit, 2005, for a review of the detailed models).

de Wit (2005) also presents further distributions generated from modified Gibrat-type processes, discussing, for example, the Waring, a generalisation of Waring, particular generalised hypergeometric, negative binomial, logarithmic, generalisation of logarithmic, geometric, Poisson, and extended Katz distributions. These are not discussed in any detail here (but are considered as potential further research, discussed in Chapter 6).

Finally, given the discussion in the following chapter on the emerging evidence in respect of firm growth rate distributions, Gibrat's Law with an assumed Gaussian (normal) distribution, initial information on some potential alternative distributions (the Subbotin and Laplace distributions) is also considered here. The Laplace distribution (also known as the double exponential distribution given its appearance as two exponential distributions joined by a location parameter) has the same mean as the normal distribution but exhibits greater mass in the distribution tails[28] so that firms experiencing very low or very high growth rates are more frequently observed

---

[27] As Caves (1997) notes, firm entry and exit can be integrated through various non-random mechanisms providing different highly skewed limiting distributions.

[28] The Laplace distribution has a p.d.f given by: $f(g) = 1/\sqrt{(2\sigma_g)}\ \exp[-2|g - \bar{g}|/\sqrt{(2\sigma)}]$, where $\bar{g}$ is the logarithm of the average firm growth rate (Axtell, 1999, p.44). The greater mass in the tails of the Laplace distribution compared to the normal distribution is a consequence of the Laplace p.d.f being represented by the absolute difference from the mean, with the normal distribution p.d.f. represented by the squared difference from the mean (Wikipedia, March, 2007).

(Teitelbaum and Axtell, 2005, p.2)[29].

In the case of employment data, these more frequent extreme firm growth rates would be exhibited as considerable "chunks" of positive (or negative) employment creation (destruction) within the examined time period, with such an observation perhaps reflecting the discrete nature of many manufacturing investments (for example, in terms of establishing a new production line or plant), and as a consequence influencing the (fixed) structuring of associated employment changes (Bottazzi et al., 2006). As Bottazzi and Secchi (2003) note, such heavy-tails tend to be supportive of some positive correlating structure within the dynamics of firm growth, with such structures not conforming to the independent stochastic process of Gibrat (1931).

When graphically represented on logarithmic scales the p.d.f. of the Laplace distribution of frequency against the logarithm of firm growth rates exhibits a "tent-shaped" distribution with straight "sides" (Teitelbaum and Axtell, 2005, p.2). The Laplace distribution is also observed under a specific shape parameter value of the more general Subbotin (1923) distribution[30], the shape parameter reflecting the properties of the tails of the distribution, thicker tails emerging for lower shape parameter values.

Given these various possible size and growth rate distributions, a reasonable question might be the extent to which the properties of these different distributions

---

[29] Teitelbaum and Axtell (2005, pp.2-3) tabulate the probabilities of extreme events for the normal and Laplace distributions, suggesting in their analysis that 2σ and more extreme events occur around 3 times more frequently in the Laplace distribution, 300 times more frequently for 4σ and greater, and one million times more frequently for 6σ and above, than for the normal distribution. For their large sample of U.S. firms, Teitelbaum and Axtell (2005, p.3) suggest that these differences are of considerable empirical importance. As Reichstein and Jensen (2003, p.14) note, the observation of distributions exhibiting leptokurtosis would tend to suggest a higher likelihood of appropriate representation by the Laplace distribution rather than the normal distribution.

[30] The Subbotin distribution is given by:

$f(X) = 1/(2ab^{1/b}\Gamma(1/b + 1))e^{-1/b|X - \mu/a|^{b}}$ where $\Gamma(X)$ is a Gamma function, a is a scale parameter, b is a shape parameter and μ is a positioning parameter (Bottazzi and Secchi, 2002, p.4). In the case of $b = 1$ the Subbotin distribution reduces to the Laplace distribution, the Gaussian (normal) distribution arising where $b = 2$, and large values of b providing the uniform distribution (Teitelbaum and Axtell, 2005, p.4). Information on the Subbotin distribution and the estimation of the distribution parameters using the SUBBOTOOLS software is provided by Bottazzi (2004).

matters. Although no attempt here is made to address this issue empirically, by comparing, for example, the relative performance in terms of goodness-of-fit of each of these distributions (see the discussion on potential future work in Chapter 6), two general points appear relevant. First, there may be preferences for the choice of underlying mechanisms within the model specifications (for example, the inclusion of relevant entry and exit processes replicating actual empirical observations[31]). Secondly, these reviewed distributions have particular characteristics in terms of their statistical properties, perhaps most relevantly the shape of the distribution tails[32]. This may be significant, if the primary aim of any work is simply to approximate a distribution to actual data behaviour.

However, the objective of this section has been to emphasise that whilst the lognormal distribution emerges as the limiting distribution of firm sizes under specific conditions (and variations to these conditions), there are alternative potential statistical distributions that may be capable of replicating the observed empirical properties of the distributions of firms or plants, which may be generated from modified Gibrat-type processes. There is no reason to suspect that, given particular industry characteristics that one distribution should in all cases be preferred. The aim of this research, is to focus only of the lognormal distribution derived from Gibrat's Law, partly given the foundation that this particular stochastic processes has provided and also in part due to the large volume of literature on this particular distribution from which to compare the properties of the dataset utilised in this research. Such an analysis of the lognormal plant size distribution is presented in Chapter 4.

---

[31] In fact, as Bottazzi (2007b, p.2) considers, in the absence of such entry and exit processes, Gibrat's Law cannot give rise to a stationary Pareto distribution. Difficulties in reconciling Pareto firm size distributions with the Gibrat's process and a Laplace distribution of firm growth rates have been discussed by, for example, Fujiwara et al. (2003) and Bottazzi (2007b).

[32] In his review of these distributions de Wit (2005, p.28) notes that the Pareto, Yule, Waring and particular generalised hypergeometric all exhibit tails of the distribution which provide a straight line of slope $-(1 + \rho)$ on a double-logarithmic scale, where ρ is the parameter of the distribution, with the negative binomial, logarithmic, generalisation of logarithmic and geometric distributions all providing tails which are exponentially decreasing on the same scale, the Poisson decreasing more rapidly than exponentially. The tails of the extended Katz and generalisation of Waring distributions depend on the parameters and growth specifications, respectively.

1.4. Conclusions

This chapter has presented the theoretical foundations of a particular stochastic model of firm growth known as Gibrat's Law, setting out the assumptions underlying the model. Gibrat's Law postulates that firm growth follows a random walk (so that firm growth is independent of initial size), with the variance of growth also independent of firm size and with no persistence in growth. The outcome of this process for a fixed population of firms is a lognormal limiting distribution, a right-skewed statistical distribution consistent with empirically observed firm size distributions. Under such a random walk, shocks to firm sizes have permanent effects, with the size of firms at any point in time being the cumulative sum of these shocks to the initial firm size.

The discussion has also considered the theoretical plausibility of the assumptions of Gibrat's Law, noting particularly the implied process of increasing industrial concentration under unmodified assumptions. The consistency of the postulated propositions with other models of firm growth has been assessed, examining in particular the effects of potential economies of scale, relationships between firm size and the probability of obtaining growth opportunities and the potential moderation of growth rate variance through negatively correlated product portfolio effects and firm ageing and learning processes.

A number of plausible modifications to incorporate firm, industry and market characteristics have been appended to Gibrat's Law such as firm entry and exit processes, the inclusion of mergers and acquisitions and the incorporation of serial correlation in firm growth. Although under specific conditions the lognormal distribution may still emerge as the limiting distribution, such modifications can give rise to a wide range of highly skewed distributions consistent with observed distributions of firm sizes (at least over some ranges of the distributions) including the Pareto, Zipf and Yule distributions.

Additionally, whilst Gibrat's Law assumes a random sampling of growth opportunities from a normal distribution, recent research has also examined alternative growth rate distributions, such as the Laplace and Subbotin distributions. The properties of some of these statistical distributions have been briefly described as a prelude to the survey of empirical evidence presented in the next chapter.

2

GIBRAT'S LAW: A SURVEY OF EMPIRICAL EVIDENCE

2.1. Introduction

This chapter provides a detailed survey of previous empirical studies testing the validity of Gibrat's Law and the lognormal distribution as a reasonable approximation to the dynamics of firms (and plants). Sections 2.2 and 2.3 focus on the methodologies employed in investigating the appropriateness of the lognormal and alternative skewed statistical distributions as suitable representations of firm sizes, in both unconditional and conditional forms, and the resultant empirical evidence. The emerging evidence regarding the effects of ageing and selection processes on the evolution of observed firm size distributions is also considered. Some comments on the use of statistical tests of extreme hypotheses are presented in these sections.

Section 2.4 reviews the evidence relating to the first proposition of Gibrat's Law from cross-section regression approaches, with emerging results from panel data studies discussed in Section 2.5. In examining this previous research particular attention is given to the developing econometric methodologies, statistical correction procedures, especially sample selection bias (Section 2.6), and econometric estimators. Evidence on the effects of the inclusion of additional explanatory variables into the standard Gibrat model is also provided. Sections 2.7 and 2.8 examine the evidence on the consistency of the empirical properties of the variance of growth and persistence in firm growth rates, respectively, with that postulated by Gibrat's Law. Conclusions are drawn in Section 2.9.

2.2. The Lognormal Firm Size Distribution

Gibrat (1931) examined the size distribution of manufacturing plants using graphical methods plotting the logarithm of plant size-class (or more specifically $\ln(X_c - 1)$, where X_c is the relevant size-class) on the horizontal axis versus the number of plants of size z or greater estimated as:

$$R(z) = (1/\sqrt{\pi}) \int_{z}^{\infty} e^{-z^2} \, dz \qquad (2.1)$$

(Gibrat, 1931, p.53 and p.256; Sutton, 1997, p.41), the graphical plot observed as a straight line if the distribution of plant sizes is approximately lognormal (Sutton, 1997, p.41).

Using an employment measure of plant size and data covering the period 1886-1921, Gibrat found support for the lognormal distribution, with analysis of selected regional, industrial and incomes data also broadly supporting the lognormal hypothesis (Sutton, 1997, p.41) (see also Kleiber and Kotz, 2003, p.9, for further economic data and geographies examined by Gibrat). Despite raising some reservations about the theoretical assumptions and implications of the Gibrat model, Kalecki (1945) reached similar conclusions using employment data on U.S. manufacturing plants for the year 1937.

In an influential study, Hart and Prais (1956) provided a statistical analysis of the distribution of the sizes of U.K. quoted companies in the mining, manufacturing and distribution industries for the period 1885-1950 using market valuations and net assets measures of company size. A lognormal distribution was fitted to the frequency distribution of firm sizes for each of the years 1885, 1896, 1907, 1939 and 1950, finding some initial support for the lognormal distribution.

Further support was generated by using Fisher's (1941) g_1 and g_2 tests of symmetry and kurtosis (these statistics and their relation to the skewness and kurtosis measures used in this research are discussed in D'Agostino et al., 1990), albeit with some statistically significant deviations from the strict lognormal detected for all examined years other than 1885 and 1896, with some degree of positive skewness and leptokurtosis and an excess of firms in the right tail of the distribution (Gorman, 1956). Further corroboration was provided by Hart and Prais' (1956) analysis of firm growth patterns using transition matrices which showed that the distribution of firm growth rates across constructed firm size-classes exhibited generally similar frequency distributions approximating to the normal distribution.

Quandt (1966) employed non-parametric methods of estimating the fit of competing theoretical statistical distributions and statistical tests of the goodness-of-fit for selected statistical distributions against the actual size distributions of firms sampled across 32 U.S. industries at the 4-digit level classification. The data covered the period 1955-1960 with firm size measured by asset values. The analysis indicated that the three examined Pareto distributions tended to perform less well than the two-parameter lognormal distribution, with the latter providing a reasonable approximation in about two-thirds of the samples, a similar frequency of acceptable statistical goodness-of-fit as the double exponential and composite distributions (see, Quandt, 1966, p.421, for a discussion of these distributions). However, none of the distributions could be considered as providing an adequate description of all industries.

Silberman (1967) conducted an examination of the lognormal distribution for both firm and plant sizes in the U.S. manufacturing sector at the 4-digit industry-level using data for 1947 and 1958. A statistical method was employed, fitting and testing differences in expected and observed concentration ratios and focusing on the upper tail of the size distribution given that deviations are generally most noticeable in this area (Silberman, 1967, p.819).

At the firm-level, for the 1958 sample, the lognormal distribution was found to provide a reasonable description in 42 of the 90 industries using the value of shipments data as the size measure, with further support for 35 of the samples when examining the data for 1947. When both the value of shipments and value added data were used only 36 of the 90 samples were found to be adequately represented by the lognormal distribution. However, over seventy percent of industries retained their classification of being either lognormal or not lognormal in both years. Further analysis provided some support for firm-level size distributions generally exhibiting a greater likelihood of being appropriately represented by the lognormal distribution than the plant-level data.

In another early study, Steindl (1965) found supportive evidence for the lognormal distribution for plant-level manufacturing data from a number of different countries, albeit with some heterogeneity in the observed deviations from the lognormal across 2-digit level industries (Clarke, 1979, p.419).

In a U.K. study, Clarke (1979) tested the lognormal prediction of Gibrat's Law using employment data on the manufacturing and mining industries at both the firm and plant-levels. The analysis fitted normal curves using the method of areas together with tests of a compound null hypothesis (Kendall and Stuart, 1973) and the skewness and kurtosis tests of Geary and Pearson (Clarke, 1979, p.421)[33]. At the firm-level only 9 of the 133 tested distributions did not reject either the compound null hypothesis or the test of skewness and kurtosis. The results, however, tended to provide somewhat more support for lognormality of plant-level size distributions. In neither case though was the lognormal found to be a generally acceptable distribution. The detected deviations from the lognormal distribution were generally in the form of positive skewness and

---

[33] Details of these statistical tests are presented in the appendix of Clarke (1979, pp.426-427).

platykurtosis for both firms and plants, but with a more significant pattern of excess mass in the right tail of the distribution in the firm-level data.

Stanley et al. (1995) tested the lognormal distribution using Compustat data on a sample of 4,071 publicly traded manufacturing firms in 1993 by comparing a Zipf plot of the logarithm of the rank of a firm by its size versus the logarithm of sales (the measure of firm size) versus the expected Zipf plot from fitting a lognormal distribution[34]. Comparison of the predicted lognormal distribution to the actual firm sizes on a double-logarithmic scale supported the lognormal as a reasonable approximation for much of the observed size distribution. However, there was insufficient mass in the upper tail, with the largest firms reporting sales statistically significantly lower than would be expected under the lognormal distribution. Similar results were also observed in further analysis of data relating to the years 1975, 1979, 1980 and 1984.

Further evidence of statistical deviations from the lognormal distribution have also been reported by, for example, Hart and Oulton (1997) for a large sample of U.K. firms in 1993, finding too much mass in the upper tail compared to the lognormal (Kleiber and Kotz, 2003). Hart and Oulton (1996) also found evidence of deviations from the strict lognormal distribution for three measures of firm size (employment, sales and net assets) with evidence of the distributions exhibiting positive skewness (with longer right tails) and leptokurtosis, deviations from the lognormal which they suggest could be adequately represented by a Galton-Markov mean reversion growth process[35]. Whilst the development of more complex models such as the four-parameter lognormal[36] distribution used by Saving (1965) can be employed to address such observed deviations,

---

[34] More specifically, following Stanley et al. (1995) the methodology plots $\ln X_i$ versus $\ln r$, where X is the firm size and r is the rank of firm size, which given that $r/N = 1 - F(X_i)$, where F(X) is the c.d.f, provides $\ln r = \ln[1 - F(X_i)] + \ln N$. The Zipf plot of the lognormal distribution is provided by $\ln r = \ln[1 - \Phi\{(\ln X_i - \mu)/\sigma\}] + \ln N$, where μ and σ are the respective mean and standard deviation of $\ln X_i$ and Φ is the standard normal c.d.f. (Stanley et al., 1995, p.454).

[35] Such a first-order Galton-Markov relationship can be represented by $\ln X_{it} = \beta \ln X_{it-1} + \varepsilon_{it}$, where $\beta < 1$.

[36] The mathematical properties of the three-parameter and four-parameter lognormal distributions are discussed in Kleiber and Kotz (2003, pp.121-124).

in doing so some of the advantages of the simplicity of the simple two-parameter lognormal distribution are somewhat compromised.

A number of more recent studies have attempted to move beyond the analysis of static firm size distributions. In such a study, Cabral and Mata (2001) examined the evolution of the size distribution of Portuguese manufacturing firms in 1983 and 1991 using data from two sources: a sample of 587 large firms and an official employment survey providing a sample of over 33,000 firms. An initial statistical analysis of the logarithms of firm size, measured by employment, using a Jarque-Bera test (which examines weighted sample moments relating to excess kurtosis and skewness (Verbeek, 2000, p.174) to test for deviations from the hypothesised normal distribution (Kennedy, 1998, p.79) and which under the null hypothesis follows a χ^2 distribution with two degrees of freedom) on the sample of large firms did not reject the hypothesis of a lognormal distribution of firm sizes. Non-parametric kernel density estimation techniques provided further support for the lognormal distribution in the case of the sample of large firms. However, deviations in the form of a higher degree of skewness than would have been expected under the lognormal distribution were detected for the more comprehensive official dataset.

To consider the role of firm age in shaping the observed size distribution the cross-sections of firms were classified into six age-classes with kernel density estimates produced for each of these classes. The analysis identified a shift in the firm (log)size distribution with increased firm age, from being highly skewed to the left towards a more symmetrical distribution, giving rise to thicker right tails of the distribution, reduced mass in the left tails and an increased mean average firm size (Cabral and Mata, 2001, p.6). Cabral and Mata (2001) employed the extended generalised gamma distribution (which reduces to the lognormal distribution under an appropriate skewness parameter, see Cabral and Mata, 2001, pp.6-7, for the assumed distribution of firm sizes and specific parameter representations) to more formally quantify the change in the properties of the firm size distributions, finding that even for those firms that had existed for a time-span of over 30 years the size distribution could not be adequately described by the lognormal distribution for the majority of firms.

The study also tracked a cross-section of new entrant firms in 1984 through to 1991 to examine the roles of firm ageing and selection effects on the evolution of the firm size distribution. A comparison of the distributions of all firms in 1984 and the survivors in both 1991 and their respective sizes in 1984 suggested that the reduced skewness of the firm size distribution was a consequence of both selection and ageing effects, with ageing providing the primary influence.

Lotti and Santarelli (2001a) have also provided a recent examination of the evolution of the firm size distribution focussing on four Italian industries: electrical and electronic engineering, instruments, footwear and clothing, and food. The study used official quarterly data on twelve cohorts of Italian firms entering in each month of 1987, tracking employee numbers (the measure of firm size) until 1992. The evolution of the firm size distributions was tested against the lognormal distribution using a range of summary statistics, including skewness and kurtosis indices. In addition, formal statistical tests of the size distributions using the D'Agostino et al. (1990) tests of skewness and kurtosis, the omnibus D'Agostino and Pearson (1973) K^2 test, the Shapiro-Wilk W (1965) test, and the Kolmogorov-Smirnov test were also applied.

These statistical tests provided evidence of some degree of deviation of the observed firm sizes from the lognormal distribution across all four industries in the initial time periods following entry. However, for the electrical and electronic engineering and the instruments industries the size distributions were found to evolve towards the lognormal distribution over time, with such a tendency emerging more quickly in the case of the former. In the other two industries there was no conclusive support for an evolution towards the lognormal distribution over the examined time period, a finding suggesting the presence of industry-level heterogeneity, a feature also identified, for example, by Hymer and Pashigan (1962).

Extending this analysis, Lotti and Santarelli (2001b) again considered the evolution of the firm size distribution using the same dataset and methodology as in Lotti and Santarelli (2001a). In addition to the statistical tests discussed above, they made use of non-parametric kernel density estimation methods using a Gaussian kernel function (with some sensitivity analysis using the Epanechnikov kernel function suggesting that the generated results were not highly sensitive to the choice of kernel function) and an automatic band-width parameter (the properties of these kernel functions are discussed in Chapter 4). The analysis confirmed the results of their earlier study.

The statistical studies reviewed to this point have considered unconditional firm size distributions. However, in a recent econometric analysis Machado and Mata (2000) employed a flexible Box-Cox quantile regression approach to examine the conditional firm size distribution, the distribution that would have arisen under the homogeneous conditions implied by Gibrat's Law[37]. The approach models the conditional quantiles of the firm size distribution under a Box-Cox transformation (Box and Cox, 1964) as a linear function of a vector of explanatory covariates reflecting industry characteristics permitting the analysis of the marginal effects of the selected covariates on different points of the firm size distribution (Machado and Mata, 2000, p.255). Two samples of Portuguese manufacturing firms were examined, one relating to the year 1983 providing a sample of 18,552 firms, and a second similar sample of 26,515 firms for the year 1991.

The effects of a range of covariates on the properties of the conditional firm size distribution were examined for each year, with these covariates including industry-level employment growth, a measure of turbulence based on the employment associated with

---

[37] Mata and Machado (1996) also employed the quantile regression approach to analyse the conditional start-up sizes of Portuguese manufacturing firms. Jarque-Bera tests rejected lognormality for the unconditional firm size distribution.

firm entry and exit, measures of the minimum efficient scale of production and market size, the ratios of patents to production and of exports and imports to output, the proportion of employment in state owned firms and a firm age variable based on maximum recorded job tenure length.

Using an employment measure of firm size, an initial analysis of the unconditional firm size distributions employing the Jarque-Bera test rejected the lognormal distribution for both samples. However, the results of the quantile regression analysis provided some support for the lognormal in the case of the conditional distribution for 1983, but not for the 1991 sample. Even though the estimated coefficients were found to exhibit some degree of variability, the analysis tended to indicate that industry growth, economies of scale, exports and patent measures were associated with a movement of the firm size distribution to the right, with turbulence and import measures providing a shift in the distribution in the opposite direction. These industry-level factors were found to be generally more significant for larger firms (although some effects were detected in relation to the dynamics of smaller firms). The effect of firm age was identified as being inconsistent with the proposition that the firm size distribution would tend to lognormality with increased time duration.

2.3. Alternative Statistical Firm Size and Growth Rate Distributions

As discussed in Chapter 1, perhaps the most frequently empirically tested alternative to the lognormal distribution, in respect of firm and plant size distributions, is the Pareto distribution. Steindl (1965), employing graphical methods and analysing data on firms in both manufacturing and a number of services industries in West Germany during the 1950s, found some support for the Pareto distribution as an appropriate broad description of firm sizes in the upper tail of the size distribution, with α coefficients (as in equation 1.23) for a range of countries and industries, being of the order 1.0 to 1.5, indicative of a relatively unequal distribution of firm sizes (Kleiber and Kotz, 2003, p.62 and p.93).

Ijiri and Simon (1971) using a sample of large firms in the U.S. also provided broad support for the Pareto distribution, though with some significant deviations in the form of an observed downward curvature of the plotted distributions. The consequences of mergers and acquisitions on the firm size distribution were examined by developing a model with a two-stage process of reducing firm numbers through acquisition or merger and the associated reallocation of market shares. Their initial analysis indicated that merger and acquisition processes provided only a relatively modest effect on the acceptability of the Pareto distribution with both internal growth and growth through mergers and acquisitions conforming to Gibrat's Law.

To further examine the observed downward curvature of the actual firm size distribution, Ijiri and Simon (1974) examined the effects of both serially correlated firm growth and mergers and acquisitions. The mergers and acquisitions data were sourced from a U.S. Federal Trade Commission Report examining mergers and acquisitions in selected production industries during the period 1948-1969. The effects on the size distribution of the actual data post mergers and acquisitions were compared with an estimated expected distribution had the initial firms not merged or been acquired. Their analysis suggested that mergers and acquisitions did influence the observed downward curvature of the actual firm size distribution, a result contrasting with that found by Hannah and Kay (1977) who concluded that mergers tended to positively support growth rate distributions in following a Gibrat-type process.

Interestingly, Ijiri and Simon (1977, pp.211-212) also reported some evidence that other statistical distributions may have provided a more suitable fit of the observed data than the Pareto distribution. Supporting this position, it is of note that numerical simulations of the model developed by Ijiri and Simon (1964) based on a Gibrat-type process supplemented with an assumption of serially correlated growth rates provided support for a limiting firm size distribution approximating to the Yule distribution.

A study by Freeman (1986) has also provided an analysis of the firm size distribution, investigating the one hundred largest firms in the U.K. in the years 1970, 1975 and 1980 using data compiled by *The Times* publication. Using graphical methods some downward curvature of the plot of the actual firm size data compared to the linear Pareto distribution relationship was observed for the 1980 data, indicating that the mid-sections of the firm size distribution were not appropriately represented by the Pareto distribution. To explore this observation the possible role of serially correlated growth was examined using linear regression techniques. The results provided evidence of both autocorrelation and heteroscedasticity in the residual terms. A piecewise regression analysis did, however, provide some support for a Gibrat-type process but based on weakened assumptions.

Axtell (2001) using U.S. Census data covering the full range of firm sizes and OLS regression to estimate the relationship between the frequency of sizes and firm size (on a double-logarithmic scale) found evidence supporting the Pareto distribution with an exponent close to that of the Zipf distribution (Teitelbaum and Axtell, 2005, p.5). The specific empirical properties of the Census data also suggested that the lognormal could not provide an adequate description of some components of the distribution. Gabaix (2006) also adds support, suggesting that in many cases that Zipf's Law provides a good approximation for various measures of firm sizes (see, Axtell, 2001; and Fujiwara et al., 2004).

Turning briefly to the evidence on firm growth rate distributions, Stanley et al. (1996) using data for the period 1980-1998 drawn from the Compustat dataset found evidence to support the Laplace distribution as an appropriate statistical description of their data rather than the normal distribution. Their analysis also found support for a power law (or scaling) relationship for the variance of firm growth rates.

Bottazzi and Secchi (2002) also examined the distribution of firm growth rates using official data on Italian firms with 20 or more employees covering the period 1989-1996. Using data at the 3-digit industry-level, in order to alleviate potential concerns regarding the generation of apparent statistical regularities through aggregation, and pooled across the full time period, support was found for the Laplace distribution for a number of industries. Analysis of the shape parameters of the Subbotin distribution (which reduces to the Laplace and normal distributions under appropriate parameter values) estimated using maximum likelihood suggested that the Laplace distribution provided a reasonable representation for around four-fifths of the industries examined.

In a further study, Bottazzi and Secchi (2003a) again considered the properties of firm size and growth rate distributions using sales data as the measure of firm size (drawn from the Compustat database) for 15 U.S. manufacturing industries over the period 1982-2001. A balanced and an unbalanced panel of firms were constructed with the former employed in a parametric analysis and the latter using non-parametric techniques. At the aggregate-level evidence was found supporting a lognormal distribution of firm sizes with the Laplace distribution providing an appropriate representation of the distribution of firm growth rates. The variance of growth rates was observed to correspond to that of a power law. Significant heterogeneity in the size distributions was detected across 2-digit level industries, finding limited support for the general applicability of the lognormal distribution, with both observed skewness and cases of non-uni-modal distributions (Bottazzi and Secchi, 2003a, p.9). Growth rates at the industry-level were often found to reasonably correspond to that of the Laplace distribution. Growth rate distributions with "fat tails" also appeared in the study of the pharmaceutical industry conducted by Bottazzi et al. (2001).

Reichstein and Jensen (2003) have presented a statistical investigation of the distribution of firm growth rates over 1995-1996 for the iron metal, machine, pharmaceutical and textile industries in Denmark, finding cases of both positive and negative skewness, but in general a higher degree of kurtosis than implied by the lognormal distribution in aggregate and across industries. Graphical plots of the industry-level size distributions tended to indicate some potential bi-modality[38] in the iron metal, machine and textile industries with a multi-modal distribution exhibited in the case of the pharmaceutical industry. The observed concentration of firm growth rate observations around the mean and median firm sizes was taken to indicate some support for the Laplace distribution (Reichstein and Jensen, 2003, p.5). Kolmogorov-Smirnov tests for normality suggested that the growth rate distributions could not be adequately described by a normal distribution (in some cases with growth rates exhibiting greater mass in the distribution tails).

Teitelbaum and Axtell (2005) examined data on plants existing between 1998-1999 sourced from the U.S. Census Bureau's Statistics of U.S. Business. Using OLS regression methods support was found for the Pareto distribution for the size distribution, with an exponent near unity (suggestive of a Zipf distribution) with statistical fitting procedures and Kolmogorov-Smirnov statistical tests suggesting that the Laplace distribution provided an appropriate representation of the distribution of firm growth rates, the data exhibiting significantly more mass in the tails of the distribution than would be expected under the normal distribution. However, the Subbotin distribution (with shape parameter below unity and heavier tails than the Laplace distribution) was also found to provide an appropriate description, and in some cases (such as for the Finance industry) a superior fit to that of the Laplace distribution which

---

[38] A similar finding of bi-modality was also observed for some French manufacturing industries by Bottazzi, Coad, Jacoby and Secchi (2005) with non-parametric multi-modality tests rejecting uni-modality in 18 out of 20 2-digit industries examined.

under-represented some of the mass of the centre of the firm growth rate distribution. Use of the Kolmogorov-Smirnov test of goodness-of-fit suggested greater support for the Laplace distribution than the normal distribution across 2-digit level industries[39].

In concluding this review of the identified literature on firm (or plant) size and growth rate distributions, however, it is worth making a few remarks on the interpretation and limitations of some of the employed statistical testing procedures. An important issue in this context is the extent to which deviations from a specific distribution should be seen as invalidating a particular parametric model. A strong argument can in fact be made against the over-interpretation of strict statistical goodness-of-fit tests and tests of extreme hypotheses (see, for example, Berkson, 1938; and Savage, 1954, for such a discussion), with such tests more appropriately viewed as providing an assessment of a broad benchmark rather than a specific test of precise form (Ijiri and Simon, 1977; and Hart and Oulton, 1996, p.1244). This may be particularly of relevance with regards to the point that over a given time period the particular stochastic process may not have been operating for a sufficient time duration to evolve to the theoretical limiting distribution (Hart, 1980).

More practically, as both Ijiri and Simon (1977, pp.113-114) and Curry and George (1983) both reflect, goodness-of-fit type statistical tests often have limited power in determining whether one particular statistical distribution could be preferred to other similar statistical distributions. Additionally, in many of these studies the available data provided only an incomplete distribution of firm sizes, a factor which could potentially

---

[39] Bottazzi, Coad, Jacoby and Secchi (2005, p.2) suggest that the evidence supporting the Laplace distribution across different industry aggregations provides a "stylised fact" regarding firm growth rates (although Dosi, 2005, goes somewhat further suggesting that the relevant stylised fact may be that firm growth rate distributions are consistent with Laplace distributions or even greater mass in the distribution tails). However, analysis of official sales data on French manufacturing firms surviving between 1996-2002 indicated some greater mass in the tails of the growth rate distributions than predicted by the Laplace distribution, a finding supported by maximum likelihood estimation of the parameters of the Subbotin distribution which provided a shape parameter coefficient of 0.77. A weak negative relationship between firm size and growth rate variance was found. Negative serial correlation was detected, except for larger firms where a positive relationship was observed, so that despite the estimated regression coefficients suggesting independence between firm size and growth Gibrat's Law could be rejected.

affect the statistical properties of the examined distribution of firm sizes. Given these qualifications, it is perhaps not unreasonable to conclude that the evidence reviewed in this section is not unsupportive of firm size distributions broadly consistent with a Gibrat-type process, although as both Schmalensee (1989) and Sutton (1995a) conclude, no single statistical distribution appears capable of providing a universally adequate representation of either firm (and plant) size or growth rate distributions.

2.4. Gibrat's Law: Cross-Section Regression Studies

Complementing the statistical approaches to testing the lognormal distribution, regression methods have also been frequently utilised to test the propositions of Gibrat's Law, with the focus primarily on the examination of whether the firm size coefficient, β in equation (1.6), is statistically significantly different from unity, or from zero in the growth equation (1.10). Repeating these equations:

$$\ln X_{it} = \alpha + \beta \ln X_{it-1} + \mu_{it} \qquad (1.6)$$

and $$\Delta \ln X_{it} = \alpha + \beta_1 \ln X_{it-1} + \nu_{it} \qquad (1.10)$$

where $\Delta \ln X_{it} = \ln X_{it} - \ln X_{it-1}$.

The following sections provide a survey of empirical studies examining Gibrat's Law applying the regression approach. As Goddard et al. (2002) note, the estimation of these regression models of Gibrat's Law have generally been conducted using OLS estimation on cross-sections of data with evidence accumulated across a wide range of industrial samples (but particularly for production industries), time periods and geographical areas.

Hart and Prais (1956) tested Gibrat's Law using data on U.K. quoted companies in the Breweries and distilleries, Commercial and industrial, and Iron, coal and steel industries between 1885-1950. Estimation the standard logarithmic model provided a β coefficient close to unity for the periods 1885-1896, 1896-1907, 1907-1924 and 1924-1939, although the sampling errors for the first two of these periods did not provide sufficient confidence that the β coefficient could be considered as unity. Despite broad consistency, the third time period exhibited a coefficient somewhat higher, and the fourth period somewhat below unity. For the 1939-1950 period the β coefficient was significantly below unity, rejecting the first proposition of Gibrat's Law.

Several other regression based studies have found support for the first proposition of Gibrat's Law. Aaronovitch and Sawyer (1975) using data on a sample of 233 U.K. firms with at least £5m in assets, and estimating the logarithmic regression specification for the period 1958-1967 with corrections for heteroscedasticity found a β coefficient not statistically significantly different from unity.

Hymer and Pashigan (1962) also provided some further support for Gibrat's Law for large U.S. manufacturing firms, a finding, consistent with the evidence from the research studies of Hart (1965), Marcus (1969), Ijiri and Simon (1971) and Hannah and Kay (1977). A further U.K. investigation by Hart (1962) using samples comprising of 40 quoted Brewing companies, 36 Cotton-spinning companies, 124 firms in the Drinks industry and 229 unquoted firms with data across different periods in the 1930s to the 1950s also reported some broad support for the first proposition of Gibrat's Law.

However, a number of other early studies found less support for this proposition. Utton (1971) examining the growth of firms in manufacturing industries over the period 1954-1965 using the standard logarithmic regression specification, identified a statistically significant (at the 5 percent level) β coefficient exceeding unity for the full sample. For 10 of the 13 industry samples the β coefficient was observed to be above unity, indicating that larger firms were growing proportionally faster than smaller firms. Samuels and Chesher (1972), Utton (1972) and Prais (1976) also detected a positive size-growth relationship with larger firms growing faster then smaller firms. Further evidence provided by Samuels (1965) examining a stratified sample of 322 quoted companies over the period 1950-1960 further identified a statistically significant β coefficient of 1.07. Additional analysis suggested support for the first proposition of Gibrat's Law in the first half of the examined time period with a positive relationship exhibited thereafter.

Yet other studies have indicated potential non-linearities in the firm size-growth relationship with Fillipi and Zanetti (1971) finding that both smaller and larger firms grew relatively faster than intermediate-sized firms. A similar robust "U-shaped" relationship between firm size and growth, with smaller and larger firms growing faster than medium-sized firms was also found more recently by Bottazzi, Coad, Jacoby and Secchi (2005) using a regression approach employing a Least Absolute Deviation (LAD) estimator (the maximum likelihood estimator in the presence of Laplace distributed error terms, Kennedy, 1998, p.307) and employing a four-stage instrumental variable methodology (Ljung, 1987) to alleviate potential heteroscedasticity and autocorrelation.

In an important early study, Mansfield (1962) considered three interpretations of Gibrat's Law: for all firms, for surviving firms and for firms above the minimum efficient scale of production (MES) noting that if the probability of firm failure is related to size, then examination of surviving firms only could give rise to sample selection bias in the estimated relationship, and a bias towards finding faster growth in smaller firms.

These alternative versions of Gibrat's Law were empirically examined using ten samples of firms in the U.S. Steel, Petroleum refining and Rubber tyre industries over the period 1916-1957. Application of χ^2 tests and regression methods rejected the first proposition of Gibrat's Law in seven of the samples relating to the version covering all firms, with the independence of growth and size also being rejected in four of the industry samples at the 5 percent level of significance. The second version gave rise to three samples rejecting the proposition of independence between firm size and growth. The third specification referring to only those firms above the MES provided support for the first proposition of Gibrat's Law in all 10 samples, but with the variance of growth declining in six of these samples. In aggregate, Gibrat's Law provided a reasonable description in less than half of the samples tested with smaller firms exhibiting higher probabilities of failure, but higher average growth rates[40] for those surviving through the examined time periods.

In a more recent study employing a similar regression approach, Hart and Oulton (1996a) examined a sample of over 87,000 U.K. firms surviving across the period 1989-1993. The model specification controlled for both persistence and heteroscedasticity (following White, 1980) but not for potential sample selection bias. Three measures of firm size: employment, sales and net assets, were considered. The models were estimated on cross-sections of firm-level data using OLS procedures. The results suggested a consistent pattern of mean reversion for all size measures, although as Hart and Oulton (1996) noted the very short time period considered may have given rise to downward bias in the estimated β coefficients due to potential transitory factors in the measurement of initial firm size (a measurement error problem). Further analysis using graphical methods plotting geometric growth across size-classes and employing OLS estimated sequential regressions adding larger companies identified more significant deviations from Gibrat's Law for smaller firms than for larger firms.

---

[40] Similar evidence that smaller firms grew faster than larger firms was also found by Wedervang (1965) and is consistent with the model of Cabral (1995) where smaller firms choose to invest over longer time-spans giving rise to extended periods of higher growth rates for surviving small firms.

A number of published studies have extended the standard logarithmic model of Gibrat's Law to incorporate a range of additional explanatory variables, providing a conditional estimate of β (Goddard and Wilson, 2001, p.329). In such a study, Hart and Oulton (1998b) used a sample of 8,103 firms surviving over the period 1986-1995 and a model specification including firm size, age, their quadratic terms, a size-age interaction term and industry-level dummy variables. Cross–section estimates again suggested a process of mean reversion across the three three-year time periods considered with a broadly negative relationship between firm age and growth reported. Similar results were found at the industry-level and when using a translog form of the regression model. Hart and Oulton (1999) also produced further consistent results.

Elston (2002) has also provided a recent investigation of firm growth using two samples of German firms over the period 1970-1986, the first sample comprising of 820 firms with data constructed from a number of sources, and the second covering 295 firms listed on the German Neuer Markt. Several firm growth models were tested, including specifications with age and non-linear size terms and with controls for heteroscedasticity, multicollinearity and additionally industry, accounting period and year dummies. Although general support for Gibrat's Law was not found, more supportive evidence was produced when controlling for firm cash flow and in high-technology industries. In the case of traditional industries there was some evidence of larger firms growing more quickly, although such a finding was reversed for firms classified as operating in "new economy" activities.

In summary, these initial cross-section regression based studies, primarily focussing on larger production firms, provide rather mixed evidence on the consistency of Gibrat's Law, but perhaps with Gibrat's Law providing a more reasonable description of larger firms close to or above the MES. Some industry-level variation has been reported, with Audretsch et al. (2002) recently noting systematic effects across industries, with capital-intensive industries subject to larger sunk costs and economies of scale giving rise to higher growth rates and exit rates of smaller firms, and with industries with less significant sunk costs and economies of scale appearing to exhibit some greater degree of consistency with Gibrat's Law.

To some extent these differences in results may have been, at least partly, a consequence of the choice of the measure of firm size employed in the analyses. While most reasonable measures of firm size tend to be highly correlated (Blair, 1972; Hart and Oulton, 1998b) such correlation may be an insufficient condition for the inter-changeability of size measures, with any non-linear relationships between the logarithms of alternative measures potentially giving rise to different views on the acceptability of Gibrat's Law for alternative size measures (Smyth et al., 1975, p.112). Curry and George (1983, p.213) have also identified potential measurement issues when using financial and employment measures in examining Gibrat's Law across capital-intensive or labour-intensive firms, with some financial measures also subject to greater potential smoothing from accounting procedures (see, for example, Geroski et al., 2003, p.51).

Further influences on the results could also have arisen from the length of the time period examined and the use of either initial year or averaged (over a short time period) firm sizes as the measure of size (Hart, 1995, provides a discussion of the incorrect classification of firms by final rather than initial size employed by Davis et al. (1993) in examining firm growth in U.S. manufacturing plants between 1973-1988). The choice of the version of Gibrat's Law tested also appears significant, at least in some studies, in the observed acceptability of Gibrat's Law.

2.5. Gibrat's Law: Panel Data Studies

Although appropriate panel data methods offer the opportunity for improvements in the efficiency of econometric estimates and inferences, for the appropriate control of individual heterogeneity and the possibility of obtaining more detail on firm dynamics (Hsiao, 2003; Baltagi, 2001; Goddard et al., 2002), to date only a limited number of studies have examined Gibrat's Law using panel data estimation methods and estimators.

One such panel data study is provided by Goddard et al. (2002) who presented an analysis of several estimators for the growth regression model of Gibrat's Law: a cross-section estimator, a pooled OLS estimator with time dummies, a fixed effects estimator with both time and cross-section dummy variables (the latter enabling the possibility of heterogeneous firm-level effects) and the estimator of Breitung and Meyer (BM) (1994), described below, which is suitable for panels of short time dimension (Breitung and Meyer, 1994)[41]. The BM approach applies an OLS estimator to a transformed long-differences model with the test statistic provided by a modified Dickey-Fuller statistic (Oliveira and Fortunato, 2003).

The panel data econometric analysis of Goddard et al. (2002) employed a re-parameterised version of equation (1.10) with individual firm, α_i, and time effects, ϕ_i, given by:

$$\Delta \ln X_{it} = \alpha_i(1 - \rho) + (\phi_t - \rho\phi_{t-1}) + \beta_1 \ln X_{it-1} + \rho(\ln X_{it-1} - \ln X_{it-2}) + \eta_{it} \qquad (2.2)$$

where $\Delta \ln X_{it} = \ln X_{it} - \ln X_{it-1}$, $\eta_{it} = \varepsilon_{it} + \rho(1 - \beta)\ln X_{it-2}$, and under H_0: $\beta = 0$ that $\eta_{it} = \varepsilon_{it}$ (Goddard et al., 2002, p.422). The BM estimator is formulated by subtracting the initial firm size observations, $\ln X_{i0}$, from the right-hand-side of equation (2.2). The individual firm-effects are then added into the error term (Goddard et al. 2002, p.423), providing the model:

---

[41] The Anderson and Hsiao (1981) and Arellano and Bond (1991) instrumental variable estimators would be inappropriate in dynamic panels where $\beta = 1$, the hypothesis in testing the standard version of Gibrat's Law (Goddard and Wilson, 2001).

$$\Delta \ln X_{it} = (\phi_t - \rho\phi_{t-1}) + \beta_1(\ln X_{it-1} - \ln X_{i0}) + \rho(\ln X_{it-1} - \ln X_{it-2}) + \xi_{it} \qquad (2.3)$$

where $\xi_{it} = \eta_{it} + \alpha_i(1 - \rho)\beta_1 \ln X_{i0}$.

Reviewing these estimators, Goddard et al. note that the first two estimators provide unbiased and consistent estimators in the case of homogeneous individual effects[42]. The fixed effects panel data estimator is downward biased (Nickell, 1981), although the bias reduces as $T \rightarrow \infty$, with the power of the test statistics not weakened in the presence of heterogeneity in the individual firm-effects (reflected in the constant term, α_i, so that $\alpha_i \neq \alpha$, see equation 5.2). The BM estimator provides an unbiased estimator under acceptance of the null hypothesis H_0: $\beta_1 = 0$, but exhibits upwards bias in the case of H_0: $\beta_1 < 0$. This latter estimator is unaffected by heterogeneity in the individual firm-effects.

The outcomes of the application of these four estimators on the validity of Gibrat's Law was examined using assets data on a panel of 443 quoted Japanese manufacturing firms surviving over the period 1980-1996, with the analysis covering firm growth rates over the full 15-year duration and a number of shorter time-spans. The initial observation year 1980 was used to provide the lagged term reflecting growth persistence. Industry dummy variables were included in the cross-section regression with the panel data models additionally specified with macroeconomic and industry-time period interaction dummy variables. Heteroscedasticity was found to not be a significant problem with the use of White's (1980) correction having only a minor effect on the acceptance or rejection of Gibrat's Law.

Estimation of the model for the full 15-year time period rejected the null

---

[42] Although the cross-section estimator is unbiased where $\alpha_i = \alpha$, with the t-statistic providing a suitable test of H_0: $\beta_1 = 0$ versus H_1: $\beta_1 < 0$, under heterogeneity the cross-section estimator is inconsistent with β_1 generally over-estimated in the case of $\beta_1 < 0$ (Goddard et al., 2002, p.420).

hypothesis of proportional growth at the 1 percent level of significance for all four estimators with the reported model coefficients supporting a process of mean reversion and positive growth persistence. The cross-section estimator and the pooled panel data estimator suggested a coefficient closer to unity than that of the BM approach, a finding resulting from bias induced by failure to appropriately allow for individual heterogeneous effects in the former estimators (Goddard et al., 2002). Similarly, the first proposition of Gibrat's Law was also rejected using the annual data for each of the 5-year sub-periods, although at the industry-level the results were rather less conclusive. The hypothesis of no persistence in firm growth was rejected for over half of the estimates with most of these tending to suggest negative persistence.

Oliveira and Fortunato (2003) tested Gibrat's Law using an unbalanced panel of 8,814 Portuguese manufacturing firms between 1990-1999 employing the logarithmic growth specification without persistence term. The panel tests of Breitung and Meyer (1994) and Harris and Tzavalis (HT) (1999) under the hypothesis H_0: $\beta_i = 0$ for all i, versus H_1: $\beta_i < 0$ allowing for heterogeneity across firm groups, and the more restrictive test of Bond et al. (BNW) (2002) with the hypotheses H_0: $\beta_1 = 0$ versus H_1: $\beta_1 < 0$ were applied. The HT test uses a bias-adjusted least squares dummy variable estimation procedure, but is inappropriate in the case of heteroscedasticity. The BNW test utilises OLS regression (in levels) regressing the explanatory variable on its lagged term with time dummies, but is limited in the case of large variances[43] (Oliveira and Fortunato, 2003).

Supplementing the firm size variable (measured by the average number of employees in each year), firm age and a number of financial accounting ratio variables

---

[43] Hall and Mairesse (2001) have also conducted an examination of the univariate time series properties of high-tech manufacturing firms in France, Japan and the U.S. using real and simulated data covering the 1978-1989 period, and specifically examining the performance of a number of unit root tests including the HT and BNW tests, and those of Im et al. (1997), Kruiniger (1998) and conditional maximum likelihood and Bayesian maximum likelihood estimators. The presence of a unit root was rejected, with the data suggesting stationarity with a high autoregressive component rather than the random walk of Gibrat's Law (Hall and Mairesse, 2001, p.23).

were included in the model. The null of a unit root was rejected using all three tests for each variable (with the exception of firm age where a unit root could not be rejected using the BNW test) with support provided for mean reversion in the firm size-growth relationship, rejecting Gibrat's Law.

In a further recent study, Geroski et al. (2003) employed a panel data approach to examine Gibrat's Law using continuous observations on a non-random sample of 147 large, quoted U.K. firms over the period 1955-1985. Deflated net assets were used as the measure of firm size. The logarithmic growth specification of Gibrat's Law was employed. The model was initially tested using a fixed effects panel data estimator which provided some evidence supporting a process of mean reversion. However, the annual cross-section estimates and long-first differences estimates tended to suggest somewhat weaker patterns of convergence (suggesting that the treatment of firm heterogeneities may have influenced the estimated regression coefficients). Regressions on the individual firm time series to estimate the distribution of the β_1 size coefficients (relating to the model specification of Gibrat's Law given by equation 1.10) using a feasible generalised least squares estimator (Swamy, 1971) to allow for potential cross-equation covariance (Geroski et al., 2003, p.53) again supported the view of overall mean reversion, but with heterogeneous convergence rates. The analysis of firm sub-groups indicated some potential tendency for greater convergence in the case of smaller firms.

To further examine whether firm growth was consistent with a random walk process Geroski et al. (2003) applied the Augmented Dickey-Fuller (ADF) (Dickey and Fuller, 1979, 1981) unit root test with and without deterministic trends (with the null hypothesis of a unit root H_0: $\beta_1 = 0$ versus H_1: $\beta_1 < 0$ for equation (1.10) examined using appropriate critical values, Wang, 2003, p.16) along with the Im et al. (IPS) (1997) panel unit root test and the test of Maddala and Wu (MW) (1999), given the potentially increased power of such tests)[44]. In both panel unit root tests the null hypothesis is that H_0: $\beta_i = \beta_1 = 0$ for all firms with the alternative hypothesis permitting some heterogeneity in β_i with H_1: $\beta_i < 0$ for all i, avoiding the more restrictive assumption that H_1: $\beta_1 < 1$ when testing for convergence, an assumption implying common rates of convergence (Oliveira and Fortunato, 2002).

The IPS unit root test, which uses averages of Lagrange Multiplier statistics, was applied across each firm for models with and without constant and trend terms. The MW test based on the non-parametric exact test of Fisher (1932) with a test statistic given by $F = -2\Sigma \log_e p_i$ and distributed as χ^2 with 2N degrees of freedom, was employed using the Schwartz Bayesian Information Criteria (Schwartz, 1978) to determine the appropriate lag structure[45].

The estimation procedure was conducted as a repeated process removing all observations where β_i was less than zero. Their analysis suggested that the presence of a unit root could not be strongly rejected using the ADF, IPS or MW tests, supporting the random walk proposition of Gibrat's Law. Residual-based tests for cointegration using the ADF test also supported the hypothesis of non-stationarity for the differences between firm sizes, offering further support for Gibrat's Law.

However, although providing an interesting emerging theme in the literature on Gibrat's Law it is at this stage perhaps premature to over generalise the results from such panel data analyses based on so few studies. To this point, such panel data studies do not provide the same breadth of industry, country and time period coverage as in the cross-section analyses, and have not employed panel data sample selection models which may

---

[44] Im et al. (1997) found some superiority of performance of their test over the Levin and Lin (LL) (1992, 1993) tests, a finding supported by Karlsson and Löthgren (2000) and Maddala and Wu (1999), the latter concluding that the Fisher (1932) test was more powerful and robust than the IPS and LL tests.

[45] Such information criteria provide a means of choosing between alternative models through a trade-off between goodness-of-fit (measured by likelihood value) and the parsimony (measured by the number of free parameters) (Verbeek, 2000; Kennedy, 1998).

have some further influence on the estimated validity of Gibrat's Law.

2.6. The Role of Sample Selection Bias

As Sutton (1997) notes, recent assessments of Gibrat's Law have increasingly focussed on providing detailed econometric treatments of a range of modelling issues such as sample attrition and censoring, heteroscedasticity and functional form. Although the role of heteroscedasticity and functional form are both considered in this section, the primary issue of focus here is the role and analysis of potential sample selection bias.

The use of sample selection models to (asymptotically) alleviate potential sample selection bias arising from the non-random exit of firms observed in the initial year has provided a significant theme in the regression approach to examining Gibrat's Law. The standard sample selection correction procedure has involved the construction of a model with a two non-independent equations structure with a probit selection equation providing the inverse Mills' ratio (generated from the equation residuals) which is then included in the growth equation as an additional explanatory variable.

Making use of such a sample selection model Hall (1987) examined the growth of firms using two panels of data drawn from a sample of U.S. publicly traded manufacturing firms, the first relating to firms with continuous employment information over the period 1972-1979 and the second for firms with employment information for 1976-1983. Potential measurement error[46], which could generate a tendency towards regression-to-the-mean bias, was controlled for through the use of firm size in the year prior to the initial year of analysis as an instrumental variable (being assumed to be highly

---

[46] If in a model specification firm growth is the dependent variable and initial firm size is the independent variable (measured with error) then over the examined time period the firms with transitorily low size as a result of measurement error will appear on average to exhibit more rapid growth than those having transitorily high size (Hall, 1987, p.584). Measurement error in the independent variable provides potentially biased and inconsistent OLS estimates, with potentially inefficient estimation where the measurement error is in the dependent variables (Kennedy, 1998).

correlated with the explanatory variable but uncorrelated with the disturbance term, see Gudjarati, 1988). The effects of measurement error were also examined using a Markov-model of growth with errors-in-variables.

Estimating the size-growth relationship using the logarithm of employment as the measure of firm size and OLS procedures suggested a negative relationship between firm size and growth. Measurement error provided only a limited contributory effect to this estimated relationship. A generalised Tobit sample selection model estimated using maximum likelihood with corrections for heteroscedasticity provided similar results to the OLS estimates, with selection bias not providing a significant influence on the observed negative firm size-growth relationship. Some non-linearity in the firm size-growth relationship was detected, so that although Gibrat's Law did not provide a reasonable description for smaller firms it was more acceptable for larger firms (Hall, 1987, p.583). In addition, the second proposition of Gibrat's Law, that the variance of firm growth rates is independent of firm size, was also rejected, with larger growth rate variances observed for smaller size-classes.

In the first of two influential papers, Evans (1987a) examined a sample of 42,339 manufacturing firms from 100 randomly selected 4-digit industries drawn from the Small Business Data-Base (SBDB) of the U.S. Small Business Administration, analysing firm growth patterns over the period 1976-1980. Explanatory variables relating to firm size (measured by employment), age and the number of plants, supplemented by non-linear size and age terms, and interaction terms between both firm and plant size and age were included in the model specification. A sample selection model was estimated using maximum likelihood with robust standard errors following White (1980). The results indicated a positive, but statistically insignificant, correlation coefficient between the growth and selection equations. Firm growth was found to be negatively related to firm size and age, with the probability of survival also positively related to these factors. The observed rejection of Gibrat's Law was more severe for firms comprising of larger numbers of plants and for smaller firms.

In the second study, Evans (1987b) drew a stratified sample of 27,046 firms operating in 1976 from the SBDB and examined firm growth over the period 1976-1982. Separate regression and sample selection models (the latter estimated using maximum likelihood) were estimated for four constructed age-classes, 0 - 6, 7 - 20, 21 - 45 and at least 46 years, with non-linear firm size and age terms included in the model. Similar parameter estimates were observed under both model specifications with the correlation coefficients between the growth and selection equations not generally being statistically significant. Again, firm growth was observed to be negatively related to both firm size and age, rejecting Gibrat's Law.

In a notable variation to the standard sample selection model procedures, Dunne et al. (1989) using a sample of 219,754 plants constructed from five consecutive U.S. Census of Manufactures' covering the period 1967-1982, employed a grouping procedure whereby plant age and size-classes could be represented by 15 dummy variables. Consistent parameter estimates for the distribution of both all and continuing plants were generated through regression procedures to estimate across-cell variance with some degree of homogeneity assumed for observations within each cell. The use of this alternative procedure avoided the distributional and functional form assumptions required within standard sample selection model procedures and alleviated issues of identifying the effects of sample selection, heteroscedasticity and non-linearities in the explanatory variables, but involved some loss of information (Dunne et al., 1989; Sutton, 1997). The estimates using this approach suggested that both firm growth and exit probabilities declined with firm size. In addition, some differences in the survival probabilities and growth performance of single and multi-plant firms were detected, but with the broad regularity of declining exit probability, growth rate and growth rate variance with increased plant size and plant age holding across both single- and multi-plant structures.

Returning to the evidence employing the conventional sample selection model procedure, Dunne and Hughes (1994) conducted a study using a sample of 2,149 U.K. manufacturing companies between 1975-1985. Net assets was employed as the measure of firm size. Following initial statistical tests of the first two propositions of Gibrat's Law using the Welch-Aspin test (Welch, 1947; Aspin, 1948) for samples with unequal variances to test for differences in means and F-tests for differences in variances (both suggesting deviations from Gibrat's Law), the standard logarithmic specification of Gibrat's Law was estimated using OLS regression with the method of White (1980) used to produce robust standard errors. Separate firm size-class and industry-level regressions were estimated, the latter to avoid potential aggregation bias.

For the period 1980-1985 these regressions suggested that the first proposition of Gibrat's Law could not be generally rejected. However, rejections of Gibrat's Law were found for the period 1975-1980 including across size-classes. These results were confirmed through examining the lower β regression coefficient estimates in the separate size-class equations for 1975-1980 using formal analysis of covariance, re-estimating the equation with slope and shift size-class dummies (Dunne and Hughes, 1994). A series of sample selection models including variables relating to firm size, age and quadratic size and age terms to reflect potential non-linear relationships were estimated. Comparisons of the OLS estimates from the standard linear size-growth regression model with the sample selection models estimated using maximum likelihood generally provided similar β size coefficient estimates. Firm growth was again found to be negatively related to both size and age. However, at the industry-level, the sample selection model tended to suggest a higher rate of rejection of the first proposition of Gibrat's Law than found using OLS regression.

Harhoff et al. (1998) also employed a sample selection model, using the two-step estimator of Heckman (1976) to examine the relationship between the legal form of firms and firm survival and growth for a sample of 8,068 West German firms. Their sample covered a number of production and services industries over the period 1989-1994. The regression model utilised the firm's employment growth rate as the dependent variable, with size, age and higher-order expansion terms of these variables included as regressors. Dummy variables for industries and firm structure (legal form and ownership) were incorporated into the model along with time dummies allowing for different growth rate time periods. Three model specifications were estimated, the first being a regression model estimated using OLS, the second and third being sample selection models with industry dummy variables at the 1- and 2-digit levels, respectively, each with corrections for heteroscedasticity. In all three specifications Gibrat's Law was rejected with a regression coefficient less than unity, with similar findings provided for the industry-level samples. Further sensitivity analysis using quantile regression[47] again supported these findings. A modest negative effect of firm age was observed for the majority of samples.

Lotti et al. (1999) tested Gibrat's Law on a sample of 970 manufacturing firms entering in early 1987 and tracked until 1993 using data from the Italian National Institute for Social Security (INPS) dataset. Potential sample selection bias was addressed through a two-equation sample selection model with firm size and quadratic size variables included in the probit selection equation with the model estimated using Heckman's (1976) two-step procedure. Corrections for heteroscedasticity were applied in both the OLS regression and sample selection models. The analysis rejected Gibrat's Law for the immediate years following entry, but with these deviations moderating over time.

A similar sample selection model approach was also employed by Mata (1994) to examine data on 3,308 new entrant Portuguese manufacturing firms between 1983-1987

---

[47] See Coad (2006) for a further application of quantile regression analysing firm growth rate distributions with lagged size, growth and year dummies as the explanatory variables.

with corrections for heteroscedasticity. The first proposition of Gibrat's Law was rejected for smaller firms, but was found to provide a reasonable description for firms in the size-classes of 84 - 89 employees and above.

Blonigen and Tomlin (1999) have further examined the relationship between plant size and growth using data on Japanese manufacturing plants located in the U.S. Gibrat's Law was rejected for three-year plant-level employment growth rates for 1987-1990, with corrections for potential sample selection effects not found to be significant. A negative relationship between plant age and growth was observed.

Finally, Cosh et al. (1996) have provided a comparison of the use of the standard sample selection model estimated using maximum likelihood with alternative semi-parametric estimators. Using U.K. firm-level data from a stratified sample of 2,142 companies with broad industry coverage (although some smaller companies were excluded) covering the period 1976-1982 and nominal assets as the measure of firm size, initial cross-section estimates of the conventional linear size-growth regression model did not lead to a rejection of the first proposition of Gibrat's Law. A negative age-growth relationship was detected. Employing the standard sample selection model estimated using maximum likelihood did not significantly alter these findings.

Further examination of firm survival applying the semi-parametric Klein-Spady (1993) empirical maximum likelihood estimator suggested a negative relationship between firm age and survival but with some industry effects in the survival components. The Ichimura (1993) semi-parametric least squares estimator also suggested some consistency in the estimated firm size-age relationship, corresponding to the linear regression estimate, though with some reported significant changes in the scale and relative effects of the explanatory variables. However, their analysis suggested that the standard probit model failed to adequately capture some of the complexities of the firm size-survival relationship (and the relationship to other explanatory variables).

Overall, the reviewed evidence is perhaps suggestive of the use of such sample selection models (estimated using either maximum likelihood or Heckman's (1976) two-step estimator) employed in the empirical studies of Gibrat's Law tending not, in general, to significantly alter the implications of the standard linear growth regression models, with most recent studies suggesting a fairly consistently observed (but frequently modest – so that authors such as Geroski (1999) have concluded that Gibrat's Law remains an appropriate benchmark) negative relationship between firm size and growth robust to such sample selection corrections.

2.7. The Variance of Firm Growth

Many of the studies of Gibrat's Law have provided only a limited direct assessment of the variance of firm growth. However, the available empirical evidence suggests that whilst some support for an independent relationship between firm size and the variance of growth has been provided by studies such as Hart and Prais (1956), Simon and Bonini (1958), Hart (1962) (in the case of samples of cotton spinning firms and quoted companies, but not for brewing firms or drinks manufacturers) and more recently Teitelbaum and Axtell (2005) using data from the U.S. Census Bureau for plants surviving between 1998-1999, more frequently deviations in the form of a declining variance of firm growth with increased firm size have been reported (with early studies finding such a relationship including Hymer and Pashigan, 1962; and Mansfield, 1962).

Singh and Whittington (1975) investigating Gibrat's Law using data on 1,955 U.K. quoted companies in the manufacturing, construction, distribution and miscellaneous services industries over the period 1948-1960 also found support for a declining growth rate variance with increased firm size in each of the 21 industries considered. Similarly, Kumar (1985) examining data on U.K. companies between 1960-1976 found lower growth rate variances for larger firms.

Hall's (1987) analysis of U.S. quoted manufacturing firms covering the period 1972-1983 provided further support for such a negative relationship between firm size and growth rate variance, and therefore rejection of the second proposition of Gibrat's Law. Dunne et al. (1989) using large samples of plants constructed from five consecutive Census of Manufactures covering 1967-1982 also found declining growth rate variances with increased firm size (also declining with firm age) a finding that was observed for both single- and multi-plant firms. Further support for such a relationship was also detected by Dunne and Hughes (1994) examining the U.K. manufacturing sector and Cosh et al. (1996, p.18) who observed that the largest three of their constructed firm size-classes exhibited lower variances than for the smallest three size-classes.

More recently other specific relationships between firm size and growth rate variance have been proposed and empirically tested. For example, Stanley et al. (1996) examining sales and employment data on quoted U.S. manufacturing firms over the period 1974-1993 found support for the variance declining according to a power law (or scaling relationship), which can be represented generally by:

$$\sigma(g_{it} \mid X_{it-1}) = \theta X^{\gamma}_{it-1} \qquad (2.4)$$

where X denotes firm size, g is the firm growth rate, θ is a constant term and σ is the standard deviation of the logarithm of firm growth rates[48]. The employment data generated γ parameter estimates of 0.16 ± 0.03 with the sales measure giving an estimate of 0.15 ± 0.03, indicative of some degree of performance correlation across units (see

---

[48] A linear model $\ln(\sigma(g_{it} \mid X_{it-1})) = \alpha + \gamma X_{it-1}$ can then be estimated (Bottazzi and Secchi, 2003a, p.6).

also Dosi, 2005, who suggests that the recently reported evidence suggesting a variance declining less than proportionately with increased firm size is consistent with the notion of the diversification of multi-product firms)[49].

Similar power law relationships between firm size and growth rate variance have also been observed by Amaral et al. (1997), Sutton (2001) using a dataset of Japanese firms covering 45 industries over the period 1973-1997, and Bottazzi and Secchi (2003a) examining 15 U.S. manufacturing industries over the period 1982-2001 drawn from the Compustat dataset. Such a power law relationship was also derived within the maximisation models of Axtell (1999) estimating a power law exponent of 0.174 ± 0.004 using OLS procedures.

2.8. Serial Correlation and Firm Growth Persistence

Persistence in firm growth has been investigated in a number of studies of Gibrat's Law. In an important study, Chesher (1979) demonstrated that the presence of autocorrelation in the regression model disturbance terms would produce inconsistent and biased OLS estimates of the β regression coefficient corresponding to the first proposition of Gibrat's Law (but with the magnitude of this bias diminishing with increased time duration) due to the dependence between measures of lagged firm size (or measures of the deviation of firm size from the mean as used in Chesher, 1979) and the lagged disturbance terms. To allow for potentially autocorrelated disturbance terms Chesher (1979) generated a model of the form:

$$z_{it} = \beta z_{it-1} + \varepsilon_{it} \qquad \text{where} \qquad \varepsilon_{it} = \rho\varepsilon_{it-1} + \upsilon_{it} \qquad (2.5)$$

---

[49] A coefficient of γ approximating to 0.5 would be consistent with independence across business units.

where z is the deviation (normalised) of the logarithm of firm size from the mean. The error term, ε_{it}, was assumed to follow a first-order autoregressive process with υ_{it} assumed to be not autocorrelated (Audretsch et al., 2002, p.7). Reformulating the error terms ε_{it} and ε_{it-1} in terms of z_{it} and its lagged terms generated the following model:

$$z_{it} = \gamma_1 z_{it-1} + \gamma_2 z_{it-2} + \varepsilon_{it} \tag{2.6}$$

where $\gamma_1 = (\beta + \rho)$ and $\gamma_2 = (-\beta\rho)$ with β providing an estimate of the slope coefficient and ρ being a coefficient of first-order autocorrelation. In this formulation, acceptance of Gibrat's Law requires the joint acceptance that $\beta = 1$ and $\rho = 0$. Chesher noted that the γ parameters of equation (2.6) could be consistently estimated using OLS procedures applied to cross-section data and then used to estimate the parameters β and ρ under the condition that:

$$(\tilde{\beta}, \tilde{\rho}) = \tfrac{1}{2}\{\hat{\gamma}_1 \pm (\hat{\gamma}_1{}^2 + 4\hat{\gamma}_2)^{1/2}\} \tag{2.7}$$

The identification problem that arises from the inability to directly assess the values of β and ρ was addressed by assuming that the β coefficient was approximately unity so that $\tilde{\beta}$ could be assumed to be given by the estimated coefficient closest to unity. These coefficients were also assumed to be constant over time (Chesher, 1979, p.407). The null hypothesis of Gibrat's Law (with first-order autocorrelation) was tested using:

$$H_0: (\gamma_1, \gamma_2) = (1,0) \tag{2.8}$$

$$H_1: (\gamma_1, \gamma_2) \neq (1,0)$$

assuming that the OLS estimates of the parameters in equation (2.6) were asymptotically normally distributed so that the test statistics could be considered as approximately distributed as χ^2 under the null hypothesis, with 2 degrees of freedom (Chesher, 1979, p.408; Audretsch et al., 2002, p.7; see also Malinvaud, 1966). Using information from a sample of 183 U.K. quoted firms surviving over the period 1960-1969 (with capital employed as the measure of firm size) even though β was found to be close to unity the estimate of ρ strongly supported positive first-order autocorrelation, so that Gibrat's Law could be rejected.

Tschoegl (1983) analysed the size-growth relationship of large transnational banks between 1969-1977 following the Chesher (1979) specification with assumed first-order autocorrelation. Using a number of financial measures of bank size and including several dummy variables to represent countries with multiple bank organisations and for macroeconomic factors, support was found for a β coefficient close to unity for the two-year growth rates examined, with modest persistence in growth rates and a declining variance of growth rates with increased bank size.

In a further study, Audretsch et al. (2002) examined a sample of over 1,170 service sector firms in Denmark between 1987-1991. In addition to testing for the independence of growth rates and firm size using the χ^2 statistic (an approach also employed by, for example, Hymer and Pashigan, 1962; Singh and Whittington, 1975; and Acs and Audretsch, 1990), Gibrat's Law was also tested using the Chesher (1979) specification.

Only minor persistence effects were identified through the examination of the estimated autocorrelation coefficients with the inclusion of second-order and third-order autocorrelation terms not adding significantly to the estimated first-order process for either the whole sector or for the individual services industries in the case of the dependent variable relating to 1991. However, some industry-level variation was observed for the dependent variable relating to 1990 with the second-order autocorrelation coefficients found to be statistically significantly different from zero (and negative) for the cafeterias, cafes and hotels industry, but not for the other three industries examined. Overall, the analysis tended to suggest some general support for Gibrat's Law but with variation across both industries and time periods.

Piergiovanni et al. (2002) also conducted an analysis using the Chesher (1979) specification with assumed first-order autocorrelation for official Italian hospitality services data covering the period 1989-1994, providing a sample of 9,051 new firms. Gibrat's Law was rejected in the majority of the samples using preliminary χ^2 tests of independence between firm size and growth[50] for the first two versions, with the joint hypothesis that $\beta = 1$ and $\rho = 0$ in the regression model being rejected in around 80 percent of samples.

An alternative approach to the examination of growth persistence has been through the direct incorporation of a persistence term into the logarithmic firm size-growth model, such as in equation (1.20), or with the dependent variable specified in growth terms, as in equation (5.3). Applying such a growth specification, Singh and Whittington (1975) considered the effects of persistence on the validity of Gibrat's Law using data on 1,955 quoted U.K. companies in the manufacturing, construction, distribution and miscellaneous services industries covering the period 1948-1960 with firm size measured by both net assets and the value of physical assets (unadjusted for inflation). Estimation of the standard logarithmic growth specification regression suggested a modest, but positive relationship between size and growth. A similar relationship was also found for 19 of the 21 industries analysed, albeit being statistically

---

[50] Such a result contrasts with that of Acs and Audretsch (1990) who using the same statistical approach testing the distributions of firm growth rates across four constructed firm size-classes employing data on 408 4-digit level manufacturing industries between 1976-1980 sourced from the SBDB found support for Gibrat's Law in around 60 percent of the industries examined.

significant at the 5 percent level for only three of these samples. Analysis of firm growth persistence conducted by regressing the growth rates in the period 1954-1960 on the growth rates in the period 1948-1954 for the whole sample and for individual industries suggested some evidence of persistence in growth rates. Further investigation suggested that this relationship was weaker when physical assets data was examined. Inspection of the R^2 values, however, tended to indicate that past growth was not a good predictor of subsequent growth.

Singh and Whittington also identified that the omission of an appropriate past growth variable in the standard logarithmic model specification could give rise to a potential upwards bias (although they also recognised the possibility of biases in the opposite direction arising from measurement errors in the size of firms in the initial year). Using a regression specification including a variable for past growth provided a lower estimated β coefficient for the full sample of firms, with the analysis suggesting that much of the positive relationship found in the initial model specification could be attributed to the effects of persistence in firm growth (Singh and Whittington, 1975).

Kumar (1985) also found evidence of some persistence in firm growth rates using a similar regression methodology applied to data on 1,747 U.K. quoted firms covering the period 1960-1976. Analysis of various measures of firm size (sales, net assets, physical assets, equity assets and employment) indicated positive growth rate persistence in some industries, but of a more modest magnitude than observed by Singh and Whittington (1975). A slight tendency for mean reversion of firm sizes for both manufacturing and services sectors for the majority of considered samples was detected.

Further information on persistence in firm growth rates has been contributed by Dunne and Hughes (1994) using net assets data on 2,149 U.K. manufacturing firms over the period 1975-1985 and examining persistence by regressing firm growth in the latter five-year period on that of a previous five-year period. Their analysis suggested weaker persistence than in both Singh and Whittington (1975) and Kumar (1984), with statistically significant regression coefficients in only four of the twenty-three industries analysed, with past growth providing a poor predictor of subsequent growth. The observed persistence tended to be generally negative in the case of smaller firms, but positive in the case of larger firms[51].

Several other studies have further contributed to the evidence on the existence and form of persistence in firm growth. For example, Ijiri and Simon (1964, 1967) using data on a sample of 96 large U.S. firms for the period 1954, 1958 and 1962 and studying four-year sales growth rates observed some degree of short-run persistence. Goddard et al. (2002) found some support for negative persistence in growth rates, with similar findings reported by Shen (1970) and Contini and Revelli (1989). In contrast, Wagner (1992) using data on around 7,000 manufacturing firms with at least 20 employees in Lower Saxony between 1978-1989 found evidence of some positive serial correlation in firm growth rates so that despite finding coefficients often near to unity Gibrat's Law could be rejected in around 90 percent of the samples. However, Wagner (1994), again studying manufacturing firms, did not detect significant firm growth persistence, with similar evidence also provided by Hart and Oulton (1998b) and Lotti et al. (1999). The overall evidence therefore tends to provide somewhat inconclusive empirical evidence regarding persistence in firm growth, finding cases of positive persistence, negative persistence and no statistically significance persistence.

2.9. Conclusions

This chapter has provided an overview of the main econometric and statistical studies testing Gibrat's Law. The initial sections presented a survey of the empirical literature

---

[51] Positive serial correlation could arise from the cumulative outcome of a firm's relative competitive position, negative serial correlation arising from a statistical error-correction process (Caves, 1997, p.5).

examining a range of statistical distributions derived from theoretical models underpinned by Gibrat-type processes, focussing in particular on the lognormal distribution.

Although the examined empirical evidence suggested some support for the lognormal distribution as an approximation to observed firm size distributions, a number of studies have also provided support for statistical distributions such as the various Pareto distributions (at least over some ranges of the distribution of firm sizes) and in several recent studies the Zipf distribution. However, perhaps not unsurprisingly none of the postulated distributions can be claimed to have provided an adequate description all of the examined distributions of actual firm and plant sizes across the full range of countries, time periods and industry samples, a point emphasised by Schmalensee (1989) and Sutton (1995a). Recent econometric and statistical studies have also generated some support for both the Laplace and more general Subbotin distributions as appropriate descriptions of the distribution of firm growth rates, rather than the normal distribution generally associated with Gibrat's Law.

In examining the appropriateness of particular theoretical distributions as approximations to firm sizes the statistical studies have often attempted to examine the goodness-of-fit between the observed (unconditional) and theoretical statistical distributions (although some more recent studies have used quantile regression methods to provide information on conditional firm size distributions). Such statistical testing procedures are frequently limited in their power to discriminate effectively between similar competing distributions. Further, many reported studies have been based on incomplete data on the distribution of firm (or plant) sizes (especially for smaller sizes), focussing particularly on large production firms. Such partial data may potentially affect the statistical properties of the examined distributions of firm sizes.

Additionally, limitations in attempting to test the validity of the lognormal distribution have also arisen since many of these tests have examined static firm size distributions at a single point in time so that the observed deviations may have been an artifact of the operating stochastic process not having had sufficient time to enable convergence (or otherwise) to the lognormal (or alternative) limiting distribution (Hart, 1980). In combination, these issues tend to suggest that the specific theoretical distributions should not be too easily discarded where they broadly approximate to the observed distributions of firm sizes.

This chapter has also presented a detailed review of a wide range of studies examining the empirical propositions of Gibrat's Law relating to the patterns of firm growth. The studies examined included early approaches using regression models, often considering only limited samples of larger firms in production industries, before reviewing some of the more recent studies which have focussed on a range of econometric issues such as potential sample selection bias, firm growth persistence, heteroscedasticity and the functional form of the models in developing more robust empirical tests of Gibrat's Law. In addition, the empirical evidence on the role of firm age in the observed patterns of growth dynamics has been discussed.

These empirical studies have often provided evidence of strict deviations from the three propositions of Gibrat's Law (summarised in McCloughan, 1995). However, these deviations have neither been always observed nor of a consistent magnitude. In relation to the first proposition, various estimates of the firm size-growth regression model across different samples, geographies and time periods have suggested parameters not statistically different from unity (or zero in the growth model specification) confirming the first proposition, but also frequently with parameters both significantly below and (less frequently) above this benchmark, and indeed with other results suggesting non-linearities across different firm or plant size-classes.

However, more recent studies have generally tended to suggest some degree of consistency in finding modest mean reversion (sometimes interpreted as being sufficiently modest so as to not reject the random walk implication of Gibrat's Law), even after correcting for potential sample selection bias, at least for production-based activities, but with Gibrat's Law being more appropriate for larger firms, a feature consistent with the models of Ijiri and Simon (1977). New evidence on the services sector is rather under-developed, but at least initially may be more supportive of this proposition of Gibrat's Law. Across most studies the inclusion of a firm age variable has often provided consistent evidence of a negative relationship between firm age and growth.

In relation to the second proposition, that the variance of firm growth is independent of firm size, the evidence tends to be rather more conclusive, with this condition not holding for most samples, and generally exhibiting a declining relationship with firm size. In some cases the firm size-growth rate variance relationship has appeared consistent with a power law representation. However, even on this point there are a number of studies which have provided some support for this proposition of Gibrat's Law.

The third proposition of no persistence in firm growth has produced mixed results with evidence of no persistence, positive persistence and negative persistence in different studies, and some significant variation in the magnitude of these estimated persistence effects. Such persistence may be, at least in part, a consequence of the time-span of the data considered and also the measure of firm size employed with some accounting-based measures being smoothed over a period of time.

Building on this empirical evidence regarding firm (or plant) size and growth rate distributions and the growth dynamics of firms (and specifically the coherence with the propositions of Gibrat's Law), the later chapters presented in this research aim to provide new estimates and further detail on such observations. In doing so, the analysis will focus on taking account of many of the statistical and econometric methodology issues identified and discussed in this chapter.

3

AN OVERVIEW OF THE MICRO-DATASET: PROCEDURES AND EMPIRICAL PROPERTIES

3.1. Introduction

This chapter provides an introduction to the business micro-database employed in this research, the Welsh Register of Manufacturing Employment. Section 3.2 reviews the structure of the database, the data collection processes and the methods employed to extract and clean the dataset for analysis. An overview of the variables available from this dataset and their empirical properties is presented in Section 3.3, providing further context to this analysis of manufacturing plant dynamics. In particular the properties of the static, aggregate plant size distributions, the empirical patterns of employment flows across plants, and plant entry and exit, which in combination contribute to the shaping of plant size distributions, are described. The chapter also examines the dynamics of different sub-samples of plants, including by industry and ownership structure, given that the examination of specific dynamic processes at these levels may assist in understanding the empirical findings in the subsequent econometric and statistical research.

Based on the analysis of this data and supplementary evidence relating to U.K. economic cycles, the discussion in Section 3.4 leads to the determination of a number of time periods of interest over which the propositions of Gibrat's Law are explored, with the methodology employed in constructing the samples for the alternative versions of Gibrat's Law to be tested also being presented. Consideration is further given in Section 3.5 to issues relating to the measurement of business size and the choice of the unit of analysis, firms or plants. A summary of the chapter and conclusions are provided in Section 3.6.

3.2. Database Structure and Data Procedures

The data employed in this research is drawn from the Welsh Register of Manufacturing Employment (WRME), a business micro-database managed until recently by the Economic Advice Division of the Welsh Assembly Government (formerly the Welsh Office). The WRME replaced the Department of Trade and Industry's Regional Data System in February 1995 (Moyes and Thomas, 1996), which had itself developed from the earlier Record of Movement, recording new manufacturing plant (establishment) openings over the period 1945-1965, and the subsequent Record of Openings and Closures from 1966.

The WRME maintains a record of manufacturing plants, covering Standard Industrial Classifications SIC (1992) 15 – 36, with more than 10 employees at a single site in Wales (DTI, 1988, p.8). In addition, some information on business units (termed here also as "plants") in the Mining of coal and lignite (SIC 10), Other mining and quarrying (SIC 14), Recycling (SIC 37), Wholesale trade (SIC 51), Retail trade (SIC 52), Activities auxiliary to financial intermediation (SIC 67); Computer and related activities (SIC 72) and Other business activities (SIC 74) is also recorded. Some records are also retained under a provisional holding code status. The overall number of such business units is small, usually representing at most around 1 percent of recorded plants in each year (although in very recent years this has increased to around 4.5 percent).

The information recorded on WRME is collected over a short period of time at the end of each calendar year using a postal survey and telephone follow-up, with some updating in-year, at various points having also made use of a range of statistical and administrative sources, such as the Census of Employment, Census of Production, Inter-Departmental Business Register (IDBR) and from the commercial information provider Dun & Bradstreet.

The primary variable recorded on WRME is a headcount measure of employee numbers. Further information on the industry (at both the SIC 2- and 4-digit industry-levels), initial country of ownership, whether the plant is part of a multi-plant enterprise, and opening and closure dates is also recorded. In some cases (partial) data on the type of opening, being a New Branch if the new plant is opened by an enterprise with other manufacturing units already in operation, a Transfer if the formation of the new plant corresponds to a closure of one or more existing plants, and an Enterprise New to Manufacturing (ENM) if the new plant is opened by an enterprise previously not having other plants within the manufacturing sector, is also available (this information is presented in Table 3.2., but is not discussed in detail due to this partial nature of the recorded data). In addition, information is provided on plant postcode location and whether this address is linked to some form of assisted-area status. The WRME microdata are electronically stored using standard database software.

To make full use of the time series and cross-section dimensions of the WRME data a complete list of records were extracted providing records on 6,372 plants with at least one year of data. However, for a number of plants the annual records of employee numbers were incomplete. Where only one or at most a few consecutive observations were missing these values were imputed using a simple straight-line averaging procedure estimated using the immediately adjacent observations[52]. For a very small number of records significant elements of the data were absent so that no such estimates could be reasonably imputed. These observations were necessarily discarded.

---

[52] Moving average methods for imputing missing employee numbers values were also investigated. The straight-line approach was preferred due to its simplicity and since the moving average methods could not be consistently applied in the case of plants with very short recorded survival durations.

Although the straight-line approach may be considered somewhat simplistic, given the relatively small number of units for which such adjustments needed to be made such a procedure would seem unlikely to induce significant bias. The straight-line method may be further justified since the observations are for employee numbers, a discrete headcount measure, rather than a more continuous full-time equivalents type measure, so that maintaining a fairly simple approach also giving rise to discrete integers, or half values does not seem unreasonable.

To check the consistency of the plant employment series a process of matching the individual employment series with the relevant plant opening and closure dates was conducted. This analysis suggested that in some cases the employment data and opening or closure data were slightly out of phase. Since the dataset is constructed by an annual postal survey without a strictly fixed identical collection point in each year (varying by a few weeks) the approach to this observation was to follow the phasing of the employment data as recorded in the employment series, rather than attempting to make adjustments to coincide with the reported opening and closure dates. In part, this decision was based on the judgement that the official reported opening date of a plant may not precisely coincide with the recruitment and commencement of production using the full reported initial employment levels at that plant, or similarly for the closing of an operation, where some employees might be retained beyond the official closure of the operating activities.

For a very small number of plants employee numbers information was not available for periods of several years following reported plant opening dates or the employee numbers information extended for several years beyond the recorded plant closure dates. For such observations the data series was retained but employment data extending beyond three years either prior to the reported opening date or after the reported closure date was discarded. Even though there is no strong *a priori* reason to support the use of a three-year period, extending this bound further seemed difficult to justify on the grounds of consistency with the opening and closure date information. However, since this issue primarily affects the most recent years of data, which are not used in the analysis, this would be unlikely to have any significant effect on the conducted econometric analysis.

A further feature of the WRME dataset is the truncation of recorded plant employment, with data only collected systematically for manufacturing plants with more than 10 employees. However, data on some plants with 10 or fewer employees have been recorded. Despite the likelihood they may not be fully representative of the population of "micro" plants, this information is as far as possible (subject to the cleaning processes described above) retained for the empirical analysis, in part because of the generally limited opportunities to conduct analysis of the empirical dynamics of such micro-plants (or firms) in a U.K. regional context using official business microdata.

3.3. Empirical Properties of the Micro-Dataset

In this section some of the broad empirical properties of the cleaned dataset of manufacturing plants in Wales are discussed, providing context to the research presented in following chapters. The application of the data cleaning procedures described above resulted in usable records relating to 6,115 plants, representing 96.0 percent of the original extracted sample. The composition of this dataset in terms of the total number of plants having been recorded in any year of the dataset by industry, the country of ownership and organisation structures are presented in Tables A3.2 and A3.3.

Of particular interest in examining this constructed data is an initial assessment of the broad consistency with the model of Gibrat's Law, and the implied lognormal size distribution of plants. Descriptive statistics on the number of plants, their associated employment and the moments of the distributions (mean, variance, skewness and kurtosis), as well as distribution percentile measures, of the sizes of these plants for each of the annual cross-sections of the samples, covering the period 1966-2003 are presented in Table 3.1. In these tables a skewness index of less than zero indicates skewness towards the left and a value greater than zero skewness to the right. The kurtosis index provides a measure of the peakedness of the distribution, with a kurtosis index above a value of three indicating higher peaks of the centre of the size distribution and thicker tails than the normal distribution (leptokurtosis) and a value of less than three exhibiting platykurtosis with less mass in the tails of the distribution and a flatter central peak than under the normal distribution.

Consistent with the broad hypothesis of Gibrat-type stochastic models of plant growth, across the full time-span of the data, there is evidence of a highly skewed distribution of plant sizes. Specifically, in each year analysed these static size distributions are seen to exhibit considerable and consistent positive skewness, with the tail of the distribution extending towards the right and the mass of the distribution to the left, with skewness index values ranging between a low of 6.98 in 2000 to a high of 16.96 in 1978, with a mean skewness index value across the full time period of 11.68.

For each aggregate cross-section the size data also show positive kurtosis, with a minimum kurtosis index of 72.50, again in 2000, a maximum of 384.06, again also in 1978, and a mean value of 207.51. Closer examination of the kurtosis indices suggests some greater volatility than the skewness index over time, with several distinct phases of stable kurtosis indices interspersed with rapid changes, increasing significantly in 1975, before a sharp decline in 1981. After some stability until around 1993, further volatility in the kurtosis index can be observed. Whilst formal testing of the consistency of such plant size distributions is conducted in the next chapter, clearly this initial overview provides evidence of the highly skewed plant size distributions, consistent with such stochastic processes.

Interestingly, the data also indicates a general decline in plant sizes across the 10th, 25th, 50th, 75th and 90th percentile measures until the mid-1980s before some degree of reversal of this trend. A comparison of the percentile measures for the most recent year with earlier data suggests that the 75th and 90th percentile measures remain notably below that of the initial years of the dataset, although less difference is observable for the other percentile measures relating to smaller plants. In part, these observations have been influenced by the significant decline in the maximum plant sizes across much of the period of the dataset (albeit with some more modest increases in more recent years) primarily through the gradual erosion of employee numbers, rather than the exit, of these largest plants.

In terms of the more general dynamics of the manufacturing sector, the data on the recorded stock of plants shows an increase from 1,432 plants in 1966 to a peak of 2,812 in 1977. Since then there has been a fairly persistent decline reaching 1,713 plants by 2003. The data indicate a rapid expansion in plant numbers between 1974-1975, increasing from 2,245 to 2,680 (or 19.4 percent), in part the result of an unusually large number of recorded plants below the official size threshold. This break in trend continues until a similar sharp fall in plant numbers between 1977-1979, falling from 2,812 to 2,460, a decline of 12.5 percent. A further sharp decline is also observed around the mid-to-late 1980s, with plant numbers falling from 2,545 in 1985 to 2,395 in 1988 before adjusting upwards to 2,509 in 1989.

The aggregate time series of employee numbers also exhibits a general increase from the initial 1966 base period of 292,760 employees (but falling to 286,706 in 1967) to peaks of 327,245 employees in 1970, followed by a modest decline until 1972, before rising again to a peak of 332,353 employees in 1974. A generally modest rate of decline then follows until around 1979, followed by a significant sharp fall in recorded employment reaching a level of 207,137 employees during the early 1980's recession, representing a decline of 37.7 percent between the peak of 1974 and 1983. More recent data from 1983 onwards exhibits some greater stability, with a period of employment recovery from 1986 to around 1989, reaching 227,656 employees, before reverting towards a more general pattern of decline, which according to this data had perhaps slightly accelerated since around the end of the late 1990s.

The mean average size of plants (as measured by employee numbers) has shown an almost continuous decline during the period 1966-1985 (the exception being a slight rise in mean plant size in 1978 and 1979) declining from 204.4 employees to 79.9 employees (a fall of 60.9 percent). However, since 1985 there has been some trend towards a slight increase in mean plant size reaching 99.2 employees in 1998 and remaining broadly at this level until 2003. The data on the standard deviation of the annual plant employee numbers suggests broadly similar trends, with particularly significant declines during the period until the early 1980s. The median plant size showed some variation over the 1966-2003 period, declining from 42 employees in 1966 to 20 employees in 1977, followed by a broadly stable period until 1986 (but with a slight rise in 1979 and 1980). Since the mid-1980s the trend has been towards a larger median plant size reaching 40 employees in 1998, 1999 and 2001. In all years the mean exceeded the median, again indicating a distribution skewed to the right.

Of course, as discussed earlier in this research, the deterministic approach to firm dynamics suggests that there may be industry-specific factors, such as economies of scale, that may influence observed firm or plant size and growth patterns. It is therefore useful at this stage to set out some of the empirical properties of disaggregated industry-level sectors, given that such information on employment and plant changes is likely to provide valuable context in understanding the later provided results.

Examination of the industrial composition of the cleaned dataset across the full

Table 3.1: Manufacturing Plants and Employment: Summary Statistics

	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984
Plants (N): All	1,432	1,490	1,601	1,737	1,857	1,935	2,044	2,150	2,245	2,680	2,726	2,812	2,665	2,460	2,506	2,475	2,516	2,513	2,514
Domestic owned	1,293	1,344	1,443	1,564	1,671	1,743	1,843	1,931	2,010	2,437	2,475	2,561	2,400	2,181	2,213	2,193	2,222	2,213	2,206
Foreign owned	139	146	158	173	186	192	201	219	235	243	251	251	265	279	293	282	294	300	308
Employment†: All	292,760	286,706	296,347	312,348	327,245	316,821	312,771	324,566	332,353	319,483	305,401	308,846	300,151	295,629	276,589	235,576	219,468	207,137	205,173
Domestic owned	221,602	218,815	225,653	236,291	244,414	236,881	231,934	239,831	243,624	236,780	223,829	223,685	213,593	208,525	190,572	158,861	148,036	140,159	138,189
Foreign owned	71,158	67,891	70,694	76,057	82,831	79,940	80,837	84,735	88,729	82,703	81,572	85,161	86,558	87,104	86,017	76,715	71,432	66,978	66,984
Mean	204.44	192.42	185.10	179.82	176.22	163.73	153.02	150.96	148.04	119.21	112.03	109.83	112.63	120.17	110.37	95.18	87.23	82.43	81.61
Standard Deviation	769.65	732.16	708.38	680.32	660.13	613.29	580.79	572.83	556.61	509.49	470.07	471.59	467.24	470.81	376.01	283.86	259.87	237.63	229.01
10th Percentile†	11	11	10	10	10	9	8	8	8	4	4	3	4	6	6	6	5	5	5
25th Percentile†	17	16	16	16	16	15	14	14	14	8	9	8	9	12	12	11	10	10	10
50th Percentile†	42	40	39	38	38	36	33	32	31	22	22	20	23	26	25	23	21	20	21
75th Percentile†	138	133	128	120	117	113	109	104	98	73	70	69	72	81	74	68	62	61	61
90th Percentile†	403	378	346	354	350	324	286	294	294	237	219	214	224	236	234	198	180	180	179
Skewness	13.20	13.67	14.02	13.91	13.87	13.97	14.35	14.24	14.19	16.51	16.57	16.61	16.96	16.39	15.00	9.86	10.12	10.09	9.76
Kurtosis	230.82	248.81	262.20	257.57	256.18	256.38	270.69	268.63	269.84	359.51	368.92	370.41	384.06	360.05	363.55	147.02	155.98	154.85	146.51

Table 3.1: Manufacturing Plants and Employment: Summary Statistics, continued

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Plants (N): All	2,545	2,505	2,470	2,395	2,509	2,553	2,504	2,366	2,264	2,193	2,151	2,123	2,072	2,002	1,982	1,975	1,883	1,805	1,713
Domestic owned	2,221	2,169	2,121	2,050	2,138	2,164	2,108	1,981	1,882	1,813	1,775	1,740	1,693	1,635	1,614	1,592	1,502	1,411	1,339
Foreign owned	324	336	349	345	369	388	394	383	380	378	374	378	375	364	365	368	356	331	315
Employment†: All	203,426	203,062	209,557	215,900	227,656	227,069	218,851	206,962	195,077	194,888	196,520	200,518	200,323	198,604	194,940	189,754	184,994	174,029	165,484
Domestic owned	136,166	136,721	141,066	143,919	149,625	148,500	142,113	132,598	124,021	122,862	122,392	124,569	124,756	122,632	120,592	117,009	113,589	105,538	99,042
Foreign owned	67,260	66,341	68,491	71,981	77,473	78,567	76,725	74,350	71,042	72,012	74,117	75,923	75,540	75,944	74,316	71,574	68,770	65,402	62,400
Mean	79.93	81.06	84.84	90.15	90.74	88.94	87.40	87.47	86.16	88.87	91.36	94.45	96.68	99.20	98.35	96.08	98.24	96.41	96.60
Standard Deviation	224.54	219.69	220.23	226.04	222.77	221.29	216.06	214.97	195.61	200.90	204.85	212.33	209.51	214.21	206.23	193.78	206.76	209.23	202.30
10th Percentile†	5	6	8	8	9	9	10	10	10	10	11	12	12	12	12	11	11	11	11
25th Percentile†	10	11	13	14	14	15	15	14	15	15	16	17	18	19	20	20	20	19	17
50th Percentile†	21	23	26	30	30	30	30	30	30	32	35	36	38	40	40	38	40	38	37
75th Percentile†	63	65	70	79	78	80	78	80	80	85	85	90	94	97	100	100	97	92	96
90th Percentile†	171	174	185	200	194	193	194	197	197	200	205	206	220	220	212	210	224	221	227
Skewness	10.14	10.35	10.06	9.81	9.53	9.85	10.23	10.35	9.85	9.66	9.56	9.27	8.65	8.45	7.53	6.98	9.96	10.73	9.48
Kurtosis	157.42	165.69	157.75	149.96	143.55	150.00	161.41	162.74	162.64	152.79	148.94	133.81	115.97	108.46	82.28	72.50	175.48	200.46	151.62

Notes: † The employee numbers totals and percentile measures have in some cases been rounded-up.
Although in some cases the difference between the sum of the number of domestic and foreign owned plants and the reported annual totals is less than 10 plants, these figures are presented in the table given that this value represents incomplete data on some plants rather than identifying any particular plants or class of plant.
Totals may not sum due to missing ownership classification details.

time period (shown in Table A3.2) suggests a concentration of plants in the Fabricated metal products industry (769 plants, 12.6 percent of plants), Other machinery and equipment (634 plants, 10.4 percent), with other concentrations in Food and beverages (8.8 percent), Furniture (7.5 percent), Rubber and plastic products (7.0 percent) and Non-metallic mineral products (6.9 percent). However, closer inspection of the dynamics of industry-level plant numbers (at the 2-digit industry-level) indicates some notable changes in the composition of the dataset over time, with considerable heterogeneity in the observed trends.

Although industries such as Tobacco products and Coke, petroleum and nuclear fuel have maintained a relatively low number of plants over the 1966-2003 time period, other industries have clearly played an important role in shaping the observed aggregate trends (information on plant numbers, employment and mean plant sizes by industry are presented for selected years in Table A3.4[53]). The sharp rise and then fall of plant numbers in the mid-1970s, followed by further subsequent declines in plant numbers over time, is a pattern broadly characteristic of, for example, the Food and beverages, Publishing and printing, Fabricated metal products, and Furniture industries. Other industries have exhibited some differences in the phasing of changes in plant numbers. For example, the Rubber and plastic products industry showed sustained growth in plant stock until around 1990, with an increase from 42 plants in 1966 to 213 in 1990, before declining to 144 plants in 2003. Similar patterns peaking around the late 1980s to early 1990s (a time of U.K. recession) but with greater continued stability in plant numbers are also exhibited by industries such as Chemical and chemical products increasing from 53 plants in 1966 to 117 in 2003 (but with plant numbers between 120 and 140 for the whole of the period between 1984-2002).

In terms of employment the most noticeable feature is the change within the

---

[53] The complete tables for plant numbers, employment and mean sizes are not presented due to the size of such tables and due to potential disclosure issues.

Basic metals industry, where employee numbers declined from around 92,000 employees in the late 1960s and early 1970s (peaking at 93,434 employees in 1970) to 12,513 employees in 2003, a decline of 80,921 employees. Within this industry, a particularly sharp decline in employment can be observed in the period 1978-1981, falling from over 74,000 employees to just over 40,000 employees. Other industries exhibiting significant employment changes include Other machinery and equipment with employment peaking in 1974 at 28,871 employees before declining to 8,636 employees in 2003, Wearing apparel, peaking at 14,642 employees in 1975 and declining to 1,301 employees in 2003, and Fabricated metal products with employment declining from 22,940 employees in 1974 to 10,450 employees in 2003.

However, whilst some industries such as those mentioned above have reduced total headcount labour inputs within Wales, others such as Radio, television and communications had until recently shown growth in employee numbers from 13,133 in 1966 to 18,505 in 1999, although sharply declining thereafter to 2003. Food and beverages has been another industry providing employment growth in Wales, in this case from 15,429 employees in 1966 to just fewer than 20,500 in 2003, peaking at just over 22,000 in 2001[54].

Although the estimated mean plant size and standard deviations (based on the employee numbers information) suggests that for most industries the mean plant sizes across industries were largest during the late 1960s or early 1970s, it is interesting to note that for four industries (Food and beverages, Wood and wood products, Publishing and printing, and Other transport equipment) the largest mean plant size occurred in either 2002 or 2003. In combination, such differences in plant and employment dynamics at the

---

[54] Analysis of the spatial distribution of plants across Unitary Authorities (U.A.s) also suggests some differences in performance between 1966-2003, with significant employment declines in, for example, Cardiff, Neath Port Talbot, Newport, Rhondda Cynon Taff, Swansea, Carmarthen, Torfaen and Blaenau Gwent, and the Bridgend, Wrexham, Pembrokeshire, Powys, Anglesey, and Ceredigion U.A.s providing a net increase in manufacturing employment. The differences in employment patterns and mean plant sizes appear, in significant part, to be related to the industrial composition within these geographies.

industry-level, clearly suggest a need to investigate the propositions of Gibrat's Law at levels below the aggregate sample.

It is also informative to turn at this point to the examination of disaggregated information on the organisational structure[55] and ownership patterns of the constructed sample of manufacturing plants. Given the potential issues regarding the applicability of Gibrat's Law to different types of business units, the later analysis examines separate samples of plants, a non-representative sample of firms, and plants exhibiting different ownership structures, in particular comparing the properties of foreign owned plants. Some context to these investigations is therefore useful.

Examination of the data (aggregate-level data is presented in Table A3.3) across the period 1966-2003 indicates that the largest category is that of Welsh "owned" single-plants, representing 2,738 plants (44.8 percent of all plants), with U.K. owned firms with a single-plants in Wales providing a further 1,310 plants (21.4 percent). In terms of foreign ownership, the largest representation is from U.S. owned plants (234 plants, 3.8 percent of all plants, and almost two-fifths of all foreign owned plants). The next largest representations are Germany (66 plants), Japan (47 plants), France (33 plants) and Eire (29 plants). In terms of broader geographical classifications, the E.U. and non-E.U. European countries provided a similar number of plants (255 plants) to North America (253 plants). Excluding Wales and the U.K., 36 different countries or states owned at least one plant recorded on WRME.

Analysis of these different ownership and structure classifications indicates that the domestic plant stock (defined to include both Welsh and U.K. owned plants in Wales) increased over the time period 1966-1977 from 1,293 to 2,561 plants, followed by a period of relative stability before declining from 1990 onwards to 1,339 plants in 2003.

---

[55] There is one code without information on the country of origin of plants, representing around 14 percent of total recorded plants. Inspection of the individual records within this coding did not provide any clarity on the commonality of the ownership of these plants.

Of these domestic plants Welsh owned single-plants declined to 880 plants in 2003 from a peak of 1,325 in 1991 (having risen fairly continuously from 552 in 1966), with plants in Wales which are part of U.K. owned firms with single-plants in Wales (termed here as U.K. owned single-plants in Wales) peaking at 703 in 1975 (falling steadily to 265 plants by 2003) and with those plants in Wales being part of U.K. owned firms with multiple Welsh plants (here referred to as U.K. owned with multiple Welsh plants) having peaked at 323 plants in 1980 (149 plants in 2003).

The number of foreign owned plants increased from 139 in 1966 to a peak of 394 by 1991. However, since then the number of foreign owned plants in Wales has tended to decline slightly, with sharper recorded falls in the years 2002 and 2003. Across all years of the dataset the primary geographies of origin for the recorded foreign owned plants are from the European Union (E.U.) and North America (in 2003 providing 34.9 percent and 43.2 percent of foreign owned plants, respectively) with plants owned by Asian companies additionally providing a notable component, growing to 49 plants by 2000. The number of E.U. originating plants declined from 152 in 1991 to 110 in 2003 (a fall of 27.6 percent), with some, albeit less dramatic reductions in North American owned plants. However, despite these changes in ownership composition, over the full time period, there is some trend towards an increased relative representation for foreign owned manufacturing plants rising from 9.7 percent of total plants in 1966 to over 18 percent since 1997.

The aggregate trend in the employee numbers of domestic owned plants in Wales has shown a decline from the peak of 244,414 employees in 1970 to 99,042 employees in 2003 (a decline of almost 60 percent). In addition, despite the increases in foreign owned manufacturing employment in the 1970s (being over 80,000 employees for most of the period between 1970-1980, and again nearly reaching these levels in the late 1980s and early 1990s) foreign owned manufacturing employment in Wales has also been in gradual decline more recently, falling to 62,400 in 2003.

Examination of the data relating to domestic owned plants points to some divergence in performance between Welsh owned single-plants and the two other primary contributing classes (see Table A3.5 for data on plants, employee numbers and mean plant sizes for selected years). Whilst both U.K. owned single-plants in Wales and U.K. owned firms with multiple Welsh plants have exhibited a pattern of a general decline in recorded employment, the Welsh owned single-plant class has provided a somewhat more positive aggregate employment performance. Although exhibiting some cyclical patterns, employment associated with Welsh owned single-plants increased from just under 25,000 employees in 1966 to just under 40,000 employees in the early 1990s, and despite some subsequent downturn, this level of employment was again broadly reached in the late 1990s and early 2000s, though declining in the most recent data. It is noteworthy that the decline in employment in the class of U.K. owned firms with multiple Welsh plants has been particularly dramatic, falling from over 125,000 employees in 1974 to 35,444 employees in 2003 (71.7 percent). Employment in U.K. owned single-plants in Wales fell from 85,542 employees in 1970 to 25,949 in 2003 (69.7 percent).

The data on employment in foreign owned plants shows that the primary contributions were from plants owned by North American companies, followed by E.U. and Asian companies, respectively. Although employment from North American owned plants was broadly stable at around 50,000 employees during the 1970s to the early 1980s, the data points to a subsequent sharp decline in aggregate employee numbers to under 40,000 employees by 1982, with further more gradual declines (though with some increases in the late 1980s) to a level of under 30,000 by 2001. Employment associated with E.U. owned plants increased during the period 1967-1974 to a level just fewer than 30,000 employees from 16,510. However, following a relatively stable period until 1980, employment in E.U. owned plants had fallen to 17,421 by 2003. Employment associated with Asian owned manufacturing plants in Wales shows a slight upwards trend until the mid-1980s, with a more rapid increase in employment until the early 1990s before further slower growth to a peak level of 17,804 employees in 1998. There has since been some decline to just over 12,000 employees recorded in 2003, still some 182.8 percent higher than in 1967.

Examination of the estimated mean average size (by employee numbers) suggests that Welsh owned single-plants have tended to have the smallest mean size. The data also points to a trend for a declining mean plant size of both domestic and foreign owned plants during the period from 1966 to the mid-1980s, in the case of domestic owned plants falling from 171.4 employees in 1966 to 61.3 employees in 1985 (a fall of 64.2 percent) and for foreign owned plants from 511.9 in 1966 to 196.2 in 1987 (a fall of 61.7 percent) before both stabilising at a mean plant size of around 70 employees and 200 employees, respectively. The similarity in the percentage decline of mean plant sizes is an interesting feature, perhaps suggestive of wider factors in manufacturing technologies and production changes, though this is a question requiring more detailed research not specifically addressed within this study.

Within domestic owned plants there are some notable divergences in the trends in mean plant size, declining for both the classification of U.K. owned firms with multiple Welsh plants and (but less dramatically) for the class of U.K. owned single-plants in Wales. For both single-plants in Wales and firms with multiple Welsh plants the mean plant sizes remained broadly stable between 1966 and 2003. The broad pattern of declining mean plant sizes is also generally a feature of most foreign owned classes, being particularly significant for Asian owned plants, falling from around 1,700 employees in the late 1960s and early 1970s to around 800 employees in 1974 and then declining more gradually to around 315 employees in 2003. However, it is important to recognise that these mean averages mask a significant amount of heterogeneity in plant sizes within these groups.

A similar pattern of falling standard deviations of plant size over time also emerges. However, whilst the standard deviations for the broad classes of domestic and foreign owned plants have been declining, there are significant differences across the ownership and structure classes, with the standard deviation of plant sizes falling significantly, for example, in the class of U.K. owned firms with multiple Welsh plants, Asian owned plants, and North American owned plants. Examination of 2-digit industry-level plant size standard deviations suggests that for the majority of, but by no means all, industries the standard deviations of plant size tended to decline fairly consistently across the time period. However, this decline appears particularly significant in the case of the Basic metals industry. Such observations are clearly relevant to the empirical properties of plant dynamics postulated by Gibrat's Law.

As noted earlier, in addition to the growth or decline of employment within incumbent plants the observed size distribution is also likely to be influenced by the entry of new plants and exit of existing plants. The empirical properties of entry and exit are therefore described in this section. From the cleaned WRME dataset information on opening dates was only available for 4,662 plants (of which 739 plants were recorded as having opened prior to 1966) with records for 3,252 plant closing dates. To address these data gaps the available information was supplemented by additional estimates using the observed beginning and final years of employment data on each individual plant time series as proxies for opening and closing dates. Plants without a closure date and with employment information in 2003 were assumed to remain in operation. Where no such information was entered for a plant it was assumed that the plant has closed within 2003. Although it may be that for some plants this simply represented a single point of missing data in 2003 rather than closure, no other appropriate treatment is obvious in the absence of further information.

Examination of the constructed plant entry data, presented in Table 3.2, suggests some volatility, peaking at a high of 492 newly recorded plants (a year with an unusually high number of recorded plants with employee numbers below the more than 10 employees threshold) in 1975 and reaching a low of just 14 recorded new plants in 2003. Although there is no documentation to support this, the exceptionally high number of new plant entries in 1975 seems highly likely to have been a consequence of the data collection process, perhaps attempting to pick up a more complete record of manufacturing plants and departing from the stated threshold. Since other official datasets are also generally incomplete for such smaller plants it is not possible to provide a direct comparison of this data with other sources to test the reliability of the coverage of these plants below the official size threshold.

However, whilst this unusual peak may be due to administrative survey procedures, it does suggest the likelihood that in many other years the plants that are recorded as being under the size threshold are likely to be incomplete, and therefore that some caution should be exercised in interpreting any such analysis (some further analysis is conducted within this research to examine the effects of the inclusion of such observations below the official plant size threshold on the plant size-growth relationship). Whilst the data is likely to be reliable (given the survey recording procedures), it cannot be concluded that this element of the data is necessarily representative of such micro-plants. In the case of the latter figure relating to plant exit in 2003, this also seems likely to reflect specific data collection issues within the final year. However, the plant entry figure has been consistently below 100 since 1991, only having previously been below this level in one other year, 1967. Excluding the year 1975, the entry rate (measured as a percentage of plant stock) was generally stable throughout the 1970s to the mid-1980s at around 6 - 8 percent. However, from around 1987 onwards a

Table 3.2: Plant Entry and Exit

	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984
Entry (Plants)	-	90	123	160	133	115	143	170	144	492	194	212	135	156	153	149	176	155	185
Entry Rate (%)	-	6.3	8.3	10.0	7.7	6.2	7.4	8.3	6.7	21.9	7.2	7.8	4.8	5.9	6.2	5.9	7.1	6.2	7.4
Exit (Plants)	-	30	13	30	32	47	50	48	65	150	120	266	371	103	154	157	148	164	178
Exit Rate (%)	-	2.1	0.9	1.9	1.8	2.5	2.6	2.3	3.0	6.7	4.5	9.8	13.2	3.9	6.3	6.3	6.0	6.5	7.1
Exits of which (plants):																			
Transfer	-	14	..	..	..	..	..	..	..	..	..	..	..	10	..	11	..	11	..
Branch plant closure	-	..	..	..	14	..	23	21	..	40	32	44	44	50	81	75	69	58	48
ENM closure	-	..	..	..	..	20	..	..	31	34	34	48	44	31	51	52	52	65	67
Net Change (Plants)	-	**58**	**111**	**136**	**120**	**78**	**109**	**106**	**95**	**435**	**46**	**86**	**-147**	**-205**	**46**	**-31**	**41**	**-3**	**1**

Table 3.2: Plant Entry and Exit, continued

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Entry (Plants)	207	182	152	140	118	130	65	87	60	58	50	39	43	17	30	65	50	47	14
Entry Rate (%)	8.2	7.2	6.1	5.7	4.9	5.2	2.5	3.5	2.5	2.6	2.3	1.8	2.0	0.8	1.5	3.3	2.5	2.5	0.8
Exit (Plants)	257	182	110	115	82	129	190	158	116	127	82	88	57	80	49	181	113	153	45
Exit Rate (%)	10.2	7.2	4.4	4.7	3.4	5.1	7.4	6.3	4.9	5.6	3.7	4.1	2.7	3.9	2.4	9.1	5.7	8.1	2.5
Exits of which (plants):																			
Transfer	..	..	..	..	..	12	..	10	..	..	..	..	..	..	..	..	..	..	..
Branch plant closure	63	46	32	49	27	53	79	47	41	58	26	27	23	37	34	141	56	126	..
ENM closure	74	59	46	48	35	34	97	86	56	44	21	27	18	31	..	19	21	..	..
Net Change (Plants)	31	-40	-35	-75	114	44	-49	-138	-102	-71	-42	-28	-51	-70	-20	-7	-92	-78	-92

Notes: 1966 values not reported due to the absence of end of 1965 year employment data.
Entry and exit rates calculated as a percentage of the stock recorded at the start of each year.
.. Not reported given small sample < 10 plants, or values have been removed to avoid disclosure.
Some exit forms are unknown so that columns may therefore not sum.
Net change totals may not directly correspond to the sum of reported annual plant entries and exits due to differences in the timing of the recording of plant openings and closures and initial or final employment information.

notable decline in entry rates is observed falling from 6.1 percent to around 1.5 - 3.5 percent, perhaps reflective of changing technologies and market conditions within some parts of the manufacturing sector in Wales.

Plant exits showed some increase from the late 1960s to a peak of 371 plants exiting in 1978, since when there has been some volatility with further peaks in 1985 at 257 plant exits, 190 plant exits in 1991 and 181 plant exits in 2000. Exit rates also exhibit volatility ranging from a low of 0.9 percent in 1968 to highs of 13.2 percent in 1978 and 10.2 percent in 1985. The data points to greater volatility in plant exit rates than entry rates, which may again to some extent reflect the structural changes within the manufacturing sector experienced over this time period. From 1966 until 1977 the recorded annual plant exits were lower than plant entries, providing a positive net effect on the stock of plants. Between the years 1978-1990 the picture was more mixed in terms of the net contribution to plant stock with particularly significant net plant losses in 1977 and 1978. Post-1991 in every year the recorded stock of manufacturing plants has been in decline.

Closer inspection of plant entry and exit by size-classes (the methodology and construction of these particular size-classes in explained in Chapter 5, with the summary data shown in Table A3.6) indicates that throughout the time period 1967-2003 there were no new manufacturing plant entries into any size-class with more than 1,024 employees[56]. Plant entry was primarily into the smaller size-classes, including into the size-classes below the WRME recording threshold, especially in the period 1975-1977. The constructed data on plant exits by size-class indicates a very limited exit of larger plants with most exits (in absolute terms) occurring within the smaller size-classes, particularly around or just above the official recording size threshold.

---

[56] The plant entry and exit data discussed here corresponds to the year 1967 onwards given issues relating to the appropriate identification of plant entry timings in 1966. Additional tabulations presented in Table A3.7 also include the aggregation of all plant entries recorded on WRME, including those prior to 1966.

Examination of industry-level entry and exits across the full time period (post-1966) of the dataset (Table A3.7) shows a relative concentration in Fabricated metal products, providing 588 new entrants, Other machinery and equipment, 484 new entrants, Furniture, 383 new entrants, and Rubber and plastic products, 381 new entrants. Other industries contributing significant numbers of new entrants include Food and beverages and Non-metallic mineral products. Similarly, plant exits were highest in Fabricated metal products, where 569 plants exited, and Other machinery and equipment where 470 plants exited, with other significant contributions from Food and beverages, Non-metallic mineral products, Furniture, Rubber and plastic products, and Wearing apparel. Such information clearly suggests some consistency of those industries with the highest levels of entry and exit. Constructed data on entry and exit rates by industry in many cases provided very small samples, precluding any significant analysis from being presented.

Bringing together the data on plant entry and exit and the changes in employment within incumbent plants, gross and net employment flows were constructed for the time period 1967-2003 (the year 1966 not being examined due to the absence on information on the initial stock of plants and employment for the beginning of that year). Given the issue in some cases of precisely matching opening and closure dates with initial and final employment records, a full components of change decomposition into the effects of entry, exit, expansion and contractions is not conducted, with the discussion here considering gross and net flows - gross gains in employment arising from the contribution of entry and plant expansions, and gross losses from contractions and plant exits, the net flow being the sum of these effects[57].

---

[57] It could also be considered that attributing the effects of exit to only the final year of data may significantly understate the employment effects of exit compared to that of contractions where the closing period is extended over a longer time period. For some empirical evidence on firm size changes prior to exit see Troske (1992) who examines the time series patterns of firm growth in selected production and services industries.

Table 3.3: Gross and Net Employment Flows

	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978
Gross Gainers (Plants)	-	574	799	928	967	861	899	1,094	1,114	1,264	1,250	1,261	1,133
Gross Losers (Plants)	-	527	433	403	542	760	799	684	778	1,018	1,124	1,021	1,110
Gross Gains (Employees)	-	10,778	20,680	26,756	27,960	16,093	21,039	26,436	22,411	20,038	17,105	20,957	19,115
Gross Losses (Employees)	-	16,832	11,039	10,755	13,062	26,517	25,089	14,641	14,624	32,908	31,187	17,512	27,809
Net Change	-	-6,054	9,641	16,001	14,897	-10,424	-4,050	11,795	7,787	-12,870	-14,082	3,445	-8,695

	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991
Gross Gainers (Plants)	972	973	813	1,023	1,048	1,167	1,206	1,185	1,248	920	1,136	958	809
Gross Losers (Plants)	904	912	1,019	988	910	844	912	828	665	499	599	645	688
Gross Gains (Employees)	16,703	14,464	9,406	11,539	13,881	16,078	14,259	16,953	19,187	16,414	24,312	16,487	12,589
Gross Losses (Employees)	21,225	33,503	50,419	27,647	26,212	18,043	16,006	17,316	12,692	10,071	12,556	17,075	20,807
Net Change	-4,522	-19,040	-41,013	-16,108	-12,331	-1,964	-1,747	-364	6,495	6,343	11,756	-587	-8,218

Table 3.3: Gross and Net Employment Flows, continued

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Gross Gainers (Plants)	882	765	894	905	796	591	819	464	618	639	353	494
Gross Losers (Plants)	931	766	605	477	470	484	711	521	725	729	410	770
Gross Gains (Employees)	13,490	13,069	17,325	13,524	14,267	10,733	16,751	9,833	11,957	24,148	11,988	11,461
Gross Losses (Employees)	25,379	24,954	17,513	11,892	10,269	10,928	18,470	13,498	17,143	28,908	22,953	20,006
Net Change	-11,889	-11,885	-189	1,632	3,998	-195	-1,719	-3,664	-5,186	-4,760	-10,965	-8,545

Notes: 1966 values not reported due to the absence of end of 1965 year employment data.
Columns may not sum due to rounding.

The data, set out in Table 3.3, indicates significant annual gross gains and losses in employee numbers with the gross employment flows being considerably larger than the net flows recorded for each year, both gross flows generally being above 10,000 employees (except for gross gains in 1981 and 1999). Over the whole time period the average employment gains were 16,762 per annum with average losses of 20,202, a net average loss of 3,440 jobs each year. However, it is not the case there has been an even and continuous process of net employment erosion within the manufacturing sector in Wales. Interestingly, however, in almost every year of the data (except for 1981, 1992 and each of the years after 1999) the number of plants experiencing increased employee numbers exceeded the number of plants exhibiting reduced employee numbers.

Comparison of the gross gains and losses data point to short periods of intensive employment shakeout in 1971, 1975-1976, 1980-1983, in the early 1990s and again in the early 2000s. The net employment effects were over 100,000 job losses in the period 1978-1983, a further net loss of almost 32,000 jobs between 1991-1993 and over 33,000 jobs lost in the period 1999-2003. The fall in gross gains to a low of 9,406 employees in combination with the increase in gross losses of over 50,000 employees in 1981, had a particularly significant effect, reducing overall recorded manufacturing employment by over 41,000 employees in that single year.

3.4. Microdata Panel Selection Methods and Samples

An important context to the empirical properties of the above described aggregate statistics are the background U.K. and international macroeconomic conditions, not least since this should provide some indication of the overall market demand for (manufactured) outputs (the U.K. data being more likely to be reflective of these conditions than regional output or Gross Value Added (GVA) data). In this study the approach followed in empirically testing Gibrat's Law, as employed in a number of empirical industrial economics studies using business microdata, is to examine the changing patterns of industrial dynamics across time periods broadly consistent with the U.K. economic cycle (see, for example, Oulton, 1998).

A number of studies have attempted to date the U.K. economic cycle, amongst these being HM Treasury (2005), Artis (2002, 2003), and Chadha et al. (2000a, 2000b). Analysis of alternative economic cycle dating approaches (including on-trend points using cyclical indicators and statistical filtering techniques, see for example, Hodrick and Prescott, 1997; and Baxter and King, 1999) by HM Treasury (2005) suggests complete U.K. economic cycles covering 1972Q4 to 1978Q1, 1978Q1 to 1986Q2, and 1986Q2 to 1997H1, with more recent evidence from revised non-oil GVA data, information on compensation of employees as a percentage of nominal GVA and experimental data on market sector GVA suggesting less support for any short economic cycle between 1997H1 and the middle of 1999 (HM Treasury, 2005).

Whilst some differences emerge in the estimated timing of economic cycles across studies due to the methods employed and slight variations in definitions, there appears to be a reasonable consensus on the broad duration and approximate start and end-points (HM Treasury, 2005). For example, analysis of HM Treasury estimates of U.K. annual output gaps and those produced by the IMF, OECD and EC (presented in HM Treasury, 2005) suggest broadly comparable properties, although the IMF data indicates some more rapid elimination of excess capacity around the mid-1980s, suggesting this occurred two or three years earlier. Artis (2002, 2003), discussed in HM Treasury (2005), making use of statistical filtering approaches, despite finding some greater frequency of trend passing and timing differences to the HM Treasury on-trend points estimates, found patterns of estimated economic cycles exhibiting the same broad major movements, supporting the view of three full economic cycles between the early-

1970s and mid-to-late-1990s.

Given the annual nature of the WRME data, the slight differences in precisely dating the U.K. economic cycle across studies and that there may have been some variation in the phasing (and observed effects) of aggregate economic performance in Wales, the periods used here attempt to provide an approximation to this information. In the case of 1978Q1 this seems most reasonably to correspond with the 1977 annual WRME data since the data was collected around the end of 1977 and beginning of 1978. Artis (2003), for example, also tends to suggest movement through trend in 1976-1977, so the choice of 1977 seems consistent with this broad evidence. Judgement, using the observed evidence, has also been applied in the selection of the other start and end points of the U.K. economic cycle, consistent with the above reported economic evidence, providing the periods of 1972-1977, 1977-1986 and 1986-1997 as the basis for the empirical research in this study. In addition, use is also made of shorter time periods in order to draw out greater detail on industrial dynamics and plant size distributions.

Further to the periods covering one economic cycle a long period covering all three cycles between 1972-1997 is also considered to permit the analysis of the long-run dynamics of manufacturing plants, an area for which only limited research evidence is available. Whilst such a constructed panel has the advantage of significantly extending the time dimension of analysis it does however limit the cross-section dimension (number of plants). Such concerns should though be alleviated through the use of the samples covering the periods of shorter time dimensions described above.

These economic cycle periods closely correspond to particular phases of aggregate employment movements within this dataset. In terms of plant employment, the first period of 1972-1977 exhibits fairly stable employment with over 300,000 employees recorded for the whole period. This is followed by a period of significant decline in employment to just over 203,000 employees in 1986, with decline in every year during this cycle. The final period 1986-1997 exhibits some employment stability, with the final years showing plant employment only about 3,000 lower, but with some upturn peaking at over 227,000 employees in 1989 and 1990 and a low point within this period of just under 195,000 employees in 1994.

These cycle periods also show consistency with changes in plant numbers, increasing during 1972-1977 from 2,044 to 2,812. Between 1977-1986 the level of recorded plants showed some decline from 2,812 to 2,505. The final period, 1986-1997 is generally a period of some stability until around 1991 but then with declining plant numbers, reaching 2,072 in 1997.

The mean average plant size (measured by employee numbers) declined from 153.0 to 109.8 employees in the period 1972-1977, declining further during 1977-1986 (reaching the lowest level in 1985 at 79.9 employees) before trending upwards during the 1986-1997 period, calculated as 96.7 employees in 1997. The median and mode sizes show similar trends, but with the mode being at it's lowest during the period 1975-1978, in low single-digit figures (again due to the recording of significantly greater numbers of very small plants).

Using this assessment of U.K. economic cycles, samples were constructed for each of these time periods using employee numbers as the measure of plant size. Given such time periods, it is of note that the specific issue of the unusual peak in new plant entries below the official size threshold in 1975 is not a factor in any of these analyses. However, the existence of the presence of (partial records of) plants below this threshold suggests a continuing issue regarding truncation more generally, which is indirectly examined through a range of procedures in the next chapters.

Consistent with the three versions of Gibrat's Law identified by Mansfield (1962) (Acs, Audretsch and Thurik, 2002, have also suggested a distinction between static and dynamics versions), separate samples for each of these periods were developed: first, for all plants in the initial sample; second, for surviving plants; and third, for plants above the minimum efficient scale (MES) of production. Samples consistent with the first version of all plants provided sample sizes of 2,044 plants for the two periods commencing in 1972, 2,812 plants for the 1977 initial year, and 2,505 plants for the 1986 commencing sample. Corresponding samples of 1,754 and 693 surviving plants were developed for the 1972-1977 and the 1972-1997 periods, respectively, with 1,444 and 1,417 plants for the 1977 and 1986 commencing time periods, relating to the second version of Gibrat's Law.

Although it can be argued that measures of the MES may be most appropriate at the level of the firm, a reasoned position can also be taken supporting the plant, particularly in the case of products being manufactured without significant component inputs from other parts of the firm, and with production located at single-sites. Accepting such arguments, and given the nature of the data, a plant-related MES measure was generated following Audretsch et al. (2002) with the MES defined as the minimum size of the largest plants representing half of the 2-digit SIC industry-level employee numbers[58]. Such an estimate is somewhat crude, reflecting only labour inputs. It is also arguable that the U.K. industry-level MES (or for some industries an international MES) would have been a more appropriate benchmark, but there is no specific information provided by WRME that would enable such an estimate to be produced. Although it is noted that there may be a further potential limitation from using such a plant-level estimate if a number of plants below the estimated measure of the MES may in fact be components of multi-plant firms above the MES, without further information no such adjustments can be made.

Application of the measure of Audretsch et al. (2002) estimated for each 2-digit

---

[58] This measure is slightly different to Comanor and Wilson (1967) who defined the MES as the mean size of the largest plants associated with half of the industry-level employment (Audretsch, 1991, p.446). However, given the small sample sizes of plants exceeding the MES, without a strong theoretical justification between these definitions the measure of Audretsch et al. (2002) is used. This provides slightly larger samples for consideration.

industry (given the likely different cost structures and economies of scale across different industrial activities) provided initial samples of 191 plants in 1972, 211 plants in 1977 and 243 plants in 1986. Although it could be reasonably argued that the application of the MES estimation procedures at the 2-digit industry-level may still be too broad to appropriately reflect heterogeneity within groups, for this database further disaggregations gave rise to very small samples which were considered potentially more likely to result in misleading MES estimates. Examination of the data suggests some considerable variation in the calculated MES across industries and indeed some significant changes over time. In addition, there is evidence of larger MES (and smaller numbers of plants above the estimated MES) in those industries that might be expected to be particularly capital-intensive and with greater opportunities for economies of scale providing some broad confidence that these estimates are at least consistent with prior expectations.

Comparisons between the samples of all plants and surviving plants for each of the four periods 1972-1977, 1972-1997, 1977-1986 and 1986-1997 are given throughout the following chapters, and so are not replicated in detail here. However, several broad points are of note. First, the data point to both a higher mean size and standard deviation of plant sizes for the surviving plant samples. Secondly, although the distributions are observed to be highly skewed and exhibit significant leptokurtosis in each time period, both for all plants and surviving plants, there is a consistent indication of lower skewness and kurtosis amongst the surviving plants. Third, in respect of the composition of the samples (the details are presented in the Tables in Chapters 4 and 5) the samples of all plants and surviving plants appear broadly similar across the initial years, including (although with a modest number of exceptions) in terms of industry and ownership structures. Plants are however observed, on average, to be slightly older for the surviving plants in the initial year of analysis than for all plants. For all four periods, the surviving plant samples generally exhibit slightly higher percentages of foreign owned plants.

3.5. Size Measures and the Unit of Observation

A range of potential measures of the size of plants or firms have been examined in the empirical literature including employment, assets, value-added and sales-based measures. Curry and George (1983, p.213) suggest that the technically appropriate measure of a firm's activities is generally net outputs or value added (except when examining firms in the same market where vertical integration will affect this variable). Measures based on sales will have the effect of tending to inflate the value of firms (or plants) conducting distribution activities (Curry and George, 1983, p.213), with assets based measures subject to a range of measurement issues, including, for example, the determination of the appropriate deflator and accounting conventions (Curry and George, 1983, p.213) (for example, in relation to the treatment of stocks and the depreciation of fixed assets). Such procedures and conventions may give rise to the inappropriate smoothing of financial measures of firm or plant size, removing some of the actual variation in growth rates. As Geroski et al. (2003, p.51) observe, such potential smoothing could gives rise to a bias against the null hypothesis in this study of a random walk, and against the first proposition of Gibrat's Law. Additionally, many smaller firms may not produce formally audited accounts, potentially giving rise to further selection issues with the systematic non-inclusion of smaller firms (though other measures may also be similarly constrained).

An alternative measure frequently employed in studies of Gibrat's Law is that of employment. Although the use of employment measures may raise some issues when investigating plant growth across industries with differing capital and labour input intensities (Curry and George, 1983, p.213), employment based plant size measures have the significant benefit of providing a measure which in all cases is non-negative; negative size measures, for example, being feasible in the case of profit-type measures. Reflecting the point above, and importantly within the context of this study, employment measures also have the distinct advantage of representing a freely variable input (in the terminology of Geroski et al, 2003, p.51). Further, the use of an employment measure may also be seen as being of primary policy relevance in this regional context, relating to the relative employment generating propensities of plants[59]. Although financial measures may have particular merit at the UK level in terms of assessing market dynamics and concentration, such concentration issues are not particularly meaningful, in general, at such geographies.

A case for the use and analysis of employment based measures can clearly be made, and indeed the measure of plant size available from the WRME database is employment based, a headcount employee number measure. It is fair to reflect, however, that, despite the advantages set out above this specific measure does have some limitations compared to continuous, full-time equivalents based measures, particularly in respect of appropriately representing the sizes (and therefore growth) of smaller firms (Hart and Oulton, 1996, p.1243), or if there are systematic differences in the utilisation (and changes in utilisation) of part-time and full-time employees over time across plants of different true sizes. However, whilst a full-time equivalents measure may be preferable in terms of reflecting both plant size across the full size distribution and employment change (growth), the headcount measure of employment available here would still seem to provide a useful measure in relation to proportionate employment generation. As such this measure forms the main focus of the subsequent empirical research.

In using this single measure of plant size, there is one point that should be recognised: that there may be different observed effects through the use of alternative size measures. Whilst, as mentioned earlier, many of the alternative measures of firm (or

---

[59] See for example, Birch (1979), Davis, Haltiwanger and Schuh (1993, 1996) and Hart and Oulton (1998a) for investigations of the employment generating propensities of firms of different sizes using a similar framework.

plant) size tend to be highly correlated (Blair, 1972; Hart and Oulton, 1998b), in addition to the possibility of non-linear relationships affecting the inter-changeability of size measures (Smyth et al., 1975, p.112), there is also the potential for different estimated effects when examining these measures over time. For example, the introduction of new technologies may permit firms to produce a given level of output using less factor inputs, for example, less labour (or capital), or both (effectively a standard isoquant line shifting downwards to a cost minimising point on the (lower) isocost line, and as such reflecting the lower required total costs of production; Katz and Rosen, 1994, p.287).

Given the greater opportunity to adjust inputs and organisational structures and practices to efficiently use new technologies over longer time periods, such an issue may be particularly relevant over the longer time-span element of analysis considered in this research. The implication is though that such technological effects may give rise to different empirical results in respect of Gibrat's Law when using employment and non-employment-based measures of size[60]. Unfortunately, given the general lack of previously analysed datasets with the considerable time dimension available here there is little empirical evidence to establish the magnitude, and consequences, of this possible effect. However, it can be reasonably argued that such differences across measures arising from technological effects may be much more modest in the shorter time periods which form a large part of this study, and therefore that the use of an employment-based measure may be less affected by such issues during these periods.

A further measurement issue, concerning the appropriate unit of analysis: either firms or plants, has received relatively little discussion within the existing literature, with previous studies of Gibrat's Law seemingly based on the nature of the available data.

---

[60] It may also be plausible that new technologies are not generally taken up by all firms (at least at the same time and rate) so that firms closer to the technological frontier (and hence more able to take advantage of such technologies) may effectively utilise new technologies more quickly and intensively, improving productivity, so that in a given period they may be observed to decline in size by the employment measure relative to other plants; which would not be the case if using, for example, a financial measure of size.

Many deterministic theories of business dynamics have been conceptualised around the firm as the unit of interest given the decision making responsibilities and capabilities at this level. However, the model of Gibrat (1931) is less explicit about the appropriate unit of observation.

Whilst it is possible to interpret many feasible factors with stochastic elements as potentially influencing the firm, these factors will clearly also influence the plant, perhaps also arguably with the possibility of additional decision making uncertainty and mistakes by plant management. Further, the erratic outcomes of product innovation and development and the (uncertain) effects of recruitment and of advertising (Scherer, 1970, p.148) can also be suitably interpreted as relating to a particular product, which may plausibly be produced in single-site plants. Further, some of the factors which may contribute to the stochastic appearance of business dynamics (discussed in the Chapter 1) such as industrial strikes (Clarke, 1985, p.33) and natural phenomena and disasters (Hart and Oulton, 1996, p.1244) seem at least, if not more plausibly, to be likely to apply to single-site plants or firms, as to multi-site (particularly multi-national) operations.

It therefore seems reasonable arrive at the view that Gibrat's Law seems at least as relevant at the plant-level as it is for firms. And indeed, in his original analysis Gibrat (1931) made use of establishment (plant) level data to test his hypothesis. As with the data available to Gibrat, the WRME data is provided at the level of the plant and for the manufacturing sector. Using this data source is also possible to identify single-plant firms, which are separately tested in the econometric analysis, providing a link to the previous firm-level analyses of Gibrat's Law. However, since it is not possible to create the data for the all components of firms in the case of UK-multi-sited or multinational firms, such a sample of firms cannot be considered to be fully representative of all firms.

Given the arguments reviewed in this section, and of course data availability, the subsequent analysis presented in the later chapters makes use of the employment-based measure of size available in WRME, with a wide range of short time periods covering annual time-spans, periods of macroeconomic expansion and contraction being examined (as well as full economic cycle time periods), hopefully alleviating some concerns regarding the differential effects of technological changes on alternative size measures, and conducted at the level of the plant.

3.6. Conclusions

This chapter has provided an introduction to the Welsh Register of Manufacturing Employment (WRME) database. The first sections described the historical development of the dataset, data coverage, collection processes and the available variables. The procedures involved in the construction of the full microdata panel and in cleaning the WRME data have been documented.

In addition, this chapter has provided an aggregate overview of some of the main trends in the data, including in relation to aggregate plant stock and employment and also disaggregated samples covering both industrial composition and ownership structures. This chapter has also presented an initial overview of the properties of the aggregate cross-section distributions of manufacturing plant sizes, with the statistical data on the mean and median, the former exceeding the latter in all years, percentile measures and constructed skewness and kurtosis indices pointing strongly to the highly skewed nature of plant sizes across each of the examined aggregate cross-sections spanning a period of over three decades, an observation at least initially consistent with a Gibrat-type stochastic process.

Consideration was also given to the employment flows within incumbent plants, and plant entry and exit, which would, in combination, influence both broader patterns of industrial dynamics and more specifically the evolution of plant size distributions. The data suggest that although there was some growth in the stock of plants over the early part of the dataset until the mid-1980s, since then the recorded stock of plants has shown a fairly persistent declining trend. Similarly, following a peak in recorded aggregate employment in 1974 for the whole time series, thereafter the trend has been generally downward, pre-dating the reported most significant downward trend in the stock of recorded plants by over a decade. Some notable differences in performance across industries and ownership classes over the time-span of this dataset were also identified, with the documentation of such features providing useful context for the plant size and growth dynamics examined in the following chapters.

The plant entry data suggested some volatility over time, peaking in 1975, before declining to less than 100 new plants in every year since 1991. Plant exits were lower than entries until 1977, providing a positive net effect on plant stock. However, after a mixed picture during the period to 1990 the stock of plants has fallen in each subsequent year. In conjunction with plant expansions and contractions the effect of these plant entries and exits has provided annual gross in-flows and out-flows of employment (considerably larger than the net flows) generally both being above 10,000 employees, though exhibiting some degree of volatility particularly in gross employment losses flow between the mid-1970s and 1980s with a number of short periods of particularly intensive shedding of manufacturing labour. The mean plant size fell consistently throughout the period 1966-1985, although followed by some reversion of this trend. The tendency was also for a reduction in the standard deviation of plant sizes, partly reflecting a shift away from very large manufacturing operations.

The rationale for the selection of time periods employed in the analysis of Gibrat's Law and the broad properties of the constructed samples were discussed. In conjunction with supplementary information on the U.K. economic cycle, three primary time periods for analysis were determined 1972-1977, 1977-1986 and 1986-1997 with an additional longer period covering the full time-span of these economic cycles. Such periods correspond reasonably closely to observed aggregate trends within the WRME data. The selection and analysis of this range of periods (in addition to a range of sub-periods examined in the later analysis) permits maximisation of both the cross-section and the time-span dimensions of the dataset for different samples, whilst providing a reasonable economic basis for the selection of such panels, and in the case of the long panel the opportunity to examine the industrial dynamics of a particular group of long-surviving plants.

This chapter also considered the issues around the choice of measures of plant or firm size, and the appropriate unit of analysis, either firms or plants. Given the arguments put forward and data availability, an employee numbers measure of size provides the specific measure used in this research. Whilst there are some limitations to this measure of size, as there are with all measures, use of this employment-base measure also provides a direct link to broader interests in the relative employment generating propensities of businesses. Arguments were also presenting supporting the plant as an appropriate unit of analysis.

4

PLANT SIZE AND GROWTH RATE DISTRIBUTIONS: A NON-PARAMETRIC ANALYSIS

4.1. Introduction

This chapter provides a new empirical assessment of the validity of the lognormal distribution as representative of the unconditional size distribution of manufacturing plants. The analysis examines the moments of the plant size distributions, presents formal statistical tests of the lognormal distribution and makes use of non-parametric kernel density estimation procedures to provide a detailed assessment of the statistical form of the actual observed plant size distributions. These analytical approaches are also employed to examine the statistical form of plant growth rate distributions.

Section 4.2 provides an overview of the statistical tests and non-parametric estimators used to examine the plant size distributions. Section 4.3 then sets out an analysis testing the validity of the lognormal distribution for the aggregate static size distributions for each of the cross-sections of plants over the period 1966-2003. Section 4.4 develops this analysis to consider the effects of plant ageing on the observed size distributions, before further extending this framework to consider the role of both age and selection effects, tracking the evolution of the distributions of plant sizes across the previously selected economic cycle time periods, and additionally over a time duration spanning these periods, a time-span significantly longer than examined in previous studies.

Given the potential industry-specific conditions found to influence the form of firm (or plant) size distributions (for example, in Machado and Mata, 2000) the analysis presented in

Section 4.5 also sets out to assess the size distributions of manufacturing plants across 2-digit-level industries. Consistent with the literature on the post-entry performance of firms, consideration is also given to the evolution of the size distributions of new entrant plants for nine cohorts covering entry years relating to different phases of U.K. economic cycles.

Section 4.6 provides an analysis of the growth rate distributions for aggregate samples and for disaggregated samples (Section 4.7), in particular assessing the suitability of the Gaussian (normal) distribution as a reasonable approximation to observed plant growth rate distributions, and considers the heterogeneity of such distributions at disaggregated industry-levels. A summary of the empirical findings and conclusions are presented in Section 4.8.

4.2. Statistical Tests and Estimators

A wide range of statistical testing procedures have been employed to examine the properties of observed firm and plant size distributions, with such statistical tests generally focussing on testing for outliers or deviations from particular hypothesised statistical distributions. More recently applied non-parametric kernel density estimation techniques and quantile regression[61] approaches have also provided greater detail on the empirical properties of firm size distributions.

Consistent with several previous studies, including, amongst others, Lotti and Santarelli (2001a, 2001b), the analysis of plant size and growth rate distributions here makes use of summary statistics of the moments of the observed size distributions, including the mean, standard deviation, and particularly skewness and kurtosis indices. Such skewness and kurtosis indices, the third and fourth moments of the distribution, are defined generally as:

---

[61] Although the Box-Cox quantile regression methodology employed by Machado and Mata (2000) to examine firm size distributions conditional on industry covariates is intuitively appealing, given the limited industry variables available in WRME (with these variables being primarily fixed plant or industry characteristics) such an approach is not followed here.

$$\text{Skewness index} = \sqrt{\beta_1} = \frac{E(X - \mu)^3}{[E(X - \mu)^2]^{3/2}} = \frac{E(X - \mu)^3}{\sigma^3} \tag{4.1}$$

$$\text{Kurtosis index} = \beta_2 = \frac{E(X - \mu)^4}{[E(X - \mu)^2]^2} = \frac{E(X - \mu)^4}{\sigma^4} \tag{4.2}$$

where μ and σ are the mean and standard deviation of plant sizes, respectively, and E denotes the expected value operator (D'Agostino et al., 1990, p.317). As D'Agostino et al. (1990, p.317) note, under normality the expected values of skewness and kurtosis are 0 and $3(n - 1)/(n + 1)$, respectively (where n is the number of observations), which for relatively moderate sample sizes provides the (approximate) benchmark values of $\sqrt{\beta_1} = 0$ and $\beta_2 = 3$ for a normal distribution.

Comparison of the calculated skewness indices with benchmark values provides information on the extent of any deviation from a symmetrical distribution (being positively skewed when indices exceed 0 – and therefore with a mean above the median - or negatively skewed in the case of skewness index values below 0 – with the mean below the median). For the kurtosis index, which examines the dispersion of the variable of interest around the mean and the thickness of the distribution tails, indices above 3 indicate leptokurtosis with central peaks tending to exceed those of the normal distribution, having less mass in the "shoulders" of the distribution and generally with the tails of the distribution exhibiting greater mass ("fat tails") than observed in the case of the normal distribution. Bottazzi, Coad, Jacoby and Secchi (2005, p.8) note that kurtosis values of 6 are consistent with the Laplace distribution. In contrast, platykurtic distributions (or negative kurtosis) with index values below 3 are generally represented by flatter and broader peaks and less mass in the tails of the distribution than under the normal distribution (Lotti and Santarelli, 2001b, pp.11-12). Mesokurtic distributions exist in the case of neither positive nor negative kurtosis (Upton and Cook, 2002).

To more formally assess the fit of the hypothesised lognormal distribution to the actual distribution of plant sizes (or more precisely the logarithmic transformation of the employee numbers data[62] - providing the normal distribution as the test distribution) use is made of the omnibus test presented in D'Agostino et al. (1990)[63]. This approach tests the null hypotheses that H_0: $\sqrt{\beta_1} = 0$ and H_0: $\beta_2 = 3$ versus the alternative hypotheses that H_1: $\sqrt{\beta_1} \neq 0$ and H_1: $\beta_2 \neq 3$, testing for deviations in both measures (Lotti and Santarelli, 2001a, 2001b; Kennedy, 1998, p.79) and then combines these into an overall test statistic (following the procedure of D'Agostino and Pearson, 1973), given by:

$$K^2 = Z^2(\sqrt{b_1}) + Z^2(b_2) \tag{4.3}$$

where $Z(\sqrt{b_1})$ and $Z(b_2)$ are transformations[64] which are normally distributed under a null hypothesis that the population is also distributed normally (D'Agostino et al, 1990, p.318). These transformations provide approximations to $\sqrt{b_1}$ and b_2 (the sample estimates of $\sqrt{\beta_1}$ and β_2), for which $\sqrt{b_1} = m_3/m_2^{3/2}$ and $b_2 = m_4/m_2^2$, $m_k = \Sigma(X_i - \bar{X})^k/n$ and $\bar{X}$ is the sample mean provided by $\bar{X} = \Sigma X_i/n$, when n is the number of observations. For a normal population, the K^2 statistic follows an approximately χ^2 distribution with 2 degrees of freedom.

In practice, to apply the D'Agostino et al. (1990) test the empirical adjustments to the overall χ^2 and significance level proposed by Royston (1991) are suppressed on the skewness-kurtosis test in STATA (StataCorp LP, 2005). The test results reported here focus on the omnibus test results rather than the individual component tests, since other plausible distributions (within this context) may also be characterised by indices of such higher moments

---

[62] Although, the position of particular plants or firms within a specific size distribution may be influenced by the selected size measure, many studies have tended to suggest that it makes little difference to the qualitative characteristics of the size distribution (see, for example, Stanley et al., 1995), although a recent study by Bottazzi et al. (2006), finds somewhat contrary evidence to this observation. Given that there is only one measure available in this current dataset, employee numbers, it is not possible within this research to examine this possibility.

[63] Unlike Lotti and Santarelli (2001a, 2001b), the Kolmogorov-Smirnov test is not employed here given it's unsuitability in testing for normality (Lotti and Santarelli, 2001b, p.12; D'Agostino et al., 1990).

[64] The multi-step computational formulas for both $Z(\sqrt{b_1})$ and $Z(b_2)$ based on these sample estimates are presented in D'Agostino et al. (1990, pp.317-318).

consistent with the stated benchmark values.

The Shapiro-Wilk W test (1965), a test for outliers in the form of deviations of the ordered sample values from the specified normal distribution (Kennedy, 1998, pp.78-79; Upton and Cook, 2002, p.334), is also applied to the selected samples, a test which D'Agostino et al. (1990, p.316) suggest has good power properties. The test (for the normal distribution) is given by:

$$W = \Big(\sum_{i=1}^{n} w_i X_i\Big)^2 \Big/ \sum_{i=1}^{n} (X_i - \bar{X})^2 \tag{4.4}$$

where X_i is the ith largest observation from a sample of n observations, $\bar{X}$ is the sample mean and w_i denotes a function of the properties of the order statistics from the specified normal distribution (Upton and Cook, 2002, p.334), where $w' = (w_1, w_2, \ldots, w_n) = m'V^{-1}[m'V^{-1}V^{-1}m]^{-\frac{1}{2}}$; $m' = (m_1, m_2, \ldots, m_n)$ being a vector of expected values and V is the relevant covariance matrix (Park, 2002-06, p.8). The Shapiro-Wilk W test available in STATA makes use of the transformation of the null distribution of W suggested by Royston (1982) permitting appropriate testing for samples of up to 2,000 observations (and of at least 7 observations, StataCorp LP, 2005). The outcome of this test is a value of $W \leq 1$, with values approaching $W = 1$ indicating closer approximations to the normal distribution (Park, 2002-06, p.8). The Shapiro-Francia W′ (1972) test is applied in the case of larger aggregate samples, being appropriate for samples of up to 5,000 observations (and of at least 5 observations, StataCorp LP, 2005). This approximate test replaces w' with $b' = (b_1, b_2, \ldots, b_n) = (m'm) = \frac{1}{2}$ (Park, 2002-06, p.8).

Despite the earlier discussed (in Chapter 2) limitations of such statistical testing procedures in examining extreme hypotheses, these concerns may be somewhat moderated here both by the use of a broader range of quantitative methodologies, and the avoidance of an excessively strict interpretation of the statistical tests; the latter alleviating the unreasonable rejection of the lognormal distribution where it may provide a reasonable approximation to the actual data (Hart and Oulton, 1996, p.1244).

Such weight on the precise statistical acceptance or rejection of hypothesised distributions would seem especially inappropriate where a poor statistical fit may simply have arisen as a consequence of an insufficient timescale for the data to have reached the theoretical limiting distribution (Hart, 1980), and additionally given the incomplete sectoral coverage and the (at least partial) truncation of the WRME data - the latter of which may have important implications for the aggregate statistical properties of the dataset, removing some of the mass of the left tail of the distribution and giving rise to a higher mean and smaller variance than would be the case under the full distribution (Greene, 2000, p.899). Given these potential limitations the even more problematic task of attempting to distinguish between the lognormal and competing plausible statistical distributions, such as the Pareto distribution, is not attempted here.

It should also be further emphasised that as in the case of the deterministic models of firm growth, that such stochastic models aim to provide a relatively parsimonious representation of actual dynamic processes – in both cases, no matter how good the model, it will always be possible to identify some samples or groups that do not strictly adhere to the models predictions. It is argued here that testing for an approximately lognormal distribution matters because of the implied stochastic processes behind its generation and the consequences for the understanding of wider industrial processes, rather than because the predicted lognormal distribution might provide a "better fit" to the data than other distributions (which have also frequently been founded on the stochastic process described by Gibrat, 1931). Similarly, deviations from such a distribution might suggest further deterministic factors that need to be more adequately reflected (see Ijiri and Simon, 1977, for such development work).

Given that a broad consistency of the actual plant size distributions with the lognormal would tend to suggest some underlying Gibrat-type process, for this reason commencing with the lognormal distribution seems of significant merit.

To supplement these statistical tests non-parametric kernel procedures are also employed to examine plant size and growth rate distributions, a methodology which avoids imposing overly-restrictive parametric structures and assumptions (Kennedy, 1998; Lotti and Santarelli, 2001b). The kernel density estimation method essentially generates a (continuous) smoothed histogram[65] (Bottazzi, Coad, Jacoby and Secchi, 2005, p.13) through a procedure of weighting "local" data observations (Cabral and Mata, 2001, p.4; Kennedy, 1998), with the resulting estimate of the probability density function (Upton and Cook, 2002, p.185) exhibiting peaks where there are the densest concentrations of data (Upton and Cook, 2002, p.186).

Following Lotti and Santarelli (2001b, p.10) and Bottazzi, Coad, Jacoby and Secchi (2005, p.13) the estimation procedure requires a kernel function (denoted K) and a bandwidth (interval width) for the bin (h), with the kernel density estimator represented in general form as:

$$f(x) = 1/\mathrm{n}h \sum_{i=1}^{\mathrm{n}} K\,(x - x_i/h) \qquad (4.5)$$

where the estimated density is denoted $f(x)$, $K(x) \geq 0$, $\forall\ x \in (-\infty, +\infty)$ and $\int dx\ K(x) = 1$ (a non-negative function), x and x_i are data points and n is the size of the sample of plants (Bottazzi, Coad, Jacoby and Secchi, 2005, p.13).

A range of possible kernel functions, formulas providing the weightings to be attached to observations relative to their distances from each examined point (giving lower, but non-

---

[65] As Kennedy, 1998, p.309; see also Greene, 2003, p.455) notes the kernel density estimation methodology avoids a number of the problems of using a histogram to examine the distribution of plant sizes: that the data would not be smooth between intervals, in part since some size values would be absent and that, the choice of the start and end points of each interval may have a significant effect on the observed density. However, the problem that the width of the interval may influence the observed distribution remains an issue in both approaches.

zero, weights to observations further from the examined points, Greene, 2003, p.881), can be considered. Kennedy (1998, p.311) suggests that given that the kernel is both centred at zero and symmetric, the choice between such kernels may not be that significant. The analysis here uses the Gaussian (normal) kernel function supplemented with the Epanechnikov kernel function[66], consistent with the recent studies of Lotti and Santarelli (2001b), Cabral and Mata (1996) and Bottazzi, Coad, Jacoby and Secchi (2005). Some initial empirical investigation conducted here indicated that, consistent with previous empirical studies, the choice of kernel function did not in general significantly alter the qualitative properties of the estimates.

In generating these kernel density estimates of the size (and growth rate) distributions of plants the selection of the band-width parameter, and hence the width of examined intervals, may be potentially influential (Kennedy, 1998; Lotti and Santarelli, 2001b), with the choice of the band-width determining the extent of the smoothing of the estimated density (Silverman, 1986), and affecting the estimator variance and bias (Kennedy, 1998, p.311). Given the lack of a widely accepted method for selecting the band-width a potential issue arises regarding the possibility of under-smoothing or over-smoothing the density estimates, the former occurring when the band-width is insufficiently large with a corresponding insufficient number of observations given non-negligible weight, and the latter arising in the case of an inappropriately wide band-width.

To examine the effect of the band-width choice some initial experimentation was conducted using a range of band-widths including the default "optimal" band-width selection procedure in STATA (which minimises the integrated mean squared error of the estimate under the conditions of the data being Gaussian, using the Gaussian kernel; StataCorp LP, 2005) and the straight-forward method of employing a fixed coefficient band-width (as employed by Cabral and Mata, 2001, who found that a parameter of 0.5 provided a reasonable balance

---

[66] Following Upton and Cook (2002, pp.185-186), the kernel functions, K(x, t), are: $\{1 - (x - t)^2/5h$
$\{0$
for $-\sqrt{5}h < x - t < \sqrt{5}h$ in respect of the Epanechnikov kernel and $\exp(x - t)^2/2h^2)$ where $-\infty < x < +\infty$ for the Gaussian kernel, x being the sample values of a variable X.

between clarity and consistency of the overall distribution estimates for both small and large samples without significant loss of information from potential over-smoothing). However, for this data the choice between the various alternative band-width criteria typically did not significantly alter the broad qualitative properties of the estimates. Further experimentation suggested that additional useful information could be obtained by extending the band-widths to slightly more extreme values. Given these observations, the simple approach of Cabral and Mata (2001) was employed as the primary means of band-width selection, with an initial band-width parameter of 0.5, this value consistent with this previous study and deemed appropriate from some initial experimentation conducted on this dataset.

Throughout the analysis the number of points of estimation (the number of points at which the kernel density estimate is evaluated) is 50, the default of STATA, with some sensitivity to this selection conducted during the analysis. In practice, little difference was found in the overall properties of the kernel density estimates for larger samples through increasing the number of points of estimation. However, decreased numbers of points below 50 led to increasingly angular density estimate representations.

It is worth noting that this approach using non-parametric kernel density estimation to examine the size and growth rate distribution of plants provides some novelty, with few identified studies examining the implications of Gibrat's Law using U.K. data and employing such a technique, and none identified at the U.K. regional-level. However, in conducting this analysis, it is recognised that appropriate estimation using the non-parametric kernel density estimation method requires a relatively large sample size (Kennedy, 1998) so that the analysis is generally restricted to higher-level aggregations of plants.

4.3. The Aggregate Plant Size Distribution

As a first step in assessing the suitability of the lognormal distribution as an approximation to the distribution of this sample of plants, the aggregate cross-section samples for each year 1966-2003 were examined using the above described statistical tests and summary statistics of the moments of the size distributions. The testing of the static aggregate size distributions is consistent with a number of early statistical studies (for example, Gibrat, 1931; Hart and Prais, 1956; and Clarke, 1979) and would be consistent with the assumption that the stochastic process of Gibrat's Law had been operating for a sufficient time period for the observed plant size distribution to have evolved to the lognormal limiting distribution.

The information, presented in Table 4.1, is based on the logarithm of plant sizes to facilitate a more straight-forward analysis testing the observed (transformed) distributions against the normal distribution. Further information on the static size distributions of these plants was presented in Table 3.1 of the previous chapter, and given that this analysis makes use of a simple transformation of the data the reported percentile measures are not reproduced again.

Examination of the calculated indices points to some degree of positive skewness in the aggregate size distributions in each of the 38 annual cross-sections, with the skewness indices ranging between 0.34 and 0.64. Across the period there is no strong evidence of a trend towards a less skew distribution over time[67]. The reported kurtosis indices generally suggest some slight leptokurtosis in the distributions with a maximum index value of 3.33. However, in two of the cross-sections relating to 1977 and 2000 some slight deviations towards platykurtosis were observed. Since 1997, the distributions have been very close to being mesokurtic. In broad terms, these empirical observations regarding the skewness and kurtosis of plant size distributions are qualitatively similar to those of Hart and Prais (1956, p.159) who also observed some degree of positive skewness and modest leptokurtosis.

The D'Agostino et al. test rejected the normal distribution as a suitable statistical

---

[67] Although, Higson, Holly and Kattuman (2002) examining quoted U.S. firms detected correlations between the higher moments of the cross-sectional distribution of firm growth rates and selected macroeconomic variables (see also Marsili et al., 2004, p.5), a preliminary graphical inspection did not suggest any strong relationships for this data. However, further statistical analysis would be required here to more definitively consider this possibility.

Table 4.1: Normality Tests of the Plant (Log)Size Distribution

	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984
Plants (N)	1,432	1,490	1,601	1,737	1,857	1,935	2,044	2,150	2,245	2,680	2,726	2,812	2,665	2,460	2,506	2,475	2,516	2,513	2,514
Mean	3.97	3.92	3.88	3.86	3.85	3.79	3.72	3.69	3.66	3.26	3.24	3.15	3.29	3.48	3.44	3.34	3.28	3.25	3.26
Standard Dev.	1.47	1.46	1.46	1.46	1.47	1.46	1.47	1.48	1.48	1.60	1.57	1.64	1.55	1.45	1.44	1.41	1.40	1.39	1.38
Skewness	0.59	0.59	0.57	0.56	0.53	0.50	0.45	0.46	0.49	0.41	0.43	0.35	0.43	0.52	0.52	0.53	0.50	0.51	0.50
Kurtosis	3.17	3.22	3.20	3.19	3.20	3.22	3.23	3.22	3.20	3.02	3.06	2.98	3.11	3.28	3.25	3.24	3.25	3.17	3.18
D'Agostino	73.97*	77.74*	78.18*	83.06*	80.04*	75.80*	68.66*	74.53*	84.36*	70.28*	79.57*	55.09*	77.11*	104.36*	107.35*	109.82*	101.64*	100.83*	98.24*
Shapiro-Francia	0.97*	0.97*	0.98*	0.98*	0.98*	0.98*	0.99*	0.98*	0.98*	0.99*	0.99*	0.99*	0.99*	0.98*	0.98*	0.98*	0.98*	0.98*	0.98*

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Plants (N)	2,545	2,505	2,470	2,395	2,509	2,553	2,504	2,366	2,264	2,193	2,151	2,123	2,072	2,002	1,982	1,975	1,883	1,805	1,713
Mean	3.25	3.35	3.46	3.54	3.55	3.55	3.57	3.57	3.59	3.65	3.70	3.74	3.79	3.81	3.82	3.79	3.81	3.77	3.74
Standard Dev.	1.39	1.32	1.27	1.28	1.28	1.26	1.22	1.23	1.21	1.19	1.17	1.16	1.13	1.14	1.13	1.15	1.16	1.18	1.23
Skewness	0.47	0.50	0.51	0.47	0.44	0.44	0.54	0.48	0.50	0.51	0.52	0.54	0.64	0.59	0.60	0.53	0.50	0.45	0.34
Kurtosis	3.10	3.24	3.31	3.22	3.24	3.33	3.31	3.22	3.14	3.14	3.21	3.21	3.09	3.09	3.08	2.99	3.00	3.08	3.02
D'Agostino	85.74*	100.39*	105.31*	85.12*	80.82*	85.84*	115.39*	85.27*	88.24*	87.68*	89.43*	95.23*	122.10*	100.42*	103.34*	82.62*	70.06*	56.56*	31.53*
Shapiro-Francia	0.99*	0.98*	0.98*	0.98*	0.98*	0.98*	0.98*	0.98*	0.98*	0.98*	0.98*	0.98*	0.97*	0.97*	0.97*	0.98*	0.98*	0.98*	0.99*

Notes: Statistical data and tests of normality applied to the logarithm of plant size, measured by employee numbers.
* Statistically significant at the 5 percent level.

description of the logarithm of plant sizes in all years covering 1966-2003 at the 5 percent level of significance, with the skewness component being rejected at this level in each year, but the kurtosis component being rejected in only around one-third of samples. The Shapiro-Francia test also rejected (at the same significance level) the null hypothesis of normality of the (log)size distribution across the full range of cross-section samples, therefore providing evidence of statistically significant deviations from the lognormal distribution postulated by Gibrat's Law.

To provide additional graphical detail on the logarithmically transformed size distributions kernel density estimates were generated for each of the annual cross-sections using the Gaussian kernel function with a band-width parameter of 0.5 and 50 estimation points. Given the large number of kernel density estimates generated these are not all presented with Figures 4.1a-d selected as illustrating the general properties observed across the time period and since these also represent the initial years of the constructed plant-level samples. Normal density plots are added for ease of comparison.

These kernel density estimates suggests a number of points. First, consistent with the constructed index of skewness, in all years the plant size distributions exhibit slight positive skewness (so that the right tail of the distribution is extended), with the peak of the density estimate generally around a logarithmic plant size of 2.5, corresponding to approximately 11 - 12 employees, just above the minimum size threshold of the WRME dataset. The kernel density estimates also indicate some very slight under-represented mass in parts of the left portion of the distribution (particularly for lower band-width parameters), consistent with the truncation of the dataset[68], most notably during the 1990s. However, in the case of the cross-section for 1977 the kernel density estimate exhibits some modest excess mass in parts of the left portion of the actual size distribution.

---

[68] Although not conducted here, a natural extension to this work, given the identified properties of the data, could be to examine the suitability of truncated distributions, such as the truncated lognormal distribution.

Broadly similar results were obtained through some sensitivity analysis using band-widths of the order 0.1 - 0.2 to provide further detail on the estimated distributions, although with perhaps sharper observed rises in the density estimates at values just below those reported above, but again broadly consistent with the truncation threshold. Experimentation with the use of larger band-widths of the order 0.7 - 1.0 was generally found to produce distributions increasingly consistent in appearance with the (log)normal distribution.

Comparison of the kernel density estimates with the normal density distribution suggests some greater consistency between the distributions during the period 1975-1978. It is plausible that such an observation may be the result of the greater recording of very small plants in the WRME database during this period, providing a more complete size distribution. More detailed inspection of these aggregate plant size distributions using band-widths of 0.2, however, suggests that despite this greater consistency with the normal distribution that the slight positive skewness and leptokurtosis identified earlier continues to be observable. In addition, a further slight change in the plant (log)size distribution was observed from 1992 onwards with the kernel density estimates exhibiting lower peaks than the normal density, although a band-width of 0.2 returned the estimated distribution to a pattern of a somewhat higher peak density than the normal density.

Although there was some sensitivity to the precise form of the kernel density estimates of the plant (log)size distributions, the findings provide further consistent evidence of the aggregate size distributions being slightly positively skewed and mildly leptokurtic. Given the earlier discussion regarding the properties of this dataset and the limitations of statistical tests of extreme hypotheses such results do not seem to suggest a strong rejection of the operation of a Gibrat-type process (though not strictly supporting the lognormal distribution) albeit with perhaps some complicating factors in operation.

Figure 4.1a: Plant (Log)Size Distribution, 1972

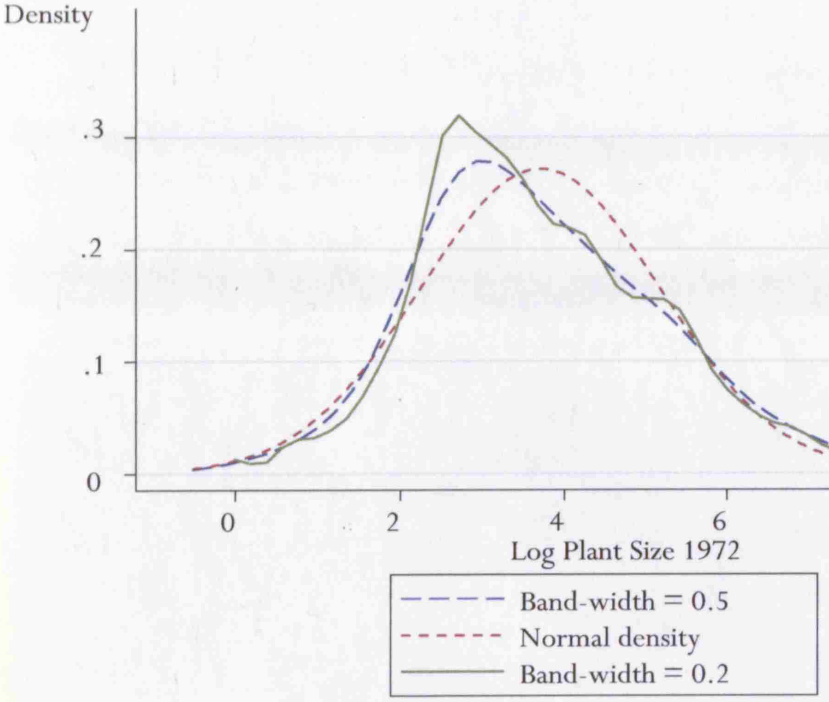


Figure 4.1b: Plant (Log)Size Distribution, 1977

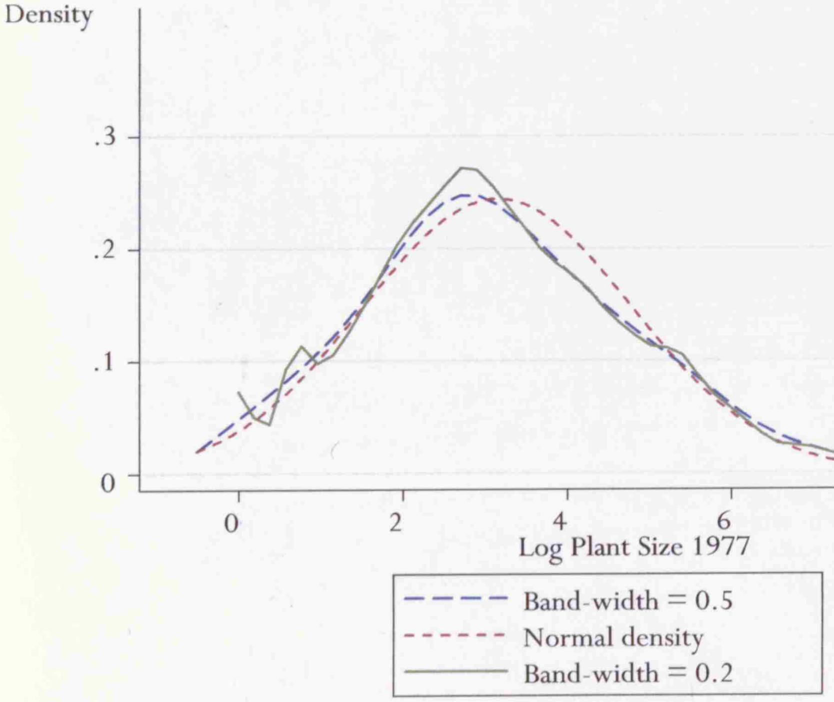


Figure 4.1c: Plant (Log)Size Distribution, 1986

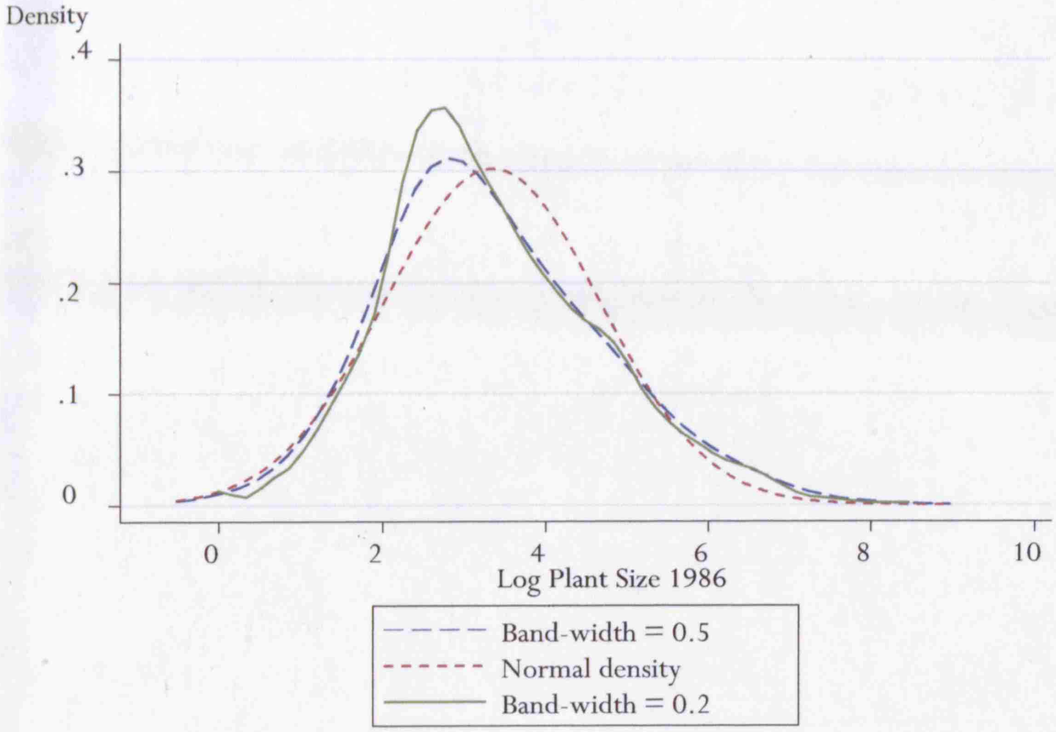
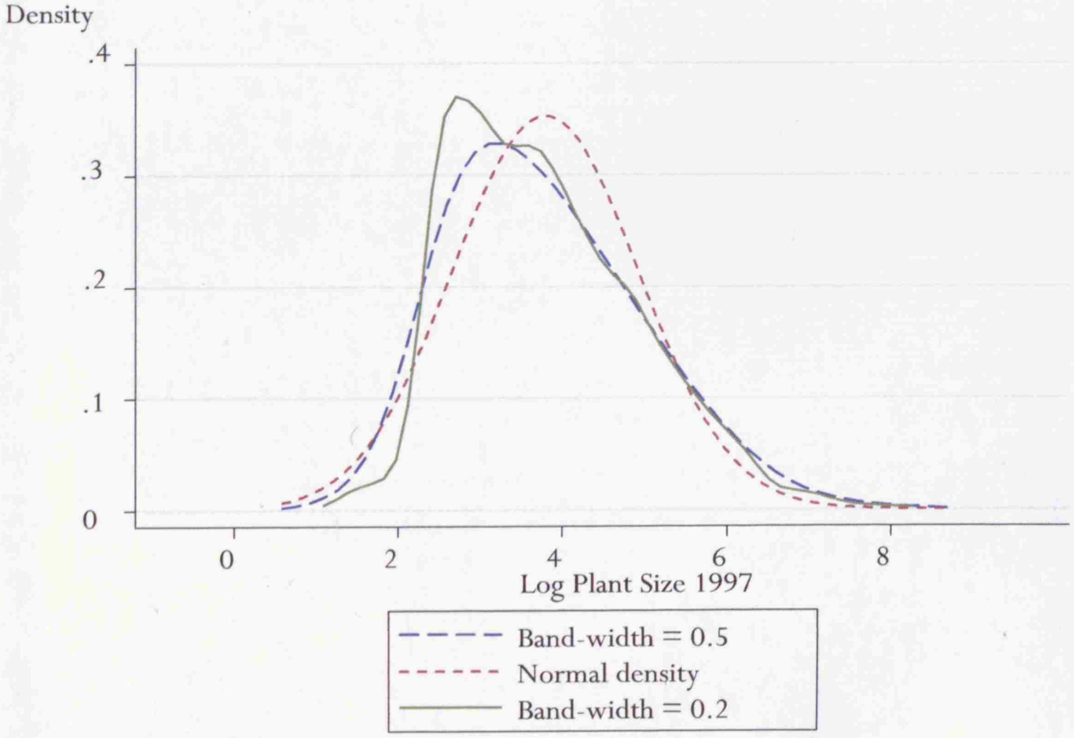


Figure 4.1d: Plant (Log)Size Distribution, 1997



4.4. Plant Age, Selection and the Evolution of Plant Size Distributions

Although Gibrat's Law abstracts from the process of plant entry[69] and exit, as de Wit (2005) sets out, theoretically such processes can have important effects on the implied limiting size distribution. The role of selection processes, observed as plant exit, and learning are also emphasised in the models of Jovanovic (1982), Ericson and Pakes (1995) and Audrestch (1995), so that both ageing and exit may have significant effects on the evolution of the observed size distribution of manufacturing plants.

To examine the role of plant age in shaping manufacturing plant size distributions a plant age variable was created using the available WRME records on plant opening dates. For those plants with no opening date information plant age was estimated using the oldest employment observation, providing a lower bound to plant age. Inspection of the distribution of the generated age values suggests a somewhat greater than expected number of plants with estimated ages corresponding to entry in 1966, an artifact of using the employment values in this year (where plant opening date information is absent) being the first recorded in the WRME database. In addition, a second peak in the distribution of plant ages also emerged corresponding to the entry year 1975, in likelihood also an artifact of the data collection methodology used in that specific year.

Following the approach of Cabral and Mata (2001) the cross-sections of plants in each of the years 1972, 1977 and 1986 were grouped into five age-classes, here being 0 - 1 years, 2 - 5 years, 5 - 10 years, 11 - 20 years and +20 years. The youngest age-class was deliberately bounded to a very early age limit to allow for the possibility that any evolution to a specific distribution may occur rapidly over a period of several months or

---

[69] The role of plant entry in shaping plant size distributions is not addressed in this work. Some information on plant entry is provided in Chapter 3, but due to the frequently small sample sizes, especially in more recent years, such data is not presented disaggregated by industry or plant size-class.

quarters (see Lotti and Santarelli, 2001a, 2001b, for such empirical evidence).

Information on the sample sizes for these age-classes, the summary statistics relating to the age-class size distributions and the D'Agostino et al. and Shapiro-Francia statistical tests (employed for the aggregate samples and for the smaller age-class samples for consistency) are presented in the first broad column (headed "All" with the respective initial year) in Tables 4.2a-d. The second broad column (headed "Survivors") provides the corresponding information on the size distributions of those plants in the initial cross-section which survived until the final year of each analysed time period. The final broad column (headed with the final year of each assessed time period) presents the associated information on the size distributions of the surviving plants in the last year of each sample.

Representing the information in such a format permits some assessment of the roles of both age and selection effects through plant exit on the evolution of the plant size distributions. Given the relatively small sample sizes for some age-classes, and for the later analysed industries and cohorts, the various percentile measures presented earlier in respect of the aggregate plant size distributions are not reported here or in the following sections.

Analysis of the statistical data on the logarithms of plant sizes associated with these samples suggests a number of points. First, for each period the mean size of plants generally increased with plant age. Second, there is a tendency for the standard deviation of plant size to also increase with plant age. However, for some periods this increasing standard deviation tapered-off after a period of greater than 6 - 10 years. Although deviations in the form of the reported skewness and kurtosis indices are observed from that consistent with the (log)normal distribution, these are often of fairly modest magnitudes.

However, inspection of the skewness indices for the two panels commencing in

1972, suggest that there are somewhat more notable deviations from the benchmark index value for the 6 - 10 years age-class in all three samples of all plants, surviving plants in the initial year and surviving plants in the final year. For the 1972-1977 sample, in all cases except for the +20 years age-class positive skewness was detected. In the case of the longer time period, in only the sample of plants aged 11 - 20 years for the surviving plants in 1972 was a non-positive skewness index calculated.

The kurtosis indices for the 6 - 10 years age-class also exhibited some relatively higher positive kurtosis in both periods and for each of these samples. However, more broadly, with the exception of this specific age-class (which corresponds to the noted age estimation issues discussed above), for the 1972-1977 data there is a clear pattern for declining kurtosis with increased age in the samples of initial and surviving plants, in all cases moving from being positively to negatively skewed. Whilst such a relationship of declining kurtosis indices with increased age-class is also generally observed for the longer time period in the case of the sample of initial plants, such a relationship is less clear for the samples of surviving manufacturing plants.

The D'Agostino et al. and Shapiro-Francia tests, conducted at the 5 percent level of significance, rejected the (log)normal distribution for the aggregate samples of all plants and both samples of surviving plants in the 1972-1977 period. For both the samples relating to 1972, all plants and surviving plants, the null hypothesis of lognormality could only be rejected for the age-class 6 - 10 years, again with less significant rejections for the surviving plants. However, whilst lognormality could again be rejected for manufacturing plants in 1977, these tests also rejected the null hypothesis for the 0 - 1 years age-class. The Shapiro-Wilk test additionally rejected lognormality for the 2 - 5 years age-class. In combination, the tests suggest modest selection effects (the results changing relatively little between the first two 1972 samples) and ageing effects at the aggregate-level although perhaps acting in opposite directions.

Table 4.2a: Normality Tests of the Plant (Log)Size Distribution by Age Group, 1972-1977

	All 1972						1972 Survivors						1977					
Plant Age (Years)	All	0 - 1	2 - 5	6 - 10	11 - 20	+20	All	0 - 1	2 - 5	6 - 10	11 - 20	+20	All	0 - 1	2 - 5	6 - 10	11 - 20	+20
Plants (N)	2,044	236	458	865	214	271	1,754	190	378	754	192	240	1,754	190	378	754	192	240
Mean	3.72	2.63	3.40	3.85	3.84	4.65	3.75	2.57	3.42	3.89	3.84	4.69	3.79	2.99	3.60	3.82	3.78	4.62
Standard Deviation	1.47	1.20	1.24	1.42	1.45	1.51	1.52	1.26	1.29	1.47	1.47	1.55	1.47	1.19	1.27	1.45	1.53	1.52
Skewness	0.45	0.10	0.24	0.73	0.18	0.02	0.42	0.16	0.26	0.70	0.17	-0.02	0.47	0.30	0.20	0.72	0.14	-0.02
Kurtosis	3.23	3.10	3.00	3.67	2.54	2.47	3.12	2.98	2.89	3.54	2.53	2.44	3.21	3.86	3.19	3.64	2.42	2.58
D'Agostino	68.66*	0.71	4.47	73.69*	3.80	5.13	50.18*	0.92	4.21	58.97*	3.36	5.25	61.07*	7.36*	3.50	62.70*	5.16	2.30
Shapiro-Francia	0.99*	1.00	0.99	0.96*	0.99	0.99	0.99*	1.00	0.99	0.96*	0.99	0.99	0.98*	0.98*	0.99*	0.97*	0.99	0.99

Notes: Statistical data and tests of normality applied to the logarithm of plant size, measured by employee numbers.
Age-classes defined in the initial year of each sample.
* Statistically significant at the 5 percent level.

Table 4.2b: Normality Tests of the Plant (Log)Size Distribution by Age Group, 1972-1997

	All 1972						1972 Survivors						1997					
Plant Age (Years)	All	0 - 1	2 - 5	6 - 10	11 - 20	+20	All	0 - 1	2 - 5	6 - 10	11 - 20	+20	All	0 - 1	2 - 5	6 - 10	11 - 20	+20
Plants (N)	2,044	365	458	865	214	271	693	67	151	259	81	135	693	67	151	259	81	135
Mean	3.72	2.74	3.40	3.85	3.84	4.65	4.11	2.87	3.71	4.21	4.26	4.87	4.10	3.83	4.15	3.93	4.20	4.41
Standard Deviation	1.47	1.22	1.24	1.42	1.45	1.51	1.63	1.39	1.40	1.62	1.61	1.56	1.24	1.09	1.13	1.30	1.26	1.25
Skewness	0.45	0.17	0.24	0.73	0.18	0.02	0.34	0.42	0.20	0.63	-0.05	0.08	0.38	0.29	0.22	0.63	0.15	0.21
Kurtosis	3.23	3.29	3.00	3.67	2.54	2.47	2.87	2.84	2.50	3.29	2.52	2.50	2.62	2.24	2.33	3.03	2.06	2.58
D'Agostino	68.66*	0.71	4.47	73.69*	3.80	5.13	13.47*	2.23	3.15	16.53*	2.56	1.92	22.00*	3.91	6.56*	15.60*	8.47*	2.12
Shapiro-Francia	0.99*	1.00	0.99	0.96*	0.99	0.99	0.99*	0.98	0.99	0.97*	0.98	0.99	0.99*	0.98	0.99	0.97*	0.98	0.99

Notes: Statistical data and tests of normality applied to the logarithm of plant size, measured by employee numbers.
Age-classes defined in the initial year of each sample.
* Statistically significant at the 5 percent level.

Table 4.2c: Normality Tests of the Plant (Log)Size Distribution by Age Group, 1977-1986

	All 1977						1977 Survivors						1986					
Plant Age (Years)	All	0 - 1	2 - 5	6 - 10	11 - 20	+20	All	0 - 1	2 - 5	6 - 10	11 - 20	+20	All	0 - 1	2 - 5	6 - 10	11 - 20	+20
Plants (N)	2,812	375	776	474	874	313	1,444	97	309	284	512	242	1,444	97	309	284	512	242
Mean	3.15	1.71	2.36	3.51	3.81	4.45	3.64	2.47	2.80	3.71	3.96	4.44	3.64	3.20	3.15	3.76	3.69	4.21
Standard Deviation	1.64	1.28	1.31	1.26	1.47	1.54	1.57	1.24	1.31	1.29	1.54	1.58	1.40	1.20	1.30	1.28	1.41	1.45
Skewness	0.35	0.70	0.39	0.30	0.62	0.01	0.41	0.50	0.10	0.39	0.61	0.11	0.30	0.55	0.04	0.21	0.48	0.06
Kurtosis	2.98	3.53	2.73	3.32	3.45	2.54	3.10	3.20	2.76	3.05	3.16	2.54	2.99	3.02	3.09	2.85	3.06	2.62
D'Agostino	55.09*	30.16*	21.71*	9.02*	54.66*	4.07	38.99*	4.73	1.20	7.24*	28.65*	3.58	21.00*	5.08	0.35	2.30	18.62*	1.80
Shapiro-Francia	0.99*	0.99*	0.99*	0.99*	0.97*	0.99	0.99*	0.97	1.00	0.98*	0.97*	0.99	0.99*	0.98	1.00	0.99	0.98*	0.99

Notes: Statistical data and tests of normality applied to the logarithm of plant size, measured by employee numbers.
Age-classes defined in the initial year of each sample.
* Statistically significant at the 5 percent level.

Table 4.2d: Normality Tests of the Plant (Log)Size Distribution by Age Group, 1986-1997

	All 1986						1997 Survivors						1997					
Plant Age (Years)	All	0 - 1	2 - 5	6 - 10	11 - 20	+20	All	0 - 1	2 - 5	6 - 10	11 - 20	+20	All	0 - 1	2 - 5	6 - 10	11 - 20	+20
Plants (N)	2,505	321	459	377	937	411	1,417	131	230	231	540	285	1,417	131	230	231	540	285
Mean	3.35	2.55	2.99	3.29	3.52	4.03	3.65	2.58	3.11	3.43	3.88	4.32	3.93	3.53	3.74	3.81	3.97	4.27
Standard Deviation	1.32	0.94	1.08	1.14	1.34	1.48	1.34	0.93	1.12	1.15	1.27	1.44	1.16	0.90	1.03	1.09	1.19	1.26
Skewness	0.50	0.42	0.59	0.50	0.35	0.04	0.57	1.08	0.86	0.62	0.61	0.03	0.58	0.95	0.74	0.77	0.52	0.17
Kurtosis	3.24	4.08	3.59	3.15	3.32	2.60	2.92	5.29	3.68	3.12	3.11	2.49	2.95	3.98	3.53	3.34	2.99	2.30
D'Agostino	100.39*	18.08*	28.56*	15.28*	21.99*	3.99	68.23*	30.91*	26.84*	14.09*	29.84*	4.84	70.51*	20.48*	21.00*	20.97*	22.59*	14.53*
Shapiro-Francia	0.98*	0.98*	0.98*	0.98*	0.99*	1.00	0.97*	0.93*	0.95*	0.97*	0.97*	0.99	0.97*	0.94*	0.96*	0.95*	0.98*	0.98*

Notes: Statistical data and tests of normality applied to the logarithm of plant size, measured by employee numbers.
Age-classes defined in the initial year of each sample.
* Statistically significant at the 5 percent level.

The statistical tests conducted on the 1972-1997 period data provided an opportunity to investigate whether stronger effects could be observed over longer time periods. In addition again to the rejections of lognormality for the aggregate distributions for all samples of all plants (although being less significantly rejected in the aggregate surviving plant samples) and surviving plants and for the 6 - 10 years age-class, the D'Agostino et al. test also rejected this distribution for the 2 - 5 and 11 - 20 years age-classes. These tests are therefore again suggestive of selection and ageing effects acting in opposite directions, with selection supporting some evolution of the aggregate distribution towards the lognormal (but not sufficiently strong enough to exhibit any changes in the statistical significance, at the 5 percent level, of the tests), and plant age processes tending to lead to divergence from such a distribution, as identified by the greater frequency of rejections within age-classes in the final panel compared to the same sample of (surviving) plants in the initial year.

For the 1977-1986 samples the skewness indices all indicated positive skewness at the aggregate-level, with some variation across age-classes, declining first with increased age-class, then increasing before decreasing again for the oldest age-class, being fairly close to the benchmark index value of 0 for this group in the samples of all and surviving plants (for the initial and final years). In these samples, the +20 years age-class in all cases provided kurtosis indices below 3, with such platykurtosis also found for the 2 - 5 years age-class for both the 1977 samples and in the case of the 6 - 10 years age-class in 1986, with in all other cases the indices being above 3.

In respect of the statistical tests, these results are more consistent with the hypothesis of convergence to the lognormal distribution. Although the null hypothesis of the (log)normal distribution could be rejected for each aggregate sample of surviving plants and all plants in the initial year using both the D'Agostino et al. and the Shapiro-Wilk tests (but with a decreasing magnitude of rejection through both selection and ageing effects), for the age-classes, the rejection of the lognormal in all age-classes except for +20 years in the initial sample, evolved to a rejection in only the 6 - 10 and 11 - 20 years age-classes for surviving plants in the initial year (a selection effect) and to only a rejection for the 11 - 20 years class in the final year (an ageing effect).

For the 1986 sample the skewness indices suggest slight positive skewness for all of the aggregate and age-class samples, perhaps being of a slightly higher magnitude in the samples of surviving plants in the initial year than the sample of all plants. A tendency for slight negative kurtosis was also detected for the aggregate samples of surviving plants, in contrast to the sample of all plants. The data is suggestive of a clear decline in kurtosis with increased age-class in all samples, exhibiting negative kurtosis in each sample for the oldest age-class and additionally for the 11 - 20 years age-class in the case of surviving plants in the final year.

The statistical tests conducted on the data for the 1986-1997 period were not supportive of the (log)normal distribution, rejecting the hypothesis for the aggregate samples for the initial years, for surviving plants and for surviving plants in the initial year, although there was some suggestion of a weakening in the strength of the statistical rejection of the (log)normal hypothesis for surviving plants (a positive selection effect towards the lognormal distribution). In addition, (log)normality was rejected for all age-classes in all three corresponding samples, except for the +20 year's age-classes of all plants in the initial year and plants surviving until 1997 in the initial year. However, even for this age-class, (log)normality could be rejected for the 1997 surviving plants, suggesting, if anything, a move away from the (log)normal for this age-class rather than any evolution towards it.

In combination, these statistical tests, applied to the different age-classes of plants and for groups of all and surviving plants, do not suggest a strong and consistent pattern of convergence to the (log)normal distribution (accepting the earlier discussed limitations of such tests of extreme hypotheses). The analysis tends to support often modest selection and ageing effects, and not always operating in the same directions, a less straight-forward result than found by Cabral and Mata (2001). Such an observation is consistent with the views of Schmalensee (1989) and Sutton (1995a) who have pointed to the lack of empirical support for any single statistical distribution as being capable of providing an adequate description of the full range of industrial size distributions. Even though Cabral and Mata (2001) note that any such convergence may take a considerable time period, the analysis of the longer panel does not strongly support any such convergence, at least to the (log)normal distribution.

To draw out further detail regarding the observed plant (log)size distribution by age-classes kernel density estimates were again employed applying the same estimation parameters for consistency to each of the initial cross-section samples. The kernel density estimates are presented in Figures 4.2a-c, being manually mapped using consistent scaling to provide some comparability of the changes in observed age-class distributions.

Consistent with the statistical analysis, the kernel density estimates do not suggest a clear evolution towards the (log)normal distribution. The increasing means and standard deviations with increasing plant age reported earlier are clearly observed for each sample. The 1972 sample points to some movement from a relatively symmetrical and single-peaked distribution, towards a slightly bi-modal distribution for the age-class of 11 - 20 years (with a higher left-peak), with the oldest age-class presenting a similar bi-modal distribution but with a slightly higher right peak. A broadly similar pattern is also exhibited for the 1977 cross-section, with a highly right-skewed and single-peaked distribution for younger plants tending to become a significantly flatter peaked distribution for increased age-classes. The 1986 cross-section is perhaps most consistent with the analysis of Cabral and Mata (2001) with the distribution remaining largely single-peaked and fairly symmetrical for most age-classes, albeit with the peak gradually

Figure 4.2a: Plant (Log)Size Distribution by Age-Class, 1972 Cross-section

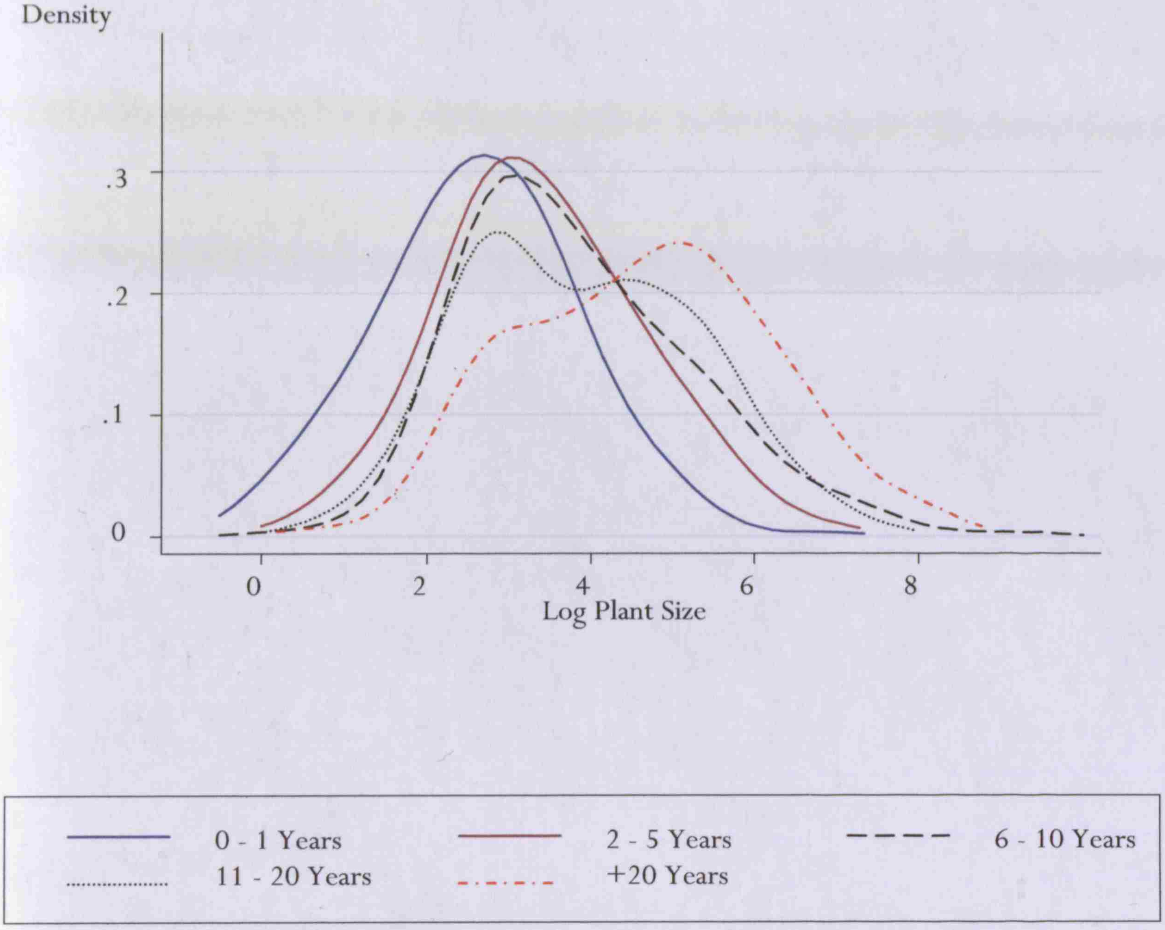


Figure 4.2b: Plant (Log)Size Distribution by Age-Class, 1977 Cross-section

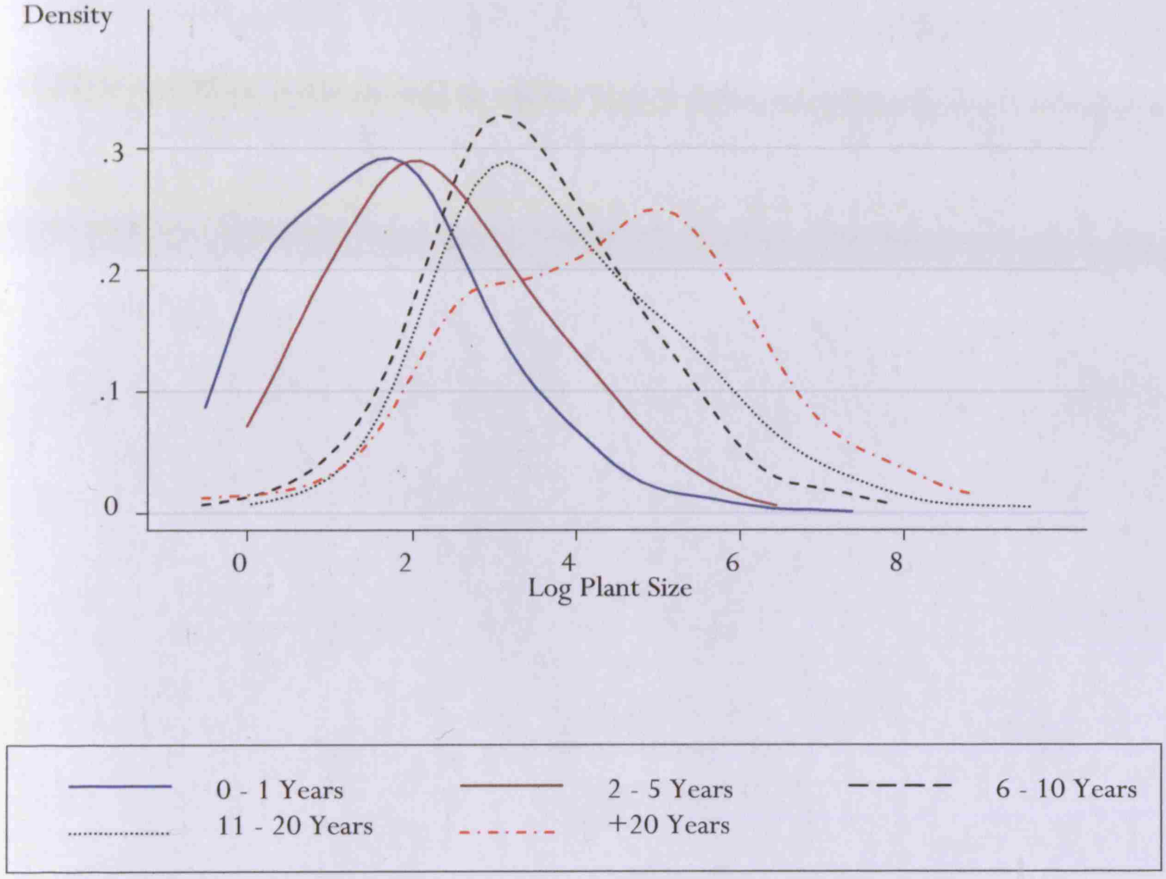
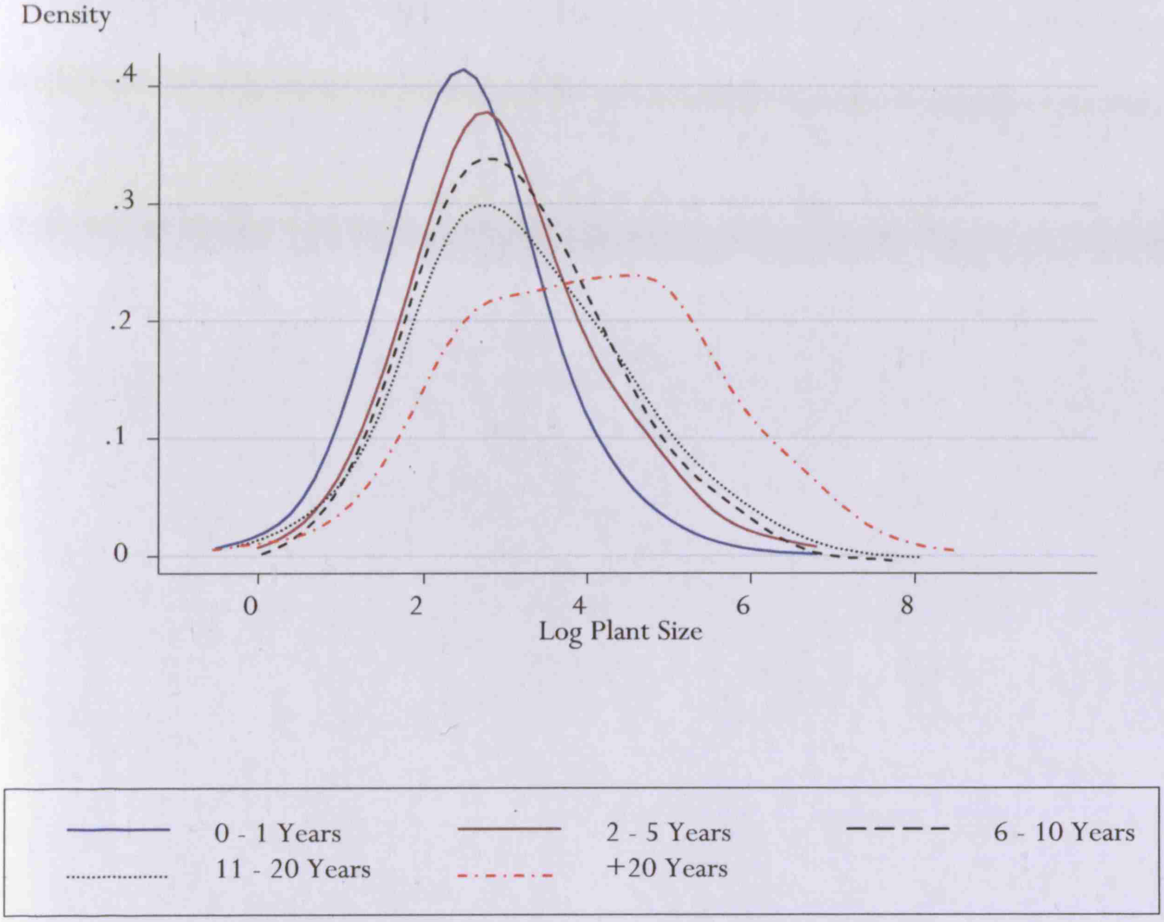


Figure 4.2c: Plant (Log)Size Distribution by Age-Class, 1986 Cross-section



declining and the base broadening with increased age. However, for the oldest age-class of +20 years a similar flat peak emerged.

As a further sensitivity analysis to consider whether there was any variation across the intermediate years between the initial and final years of the analysis, each of the cross-sections were tracked for the periods relating to each of the three economic cycles and for the long panel covering 1972-1997 (see Tables 4.3a-c to 4.5a-c). The D'Agostino et al. and Shapiro-Francia tests were again applied. In each case the cross-section samples (from which there is plant exit) exhibited some trend towards increasing mean sizes, with the mean plant size increasing from a logarithmic mean of 3.72 in 1972 to 4.10 in 1997, from 3.15 to 3.64 during 1977-1986 and from 3.35 to 3.93 in the period 1986-1997. However, for the samples of surviving plants such a pattern of increasing mean size over time is only observed for the 1986 cross-sections (increasing from 3.65 to 3.93). In all periods, for both surviving and all plants, the standard deviation of plant sizes declined over time, consistent with the "shake-out" (or convergence) of non-optimal sized plants.

In all samples and years, slight positive skewness indices are observed, in the case of the 1972 cross-section becoming very slightly more positively skewed in the initial years, before a reversal of this observation from around 1980 onwards. The kurtosis indices also exhibit fairly modest deviations from the benchmark normal values, but with lower kurtosis indices for surviving plants (than all plants) in all years for the 1972 cross-section (in all years below 3), a finding consistent with the 1986 cross-section. However, whilst again broadly following this pattern, in the 1977 cross-section the kurtosis indices are higher for the sample of surviving plants in three years (1977, 1978 and 1985).

The D'Agostino et al. and Shapiro-Francia tests (presented in Tables 4.5a-c) again rejected the (log)normal distribution at the 5 percent level in all years of each period for both the cross-section and surviving plants panels supporting the view that the above results are not an artifact of the choice of the selection of the time-span of the panels.

Kernel density estimates, shown in Figures 4.3a-d using the same estimation parameters to compare the plant size distributions in the initial and final years of each panel provided further evidence of the shift in the distribution of plants towards the right, being most notable over the longer time periods considered. Whilst three of the four distributions remained single-peaked (even if the distribution was not notably more symmetrical at the end of each respective period) for the 1972-1997 sample there was some evidence of an emerging flat peak to the (log)size distribution.

Further kernel density estimates (presented in Figures 4.4a-d) comparing only the surviving plants in the respective initial and final years of each sample again showed a fairly similar pattern of evolution. However, for the 1972-1997 sample of surviving plants the kernel density estimates suggested some evidence of an initially bi-modal (or at least flat-peaked) distribution and although such a broad distribution was also observed in 1997, there is some suggestion of a move towards a more peaked and symmetrical size distribution over time. This pattern of a slight movement towards a more peaked and symmetrical distribution for surviving plants is observed for each sample, and broadly consistent with Gibrat's Law, with a fixed population of plants evolving towards a lognormal distribution. However, it cannot be concluded that this distribution is a more appropriate representation than other statistical distributions.

4.5. Industry and New Entrant Plant Size Distributions

Given the heterogeneity of industry dynamics observed in previous studies (such as Lotti and Santarelli, 2001a, 2001b) consideration is given here to the evolution of plant size distributions at the 2-digit SIC industry-level. In analysing such industry-level plant size distributions within this section, a notable advantage here is the much longer time-span of data than has been available in previous studies. Although the use of 4-digit industry-

Figure 4.3a: Plant (Log)Size Distribution, 1972-1977

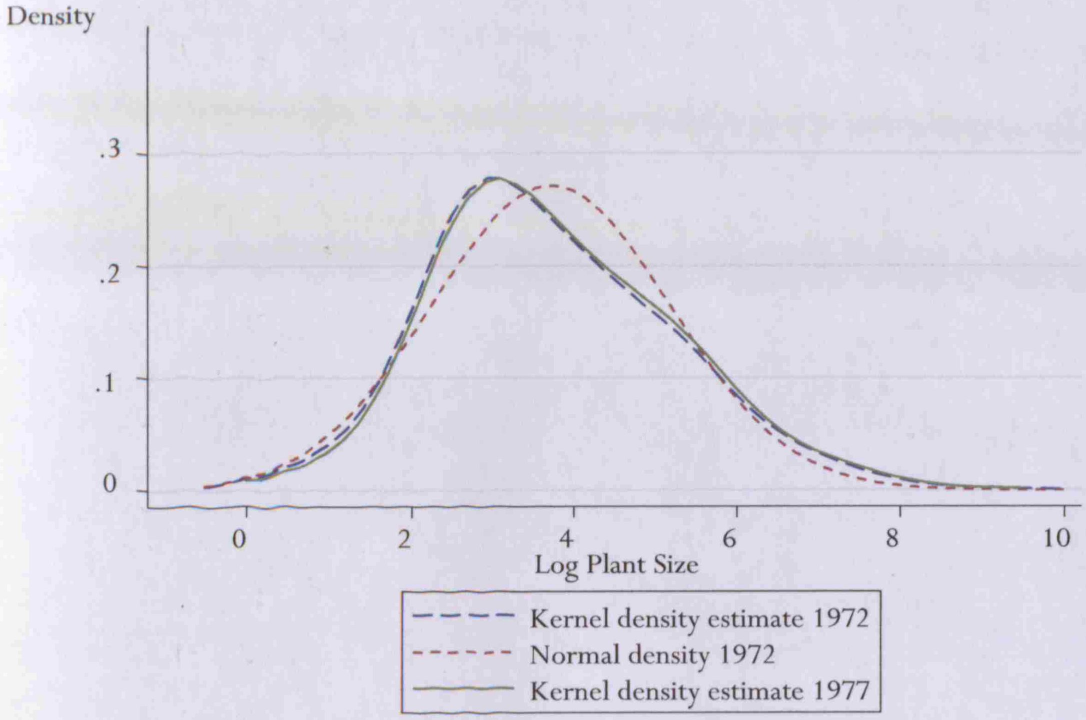


Figure 4.3b: Plant (Log)Size Distribution, 1972-1997

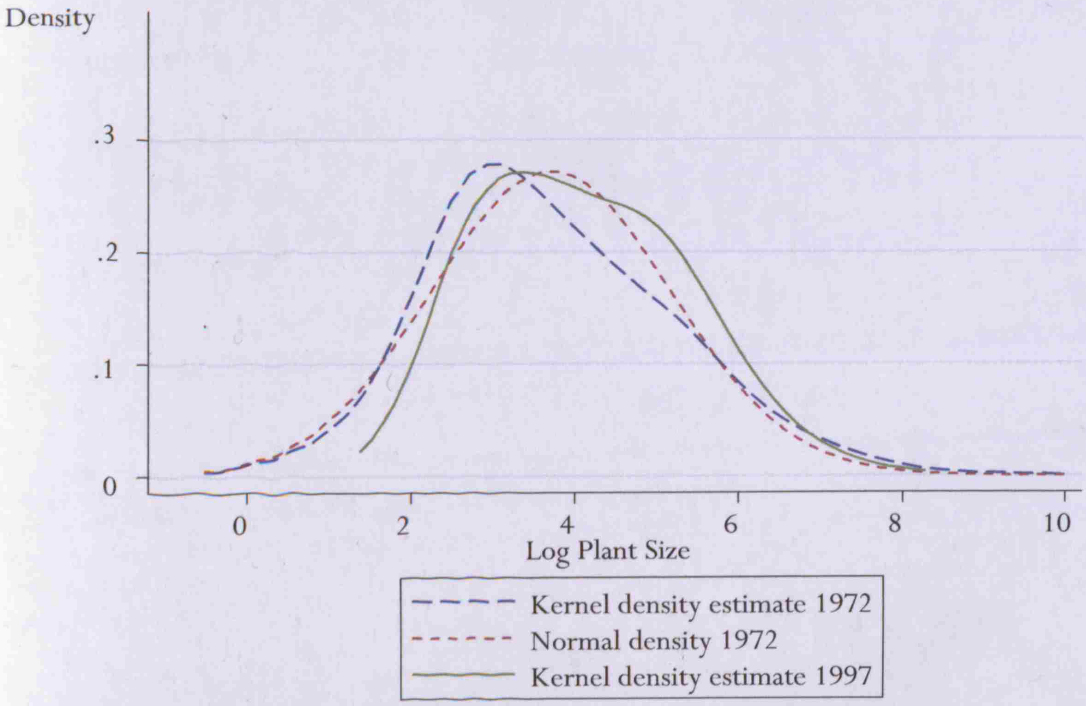


Figure 4.3c: Plant (Log)Size Distribution, 1977-1986

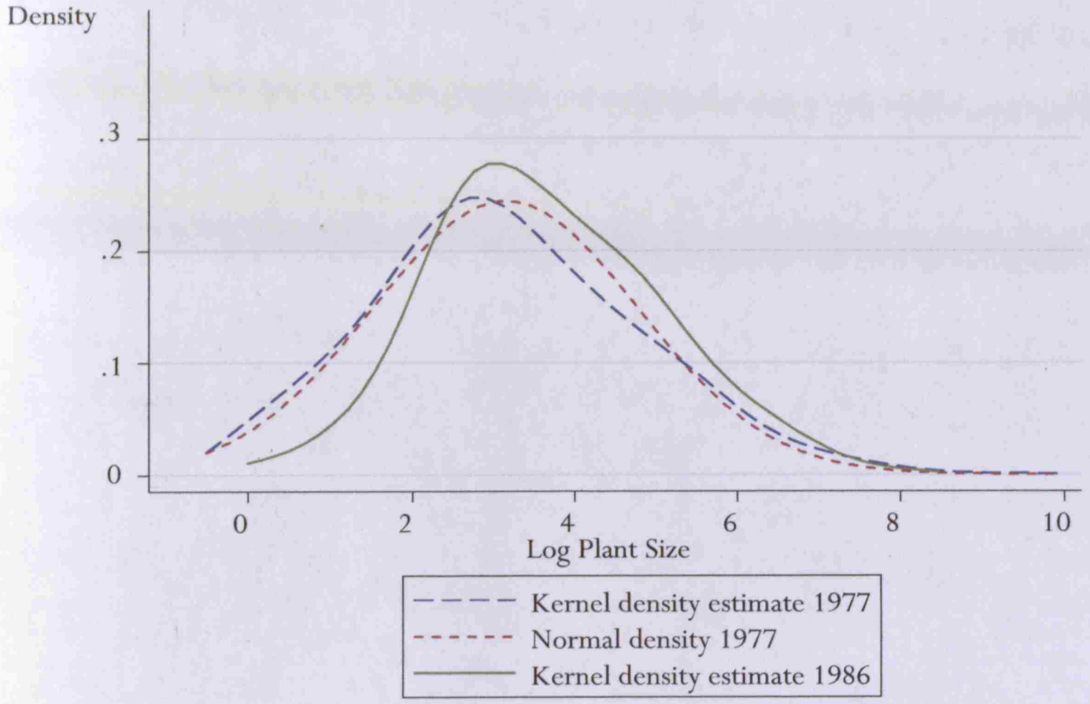


Figure 4.3d: Plant (Log)Size Distribution, 1986-1997

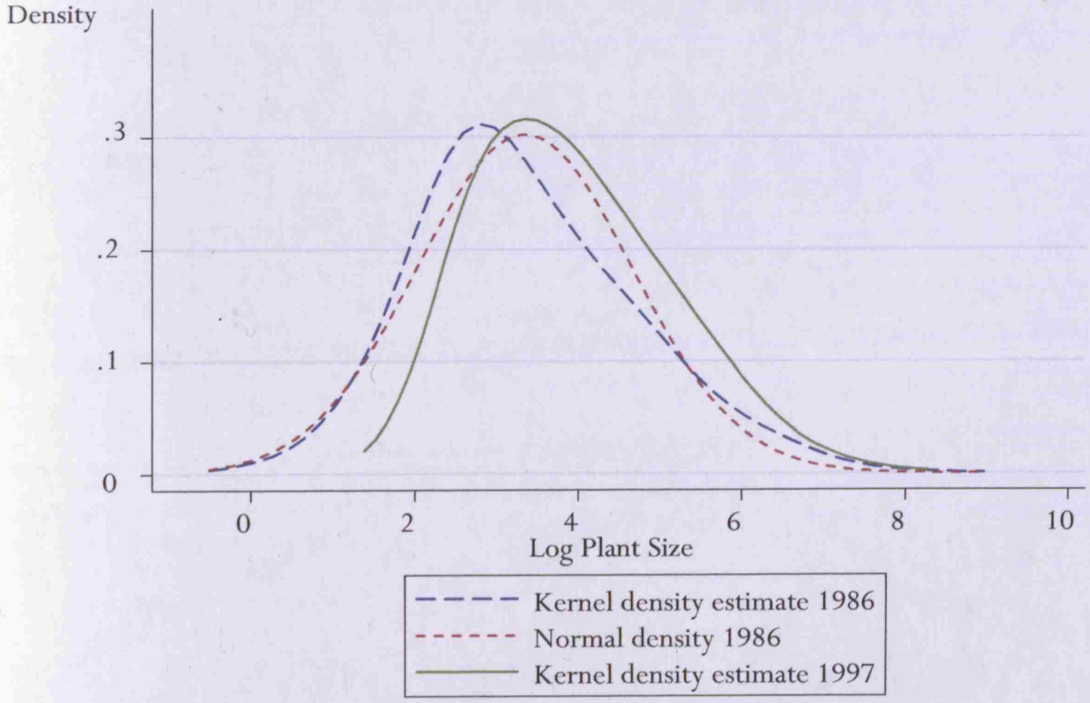


Figure 4.4a: Surviving Plants (Log)Size Distributions, 1972-1977

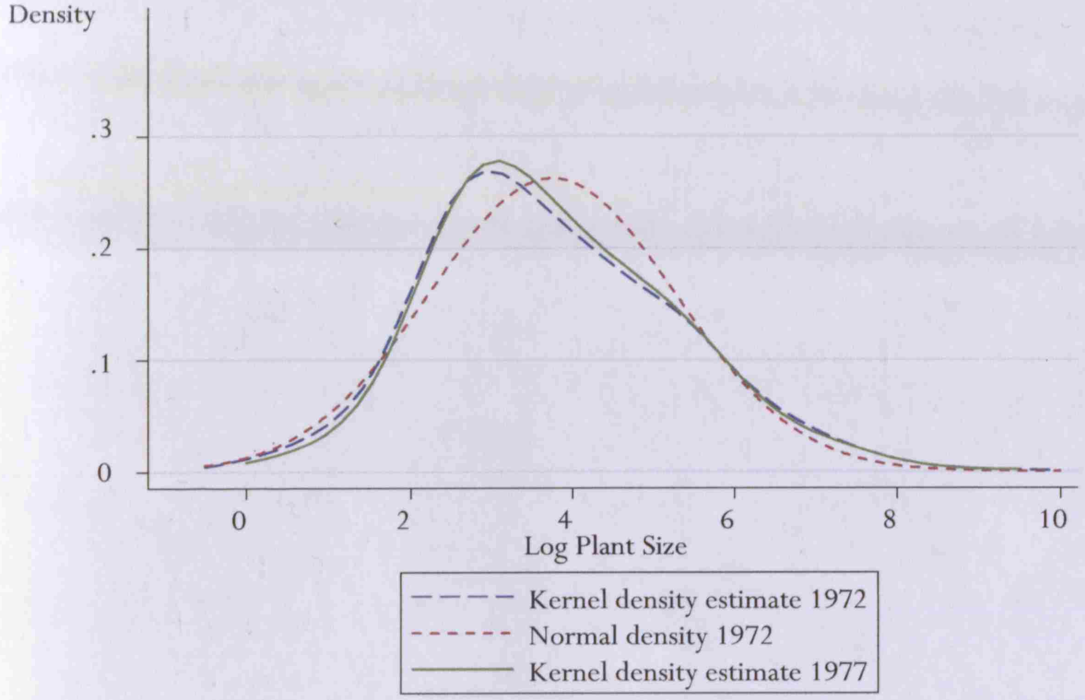


Figure 4.4b: Surviving Plants (Log)Size Distribution, 1972-1997

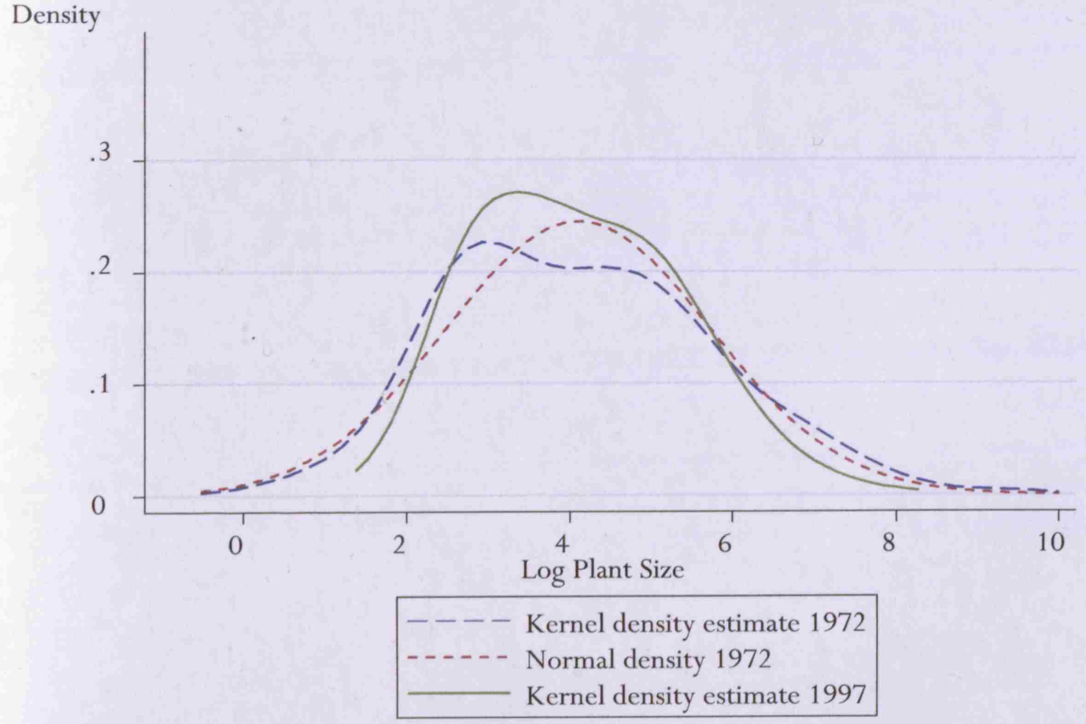


Figure 4.4c: Surviving Plants Log(Size) Distribution, 1977-1986

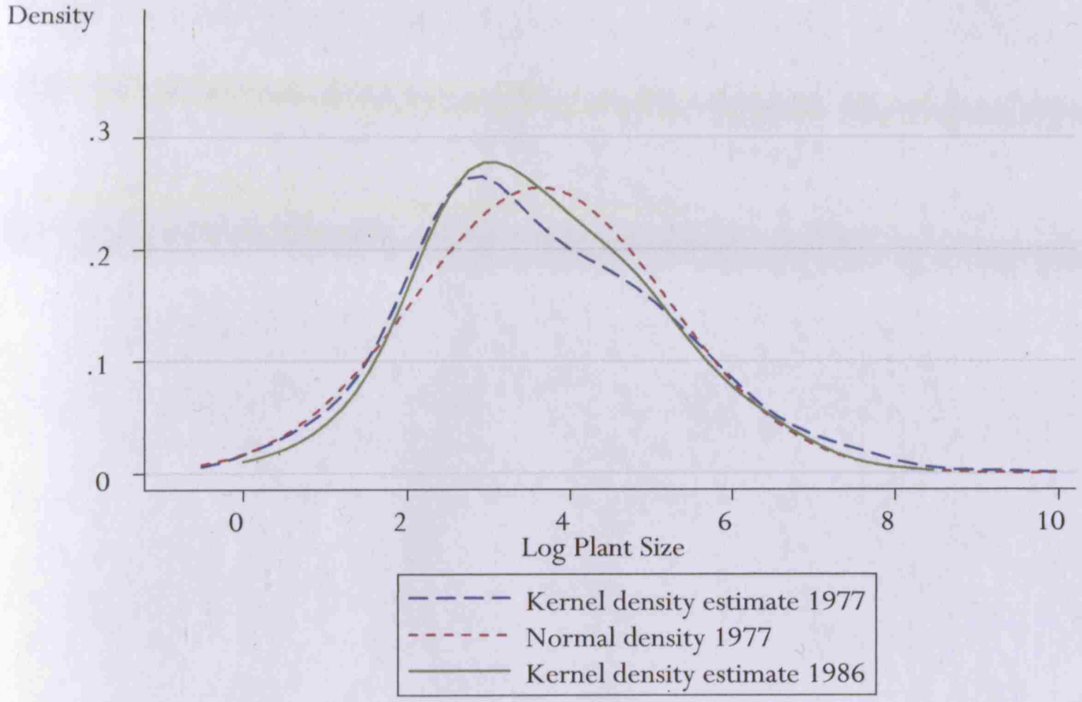
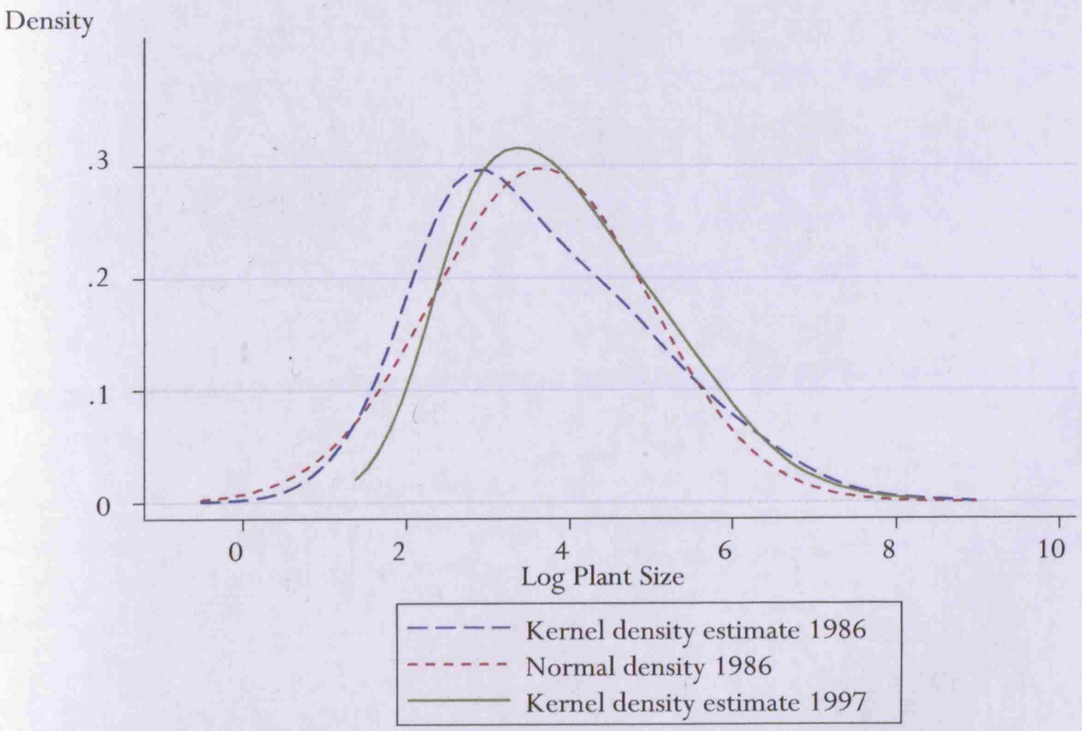


Figure 4.4d: Surviving Plants Log(Size) Distribution, 1986-1997



level information may have further alleviated the potential generation of apparent statistical regularities through aggregation (Bottazzi and Secchi, 2002) such an analysis was not feasible given the generally small sample sizes available.

The evolution of industry size distributions was examined by tracking samples from each of the three cross-sections 1972, 1977 and 1986 over the associated economic cycles, and additionally over the period 1992-1997. Although the initial cross-section would contain plants of different ages the alternative of examining industry-level new entrant cohort samples was not feasible due to the very small samples generated by that approach. Given the number of industries on which data was available, a sample of ten industries was selected for analysis, based on an informal assessment of expected characteristics relating to factors such as technology usage, capital intensity and operating in a relatively new or traditional industry. The chosen industries were Food and beverages, Wearing apparel, Publishing and printing, Chemical and chemical products, Basic metals, Fabricated metal products, Radio, television and communications, Medical and optical instruments, Motor vehicles and trailers, and Furniture.

This analysis of industry size distributions focuses on the calculated skewness and kurtosis indices and the D'Agostino et al. and Shapiro-Wilk tests, given the unsuitability of kernel density estimation in smaller samples. The distribution means and standard deviations are presented in Tables 4.3a-c. In all ten industries the mean plant size (of the surviving plants) increased, with the most rapid increase in mean sizes observed in the Medical and optical instruments, Publishing and printing and Food and beverages industries in both the 1972 and 1977 cross-sections and the Motor vehicles and trailers and Chemical and chemical products industries in the 1986 cross-section.

The standard deviation of plant sizes was also observed to decline across the full time periods of each cross-section in all industry samples, except for Publishing and printing in the 1972 cross-section where a modest increase was detected. Particularly notable decreases in the standard deviation of plant sizes were observed in the Radio, television and communications, Chemical and chemical products and Basic metals industries in the 1972 cross-section, the Radio, television and communications, Fabricated metal products and Chemical and chemical products industries in the 1977 cross-section and the Chemical and chemical products, Furniture, Basic metals and Food and beverages industries in the 1986-1997 cross-sections. Of course, the 1972 cross-section spans the 1977 and 1986 cross-section so the consistency of these results might not be unexpected.

Examination of the skewness indices (shown for all samples in Tables 4.4a-c) for the 1972 cross-section is suggestive of heterogeneity in the observed size distributions across industries over time, an observation consistent with for example, Hymer and Pashigan (1962) and Lotti and Santarelli (2001a, 2001b). Of particular note are the Wearing apparel and Radio, television and communications industries which exhibited a persistent shift towards increased negative skewness indices reaching a maximum of -1.47 for the latter industry in 1984 and -1.25 for the former industry in 1983, before both moderating fairly consistently for the remainder of the period to 1997. In contrast, industries such as Publishing and printing exhibited a general trend towards slightly increased positive skewness.

For the ten 2-digit industries the initial range of the skewness index was -0.35, with only one negative value for Radio, television and communications, to +0.55 in the Basic metals industry. However, by 1997 three industries exhibited indices which were slightly negative (Chemical and chemical products, Radio, television and communications and Motor vehicles and trailers), one of value zero (Furniture) with the indices of the remaining industries being slightly positive, and with corresponding extreme index values of -0.35 (Radio, television and communications) and 0.95 (Publishing and printing). The kurtosis indices, ranging initially between 2.05 (Motor vehicles and trailers) and 3.99

Table 4.3a: Means and Standard Deviations, 1972 Cross-Section

Sample (SIC)	N_i	N_f	1972		1973		1974		1975		1976		1977		1978		1979		1980		1981		1982		1983		1984	
			Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
15: Food and beverages	215	76	3.61	1.32	3.61	1.33	3.60	1.35	3.57	1.33	3.61	1.34	3.60	1.35	3.60	1.36	3.69	1.34	3.71	1.35	3.68	1.34	3.69	1.32	3.69	1.30	3.67	1.31
18: Wearing apparel	116	24	4.17	1.10	4.22	1.08	4.25	1.08	4.25	1.09	4.20	1.09	4.29	1.07	4.31	1.10	4.30	1.11	4.25	1.23	4.11	1.23	4.08	1.27	4.12	1.33	4.36	1.00
22: Publishing and printing	111	43	3.07	1.14	3.07	1.14	3.13	1.14	3.12	1.14	3.10	1.08	3.10	1.12	3.12	1.14	3.13	1.11	3.16	1.11	3.12	1.08	3.06	1.09	3.05	1.12	3.11	1.11
24: Chemicals and chemical products	70	35	4.19	1.76	4.30	1.74	4.28	1.76	4.28	1.79	4.41	1.66	4.40	1.62	4.37	1.63	4.37	1.61	4.30	1.61	4.27	1.58	4.22	1.55	4.19	1.56	4.16	1.54
27: Basic metals	127	52	4.62	1.85	4.59	1.85	4.65	1.84	4.60	1.86	4.62	1.83	4.60	1.88	4.57	1.83	4.54	1.77	4.54	1.75	4.38	1.73	4.36	1.69	4.34	1.72	4.39	1.70
28: Fabricated metal products	278	88	3.38	1.31	3.46	1.30	3.50	1.29	3.42	1.30	3.47	1.25	3.49	1.22	3.51	1.19	3.54	1.19	3.53	1.18	3.40	1.16	3.37	1.16	3.33	1.16	3.33	1.19
32: Radio, T.V. and communications	34	15	4.58	2.07	4.76	2.03	4.77	2.04	4.75	2.01	4.69	2.03	4.72	2.01	4.74	2.04	5.02	1.76	5.03	1.67	5.13	1.70	5.05	1.74	5.06	1.77	5.11	1.67
33: Medical and optical instruments	46	16	3.00	1.54	3.10	1.56	3.16	1.53	3.13	1.52	3.20	1.42	3.21	1.41	3.22	1.36	3.25	1.40	3.32	1.35	3.34	1.28	3.23	1.27	3.30	1.31	3.36	1.40
34: Motor vehicles and trailers	60	31	4.30	1.80	4.43	1.81	4.51	1.78	4.38	1.75	4.43	1.77	4.55	1.79	4.55	1.81	4.59	1.82	4.54	1.78	4.43	1.70	4.48	1.64	4.36	1.67	4.41	1.66
36: Furniture	115	33	3.79	1.36	3.86	1.37	3.92	1.38	3.87	1.39	3.87	1.38	3.86	1.43	3.88	1.42	3.93	1.34	3.97	1.33	3.91	1.31	3.81	1.34	3.71	1.46	3.85	1.39
All (CS)	2,044	693	3.72	1.47	3.77	1.47	3.81	1.48	3.76	1.47	3.78	1.45	3.79	1.47	3.80	1.46	3.84	1.44	3.83	1.44	3.75	1.41	3.74	1.40	3.71	1.43	3.75	1.42
All (S)	693	693	4.11	1.63	4.18	1.62	4.23	1.60	4.19	1.58	4.19	1.56	4.19	1.56	4.21	1.54	4.23	1.51	4.21	1.49	4.16	1.44	4.15	1.41	4.11	1.40	4.11	1.39

Table 4.3a: Means and Standard Deviations, 1972 Cross-Section, continued

Sample (SIC)	1985		1986		1987		1988		1989		1990		1991		1992		1993		1994		1995		1996		1997	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
15: Food and beverages	3.67	1.30	3.75	1.30	3.80	1.23	3.77	1.26	3.85	1.25	3.92	1.23	3.91	1.23	3.94	1.22	3.94	1.23	3.89	1.21	3.94	1.20	3.96	1.21	4.02	1.22
18: Wearing apparel	4.26	1.20	4.43	1.00	4.43	1.08	4.40	1.07	4.45	1.02	4.44	1.05	4.38	1.10	4.44	1.08	4.41	1.07	4.27	1.13	4.28	1.16	4.33	1.09	4.26	1.09
22: Publishing and printing	3.16	1.15	3.23	1.12	3.48	1.04	3.50	1.06	3.53	1.06	3.50	1.09	3.47	1.10	3.49	1.13	3.51	1.13	3.51	1.12	3.53	1.17	3.51	1.18	3.61	1.16
24: Chemicals and chemical products	4.18	1.44	4.13	1.43	4.24	1.34	4.24	1.39	4.32	1.34	4.26	1.35	4.26	1.39	4.26	1.37	4.35	1.32	4.32	1.19	4.34	1.25	4.42	1.26	4.53	1.15
27: Basic metals	4.40	1.71	4.47	1.66	4.59	1.59	4.64	1.63	4.67	1.64	4.70	1.60	4.64	1.58	4.57	1.63	4.56	1.56	4.63	1.56	4.69	1.49	4.75	1.44	4.73	1.44
28: Fabricated metal products	3.39	1.17	3.40	1.11	3.46	1.08	3.51	1.14	3.56	1.17	3.53	1.18	3.57	1.12	3.55	1.13	3.60	1.12	3.59	1.12	3.59	1.16	3.66	1.10	3.69	1.10
32: Radio, T.V. and communications	5.02	1.54	5.13	1.14	5.35	1.15	5.37	1.11	5.42	1.16	5.28	1.20	5.20	1.20	4.92	1.17	4.91	1.16	4.88	1.23	4.98	1.15	4.91	1.14	4.89	1.13
33: Medical and optical instruments	3.35	1.50	3.57	1.34	3.67	1.32	3.59	1.39	3.66	1.37	3.66	1.38	3.47	1.35	3.45	1.35	3.52	1.36	3.60	1.34	3.71	1.31	3.66	1.39	3.82	1.25
34: Motor vehicles and trailers	4.41	1.62	4.39	1.57	4.47	1.56	4.51	1.58	4.56	1.60	4.56	1.58	4.54	1.54	4.48	1.65	4.48	1.62	4.55	1.61	4.65	1.61	4.67	1.61	4.69	1.59
36: Furniture	3.78	1.45	3.88	1.45	4.13	1.30	4.04	1.43	4.12	1.43	4.17	1.42	4.09	1.35	4.01	1.36	3.95	1.35	4.02	1.37	4.08	1.30	4.19	1.16	4.12	1.17
All (CS)	3.75	1.43	3.81	1.39	3.93	1.34	3.94	1.36	3.99	1.36	3.98	1.36	3.96	1.33	3.93	1.34	3.94	1.32	3.97	1.29	4.00	1.29	4.05	1.28	4.10	1.24
All (S)	4.10	1.39	4.11	1.37	4.14	1.35	4.17	1.34	4.18	1.34	4.17	1.33	4.15	1.30	4.12	1.29	4.08	1.26	4.08	1.25	4.09	1.25	4.10	1.25	4.10	1.24

Notes: N_i = initial sample size; N_f = final sample size. Annual sample sizes are provided in Table A4.1a.
Mean and Standard Deviation (S.D) of the logarithm of plant size, measured by employee numbers.
CS = Cross-section; S = Surviving plants.
The sample of 1,754 plants surviving over the period 1972-1977 provided the following descriptive statistics: Mean (log): 3.75, 3.81, 3.84, 3.80, 3.80 and 3.79; Standard Deviation: 1.52, 1.51, 1.50, 1.48, 1.46 and 1.47.

Table 4.3b: Means and Standard Deviations, 1977 Cross-Section

Sample (SIC)	N_i	N_f	1977		1978		1979		1980		1981		1982		1983		1984		1985		1986	
			Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
15: Food and beverages	270	143	3.07	1.53	3.25	1.45	3.53	1.37	3.55	1.38	3.53	1.36	3.55	1.34	3.55	1.33	3.53	1.35	3.62	1.32	3.71	1.32
18: Wearing apparel	127	60	4.02	1.20	4.07	1.23	4.21	1.06	4.18	1.16	4.10	1.16	4.08	1.22	4.08	1.32	4.35	1.01	4.24	1.22	4.43	1.01
22: Publishing and printing	184	93	2.38	1.33	2.63	1.22	2.92	1.14	2.98	1.13	2.96	1.11	2.90	1.14	2.88	1.14	2.95	1.13	3.07	1.13	3.14	1.11
24: Chemicals and chemical products	97	66	3.60	1.89	3.68	1.79	3.86	1.70	3.88	1.65	3.86	1.63	3.83	1.58	3.82	1.61	3.84	1.58	3.88	1.51	3.88	1.45
27: Basic metals	148	81	4.20	1.93	4.22	1.88	4.32	1.77	4.33	1.76	4.20	1.73	4.17	1.71	4.12	1.75	4.19	1.73	4.25	1.71	4.38	1.63
28: Fabricated metal products	402	181	2.85	1.40	3.00	1.32	3.31	1.18	3.32	1.15	3.22	1.10	3.22	1.09	3.21	1.09	3.19	1.14	3.26	1.10	3.28	1.05
32: Radio, T.V. and communications	51	33	4.27	1.87	4.44	1.83	4.63	1.69	4.63	1.63	4.64	1.65	4.60	1.62	4.51	1.62	4.61	1.64	4.61	1.51	4.65	1.36
33: Medical and optical instruments	73	43	3.00	1.41	3.09	1.34	3.19	1.33	3.30	1.34	3.33	1.30	3.27	1.31	3.34	1.33	3.40	1.39	3.40	1.44	3.63	1.31
34: Motor vehicles and trailers	85	50	3.68	1.94	3.82	1.91	4.12	1.88	4.12	1.85	4.06	1.79	4.17	1.73	4.11	1.75	4.20	1.70	4.21	1.70	4.29	1.57
36: Furniture	178	82	3.24	1.57	3.42	1.48	3.74	1.27	3.78	1.26	3.77	1.27	3.72	1.26	3.65	1.37	3.78	1.30	3.73	1.39	3.83	1.39
All (CS)	2,812	1,444	3.15	1.64	3.31	1.55	3.57	1.46	3.58	1.45	3.52	1.42	3.52	1.41	3.49	1.43	3.54	1.43	3.58	1.43	3.64	1.40
All (S)	1,444	1,444	3.64	1.57	3.68	1.54	3.72	1.51	3.72	1.48	3.68	1.43	3.68	1.41	3.65	1.40	3.66	1.40	3.65	1.40	3.64	1.40

Notes: N_i = initial sample size; N_f = final sample size. Annual sample sizes are provided in Table A4.1b.
Mean and Standard Deviation (S.D) of the logarithm of plant size, measured by employee numbers.
CS = Cross-section; S = Surviving plants.

Table 4.3c: Means and Standard Deviations, 1986 Cross-Section

Sample (SIC)	N_i	N_f	1986		1987		1988		1989		1990		1991		1992		1993		1994		1995		1996		1997	
			Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
15: Food and beverages	220	125	3.42	1.39	3.60	1.27	3.63	1.28	3.73	1.27	3.78	1.27	3.82	1.25	3.87	1.25	3.92	1.23	3.92	1.22	3.95	1.19	4.02	1.19	4.04	1.19
18: Wearing apparel	115	46	4.03	1.06	4.12	1.06	4.19	1.06	4.24	1.05	4.27	1.04	4.17	1.19	4.26	1.14	4.35	1.08	4.34	1.07	4.41	1.05	4.43	1.01	4.40	1.02
22: Publishing and printing	138	72	2.94	1.08	3.18	1.02	3.24	1.02	3.31	1.02	3.30	1.03	3.29	1.01	3.32	1.03	3.34	1.04	3.40	1.01	3.45	1.04	3.44	1.05	3.48	1.03
24: Chemicals and chemical products	126	88	3.41	1.38	3.62	1.30	3.72	1.25	3.81	1.25	3.81	1.24	3.83	1.25	3.87	1.22	3.94	1.19	3.94	1.13	4.01	1.13	4.05	1.14	4.13	1.09
27: Basic metals	119	83	4.00	1.62	4.18	1.52	4.23	1.54	4.26	1.55	4.30	1.52	4.27	1.50	4.25	1.51	4.32	1.44	4.37	1.43	4.41	1.39	4.47	1.37	4.45	1.38
28: Fabricated metal products	310	190	3.05	0.97	3.17	0.95	3.21	0.98	3.32	0.98	3.34	0.98	3.36	0.94	3.35	0.96	3.37	0.94	3.42	0.93	3.44	0.96	3.50	0.92	3.51	0.92
32: Radio, T.V. and communications	67	46	4.12	1.50	4.32	1.51	4.41	1.52	4.46	1.56	4.45	1.54	4.41	1.57	4.25	1.67	4.29	1.62	4.49	1.53	4.57	1.52	4.60	1.48	4.58	1.48
33: Medical and optical instruments	76	46	3.35	1.28	3.47	1.26	3.54	1.28	3.61	1.27	3.64	1.24	3.68	1.25	3.69	1.26	3.69	1.24	3.72	1.22	3.80	1.20	3.83	1.23	3.91	1.16
34: Motor vehicles and trailers	78	45	3.83	1.66	3.97	1.61	4.04	1.63	4.15	1.64	4.19	1.62	4.34	1.55	4.32	1.63	4.43	1.54	4.54	1.51	4.62	1.51	4.73	1.47	4.75	1.46
36: Furniture	163	86	3.41	1.31	3.62	1.21	3.70	1.24	3.78	1.24	3.79	1.20	3.78	1.20	3.73	1.26	3.80	1.21	3.91	1.22	3.98	1.16	4.04	1.08	4.02	1.09
All (CS)	2,505	1,417	3.35	1.32	3.52	1.27	3.59	1.28	3.66	1.28	3.68	1.27	3.71	1.23	3.72	1.24	3.74	1.23	3.79	1.20	3.84	1.20	3.89	1.18	3.93	1.16
All (S)	1,417	1,417	3.65	1.34	3.74	1.30	3.81	1.29	3.85	1.28	3.87	1.26	3.88	1.24	3.87	1.22	3.85	1.19	3.88	1.17	3.91	1.16	3.93	1.16	3.93	1.16

Notes: N_i = initial sample size; N_f = final sample size. Annual sample sizes are provided in Table A4.1c.
Mean and Standard Deviation (S.D) of the logarithm of plant size, measured by employee numbers.
CS = Cross-section; S = Surviving plants.

(Publishing and printing), suggested some moderation in reported values between the initial and final year, with the corresponding minimum and maximum values being 1.82 (Motor vehicles and trailers) and 3.01 (Furniture) in the final year. Again, the Wearing apparel and Radio, television and communications industries exhibited some volatility in their kurtosis indices around the mid-1980s, both peaking at around values of 5, being leptokurtic, before converging to values below 2 (platykurtic) by 1997.

Analysis of the 1977 cross-section provided two industries exhibiting negative skewness indices in the initial year (Wearing apparel and Radio, television and communications), the maximum negative value being -0.55 with the highest positive skewness index being 0.60 in the Basic metals industry. In the final year these minimum and maximum values were -0.62, again in the Wearing apparel industry, and 0.63 in the Medical and optical instruments industry. The Wearing apparel and Radio, television and communications industries again exhibited negative skewness throughout most of the time period, although in the case of the latter by 1986 this had become very slightly positively skewed. In contrast, the Furniture industry started as initially slightly positively skewed in 1977 before becoming slightly negatively skewed in 1986. In both the start and final periods of the panel, and in fact through most of the period for the majority of industry samples, the kurtosis indices remained broadly around the value 3, consistent with the (log)normal distribution.

A similar pattern also emerged for the 1986 cross-section, with only one negative skewness index value in the initial year, -0.13 in Wearing apparel. The maximum positive skewness value was 0.77 in the Fabricated metal products industry. In the final year two index values were negative, again in the Wearing apparel industry, but also in the Motor vehicles and trailers industry. Although with some variation, the reported kurtosis indices ranging between 3.90 and 1.91 throughout the period are not entirely inconsistent with the (log)normal distribution given the relatively small sample sizes.

Table 4.4a: Skewness and Kurtosis Indices, 1972 Cross-Section

Sample (SIC)	N_i	N_f	1972		1973		1974		1975		1976		1977		1978		1979		1980		1981		1982		1983		1984	
			Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt
15: Food and beverages	215	76	0.16	2.81	0.25	2.67	0.25	2.71	0.23	2.68	0.27	2.77	0.18	2.76	0.22	2.72	0.24	2.61	0.18	2.40	0.14	2.43	0.13	2.54	0.21	2.31	0.29	2.33
18: Wearing apparel	116	24	0.20	2.72	0.16	2.57	0.13	2.67	0.17	2.82	-0.07	2.72	-0.16	2.57	-0.25	2.78	-0.36	2.82	-0.81	3.62	-0.83	3.39	-1.12	4.33	-1.25	4.69	-0.59	2.98
22: Publishing and printing	111	43	0.29	3.99	0.36	4.11	0.30	3.97	0.27	3.95	0.31	4.40	0.26	4.08	0.29	3.90	0.43	4.02	0.44	4.08	0.48	4.30	0.52	4.27	0.61	3.72	0.69	3.62
24: Chemicals and chemical products	70	35	0.29	2.42	0.22	2.49	0.20	2.42	0.08	2.41	0.16	2.48	0.16	2.50	0.18	2.54	0.23	2.52	0.25	2.55	0.18	2.57	0.13	2.52	-0.19	2.78	-0.21	2.80
27: Basic metals	127	52	0.55	2.74	0.58	2.75	0.58	2.73	0.60	2.80	0.59	2.78	0.59	2.68	0.55	2.83	0.70	2.93	0.61	2.60	0.57	2.50	0.66	2.57	0.51	2.46	0.43	2.53
28: Fabricated metal products	278	88	0.33	3.10	0.32	3.04	0.40	2.95	0.36	3.08	0.36	3.15	0.44	3.11	0.54	3.10	0.51	3.04	0.43	3.14	0.42	3.02	0.49	3.05	0.38	3.34	0.30	3.17
32: Radio, T.V. and communications	34	15	-0.35	2.11	-0.50	2.28	-0.50	2.24	-0.69	2.48	-0.83	2.76	-0.89	2.96	-0.86	2.82	-1.09	4.00	-1.03	4.29	-1.29	4.85	-1.20	4.48	-1.24	4.41	-1.47	5.40
33: Medical and optical instruments	46	16	0.17	2.75	0.15	2.87	0.16	3.08	0.17	3.21	0.20	2.99	0.15	2.93	0.29	2.93	0.25	2.73	0.35	2.51	0.40	2.70	0.60	2.78	0.45	2.75	0.49	2.67
34: Motor vehicles and trailers	60	31	0.38	2.05	0.34	1.96	0.31	1.90	0.27	2.10	0.13	2.31	0.08	2.33	0.07	2.34	0.11	2.27	0.23	2.01	0.34	1.91	0.32	1.92	0.21	2.02	0.29	1.84
36: Furniture	115	33	0.03	2.83	0.07	2.95	0.14	2.77	0.10	2.72	0.23	2.52	0.05	2.87	0.12	2.65	0.34	2.60	0.32	2.65	0.25	2.62	0.19	2.68	0.05	2.61	0.05	2.06
All (CS)	2,044	693	0.45	3.23	0.47	3.19	0.47	3.14	0.47	3.21	0.47	3.21	0.47	3.21	0.48	3.17	0.51	3.14	0.48	3.04	0.47	3.05	0.46	3.05	0.37	2.94	0.37	2.86
All (S)	693	693	0.34	2.87	0.34	2.82	0.37	2.80	0.39	2.86	0.40	2.89	0.41	2.87	0.46	2.88	0.50	2.82	0.50	2.69	0.49	2.69	0.50	2.66	0.47	2.58	0.43	2.54

Table 4.4a: Skewness and Kurtosis Indices, 1972 Cross-Section, continued

Sample (SIC)	1985		1986		1987		1988		1989		1990		1991		1992		1993		1994		1995		1996		1997	
	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt
15: Food and beverages	0.28	2.20	0.26	2.22	0.37	2.20	0.28	2.35	0.21	2.34	0.32	2.19	0.36	2.29	0.19	2.22	0.21	2.13	0.24	2.18	0.34	2.24	0.35	2.27	0.31	2.13
18: Wearing apparel	-1.15	5.13	-0.58	3.34	-0.71	3.59	-0.70	3.42	-0.68	3.31	-0.68	2.97	-0.50	2.57	-0.60	2.94	-0.48	2.86	-0.23	2.43	-0.23	2.21	0.06	1.86	0.21	1.94
22: Publishing and printing	0.53	3.32	0.51	3.38	0.73	2.99	0.69	2.86	0.64	2.83	0.69	2.80	0.74	2.94	0.73	3.12	0.75	3.34	0.72	3.35	0.67	3.05	0.69	2.86	0.95	2.85
24: Chemicals and chemical products	0.04	2.27	0.02	2.53	0.16	2.31	-0.09	2.77	0.14	2.20	0.20	2.12	0.05	2.17	0.03	2.03	0.04	1.96	0.01	1.93	-0.10	2.01	-0.35	2.32	-0.28	2.12
27: Basic metals	0.42	2.47	0.49	2.53	0.56	2.56	0.48	2.40	0.45	2.29	0.48	2.44	0.46	2.50	0.33	2.56	0.49	2.60	0.35	2.79	0.50	2.79	0.54	2.81	0.56	2.77
28: Fabricated metal products	0.36	2.97	0.47	3.05	0.40	3.10	0.39	2.96	0.25	2.89	0.13	3.09	0.28	2.81	0.24	2.69	0.17	2.78	0.34	2.88	0.30	2.58	0.40	2.50	0.40	2.48
32: Radio, T.V. and communications	-1.11	4.23	-0.31	2.26	-0.63	2.51	-0.42	2.27	-0.55	2.17	-0.41	2.07	-0.52	2.28	-0.14	2.12	-0.18	2.04	-0.43	2.08	-0.44	2.05	-0.41	1.87	-0.35	1.89
33: Medical and optical instruments	0.36	2.32	0.47	2.20	0.36	2.14	0.44	1.99	0.35	1.96	0.29	1.83	0.56	2.08	0.43	1.75	0.26	1.68	0.07	1.62	-0.01	1.96	0.01	2.05	0.24	1.86
34: Motor vehicles and trailers	0.30	1.93	0.33	1.96	0.25	1.86	0.18	1.79	0.15	1.74	0.13	1.81	0.17	1.85	-0.15	2.14	-0.09	1.91	-0.08	1.78	-0.13	1.77	-0.12	1.74	-0.17	1.82
36: Furniture	-0.18	2.47	-0.25	2.63	-0.07	2.16	-0.33	2.91	-0.31	2.48	-0.35	2.53	-0.35	2.73	-0.38	2.73	-0.30	2.78	-0.40	2.87	-0.43	3.28	-0.10	2.95	0.00	3.01
All (CS)	0.34	2.80	0.32	2.90	0.37	2.90	0.32	2.86	0.30	2.81	0.29	2.87	0.33	2.85	0.25	2.84	0.24	2.82	0.30	2.82	0.29	2.79	0.27	2.75	0.38	2.62
All (S)	0.40	2.57	0.36	2.65	0.39	2.59	0.37	2.54	0.36	2.54	0.36	2.56	0.39	2.62	0.36	2.68	0.40	2.65	0.40	2.70	0.39	2.66	0.38	2.64	0.38	2.62

Notes: N_i = initial sample size; N_f = final sample size. Annual sample sizes are provided in Table A4.1a.
Skewness (Skew) and Kurtosis (Kurt) indices of the logarithm of plant size, measured by employee numbers.
CS = Cross-section; S = Surviving plants.
The sample of 1,754 plants surviving over the period 1972-1977 provided Skewness indices: 0.42, 0.44, 0.47, 0.47, 0.47 and 0.47; and Kurtosis indices: 3.12, 3.10, 3.08, 3.15, 3.18 and 3.21.

Table 4.4b: Skewness and Kurtosis Indices, 1977 Cross-Section

Sample (SIC)	N_i	N_f	1977		1978		1979		1980		1981		1982		1983		1984		1985		1986	
			Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt
15: Food and beverages	270	143	0.18	2.55	0.29	2.56	0.29	2.54	0.25	2.38	0.22	2.39	0.21	2.48	0.29	2.34	0.35	2.32	0.30	2.18	0.26	2.23
18: Wearing apparel	127	60	-0.55	3.78	-0.68	3.98	-0.20	2.73	-0.64	3.49	-0.77	3.47	-1.05	4.22	-1.16	4.24	-0.61	2.90	-1.15	4.74	-0.62	3.28
22: Publishing and printing	184	93	0.31	2.94	0.44	3.33	0.42	3.70	0.44	3.81	0.42	3.90	0.37	3.91	0.50	3.72	0.58	3.60	0.56	3.39	0.50	3.39
24: Chemicals and chemical products	97	66	0.24	2.37	0.39	2.44	0.43	2.53	0.46	2.64	0.38	2.56	0.35	2.53	0.01	2.63	-0.06	2.72	0.03	2.60	0.03	2.93
27: Basic metals	148	81	0.60	2.86	0.53	2.93	0.70	3.16	0.60	2.80	0.56	2.69	0.60	2.75	0.47	2.58	0.40	2.61	0.43	2.53	0.53	2.62
28: Fabricated metal products	402	181	0.24	3.02	0.30	3.13	0.59	3.22	0.57	3.33	0.58	3.42	0.60	3.48	0.47	3.62	0.41	3.35	0.48	3.24	0.59	3.24
32: Radio, T.V. and communications	51	33	-0.51	2.52	-0.58	2.80	-0.66	3.22	-0.53	3.13	-0.50	3.04	-0.43	3.11	-0.32	3.07	-0.49	3.14	-0.27	2.78	0.08	2.12
33: Medical and optical instruments	73	43	0.33	2.86	0.47	2.88	0.47	2.82	0.58	2.62	0.65	2.74	0.78	2.85	0.69	2.84	0.64	2.70	0.51	2.44	0.63	2.47
34: Motor vehicles and trailers	85	50	0.47	2.33	0.41	2.34	0.25	2.38	0.40	2.07	0.39	2.05	0.33	2.02	0.24	2.01	0.33	1.93	0.12	2.35	0.36	1.97
36: Furniture	178	82	0.07	2.54	0.06	2.65	0.38	2.79	0.31	2.85	0.21	2.62	0.14	2.74	0.03	2.60	0.02	2.19	-0.22	2.41	-0.26	2.55
All (CS)	2,812	1,444	0.35	2.98	0.41	3.10	0.47	3.25	0.46	3.15	0.46	3.13	0.44	3.12	0.40	2.99	0.37	2.92	0.34	2.87	0.30	2.99
All (S)	1,444	1,444	0.41	3.10	0.46	3.16	0.49	3.14	0.49	3.06	0.48	3.05	0.47	3.06	0.44	2.95	0.41	2.89	0.36	2.90	0.30	2.99

Notes: N_i = initial sample size; N_f = final sample size. Annual sample sizes are provided in Table A4.1b.
Skewness (Skew) and Kurtosis (Kurt) indices of the logarithm of plant size, measured by employee numbers.
CS = Cross-section; S = Surviving plants.

Table 4.4c: Skewness and Kurtosis Indices, 1986 Cross-Section

Sample (SIC)	N_i	N_f	1986		1987		1988		1989		1990		1991		1992		1993		1994		1995		1996		1997	
			Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt
15: Food and beverages	220	125	0.23	2.44	0.33	2.45	0.27	2.47	0.25	2.36	0.26	2.35	0.33	2.34	0.18	2.27	0.24	2.22	0.16	2.16	0.32	2.22	0.35	2.16	0.34	2.09
18: Wearing apparel	115	46	-0.13	2.30	-0.27	2.50	-0.34	2.56	-0.30	2.41	-0.28	2.33	-0.47	2.65	-0.56	2.83	-0.41	2.35	-0.35	2.34	-0.36	2.40	-0.11	1.94	-0.01	1.91
22: Publishing and printing	138	72	0.61	3.49	0.69	3.45	0.74	3.29	0.68	3.20	0.75	3.26	0.88	3.43	0.90	3.65	0.93	3.84	0.88	3.90	0.90	3.65	0.98	3.71	1.18	3.63
24: Chemicals and chemical products	126	88	0.43	2.93	0.37	3.01	0.46	2.91	0.54	2.68	0.53	2.64	0.48	2.54	0.42	2.44	0.42	2.39	0.37	2.29	0.32	2.33	0.20	2.37	0.22	2.25
27: Basic metals	119	83	0.61	2.79	0.73	2.97	0.64	2.81	0.62	2.69	0.63	2.79	0.63	2.86	0.51	2.84	0.60	2.96	0.50	3.08	0.61	3.06	0.64	3.04	0.62	3.01
28: Fabricated metal products	310	190	0.77	3.86	0.63	3.73	0.66	3.73	0.56	3.56	0.46	3.63	0.61	3.40	0.55	3.30	0.51	3.49	0.66	3.60	0.53	3.34	0.69	3.15	0.71	3.16
32: Radio, T.V. and communications	67	46	0.16	2.24	0.08	2.19	0.06	2.21	0.14	2.13	0.19	2.30	0.17	2.29	0.03	2.73	-0.23	2.86	0.01	2.43	-0.05	2.57	0.06	2.50	0.20	2.28
33: Medical and optical instruments	76	46	0.67	2.79	0.64	2.73	0.66	2.61	0.56	2.53	0.52	2.53	0.46	2.43	0.39	2.33	0.27	2.32	0.11	2.51	0.06	2.79	-0.06	2.82	0.19	2.71
34: Motor vehicles and trailers	78	45	0.44	2.33	0.42	2.24	0.36	2.09	0.32	2.00	0.29	2.05	0.22	2.12	-0.03	2.16	-0.03	2.08	-0.05	1.99	-0.08	1.96	-0.11	1.98	-0.16	2.07
36: Furniture	163	86	0.13	2.49	0.16	2.44	-0.02	2.68	-0.01	2.47	0.02	2.56	0.01	2.59	-0.23	2.88	-0.13	2.65	-0.31	2.88	-0.21	2.84	0.08	2.49	0.12	2.51
All (CS)	2,505	1,417	0.50	3.24	0.50	3.29	0.46	3.21	0.44	3.17	0.41	3.27	0.55	3.19	0.45	3.14	0.42	3.09	0.45	3.09	0.45	3.13	0.48	3.08	0.58	2.95
All (S)	1,417	1,417	0.57	2.92	0.58	2.93	0.54	2.87	0.54	2.86	0.55	2.93	0.56	2.95	0.55	2.96	0.55	2.96	0.57	2.98	0.59	2.99	0.58	2.98	0.58	2.95

Notes: N_i = initial sample size; N_f = final sample size. Annual sample sizes are provided in Table A4.1c.
Skewness (Skew) and Kurtosis (Kurt) indices of the logarithm of plant size, measured by employee numbers.
CS = Cross-section; S = Surviving plants.

Throughout the periods of analysis negative skewness indices were observed in 71 out of 260 industry samples (26 in the Radio, television and communications equipment industry, 20 in the Wearing apparels industry and 12 in Furniture) in the 1972 cross-section, 22 out of 100 industry samples in the 1977 cross-section (10 in Wearing apparel and 9 in Radio, television and communications equipment) and 27 out of 120 industry samples (12 in Wearing apparel) in the 1986 cross-section, with the remainder being zero or positive. Over these time periods kurtosis indices above 3 were found in 53 industry samples in the 1972 cross-section (19 in Publishing and printing and 14 in Fabricated metal products), 34 samples in the 1977 cross- section (10 in Fabricated metal products, 9 in Publishing and printing, 8 in Wearing apparel and 6 in Radio, television and communications equipment) and 29 (with 12 in both Publishing and printing and Fabricated metal products) industry samples in the 1986 cross-section, with the remainder being exhibiting index values of 3 or below.

For the 1972 cross-section the D'Agostino et al. test (the statistical tests for this and the other samples are presented in Tables 4.5a-c) of normality rejected the (log)normal in 64 of the 260 samples (24.6 percent) at the 5 percent level of significance. The most frequent rejections of lognormality were observed in Motor vehicles and trailers (16 rejections), Publishing and printing (12 rejections), Basic metals (10 rejections) and Food and beverages (8 rejections). For the Chemical and chemical products, Medical and optical instruments and Furniture industries lognormality could not be rejected in any of the 26 industry-level samples. The tests additionally suggest some evolution towards the lognormal distribution for the Basic metals industry with no reported rejections after 1982 (the null hypothesis of normality being rejected in all but one year prior to this). For the other industry samples there was a mixed picture with lognormality being rejected in some years but not others. However, this statistical test failed to reject the lognormal hypothesis in at least 8 of the 10 samples (although rejection rates were only slightly higher until the late 1970s) in the period from 1988. There is some suggestion of increased rejection of lognormality at the industry-level during the late 1970s to mid-1980s, but only in one year, 1982, could more than half of the industry-level samples be rejected.

For the same sample and time period the Shapiro-Wilk test provided a notably higher rejection rate of the (log)normal distribution at 103 cases (39.6 percent). Again, an increased rate of rejection was observed for the early 1980s, peaking at 7 out of 10 rejections in 1981, occurring during a time period of U.K. recession. However, no more than 2 out of the 10 annual industry-level samples produced rejections of (log)normality as a reasonable description of the plant size distributions from 1990 onwards. Given the higher rates of rejection (around half of the samples prior to this point, including in the early and mid-1970s) the Shapiro-Wilk test provides perhaps some support for a more general convergence of industry cross-sections towards a lognormal distribution of plant sizes (at least for this time period) although is not sufficiently conclusive to suggest that this distribution emerges as a long-run stable distribution.

Consistent with the results of the D'Agostino et al. test, the Shapiro-Wilk test did not reject (log)normality in any of the 26 industry-level samples for the Chemical and chemical products and Furniture industries, with only one rejection in the Medical and optical instruments industry. Industries including the Basic metals and the Fabricated metal products and Motor vehicles and trailers industries did, however, show some evidence of convergence towards the lognormal. The Radio, television and communications and Wearing apparel industries also exhibited no rejection of the lognormal distribution from 1985 and 1989 onwards, respectively. However, there is no strong evidence of convergence to the lognormal distribution in the Food and beverages or Publishing and printing industries.

For the 1977 cross-section tracked until 1986, the D'Agostino et al. test rejected (log)normality in 42 of the 100 annual industry-level samples, with 58 rejections using the Shapiro-Wilk test. In the case of the former test, lognormality could not be rejected in any of the ten annual samples in Chemical and chemical products, Radio, television and communications, Medical and optical instruments and the Furniture industries. In contrast, lognormality was rejected in at least half of the annual samples in the Publishing and printing (8 rejections), Fabricated metal products (8 rejections), Food and beverages (7 rejections), Wearing apparel (7 rejections), Motor vehicles and trailers (6 rejections) and Basic metals (6 rejections) industries, but in the latter case with some tentative evidence of less significant rejections over time.

The Shapiro-Wilk test did not reject lognormality in any year for the Radio, television and communications industry, with lognormality being rejected in more than half of the samples in all of the remaining industries except for the Publishing (a very different result compared to the D'Agostino et al. test) and Chemical and chemical products industries, with rejections in all years in the case of the Food and beverages and Fabricated metal products industries.

Similar results emerged for the 1986 cross-section with 48 rejections out of 120 samples using the D'Agostino et al. test. Rejections of the lognormal distribution were reported in all years for the Food and beverages, Publishing and printing, and Fabricated metal products industries with no rejections for the Chemical and chemical products, Radio, television and communications, Medical and optical instruments and Furniture industries. There is some evidence of potential convergence towards the lognormal size distribution for the Motor vehicles and trailers and Basic metals industries, with no rejections at the 5 percent level of significance from 1992 onwards. In contrast the Wearing apparel industry exhibited rejections in the final two periods, following no previous rejections from 1986. However, it cannot be ruled out that such an observation may be only temporary (no such rejections are observed using the Shapiro-Wilk test).

Table 4.5a: Normality Tests, 1972 Cross-Section

Sample (SIC)	N_i	N_f	1972		1973		1974		1975		1976		1977		1978		1979		1980		1981		1982		1983		1984	
			D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W
15: Food and beverages	215	76	1.10	0.98*	3.28	0.98*	2.70	0.98*	2.64	0.98*	2.57	0.98*	1.36	0.98*	1.81	0.98*	2.68	0.98*	4.68	0.98*	3.58	0.98*	1.86	0.98	5.97	0.98	6.09*	0.98*
18: Wearing apparel	116	24	1.08	0.99	1.38	0.99	0.69	0.99	0.58	0.99	0.23	0.99	1.05	0.98	1.01	0.98	2.01	0.98	9.93*	0.95*	8.84*	0.95*	15.81*	0.92*	18.61*	0.90*	3.67	0.96*
22: Publishing and printing	111	43	5.46	0.97*	6.71*	0.96*	5.30	0.97*	4.92	0.97*	7.29*	0.96*	5.27	0.97*	4.65	0.97*	6.79*	0.97*	7.24*	0.97*	8.63*	0.96*	8.86*	0.95*	7.91*	0.96*	8.69*	0.96*
24: Chemicals and chemical products	70	35	2.27	0.98	1.31	0.98	1.60	0.98	1.17	0.98	0.97	0.98	0.88	0.99	0.77	0.98	1.07	0.98	1.11	0.98	0.64	0.98	0.59	0.98	0.38	0.98	0.48	0.98
27: Basic metals	127	52	6.53*	0.97*	7.16*	0.96*	6.93*	0.96*	7.22*	0.96*	6.83*	0.96*	6.99*	0.96*	5.72	0.97*	8.40*	0.95*	6.86*	0.95*	6.10*	0.96*	6.67*	0.94*	4.62	0.96*	3.21	0.97
28: Fabricated metal products	278	88	5.49	0.99*	4.83	0.99*	7.10*	0.98*	5.72	0.98*	5.75	0.98*	7.80*	0.98*	10.72*	0.97*	9.35*	0.98*	6.95*	0.98*	5.83	0.98*	7.02*	0.97*	5.09	0.98*	2.76	0.99
32: Radio, T.V. and communications	34	15	2.78	0.95	2.49	0.95	2.67	0.95	3.12	0.92*	4.02	0.91*	4.67	0.91*	4.26	0.91*	8.34*	0.92*	8.42*	0.93	11.28*	0.90*	9.44*	0.91	9.48*	0.90*	13.63*	0.87*
33: Medical and optical instruments	46	16	0.27	0.98	0.28	0.98	0.56	0.97	0.83	0.97	0.54	0.98	0.31	0.98	0.84	0.98	0.51	0.98	1.19	0.98	1.22	0.98	2.49	0.96	1.39	0.97	1.51	0.97
34: Motor vehicles and trailers	60	31	7.23*	0.95*	9.26*	0.95*	10.93*	0.95*	5.19	0.96	1.71	0.97	1.38	0.97	1.28	0.97	1.82	0.95*	5.88	0.95*	9.35*	0.94*	8.25*	0.94*	5.22	0.96	10.03*	0.94*
36: Furniture	115	33	0.04	0.99	0.12	0.99	0.48	0.99	0.34	0.99	1.90	0.98	0.05	0.98	0.52	0.98	2.24	0.97	1.66	0.97	1.20	0.97	0.61	0.98	0.22	0.98	5.13	0.97
All (CS)	2,044	693	68.66*	0.99*	69.50*	0.98*	67.67*	0.98*	67.21*	0.98*	64.99*	0.98*	61.07*	0.98*	59.98*	0.98*	64.09*	0.98*	55.02*	0.98*	49.62*	0.98*	44.57*	0.98*	28.10*	0.99*	26.80*	0.99*
All (S)	693	693	13.47*	0.99*	14.22*	0.99*	16.64*	0.99*	17.30*	0.99*	17.52*	0.99*	18.93*	0.98*	23.03*	0.98*	27.46*	0.97*	29.82*	0.97*	29.39*	0.97*	31.31*	0.97*	31.04*	0.97*	30.16*	0.97*

Table 4.5a: Normality Tests, 1972 Cross-Section, continued

Sample (SIC)	1985		1986		1987		1988		1989		1990		1991		1992		1993		1994		1995		1996		1997	
	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W
15: Food and beverages	9.33*	0.97*	8.04*	0.97*	9.64*	0.96*	4.84	0.98*	4.06	0.98	7.83*	0.97*	5.92	0.97*	5.51	0.98	7.70*	0.97*	6.21*	0.98	5.27	0.97*	4.88	0.97	6.95*	0.96*
18: Wearing apparel	16.79*	0.93*	3.96	0.97	5.60	0.94*	4.98	0.94*	4.35	0.95	3.56	0.95	1.95	0.96	2.39	0.96	1.53	0.97	0.43	0.97	0.94	0.96	3.29	0.94	2.41	0.95
22: Publishing and printing	4.84	0.97	4.64	0.97	6.31*	0.95*	5.60	0.95*	4.78	0.96*	5.50	0.95*	6.10*	0.94*	5.73	0.94*	6.25*	0.94*	5.77	0.95*	4.44	0.95*	4.03	0.93*	6.64*	0.87*
24: Chemicals and chemical products	1.60	0.98	0.27	0.98	1.48	0.97	0.09	0.99	2.10	0.97	3.12	0.97	2.15	0.98	3.75	0.97	4.76	0.96	4.85	0.97	3.53	0.96	1.58	0.97	2.44	0.96
27: Basic metals	3.31	0.97	3.63	0.96*	4.15	0.96*	3.83	0.96*	4.25	0.96*	3.35	0.96	2.83	0.97	1.52	0.98	2.70	0.97	1.29	0.98	2.49	0.97	2.91	0.97	3.08	0.96
28: Fabricated metal products	3.28	0.99	5.03	0.98*	3.72	0.98	3.33	0.98	1.38	0.99	0.64	0.99	1.64	0.99	1.43	0.99	0.62	0.99	1.43	0.98	0.62	0.98	2.02	0.97	2.20	0.98
32: Radio, T.V. and communications	8.01*	0.91	0.72	0.92	1.66	0.92	0.95	0.94	1.55	0.91	1.28	0.93	1.25	0.95	0.46	0.97	0.74	0.96	1.22	0.93	1.33	0.94	2.05	0.92	1.76	0.96
33: Medical and optical instruments	1.17	0.97	1.92	0.94	1.61	0.96	2.89	0.93	2.40	0.93	3.44	0.92	2.34	0.91	3.75	0.89*	4.06	0.90	4.18	0.92	1.14	0.96	0.72	0.96	1.80	0.93
34: Motor vehicles and trailers	7.40*	0.94*	6.20*	0.95*	7.76*	0.95*	9.62*	0.94*	11.56*	0.94*	8.65*	0.95	7.11*	0.95	2.10	0.97	4.14	0.96	6.64*	0.95	6.73*	0.94	7.43*	0.95	5.44	0.95
36: Furniture	0.92	0.98	0.78	0.97	2.48	0.96	1.09	0.98	1.22	0.98	1.27	0.98	1.08	0.98	1.22	0.98	0.78	0.98	1.27	0.98	2.07	0.98	0.26	0.99	0.27	0.99
All (CS)	22.99*	0.99*	18.27*	0.99*	22.14*	0.99*	16.72*	0.99*	15.31*	0.99*	12.90*	0.99*	16.56*	0.99*	9.76*	0.99*	8.69*	0.99*	12.23*	0.99*	11.46*	0.99*	10.42*	0.99*	22.00*	0.98*
All (S)	25.60*	0.98*	19.02*	0.98*	24.14*	0.98*	25.20*	0.98*	24.10*	0.98*	23.42*	0.98*	22.68*	0.98*	18.57*	0.98*	22.27*	0.98*	20.80*	0.98*	21.20*	0.98*	20.98*	0.98*	22.00*	0.98*

Notes: N_i = initial sample size; N_f = final sample size. Annual sample sizes are provided in Table A4.1a.
Statistical tests of normality applied to the logarithm of plant size, measured by employee numbers.
D'A = D'Agostino et al. test; S-W = Shapiro-Wilk test. Samples All (CS) and All (S) tested using the Shapiro-Francia test.
CS = Cross-section; S = Surviving plants.
The sample of 1,754 plants surviving over the period 1972-1977 provided the following statistical results: D'Agostino et al.: 50.18*, 53.56*, 59.82*, 60.55*, 61.71* and 61.07*; Shapiro-Francia: 0.99*, 0.98*, 0.98*, 0.98*, 0.98* and 0.98*.
* Statistically significant at the 5 percent level.

Table 4.5b: Normality Tests, 1977 Cross-Section

Sample (SIC)	N_i	N_f	1977		1978		1979		1980		1981		1982		1983		1984		1985		1986	
			D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W
15: Food and beverages	270	143	4.74	0.99*	6.00*	0.98*	5.21	0.98*	7.74*	0.98*	6.57*	0.98*	4.16	0.98*	8.41*	0.98*	9.39*	0.97*	13.00*	0.97*	9.83*	0.97*
18: Wearing apparel	127	60	9.35*	0.97*	12.48*	0.97*	0.97	0.99	8.28*	0.97*	9.78*	0.96*	17.08*	0.93*	18.42*	0.91*	4.42	0.96*	17.85*	0.92*	4.84	0.96*
22: Publishing and printing	184	93	3.07	0.99	6.25*	0.99	6.46*	0.98	7.17*	0.98	7.07*	0.98	6.22*	0.98*	7.25*	0.98	8.06*	0.97*	6.24*	0.97*	5.13	0.98
24: Chemicals and chemical products	97	66	3.42	0.98	4.10	0.97*	3.54	0.97	3.46	0.97*	2.68	0.97	2.48	0.97	0.22	0.98	0.11	0.96	0.28	0.99	0.08	0.99
27: Basic metals	148	81	8.49*	0.97*	8.49*	0.97*	10.26*	0.96*	7.13*	0.96*	6.00*	0.97*	6.12*	0.96*	4.12	0.97*	2.95	0.98	3.71	0.97	4.27	0.96*
28: Fabricated metal products	402	181	3.93	0.99*	5.96	0.99*	17.35*	0.97*	16.37*	0.98*	16.55*	0.97*	15.92*	0.97*	11.34*	0.98*	7.24*	0.98*	8.06*	0.98*	10.77*	0.97*
32: Radio, T.V. and communications	51	33	2.85	0.96	2.99	0.96	4.22	0.96	2.78	0.97	2.23	0.97	1.80	0.97	1.08	0.97	2.16	0.95	0.56	0.96	1.76	0.97
33: Medical and optical instruments	73	43	1.47	0.98	2.72	0.97	2.56	0.96*	3.88	0.95*	4.27	0.95*	5.76	0.93*	4.47	0.94*	3.75	0.95*	2.86	0.96	3.55	0.94*
34: Motor vehicles and trailers	85	50	5.80	0.96*	4.71	0.96*	2.25	0.96*	8.02*	0.94*	8.26*	0.94*	7.65*	0.95*	7.03*	0.96	9.27*	0.95*	1.33	0.97	7.33*	0.94*
36: Furniture	178	82	2.15	0.99	0.88	0.99	3.38	0.98*	2.10	0.98	1.60	0.98	0.50	0.98	0.54	0.99	5.09	0.98	2.40	0.98	1.60	0.98
All (CS)	2,812	1,444	55.09*	0.99*	68.32*	0.99*	78.44*	0.98*	70.86*	0.98*	65.72*	0.98*	56.60*	0.99*	42.22*	0.99*	35.07*	0.99*	29.60*	0.99*	21.00*	0.99*
All (S)	1,444	1,444	38.99*	0.99*	48.02*	0.98*	54.18*	0.98*	53.03*	0.98*	50.94*	0.98*	49.49*	0.98*	42.62*	0.98*	37.82*	0.99*	30.16*	0.99*	21.00*	0.99*

Notes: N_i = initial sample size; N_f = final sample size. Annual sample sizes are provided in Table A4.1b.
Statistical tests of normality applied to the logarithm of plant size, measured by employee numbers.
D'A = D'Agostino et al. test; S-W = Shapiro-Wilk test. Samples All (CS) and All (S) tested using the Shapiro-Francia test.
CS = Cross-section; S = Surviving plants.
* Statistically significant at the 5 percent level.

Table 4.5c: Normality Tests, 1986 Cross-Section

Sample (SIC)	N_i	N_f	1986		1987		1988		1989		1990		1991		1992		1993		1994		1995		1996		1997	
			D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W
15: Food and beverages	220	125	6.64*	0.98*	7.88*	0.98*	6.04*	0.98*	7.85*	0.98*	8.06*	0.98*	9.25*	0.98*	8.71*	0.98*	10.79*	0.97*	11.77*	0.98*	10.13*	0.97*	12.21*	0.96*	15.15*	0.96*
18: Wearing apparel	115	46	4.74	0.98	2.51	0.98	2.69	0.98	3.21	0.97	3.70	0.97	3.09	0.97	3.52	0.97	3.05	0.97	2.27	0.97	1.92	0.97	6.21*	0.96	6.74*	0.96
22: Publishing and printing	138	72	10.04*	0.97*	10.44*	0.97*	10.60*	0.96*	8.80*	0.97*	10.33*	0.96*	13.57*	0.94*	13.96*	0.94*	14.58*	0.94*	13.36*	0.94*	11.97*	0.93*	13.03*	0.91*	15.68*	0.86*
24: Chemicals and chemical products	126	88	4.04	0.98	2.92	0.98*	4.20	0.98*	5.78	0.97*	5.54	0.97*	5.14	0.97*	4.89	0.97*	5.27	0.96*	5.93	0.97*	4.58	0.97*	2.94	0.98	4.73	0.97
27: Basic metals	119	83	7.17*	0.96*	9.22*	0.95*	7.24*	0.95*	6.74*	0.95*	6.52*	0.96*	6.36*	0.96*	4.29	0.97*	5.52	0.96*	4.08	0.98	5.56	0.97*	5.88	0.96*	5.50	0.96*
28: Fabricated metal products	310	190	31.92*	0.96*	22.12*	0.97*	23.14*	0.97*	16.08*	0.98*	12.80*	0.97*	16.13*	0.97*	12.45*	0.98*	11.52*	0.98*	16.35*	0.97*	10.26*	0.97*	14.65*	0.96*	14.70*	0.96*
32: Radio, T.V. and communications	67	46	3.16	0.97	3.42	0.97	3.14	0.97	4.34	0.96*	2.29	0.96	2.18	0.97	0.02	0.98	0.59	0.98	0.57	0.98	0.16	0.98	0.34	0.98	1.65	0.97
33: Medical and optical instruments	76	46	5.76	0.95*	5.20	0.95*	5.43	0.94*	4.22	0.95*	3.63	0.95*	3.20	0.96*	3.01	0.96*	2.28	0.96	0.51	0.98	0.04	0.99	0.05	0.99	0.34	0.99
34: Motor vehicles and trailers	78	45	4.96	0.96*	5.52	0.96*	7.74*	0.96*	10.01*	0.96*	7.94*	0.96*	5.04*	0.97	3.29	0.98	3.80	0.97	5.21	0.97	5.77	0.96	5.14	0.97	3.68	0.97
36: Furniture	163	86	2.88	0.99	3.57	0.99	0.47	0.99	2.25	0.99	1.13	0.99	0.83	0.99	1.05	0.99	0.68	0.99	1.61	0.98	0.75	0.98	1.09	0.98	1.09	0.98
All (CS)	2,505	1,417	100.39*	0.98*	94.23*	0.98*	74.02*	0.98*	65.39*	0.98*	59.02*	0.98*	87.84*	0.98*	56.70*	0.98*	46.68*	0.98*	49.62*	67.96*	48.53*	0.98*	52.38*	0.98*	70.51*	0.97*
All (S)	1,417	1,417	68.23*	0.97*	69.56*	0.97*	63.01*	0.98*	62.33*	0.98*	62.86*	0.98*	64.75*	0.98*	63.49*	0.98*	62.89*	0.97*	0.98*	0.97*	71.78*	0.97*	70.56*	0.97*	70.51*	0.97*

Notes: N_i = initial sample size; N_f = final sample size. Annual sample sizes are provided in Table A4.1c.
Statistical tests of normality applied to the logarithm of plant size, measured by employee numbers.
D'A = D'Agostino et al. test; S-W = Shapiro-Wilk test. Samples All (CS) and All (S) tested using the Shapiro-Francia test.
CS = Cross-section; S = Surviving plants.
* Statistically significant at the 5 percent level.

Some differences emerged using the Shapiro-Wilk test, with for example, lognormality being rejected in every year between 1987 and 1995 for the Chemical and chemical products industry (no rejections using the alternative test). In aggregate though the Shapiro-Wilk test rejected lognormality for only four industries by 1997, with some evidence of convergence again for the Motor vehicles and trailers and Medical and optical instruments industries. In total, the Shapiro-Wilk test rejected (log)normality in 69 of the 120 tests.

In summary, this analysis, consistent with previous studies, indicates considerable heterogeneity in the evolution of industry-level size distributions over time. Whilst, the skewness and kurtosis indices in conjunction with the two formal statistical tests suggest some evidence of convergence towards the lognormal distribution over time for some industries, others show some movement away from such a distribution, with yet other industries exhibiting evidence of some stability (remaining either lognormal or not lognormal) over the time periods examined. The analysis therefore does not suggest any consistent movement to constant, stable long-run plant size distributions across industries. It is not clear whether those industry-level samples rejected as lognormal may evolve towards the lognormal over longer time durations than analysed or indeed to other distributions. However, from this evidence it cannot be concluded that the lognormal distribution always provides a reasonable description (at least satisfying these statistical tests) across all industries and in all time periods.

To further investigate the evolution of plant size distributions (and the combined effects of ageing and selection) over time nine cohorts of new entrant plants were constructed and tracked for a period of 10 years, with the entry years selected to represent different points of the economic cycle. The cohorts comprised of the base years of this analysis, 1972, 1977 and 1986, plus 1973, 1979 and 1988 as years broadly representing peaks in the U.K. economic cycle and 1975, 1981 and 1992 as trough or recession years. Using these cohorts allows some examination of the effects of entry in a particular phase of each economic cycle on the evolution of the distribution of plant sizes. Again, for this analysis the statistical approach is employed without kernel density estimation given the available sample sizes.

Somewhat mixed evidence was found regarding any observed convergence to a lognormal distribution of sizes for these cohorts of new entrant plants. Corresponding with the previous analysis the mean size of plants was observed to rise across all industries over the 10-year time periods, although consistent evidence of declining standard deviations of plant size was not found (see Table 4.6). The skewness indices (presented in Table 4.7) exhibited some degree of variability across the cohorts, with no clear pattern emerging for any of the three groups representing the different phases of the economic cycle, and no strong evidence supporting a clear movement towards skewness index values consistent with symmetrical log(size) distributions such as the normal distribution.

However, there is perhaps some greater support for convergence in the kurtosis indices of each sample towards the benchmark value of 3 (albeit again with some volatility), but again no clear pattern with the economic cycle phase of entry. Given the very small sample sizes it is, however, not possible to adequately identify to what extent these observations are due to industrial or size composition factors. Further with the limited industry-specific variables available in the dataset it is not possible to test whether such deviations may have any deterministic explanations. In interpreting such evidence though it should noted that the deviations from the expected skewness and kurtosis index values from the (log)normal distribution are in almost all cases fairly modest and not (highly) inconsistent with such a size distribution.

Another point of note is that for the entry years 1975 and 1977, the two largest entry cohorts in terms of plant numbers, there is evidence of higher positive skewness

Table 4.6: Means and Standard Deviations, New Entrant Cohorts

Sample (Cohort)	N_i	N_f	Year 0		Year 1		Year 2		Year 3		Year 4		Year 5		Year 6		Year 7		Year 8		Year 9		Year 10	
			Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
1972	121	83	2.38	1.22	2.70	1.29	2.81	1.28	2.80	1.28	2.84	1.21	2.87	1.22	2.82	1.21	2.90	1.20	2.98	1.22	2.91	1.18	3.02	1.12
1977	195	47	1.39	1.18	2.49	1.05	2.61	1.10	2.69	1.15	2.65	1.13	2.68	1.07	2.75	1.08	2.78	1.09	2.75	1.09	2.93	0.99	3.05	1.00
1986	145	74	2.44	0.98	2.75	0.94	2.89	0.94	2.97	0.96	3.08	0.96	3.12	0.96	3.14	0.96	3.18	0.94	3.27	0.98	3.34	0.96	3.43	0.94
1973	128	100	2.51	1.11	2.92	1.18	3.01	1.22	3.13	1.17	3.12	1.20	3.18	1.15	3.22	1.20	3.28	1.14	3.29	1.15	3.28	1.21	3.26	1.22
1979	119	89	2.40	0.97	2.60	1.05	2.73	1.08	2.81	1.07	2.86	1.09	2.93	1.12	3.07	1.10	3.24	1.13	3.37	1.18	3.44	1.20	3.49	1.22
1988	..	76	..	..	2.98	1.05	3.12	0.99	3.21	0.97	3.24	1.01	3.34	0.94	3.46	0.97	3.53	1.00	3.63	1.03	3.71	0.98	3.65	0.95
1975	472	97	1.58	1.03	1.76	1.03	1.79	1.15	1.87	1.14	2.40	1.33	2.43	1.33	2.43	1.34	2.38	1.29	2.32	1.27	2.32	1.31	2.49	1.36
1981	108	72	2.11	1.00	2.37	1.01	2.59	1.01	2.72	1.02	2.79	1.06	3.01	1.09	3.25	1.04	3.38	1.07	3.50	1.04	3.47	1.03	3.52	1.01
1992	61	57	2.72	1.14	3.02	1.21	3.19	1.18	3.42	1.11	3.50	1.09	3.56	1.06	3.63	1.10	3.73	1.15	3.77	1.16	3.92	1.22	3.92	1.15

Notes: N_i = initial sample size; N_f = final sample size. Annual sample sizes are provided in Table A4.2.
Mean and Standard Deviation (S.D.) of the logarithm of plant size, measured by employee numbers.
Plant numbers in the initial year may be lower than the total new entrants reported due to a time lag in recording employee numbers.
.. Not reported given small sample < 10 plants.

Table 4.7: Skewness and Kurtosis Indices, New Entrant Cohorts

Sample (Cohort)	N_i	N_f	Year 0		Year 1		Year 2		Year 3		Year 4		Year 5		Year 6		Year 7		Year 8		Year 9		Year 10	
			Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt	Skew	Kurt
1972	121	83	0.10	2.55	0.21	2.97	0.27	3.11	0.14	3.12	0.12	3.05	0.04	3.03	0.21	3.19	0.15	3.10	0.06	2.81	-0.09	2.95	0.13	2.65
1977	195	47	0.94	4.56	0.26	2.89	0.31	3.22	0.31	3.14	0.29	3.34	0.02	3.33	0.43	3.19	0.34	2.93	0.43	3.35	0.54	4.03	0.52	3.67
1986	145	74	0.16	3.53	0.04	3.54	0.09	3.52	0.44	3.33	0.56	3.47	0.62	3.48	0.71	3.51	0.94	3.66	0.89	3.68	0.64	3.23	0.96	3.58
1973	128	100	0.00	2.79	0.07	2.91	0.10	2.88	0.28	2.81	0.22	2.85	0.15	2.95	0.18	2.91	0.40	2.93	0.48	3.01	0.23	3.22	0.28	3.39
1979	119	89	0.48	3.74	0.78	5.61	0.83	5.18	0.92	5.49	0.89	5.49	0.89	5.35	1.00	5.05	0.93	4.24	0.73	3.55	0.77	3.32	0.81	3.45
1988	..	76	..	..	-0.09	3.50	0.18	3.44	0.31	3.66	0.30	3.50	0.76	3.10	0.33	2.95	0.27	2.75	0.45	2.57	0.42	2.34	0.49	2.62
1975	472	97	1.29	6.56	0.93	4.50	0.89	3.82	0.88	3.89	0.45	2.64	0.46	2.67	0.52	2.73	0.55	2.55	0.64	2.84	0.45	2.62	0.39	2.46
1981	108	72	0.29	3.43	0.26	3.36	0.63	4.31	0.55	4.58	0.66	4.35	0.75	4.09	0.80	4.31	0.58	3.92	0.49	4.16	0.52	4.21	0.53	4.65
1992	61	57	0.47	3.10	0.98	4.09	0.94	4.21	0.95	4.20	1.23	5.21	1.23	5.58	1.21	5.36	1.14	4.52	0.97	3.79	0.80	3.14	0.67	3.13

Notes: N_i = initial sample size; N_f = final sample size. Annual sample sizes are provided in Table A4.2.
Skewness (Skew) and Kurtosis (Kurt) indices of the logarithm of plant size, measured by employee numbers.
Plant numbers in the initial year may be lower than the total new entrants reported due to a time lag in recording employee numbers.
.. Not reported given small sample < 10 plants.

Table 4.8: Normality Tests, New Entrant Cohorts

Sample (Cohort)	N_i	N_f	Year 0		Year 1		Year 2		Year 3		Year 4		Year 5		Year 6		Year 7		Year 8		Year 9		Year 10	
			D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W	D'A	S-W
1972	121	83	1.36	0.99	1.12	0.99	2.05	0.99	0.83	0.99	0.48	0.99	0.18	0.99	1.34	0.99	0.68	0.99	0.07	0.99	0.18	0.99	0.49	0.99
1977	195	47	32.69*	0.95*	1.28	0.98	2.20	0.98	1.89	0.97*	2.08	0.97*	0.84	0.98	2.79	0.97	1.45	0.99	3.02	0.98	5.77	0.97	4.16	0.97
1986	145	74	2.60	0.99	2.23	0.99	2.14	0.99	5.27	0.98*	7.62*	0.97*	8.51*	0.96*	9.48*	0.96*	13.74*	0.93*	12.24*	0.94*	6.14*	0.96*	11.86*	0.93*
1973	128	100	0.12	0.99	0.14	1.00	0.29	0.99	2.08	0.99	1.24	0.99	0.53	0.99	0.66	0.99	3.15	0.98	4.28	0.98	1.46	0.99	2.34	0.98
1979	119	89	7.35*	0.98*	27.64*	0.96*	24.44*	0.96*	27.13*	0.95*	25.08*	0.95*	22.42*	0.95*	23.13*	0.95*	16.49*	0.95*	9.52*	0.97*	9.27*	0.96*	10.34*	0.95*
1988	..	76	-	-	1.97	0.98	2.23	0.98*	4.58	0.97*	3.36	0.98	9.21*	0.95*	1.81	0.97*	1.28	0.97*	3.57	0.97*	4.80	0.96*	3.52	0.96*
1975	472	97	129.14*	0.92*	64.48*	0.95*	48.52*	0.95*	42.99*	0.95*	6.08*	0.97*	5.67	0.97*	6.55*	0.97*	7.78*	0.97*	8.09*	0.96*	4.73	0.98*	4.06	0.98
1981	108	72	2.97	0.98	2.99	0.99	15.03*	0.97*	13.59*	0.97*	13.31*	0.97*	11.97*	0.97*	13.07*	0.96*	7.56*	0.97	7.16*	0.98	7.75*	0.97	9.10*	0.97*
1992	61	57	2.80	0.98	15.41*	0.94*	14.89*	0.95*	14.24*	0.95*	23.00*	0.92*	23.30*	0.92*	21.44*	0.93*	16.92*	0.92*	11.34*	0.93*	6.75*	0.94*	4.93	0.95*

Notes: N_i = initial sample size; N_f = final sample size. Annual sample sizes are provided in Table A4.2.
Statistical tests of normality applied to the logarithm of plant size, measured by employee numbers.
D'A = D'Agostino et al. test; S-W = Shapiro-Wilk test.
* Statistically significant at the 5 percent level.
Plant numbers in the initial year may be lower than the total new entrants reported due to a time lag in recording employee numbers.
.. Not reported given small sample < 10 plants.

and leptokurtosis values than in the other cohorts in the initial entry year with skewness indices of 1.29 and 0.94 and kurtosis indices of 6.56 and 4.56, respectively. However, in the case of the 1975 cohort within 5 years the skewness index had declined to 0.45 and kurtosis index to 2.64, more consistent with the normal distribution, whilst for the 1977 cohort only 1 year was required before reaching index values of 0.26 and 2.89.

For the 1972 and 1973 cohorts (log)normality could not be rejected in any year using either the D'Agostino et al. or Shapiro-Wilk tests (results presented in Table 4.8). For the 1979 cohort both tests in all cases rejected the (log)normal. In the case of the 1977 cohort there is some evidence of a very quick process of evolution towards the (log)normal size distribution with rejection by the D'Agostino et al. test in year 1, before non-rejection in all subsequent years (the Shapiro-Wilk test additionally provided rejections in years 3 and 4). In the case of the 1986 and 1992 new entrant cohorts there is however statistical evidence of a movement away from an initial non-rejection of the lognormal distribution, with rejection by both tests from year 4 onwards for the 1986 sample and rejection by at least one of the tests in all but the first year in the latter case.

Similar results emerged for the 1981 cohort (no rejections in year 0 or 1), followed by rejections using both tests, although with the non-rejection of lognormality in years 7 - 9 using the Shapiro-Wilk test. For the 1975 cohort lognormality was rejected by at least one of the tests in all but the last year. Finally, for the 1988 cohort the Shapiro-Wilk tests provided rather stronger evidence of rejection of the lognormal distribution (in 8 cases) compared to the D'Agostino et al. test (only 1 case).

As found in the industry-level analysis, these results are therefore not strongly suggestive of any consistent convergence towards a long-run stable lognormal distribution of plant sizes. To some extent it is possible that the size threshold truncation of the WRME sample may have inhibited the observation of any strong convergence process by not systematically recording the very smallest plants (who would be likely to be disproportionately younger), although such an aspect cannot be directly tested here.

4.6. Plant Growth Rate Distributions

In this section, an analysis is conducted to examine the distribution of plant growth rates. As discussed earlier, Gibrat (1931) assumed a random sampling of firm (or plant) growth opportunities from a normal distribution. However, recent research by, for example, Stanley et al. (1996), Bottazzi and Secchi (2002, 2003a, 2003b), Reichstein and Jensen (2003) and Teitelbaum and Axtell (2005) has suggested that Laplace or Subbotin distributions may provide more appropriate representations of the distribution of such growth rates given the observed greater than expected mass of the distribution within the tails and more pronounced central peak of the observed distributions than would be expected under the normal distribution.

Although suitable general Central Limit Theorems may provide a lognormal limiting distribution, or an approximately lognormal distribution where the assumption of normally distributed growth opportunities is violated (Kleiber and Kotz, 2003, p.110; Hart, 1973, 1976) this section of the analysis aims to specifically test the assumption of a normal distribution of plant growth rates. Consideration is also given the possibility of plant growth rate distributions more reasonably approximating to the Laplace distribution which would have a kurtosis index of 6 (Bottazzi, Coad, Jacoby and Secchi, 2005, p.8). However, a more formal analysis of the Laplace and Subbotin distributions is not conducted, and remains an area for further research (such an extension is discussed in Chapter 6). Using the same statistical approach and kernel density estimation, analysis was conducted to examine the distribution of plant growth rates for each of the three economic cycle time periods, and for the longer time period 1972-1997.

In conducting this analysis of plant growth rates it is recognised that there may be some limitations in comparing the WRME data to specific statistical distributions. In addition to the points made earlier in respect of the size distribution of manufacturing plants regarding the incomplete sectoral coverage of the data and the (partial) truncation of the plant size data so that not all plants are recorded, in examining the growth rates of plants in each respective time period again only a partial distribution is likely to be observed, with some plants experiencing low or negative growth rates likely to have exited the sample before the end of each period. Such an issue would be expected to be more significant over longer time-spans.

Statistical information on the aggregate distribution of (proportional) plant growth rates, given by $[(X_{it} - X_{it-1})/X_{it-1}]$, where X_{it} is the size of plant i at time t, is presented in Tables 4.9a-d. Examination of the higher moments of this data indicates some deviations of the distribution of plant growth rates from the normal distribution with observed positive skewness, exhibiting index values of 5.85, 5.47, 8.42 and 25.61 for the periods 1972-1977, 1972-1997, 1977-1986 and 1986-1997 respectively, and leptokurtosis, with corresponding index values of 53.70, 43.80, 115.86 and 818.94. The D'Agostino et al. and the Shapiro-Francia tests rejected normality at the 5 percent level of significance for all four aggregate samples. In each of the four time periods the aggregate calculated mean average growth rates were positive.

Kernel density estimates, presented in Figures 4.5a-d, (with the x-scale of plant growth rates removed to avoid the disclosure of extreme values) indicate highly peaked distributions with more mass in the centre of the observed plant growth rate distribution than would be expected under the normal distribution for each of the four time periods, although interestingly for the 1986-1997 period the peak of the kernel density estimates was not dissimilar to the normal distribution. Band-widths of the order of 1.0 were in general required to give rise to kernel density estimates approximating to the peak of the normal distribution. The reported kurtosis indices are considerably higher than the benchmark value of 6 associated with the Laplace distribution for each time period, suggestive of greater mass in the tails of the distribution (although such results are not clear from the kernel density estimates even with the use of very small band-widths and using alternative kernel functions.

As discussed above, one factor perhaps influencing the observed distribution of plant growth rates may be the truncation of negative growth rates. Whilst significant numbers of plants exhibited growth rates of 100 percent or more in each of the four periods (representing between 10.43 and 25.19 percent of surviving plant for the samples), no growth rates of -100 percent or worse were observed for these samples[70].

This may be because of the exit of plants with even lower negative growth performance, although it cannot be excluded that even allowing for such possible affects that the normal distribution would still not have provided a reasonable approximation to the actual distributions of growth rates. However, despite this apparent truncation, for the periods 1972-1997, 1972-1977 and 1977-1986 more than half of the surviving plants experienced growth rates of zero percent or below, with the period 1986-1997 showing a distribution slightly more skewed towards the positive growth rate ranges.

Given the potential truncation of the growth rate data the distributions of plant growth rates were also examined across the shortest possible growth period of one-year, the effect of this being to minimise the extent of sample exit, and especially the effect of truncation on the left-side of the distribution reflecting low and negative plant growth rates. The analysis indicated qualitatively similar empirical results with highly-peaked distributions, although exhibiting some sensitivity to the form of the kernel function and particularly the band-width choice.

Similar results also emerged from a preliminary analysis of the one-year growth

---

[70] For the periods 1972-1977, 1977-1986, 1986-1997 and 1972-1997 there were 783, 712, 463 and 363 plants exhibiting negative growth rates. Of the remaining plants, with either zero or positive growth rates, there were 183, 234, 357 and 143 plants which had growth rates of 100 percent of more over the respective time periods.

Figure 4.5a: Plant Growth Rate Distribution, 1972-1977

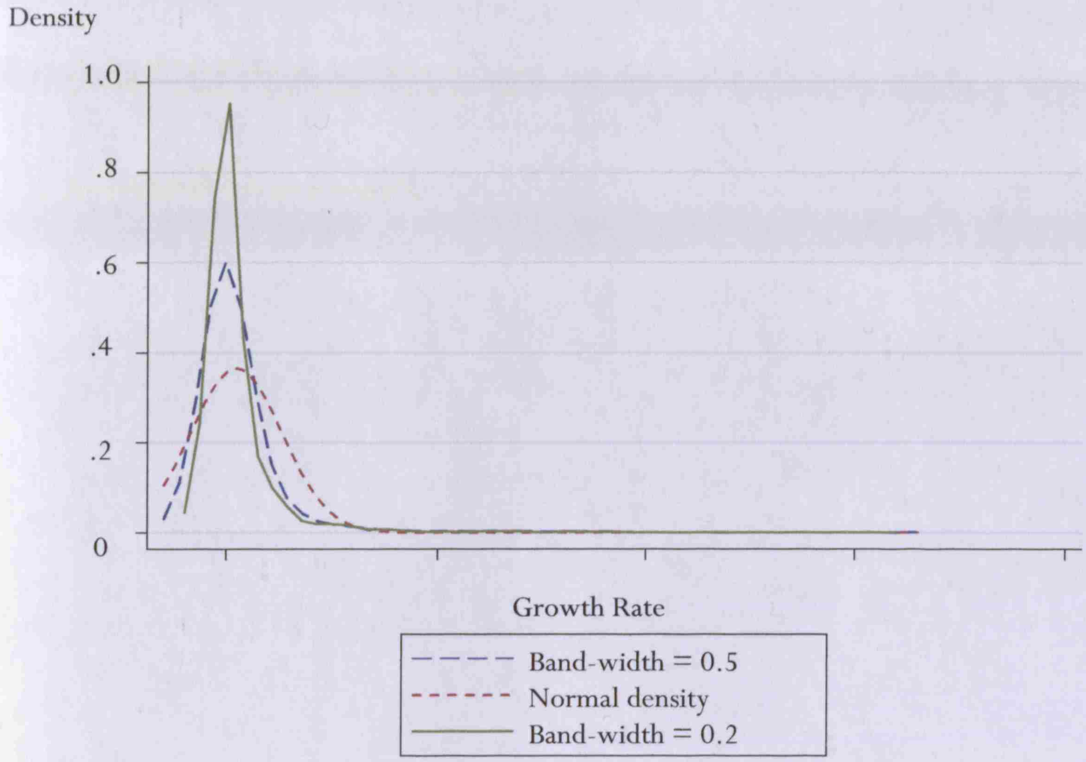


Figure 4.5b: Plant Growth Rate Distribution, 1977-1997

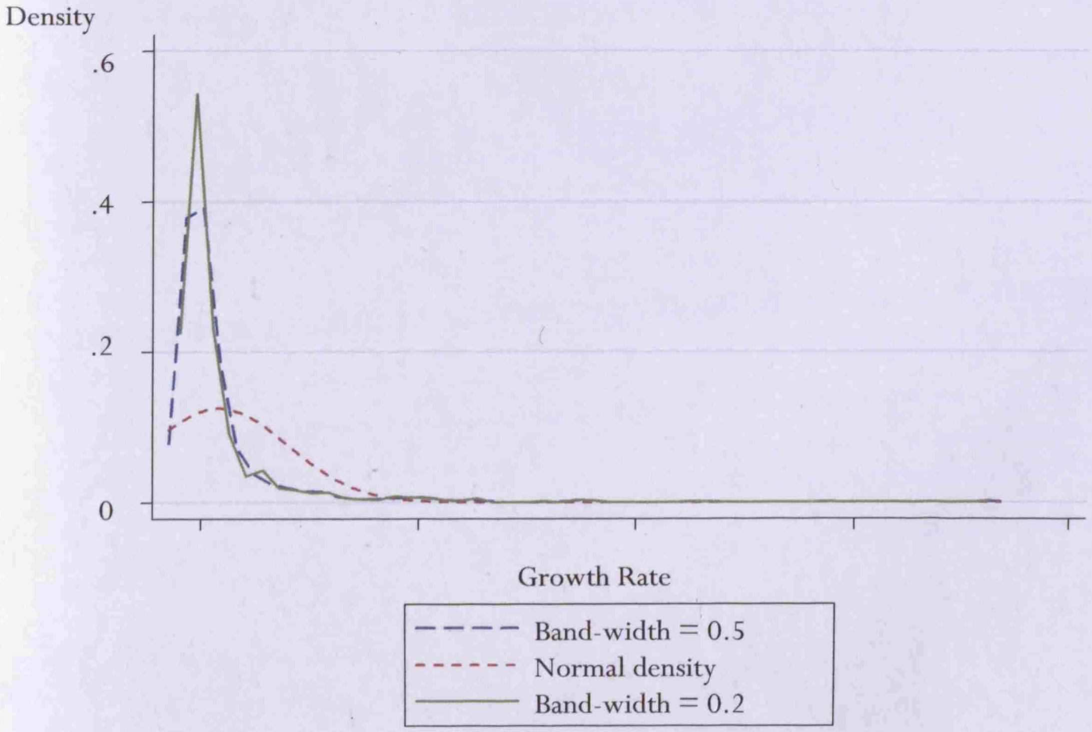


Figure 4.5c: Plant Growth Rate Distribution, 1977-1986

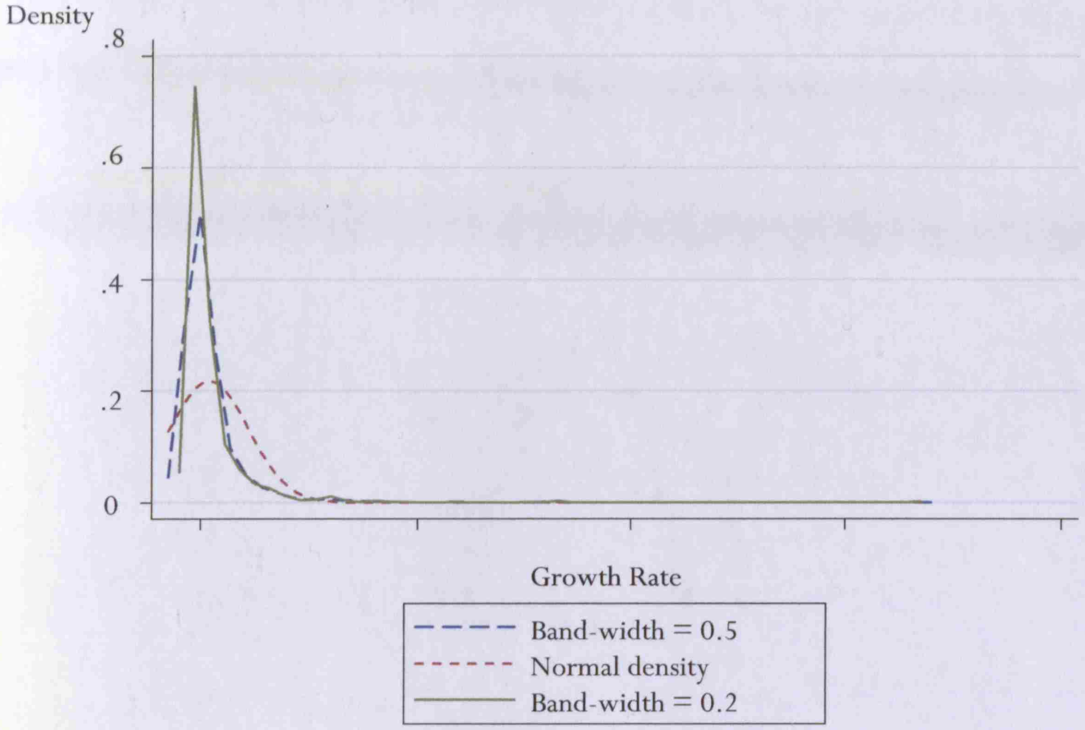
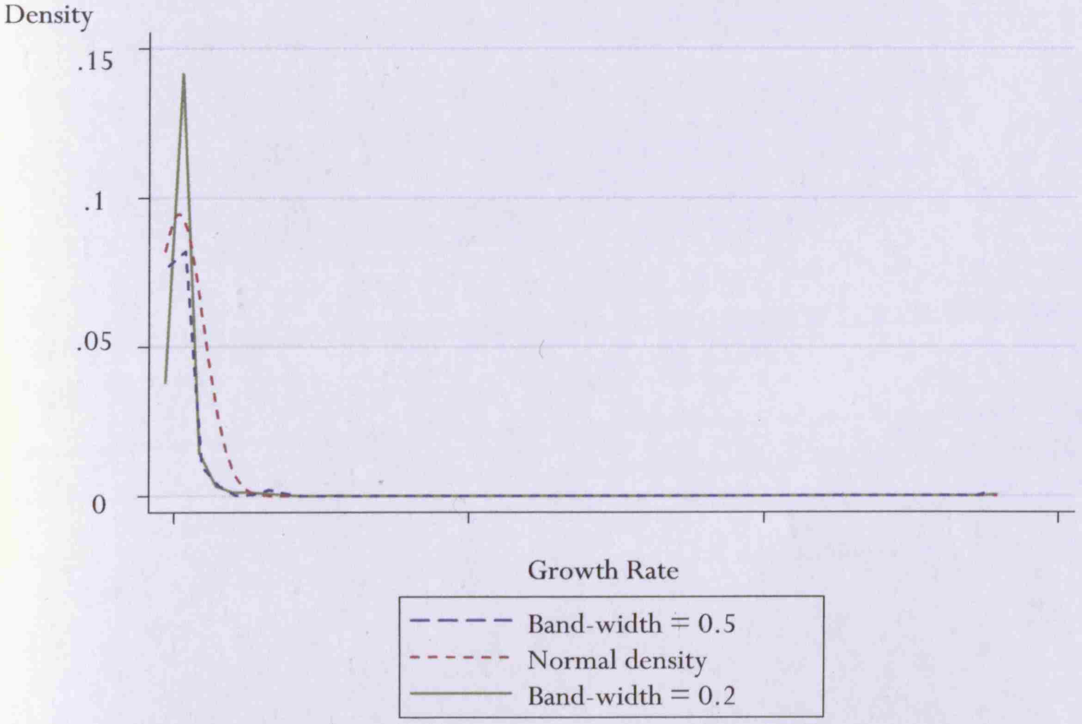


Figure 4.5d: Plant Growth Rate Distribution, 1986-1997



rates using the complete dataset cross-sections, especially under the use of the "optimal" band-width selection procedure of STATA or band-widths of the order 0.2 - 0.3. However, experimentation with a number of samples using the Gaussian kernel function and a band-width of 0.5 did provide kernel density estimates in some cases not too dissimilar to that of the normal distribution[71] (in a number of instances with insufficient mass in the centre of the distribution compared to the normal distribution), suggesting that the normal distribution may perhaps provide a more adequate representation for more complete datasets.

4.7. Industry Plant Growth Rate Distributions

The above methodology was also employed to examine industry growth rate distributions, again at the 2-digit industry-level of aggregation. Tables 4.9a-d provide information on the mean and variance of industry-level growth rates. For all industries except Leather products in 1972-1977, Radio, television and communications in 1972-1997, Leather products, Coke, petroleum and nuclear fuel and Other transport equipment in 1977-1986 and Leather products in 1986-1997, positive mean proportional growth rates were calculated.

Use of the two-sample t-test for unequal variances (following Welch, 1947; and Aspin, 1948) at the 95 percent confidence level (with a null of no difference between groups) suggested some significant differences in growth rates (compared to the rest of the sample) for a number of industries in different time periods: the Wearing apparel, Leather products, Coke, petroleum and nuclear fuel, Non-metallic mineral products and Other machinery and equipment industries in 1972-1977, the Wood and wood products,

---

[71] There was also some indication of differences in the form of the observed distributions in a number of one-year periods (for example shifts in the distribution tails, with a notable shift to the left tail in the recessionary years of 1991-1992).

Table 4.9a: Plant Growth Rate Distributions by Industry, 1972-1977

Sample (SIC)	Plants (N)	Mean	Variance	Skewness	Kurtosis	D'Agostino	Shapiro-Wilk
15: Food and beverages	183	0.14	0.65*	7.62	81.57	299.18*	0.48*
16: Tobacco products	..	..	..	..	..	..	..
17: Textiles	69	0.25	1.10	4.52	27.67	100.89*	0.54*
18: Wearing apparel	90	0.09*	0.34*	2.54	11.09	68.68*	0.75*
19: Leather products	25	-0.02*	0.09*	0.74	2.99	3.22	0.95
20: Wood and wood products	100	0.08*	0.50*	2.69	13.19	80.98*	0.76*
21: Pulp and paper products	34	0.33	2.76*	4.91	26.98	72.91*	0.37*
22: Publishing and printing	101	0.20	0.82*	4.09	21.80	118.11*	0.53*
23: Coke, petroleum and nuclear fuel	11	0.02*	0.05*	-1.03	4.32	6.92*	0.89
24: Chemicals and chemical products	64	0.35	1.55	4.39	25.82	93.81*	0.51*
25: Rubber and plastic products	77	0.49	2.58*	4.90	32.56	115.84*	0.51*
26: Non-metallic mineral products	133	0.04*	0.59*	4.28	30.46	156.50*	0.65*
27: Basic metals	116	0.10	0.89*	8.07	78.38	215.71*	0.37*
28: Fabricated metal products	236	0.38	2.11*	6.50	61.70	334.29*	0.48*
29: Other machinery and equipment	197	0.41*	1.41	3.92	23.38	195.28*	0.62*
30: Office machinery and computers	..	..	..	..	..	..	..
31: Electrical machinery	48	0.30	0.45*	2.26	7.77	36.81*	0.71*
32: Radio, T.V. and communications	29	0.05	0.29*	0.72	3.08	3.46	0.95
33: Medical and optical instruments	44	0.54	1.75*	3.38	16.59	59.77*	0.63*
34: Motor vehicles and trailers	55	0.29	0.67*	2.37	9.30	44.59*	0.73*
35: Other transport equipment	36	0.19	1.48	5.09	29.12	77.11*	0.39*
36: Furniture	94	0.34	1.86*	4.15	22.23	113.65*	0.53*
All	1,754	0.25	1.20	5.85	53.70	2,015.62*	0.52*

Notes: Difference in means: two-sample t-test for differences in means with unequal variances (differences between groups H_0: difference = 0; H_1: difference $\neq 0$).
Differences in variances: variance-ratio (F-test) test (H_0: ratio = 1; H_1: ratio $\neq 1$). Industry-level tests compare industry means and variances to the remaining aggregate sample.
.. Not reported given small sample < 10 plants. * Statistically significant at the 5 percent level.
For the aggregate sample the 10th, 25th, 50th, 75th and 90th percentiles measures were: -0.41, -0.18, 0.00, 0.30 and 1.00.

Table 4.9b: Plant Growth Rate Distributions by Industry, 1972-1997

Sample (SIC)	Plants (N)	Mean	Variance	Skewness	Kurtosis	D'Agostino	Shapiro-Wilk
15: Food and beverages	76	0.97	6.17*	3.68	20.09	90.19*	0.60*
16: Tobacco products	..	..	..	..	..	..	..
17: Textiles	22	0.48	1.97*	1.80	5.67	17.56*	0.77*
18: Wearing apparel	24	0.37	1.73*	1.94	6.00	20.17*	0.73*
19: Leather products	..	..	..	..	..	..	..
20: Wood and wood products	24	0.30*	0.82*	1.58	5.02	14.69*	0.81*
21: Pulp and paper products	22	1.99	62.02*	4.03	18.05	51.97*	0.35*
22: Publishing and printing	43	0.90	6.90	3.08	11.85	50.65*	0.53*
23: Coke, petroleum and nuclear fuel	..	..	..	..	..	..	..
24: Chemicals and chemical products	35	1.14	20.38*	3.83	17.57	58.60*	0.44*
25: Rubber and plastic products	38	1.63	18.76*	2.92	10.99	44.85*	0.56*
26: Non-metallic mineral products	43	0.08*	1.03*	2.59	10.59	43.40*	0.71*
27: Basic metals	52	0.15*	3.05*	4.48	25.96	84.29*	0.50*
28: Fabricated metal products	88	1.08	6.37*	2.00	6.32	45.23*	0.71*
29: Other machinery and equipment	73	0.42*	2.75*	2.05	6.93	42.26*	0.73*
30: Office machinery and computers	..	..	..	..	..	..	..
31: Electrical machinery	24	2.10	26.91*	3.11	12.55	40.69*	0.56*
32: Radio, T.V. and communications	15	-0.17*	0.71*	1.46	4.28	10.07*	0.81*
33: Medical and optical instruments	16	2.74	20.17*	1.43	3.38	7.98*	0.69*
34: Motor vehicles and trailers	31	0.47	4.67*	3.11	13.02	45.13*	0.58*
35: Other transport equipment	16	0.60	6.62	3.31	12.69	39.20*	0.47*
36: Furniture	33	1.49	34.18*	3.93	17.61	58.12*	0.37*
All	693	0.85	10.20	5.47	43.80	782.19*	0.47*

Notes: Difference in means: two-sample t-test for differences in means with unequal variances (differences between groups H_0: difference = 0; H_1: difference $\neq$ 0).
Differences in variances: variance-ratio (F-test) test (H_0: ratio = 1; H_1: ratio $\neq$ 1). Industry-level tests compare industry means and variances to the remaining aggregate sample.
.. Not reported given small sample < 10 plants. * Statistically significant at the 5 percent level.
For the aggregate sample the 10th, 25th, 50th, 75th and 90th percentiles measures were: -0.72, -0.48, -0.06, 0.78 and 2.82.

Table 4.9c: Plant Growth Rate Distributions by Industry, 1977-1986

Sample (SIC)	Plants (N)	Mean	Variance	Skewness	Kurtosis	D'Agostino	Shapiro-Wilk
15: Food and beverages	143	0.67	4.66*	5.28	35.62	189.17*	0.46*
16: Tobacco products	..	..	..	..	..	..	..
17: Textiles	43	0.15	1.00*	2.42	9.98	40.54*	0.74*
18: Wearing apparel	60	0.26	1.15*	4.16	24.71	86.99*	0.59*
19: Leather products	14	-0.05*	0.17*	0.03	1.72	2.15	0.93
20: Wood and wood products	67	0.17*	0.48*	1.51	5.58	27.02*	0.87*
21: Pulp and paper products	37	0.67	2.85	2.29	7.72	32.25*	0.64*
22: Publishing and printing	93	0.66	4.56*	3.75	18.12	102.24*	0.51*
23: Coke, petroleum and nuclear fuel	10	-0.03*	0.21*	0.76	2.63	1.96	0.92
24: Chemicals and chemical products	66	0.18	1.67*	2.76	10.80	58.55*	0.65*
25: Rubber and plastic products	82	0.69	14.28*	8.13	71.01	169.00*	0.24*
26: Non-metallic mineral products	123	0.30	1.09*	2.27	10.76	79.07*	0.80*
27: Basic metals	81	0.00*	0.92*	3.25	15.47	82.27*	0.65*
28: Fabricated metal products	181	0.48	2.38*	3.23	15.92	151.03*	0.64*
29: Other machinery and equipment	150	0.18*	1.06*	3.02	14.96	122.59*	0.71*
30: Office machinery and computers	..	..	..	..	..	..	..
31: Electrical machinery	47	0.80	15.12*	5.98	39.17	99.15*	0.31*
32: Radio, T.V. and communications	33	0.35	1.91*	1.61	4.57	16.17*	0.76*
33: Medical and optical instruments	43	0.72	2.40	2.31	8.76	37.15*	0.73*
34: Motor vehicles and trailers	50	0.24	1.76*	4.22	24.49	79.06*	0.57*
35: Other transport equipment	26	-0.23*	0.19*	0.39	3.12	1.37	0.96
36: Furniture	82	0.58	4.85*	5.18	35.47	125.73*	0.47*
All	1,444	0.40	3.36	8.42	115.86	2,108.59*	0.45*

Notes: Difference in means: two-sample t-test for differences in means with unequal variances (differences between groups H_0: difference = 0; H_1: difference ≠ 0).

Differences in variances: variance-ratio (F-test) test (H_0: ratio = 1; H_1: ratio ≠ 1). Industry-level tests compare industry means and variances to the remaining aggregate sample.

.. Not reported given small sample < 10 plants. * Statistically significant at the 5 percent level.

For the aggregate sample the 10th, 25th, 50th, 75th and 90th percentiles measures were: -0.57, -0.33, 0.00, 0.45 and 1.67.

Table 4.9d: Plant Growth Rate Distributions by Industry, 1986-1997

Sample (SIC)	Plants (N)	Mean	Variance	Skewness	Kurtosis	D'Agostino	Shapiro-Wilk
15: Food and beverages	125	1.16	11.15*	5.36	38.87	174.15*	0.45*
16: Tobacco products	..	..	..	..	..	..	..
17: Textiles	48	1.01	4.29*	3.07	12.82	54.85*	0.62*
18: Wearing apparel	46	0.36*	0.80*	1.44	4.41	17.02*	0.84*
19: Leather products	11	-0.23*	0.13*	0.32	2.83	0.74	0.96
20: Wood and wood products	49	0.52*	0.83*	1.29	4.18	15.01*	0.87*
21: Pulp and paper products	54	1.10	8.53*	4.78	29.47	91.14*	0.48*
22: Publishing and printing	72	0.71	4.70*	6.88	54.51	140.34*	0.36*
23: Coke, petroleum and nuclear fuel	..	..	..	..	..	..	..
24: Chemicals and chemical products	88	0.92	2.46*	2.55	10.71	67.19*	0.74*
25: Rubber and plastic products	104	0.82	3.97*	5.20	37.59	149.24*	0.52*
26: Non-metallic mineral products	72	1.08	15.69	5.48	36.69	119.58*	0.38*
27: Basic metals	83	0.38*	0.93*	1.53	4.94	29.39*	0.83*
28: Fabricated metal products	190	0.64	1.38*	2.46	11.84	122.03*	0.78*
29: Other machinery and equipment	141	0.60	2.84*	4.36	27.43	162.76*	0.57*
30: Office machinery and computers	15	1.81	12.29	2.07	6.26	19.28*	0.66*
31: Electrical machinery	57	3.37	336.65*	7.27	54.26	125.54*	0.16*
32: Radio, T.V. and communications	46	0.77	1.99*	1.37	4.46	16.28*	0.85*
33: Medical and optical instruments	46	1.13	9.16*	3.99	18.96	69.60*	0.46*
34: Motor vehicles and trailers	45	0.33*	0.63*	1.22	3.97	12.67*	0.89*
35: Other transport equipment	21	0.76	2.08*	1.67	4.98	14.67*	0.77*
36: Furniture	86	0.69	1.71*	1.64	5.51	34.57*	0.82*
All	1,417	0.86	17.77	25.61	818.94	3,535.06*	0.18*

Notes: Difference in means: two-sample t-test for differences in means with unequal variances (differences between groups H_0: difference = 0; H_1: difference $\neq 0$).
Differences in variances: variance-ratio (F-test) test (H_0: ratio = 1; H_1: ratio $\neq 1$). Industry-level tests compare industry means and variances to the remaining aggregate sample.
.. Not reported given small sample < 10 plants. * Statistically significant at the 5 percent level.
For the aggregate sample the 10th, 25th, 50th, 75th and 90th percentiles measures were: -0.42, -0.15, 0.23, 1.00 and 2.25.

Non-metallic mineral products, Basic metals, Other machinery and equipment and Radio, television and communications industries in 1972-1997, the Leather products, Wood and wood products, Coke, petroleum and nuclear fuel, Basic metals, Other machinery and equipment and Other transport equipment industries in the 1977-1986 period, and the Wearing apparel, Leather products, Wood and wood products, Basic metals and Motor vehicles and trailers industries in the period 1986-1997.

Application of F-tests for homogeneity of variances, with a null hypothesis of a variance-ratio of 1 versus a non-unity alternative hypothesis, for differences in industry-level growth rate variances, produced a large number of statistically significant results, in 16 of the 20 samples in 1972-1977, 16 out of 18 samples in 1972-1997, in 18 of the 20 tests in 1977-1986 and in 18 of the 20 tests in 1986-1997. However, only limited consistency was found in the industries exhibiting the highest growth rate variances across the different time periods.

The estimated skewness and kurtosis indices suggest some considerable variation in reported values across industries, with generally smaller positive skewness indices than observed for the aggregate distributions. In only one industry in one period (1972-1977) was the growth rate distribution negatively skewed, Coke, petroleum and nuclear fuel. Only three industries provided positive skewness indices above that of the aggregate distribution in the 1972-1977 period: Food and beverages, Basic metals and Fabricated metal products, with skewness indices of 7.62, 8.07 and 6.50, respectively. A similar feature was also found for the kurtosis indices which, in combination with the skewness indices, suggests that highly extreme growth rates are observed across industries rather than being concentrated in a small number of industries.

The industrial distributions generally exhibited kurtosis indices much higher than either the normal or Laplace distributions, with only half of the 58 reported kurtosis indices being single-digit values. As a crude comparison, if from these industries exhibiting single-digit values a central threshold of kurtosis indices of 4.5 (halfway between the normal and Laplace benchmarks) is established with a similar variation in the other direction creating a bound of 1.5 - 4.5 for the normal and 4.5 - 7.5 for the Laplace distributions, of these observations 13 could be described as being closer to the normal benchmark and 11 to the Laplace. Inspection of the industry statistics tends to suggest limited consistency in the reported kurtosis indices across the four time periods.

Application of the D'Agostino et al. and Shapiro-Wilk tests, again at the 5 percent level of significance, both rejected the normal distribution as a reasonable description in all but 6 industry samples, with half of these non-rejections of the normal distribution provided by the Leather products industry, the others being Radio, television and communications in 1972-1977 and Coke, petroleum and nuclear fuel and Other transport equipment in 1977-1986. In one further sample, Coke, petroleum and nuclear fuel in 1972-1977, the Shapiro-Wilk test did not reject normality, in contrast to the D'Agostino et al. test.

These results appear generally consistent with the recent findings of Bottazzi, Coad, Jacoby and Secchi (2005), amongst others, reporting evidence of "fat tails" of the growth rate distribution, but with considerable heterogeneity across the industry samples. Although specific statistical distributions may therefore approximate to each of these industry distributions, the results do not provide support for either of the examined distributions as being capable of consistently describing each of the actual growth rate distributions. It is of course, feasible that such heterogeneity may be even more significant at lower levels of aggregation, but this remains to be confirmed. Additionally, despite this rejection of the normal (and Laplace) distributions in general, and for many industry-level samples, it should be recognised that these results may have been affected by the exclusion of non-surviving plants, which are likely to have been drawn disproportionately from the left tail of the growth rate distribution. It is unclear whether the inclusion of such plants would in any cases improve, or not, the statistical acceptability of the normal distribution in this context.

4.8. Conclusions

This chapter has presented an analysis of the size and growth rate distributions of manufacturing plants from the WRME database, testing specifically the lognormal size distribution postulated Gibrat's Law, and the assumed Gaussian (normal) distribution of growth rates. Three principle methods were employed in this analysis: examination of the summary statistics of the moments of the observed (transformed) plant size distributions, focussing particularly on constructed indices of skewness and kurtosis, the statistical tests of D'Agostino et al. (1990) and Shapiro-Wilk (1965) (or Shapiro-Francia, 1972) comparing the observed size distributions to that of the lognormal (or normal) distributions, and the use of non-parametric kernel density estimation to graphically examine the detail of the actual plant size and growth rate distributions.

Given the sensitivity of the kernel density approach to the choice of the band-width some experimentation was conducted using a range of band-widths and band-width selection criterions. The Gaussian and Epanechnikov kernel functions were employed, the latter primarily as a check on the robustness of the results, with further sensitivity analysis conducted to examine the effects of changes to the number of points at which the kernel density estimate was evaluated.

The analysis of summary statistics relating to the moments of the size distribution of manufacturing plants and the kernel density estimates suggested a number of empirical features. First, for the logarithmic transformation of plant sizes, the aggregate cross-section plant size distributions provided evidence of generally stable plant (log)size distributions, but with consistently observed slight positive skewness and leptokurtosis, a finding consistent with a large number of empirical studies, including Hart and Prais (1956). The D'Agostino et al. test and the Shapiro-Francia test conducted on each of the annual observations between 1966-2003 rejected the null hypothesis of (log)normality of the aggregate plants size distributions at the 5 percent level of significance; strictly rejecting the lognormal distribution and Gibrat's Law in its strongest form.

However, given that the degree of reported skewness and kurtosis was generally modest, the limitations discussed earlier regarding the testing of extreme hypotheses, and the supplementary issues such as the partial truncation of the data which could alter the statistical properties of the distribution, this evidence is not highly unsupportive of the actual size distribution of plants being heavily influenced through a Gibrat-type stochastic process, but perhaps with some additional complicating factors.

The analysis of the evolution of plant size distributions provided evidence of modest ageing and selection effects, the latter functioning through the exit of plants presumably not operating efficiently or at optimal size. The results were not strongly suggestive of any consistent tendency for a simple evolution towards the single-peaked, symmetrical distribution represented by the (log)normal distribution. Examination of the oldest age-classes provided some indication of a flat peaked distribution, and possibly bi-modality in all three samples considered. However, the relatively small sample sizes for some age-classes should be recognised in the interpretation of the kernel density procedures and therefore the strength of conclusions that can be drawn on this point.

Further analysis comparing the distributions of the surviving plants in their initial and final years in each panel provided evidence of some slight tendency towards a more symmetrical size distribution over time, although the employed statistical tests rejected the null hypothesis of a (log)normal distribution in each year of each cross-section sample, even in the long 1972-1997 sample. The previous results therefore do not appear to be an artifact of any particular selection of the time-span for each dynamic cross-section.

Whilst some stability in the size distributions was observed at the aggregate-level, the statistical tests provided evidence of considerable heterogeneity in the evolution of the industry-level distributions of plant sizes. Though examination of the indices of skewness and kurtosis suggested that such deviations were generally modest, the formal statistical tests comparing the goodness-of-fit of the actual and (log)normal distributions provided a number of rejections of the lognormal distribution. However, across all samples the null hypothesis of lognormality was more often not rejected than rejected, with rejections in 32.1 percent and 47.9 percent of industry samples using the D'Agostino et al. and Shapiro-Wilk tests, respectively.

Although in some cases there was suggestive evidence of some convergence towards the lognormal distribution over time for some industries, this was not observed as a general pattern, suggesting a role for industry-specific factors influencing the emergence of the observed size distributions, and supporting the views of Schmalensee (1989) and Sutton (1995a) regarding the lack of support for any single statistical distribution in providing an adequate description of firm (or plant) size distributions across different industry samples. The available data did not, however, permit further specific testing of such possible industry-specific factors.

It is not clear whether those industry-level samples rejected as lognormal may evolve towards the lognormal over longer time durations than analysed or indeed to other distributions. However, the evidence from the long-panel suggests that such divergences may be fairly persistent, even if such an evolution towards the (log)normal distribution were to occur. Furthermore, in many cases these distributions exhibited some degree of stability remaining either lognormal or not lognormal over the periods examined.

Investigation of the size distribution of cohorts of new entrant plants, whilst suggesting some specific cohort effects, provided indices of skewness and kurtosis exhibiting deviations from those of the strict (log)normal distribution which in almost all cases were fairly modest and not inconsistent with such a size distribution (but showing larger deviations in the initial years for the 1975 and 1977 cohorts). However, again, more formal tests for normality of the evolving cohort size distributions provided mixed evidence regarding any consistent convergence towards the (log)normal distribution over time.

Across these plant-level samples, it is therefore clear that, whilst in some cases the lognormal distribution implied by Gibrat's Law may provide a reasonable broad first approximation to some actual plant size distributions (specifically for some industries), more generally, it cannot be concluded that the lognormal distribution is across all of the examined time periods and samples (even within the manufacturing sector) capable of providing a strictly statistically appropriate representation of the actual manufacturing plant data.

Turning to the distribution of plant growth rates, a number of features were observed. First, the calculated growth rate distributions were highly suggestive of potential left-sided truncation, which may plausibly be the consequence of the exit of weaker performing plants. Second, analysis of the tabulations of growth rates across each time period pointed to considerable heterogeneity of performance in terms of employment growth. Third, examination of the distribution of growth rates using kernel density estimation suggested highly-peaked distributions, with in many cases kurtosis indices pointing to considerably "fatter tails" (and therefore a higher probability of extreme plant growth rates) than would be expected in either the normal or Laplace distributions). However, for annual growth rates and specific band-widths kernel density estimates similar to the normal distribution were in some cases observed.

At the industry-level the mean growth rates were generally positive, with some statistically significant differences across industries. Statistical tests also indicated a relatively high proportion of statistically significant differences between industry growth rate variances (compared to the rest of the sample), observed in 68 of the 78 industry samples. A significant degree of heterogeneity was observed in the skewness and particularly the kurtosis indices across industries, with generally smaller positive skewness (and kurtosis) indices than observed for the aggregate distributions, and only one industry exhibiting negative skewness.

Kurtosis indices were frequently much higher than either the normal or Laplace distributions, with only half of the 58 reported kurtosis indices exhibiting single-digit values. Of those which could be crudely described as approximating to the normal or Laplace distribution benchmarks, 13 could be described as being a somewhat better approximation to the former and 11 to the latter. However, only limited consistency was found in the reported kurtosis indices across the four time periods. Application of the D'Agostino et al. and Shapiro-Wilk tests, again at the 5 percent level of significance, rejected the normal distribution as a reasonable description in all but 6 industry samples. Overall, such results are generally consistent with the recently emerging evidence of "fat tails" of the growth rate distribution, but with considerable heterogeneity across the industry samples.

5

GIBRAT'S LAW AND PLANT GROWTH: A DISAGGREGATED ECONOMETRIC ANALYSIS

5.1. Introduction

Building on the statistical tests of the lognormal distribution as a reasonable description of the distribution of plant sizes conducted in the previous chapter, this chapter aims to directly test the validity of the three primary propositions of Gibrat's Law: that plant growth is independent of initial size, that the growth rate variance is independent of initial size and, that there is no persistence in plant growth, for the same manufacturing dataset using econometric methods. A range of econometric modelling issues are considered, with a focus of this analysis being on the effects of such estimation procedures. Gibrat's Law is examined across an array of disaggregated samples and time periods in order to provide some new insights into the validity of this particular model of plant (or firm) growth.

The structure of the chapter is as follows. Section 5.2 provides a statistical analysis of the mean and variance of plant growth rates across constructed size-classes. A regression approach is then used in Section 5.3 to provide an initial test the third proposition of no persistence in plant growth rates. The first proposition of Gibrat's Law, regarding the independence of plant size and growth, is examined in Section 5.4 using the standard logarithmic growth model. The approach then addresses a number of potentially relevant econometric modelling issues such as functional form and sample selection bias, with these issues presented in stages to examine the effects of these procedures on the assessed validity of Gibrat's Law. Specifically, the econometric model is expanded in Section 5.5 to consider the role of plant age on the estimated plant size-growth relationship, and non-linearities in the size-growth relationship in Section 5.6. Section 5.7 examines the effects of sample selection bias.

Various regression and sample selection model specifications of Gibrat's Law are estimated for a range of periods covering U.K. economic cycles, and also the expansionary and contractionary periods within these cycles. The analysis also provides a detailed assessment of Gibrat's Law for industry-level samples of plants, by age-classes, and by plant structure and ownership characteristics. In examining these samples, the appropriateness of the three alternative versions of Gibrat's Law: for all plants, surviving plants and for those consistent with a measure of the minimum efficient scale (MES) of production are considered. Section 5.8 concludes the chapter by drawing together the results from this analysis and discussing the consistency of these observations with those from previous empirical studies.

5.2. The Empirics of Plant Growth Rate Variance

To provide an initial statistical investigation of the first and second propositions of Gibrat's Law, that the growth and variance of plant growth rates are independent of initial plant size, plant growth rates were calculated across size-classes constructed so that the upper bound of each size-class is double that of the size-class immediately below (this size classification framework was employed by, for example, Hart and Prais, 1956; and Dunne and Hughes, 1994) for each of the four periods. Although studies such as Alexander (1949), Dyckman and Stekler (1965), Stanley et al. (1996), Amaral et al. (1997) and Sutton (2001) have pointed to specific forms of deviations from the postulated independent relationship between firm (or plant) size and the variance of growth, this analysis aims to identify only whether any deviations from such a relationship are observed for the WRME dataset, rather than to establish any specific functional form that any such deviations might take.

Consistent with previous studies the plant (proportional) growth rates were defined as ($X_{it} - X_{it-1}/X_{it-1}$), where X_{it} is the size plant i at time t. In some cases size-classes had to be aggregated due to the small sample sizes. These growth rates for surviving plants in each size-class for each of the periods 1972-1977, 1977-1986, 1986-1997 and 1972-1997 are presented in Table 5.1. Plant growth rates are observed to be above the average for the complete sample of surviving plants in the two smallest size-classes with up to 16 employees in each period. In fact, for the very smallest (aggregated) size-class of $0 \leq 8$ employees the mean growth rate of this group was between 3.5 and 5 times that of the mean growth rate of the next highest size-class. The mean growth rates declined fairly consistently with increased size-class, negative average growth rates being observed for the three largest size-classes of more than 128 employees in all periods except for the third largest size-class in 1986-1997.

Differences in these mean growth rates between adjacent size-classes were examined using a t-test for unequal variances following the Welch (1947) procedure (see also Aspin, 1948), with the statistical tests conducted at the 95 percent confidence level. These statistical tests (presented in Table 5.1) suggested significant differences between the growth rates of the very smallest (surviving) plants in the size-class $0 \leq 8$ employees and the next smallest size-class $8 \leq 16$ employees in each period.

Although no further statistically significant differences in mean growth rates across size-classes were detected in the 1972-1977 sample, further differences were observed in the other periods. For the 1977-1986 sample statistically significant differences were additionally found between the mean growth rates of all the size-classes except for between the $16 \leq 32$ and $32 \leq 64$ and the $64 \leq 128$ and $128 \leq 256$ employees classes. In the 1986-1997 period, in addition to the smallest size-classes, statistically significant differences were found between both the 16 $\leq$ 32 and 32 $\leq$ 64 employees classes and the 128 $\leq$ 256 and 256 $\leq$ 512 classes. Finally, the long 1972-1997 period exhibited additional statistically significant differences in mean growth rates between the 8 $\leq$ 16 and 16 $\leq$ 32, the 64 $\leq$ 128 and 128 $\leq$ 256 and the 128 $\leq$ 256 and 256 $\leq$ 512 employees size-classes. In combination this data provides some initial evidence suggesting that plant growth declines with size, rejecting the first proposition of Gibrat's Law.

The growth rate variances for the same size-classes and time periods are shown in Table 5.2. Given that these plants are the survivors in each period it may be expected that these statistics under-represent the magnitude of the real variance of growth rates, with those plants experiencing significantly below average growth rates exiting the sample. The data suggests some potential inconsistency with the second proposition of Gibrat's Law with the variance of growth rates showing a declining relationship (albeit with some very minor differences for some individual sizes-classes) in each sample with increasing plant size. The growth rate variance is generally only above that calculated for the whole sample in the smallest size-class being notably higher in each time period, (especially so in the 1986-1997 period), except for the very short period 1972-1977 where this is also the case for plants with up to 16 employees.

Following Dunne and Hughes (1994) the equality of the variance of plant growth rates across the size-classes was tested using the F-test for homogeneity of variances (with sensitivity analysis conducted using the robust test for equality of variances available in STATA version SE 9.0, 2005). Statistically significant differences were found in 21 of the 28 samples examined, being found in all of the tests between the smallest three size-classes. The frequency of rejection of the null hypothesis of no difference in variances however declined with increased size-class, but still being rejected in two of the

Table 5.1: Mean (Proportional) Plant Growth, Surviving Plants

Sample (Size-class: employees)	Group	1972-1977		1977-1986		1986-1997		1972-1997	
		Plants (N)	Mean	Plants (N)	Mean	Plants (N)	Mean	Plants (N)	Mean
$0 \leq 8^{\dagger}$	1	192	1.39	206	1.94	134	4.25	58	6.73
$8 \leq 16$	2	337	0.28	272	0.55	311	0.99	116	1.37
$16 \leq 32$	3	327	0.16	264	0.22	293	0.78	108	0.64
$32 \leq 64$	4	266	0.10	199	0.19	212	0.39	84	0.25
$64 \leq 128$	5	211	0.04	182	0.01	192	0.24	106	0.19
$128 \leq 256$	6	200	-0.01	151	-0.12	129	0.11	94	-0.11
$256 \leq 512$	7	112	-0.05	84	-0.25	84	-0.13	56	-0.43
$512 \leq 32{,}768^{\dagger}$	8	109	-0.06	86	-0.41	62	-0.22	71	-0.53

Notes: † Size-classes have been aggregated.
Statistical test for differences in means: two-sample t-test for unequal variances with H_0: difference = 0;
H_1: difference ≠ 0.

Statistically significant differences at the 5 percent level:
1972-1977: G1 - G2
1977-1986: G1 - G2; G2 - G3; G4 - G5; G6 - G7; G7 - G8
1986-1997: G1 - G2; G3 - G4; G6 - G7
1972-1997: G1 - G2; G2 - G3; G5 - G6; G6 - G7

Table 5.2: Variance of (Proportional) Plant Growth, Surviving Plants

Sample (Size-class: employees)	Group	1972-1977		1977-1986		1986-1997		1972-1997	
		Plants (N)	Variance	Plants (N)	Variance	Plants (N)	Variance	Plants (N)	Variance
$0 \leq 8^\dagger$	1	192	5.52	206	10.39	134	159.07	58	50.59
$8 \leq 16$	2	337	0.67	272	5.51	311	2.12	116	9.62
$16 \leq 32$	3	327	1.00	264	0.82	293	3.29	108	3.32
$32 \leq 64$	4	266	0.32	199	0.93	212	1.46	84	0.86
$64 \leq 128$	5	211	0.22	182	0.53	192	0.45	106	1.48
$128 \leq 256$	6	200	0.17	151	0.27	129	0.56	94	0.32
$256 \leq 512$	7	112	0.15	84	0.19	84	0.15	56	0.18
$512 \leq 32{,}768^\dagger$	8	109	0.11	86	0.01	62	0.25	71	0.22

Notes: † Size-classes have been aggregated.
Statistical test for differences in variances: variance-ratio (F-test) test, H_0: ratio = 1; H_1: ratio ≠ 1.

Statistically significant differences at the 5 percent level:
1972-1977: G1 - G2; G2 - G3; G3 - G4; G4 - G5
1977-1986: G1 - G2; G2 - G3; G4 - G5; G5 - G6; G7 - G8
1986-1997: G1 - G2; G2 - G3; G3 - G4; G4 - G5; G6 - G7; G7 - G8
1972-1997: G1 - G2; G2 - G3; G3 - G4; G4 - G5; G5 - G6; G6 - G7

four time periods, 1977-1986 and 1986-1997 for the largest size-classes. These statistical tests (with similar results provided the robust test for equality of variances) point to the rejection of the second proposition of Gibrat's Law, with the variance of growth rates declining with increased plant size.

In combination with the significantly higher mean average growth rates and also higher exit rates for this smallest size-class, this evidence is broadly consistent with the existing literature finding considerable turbulence in the very smallest size-classes (Sutton, 1997; Caves, 1997) and coherent with the postulation of rapid convergence towards some industry-specific MES for those surviving plants.

5.3. The Empirical Properties of Plant Growth Persistence

Although modest persistence may not preclude the emergence of firm (or plant) size distributions approximating to the lognormal (Hay and Morris, 1990), such persistence provides a strict rejection of the third proposition of Gibrat's Law (asserting that there is no persistence in growth rates). The identification of potential persistence is also important since as Chesher (1979) observed, positive serial correlation in growth rates could give rise to bias in the OLS regression estimates of the β_1 coefficient in the growth model specification. However, such upwards bias would moderate with increased time duration of the growth process (Chesher, 1979; Dunne and Hughes, 1994, p.129).

To provide an initial analysis of this proposition, the persistence of plant growth was empirically examined using the model:

$$G_{iT-t} = \alpha_p + \beta_p G_{it-b} + \varepsilon_{iT-t} \qquad (5.1)$$

where G_i is the growth of plant i (being the change in the logarithm of plant size over the respective time periods), T is the final year of each sample, t is the initial year of analysis and b is the base-year prior to the initial year (the regression coefficient is denoted by β_p to clearly distinguish this from the latter considered size-growth model regression coefficient β_1). The models were estimated over consecutive five-year periods using OLS with significance tests conducted at the 5 percent level testing the hypothesis H_0: $\beta_p = 0$, a null of no persistence, versus H_1: $\beta_p \neq 0$.

In choosing to use five-year time periods (consistent with Dunne and Hughes, 1994), the analysis exactly covers the shorter 1972-1977 panel but does not correspond exactly to the full length of the other examined time periods. Extending the span of the examined growth periods would, however, only be expected to give rise to observed weaker persistence given that many carry-over factors are likely to have moderated over longer time periods. Since there may be industry-specific factors, such as new technologies and innovations, market demand changes etc., these persistence effects were also estimated at the 2-digit industry-level, and also by age-classes (as defined in the previous chapter) and by size-classes given the possibility that growth rate persistence may vary over plant sizes[72].

A range of regression diagnostic tests were conducted (and are similarly conducted on the latter OLS regressions), including White's (1980) test for heteroscedasticity and Ramsey's (1969) RESET test for omitted variables. To test the normality of the error terms, the estimated OLS residuals were examined (see White and McDonald, 1980, for support of this procedure) applying both the D'Agostino et al. (1990) test and the variant of the Shapiro-Wilk (1965) test (recommended in this context by Maddala, 1977, pp.305-308), the Shapiro-Francia (1972) test suitable for the larger (aggregate) sample sizes encountered here. The Jarque-Bera (1980) test is not employed

---

[72] Dunne and Hughes (1994) and Bottazzi et al. (2005), for example, report negative serial correlation in the growth rates of small firms and positive serial correlation for larger firms.

due to its poor performance in small samples (Kennedy, 1998, p.306). The results of these diagnostic tests are presented in Table A5.6.

Briefly reviewing these diagnostics, White's test for heteroscedasticity did not reject the null hypothesis of no heteroscedasticity at the 5 percent level of significance in any of the aggregate samples. For the age-classes, size-classes and industry-level samples, White's test only rejected the respective null hypothesis in 8, 10 and 6 samples, respectively, in total in 24 out of the 89 samples. The uncorrected regression estimates are therefore reported (maintaining a consistent approach across all samples), but implementing White's correction for the samples where heteroscedasticity was detected did not alter the properties of the reported results. Ramsey's RESET test was rejected in less than half of the samples, in 15, 14, and 11 samples in each of the periods, respectively[73]. However, noting this potential limitation in these specific samples, the nature of persistence, and specifically the time-span of the persistence variables, is considered further in the later serial correlation models.

The result of these regressions, summarised in Table 5.3, suggest generally weak persistence in plant growth rates. In all three aggregate samples the β_p coefficients, estimated as -0.01, -0.00 and -0.12, respectively, were not significantly different from zero, and therefore did not reject the null hypothesis of no persistence in growth. The R^2 value was in all three cases close to zero (0.00 to two decimal places), indicating that the past growth variable explained very little of the subsequent growth performance of plants.

---

[73] The D'Agostino et al. and Shapiro-Francia tests indicated that in all cases except for the Textiles, Leather products, Pulp and paper products, and Coke, petroleum and nuclear fuel industries in the 1972 sample, Leather products and Radio, T.V. and communications in the 1977 sample and Leather products in the 1986 sample, normality of the residuals could be rejected by at least one of the tests. Closer inspection of the associated skewness and kurtosis indices suggested some tendency towards positive skewness (only being negative, -0.49, for Leather products in the 1986 sample), and positive kurtosis, being in single-digits in only 15, 11 and 9 of the samples, respectively. However, subject to meeting the other assumptions of the Classical Linear Regression model (discussed later in this chapter), the OLS estimates would still remain the best linear unbiased estimator of the parameter coefficients (Kennedy, 1998, pp.43-44). In each case the residuals were suitably described by a mean zero.

The regression results also suggested statistically significant persistence for only 5 of the 15 age-class samples - for the 11 - 20 years age-class in each of the three time periods, for the 6 - 10 years age-class in 1977 initial-year sample and for the +20 years age-class in the 1986 initial-year sample (where the β_p coefficient was notably larger than observed in the other age-class estimates), in no cases being statistically significant in either of the under 5 years age-classes. In all but two of the 15 age-class samples the estimated persistence coefficient, β_p, was negative (the other two being positive but not statistically significant). For all 15 samples, the low R^2 value again indicated that the past growth variable explained little of the subsequent growth performance of these plants.

For the size-classes, in both the 1972 and 1977 initial-year periods, statistically significant persistence was detected only for Small plants, whilst for the 1986 initial-year such an observation was only found for the Large plants size-class, which exhibited a much larger negative coefficient than in the other estimates. For all other size-class estimates the null hypothesis of no persistence could not be rejected at the 5 percent level of significance. Again, low R^2 values were observed.

Examination of the industry-level results also supported the finding of modest persistence in growth with the β_p coefficients found to be statistically significant in only 11 of the 59 samples considered. Specifically, the Pulp and paper products industry exhibited statistically significant persistence in all three time periods, with the Chemical and chemical products industry showing persistence in both the 1972 and 1986 initial period samples and the Non-metallic mineral products industry providing statistically significant persistence in the two earliest periods. The Basic metals, Fabricated metal products, Electrical machinery and Motor vehicles and trailers industries exhibited statistically significant persistence in one of the time periods only.

Consistent with the existing literature, the estimated regression coefficients

Table 5.3: Summary of Five-Year Plant Growth Persistence

Sample	1972					1977					1986				
	Plants (N)	α_p	β_p	$t\beta_p$	R^2	Plants (N)	α_p	β_p	$t\beta_p$	R^2	Plants (N)	α_p	β_p	$t\beta_p$	R^2
All	1,210	0.05*	-0.01	(-0.71)	0.00	1,367	0.04	-0.00	(-0.20)	0.00	1,444	0.40*	-0.12	(-1.08)	0.00
Age: 0 - 1 years	640	0.05*	0.01	(0.36)	0.00	138	0.30*	-0.00	(-0.13)	0.00	126	0.60*	-0.01	(-0.20)	0.00
Age: 2 - 5 years	155	0.05	-0.04	(-0.84)	0.00	295	0.22*	-0.09	(-1.09)	0.00	226	0.33*	0.02	(0.54)	0.00
Age: 6 - 10 years	113	0.12*	-0.12	(-1.22)	0.01	575	-0.06*	-0.11*	(-2.30)	0.01	271	0.41*	-0.06	(-0.78)	0.00
Age: 11 - 20 years	225	0.07	-0.20*	(-2.13)	0.02	154	0.09	-0.21*	(-2.38)	0.04	569	0.18*	-0.30*	(-2.17)	0.01
Age: +20 years	77	-0.03	-0.04	(-0.42)	0.00	205	-0.06	-0.13	(-1.68)	0.01	252	0.57	-3.76*	(-2.45)	0.02
Size: Micro	81	0.61*	-0.03	(-0.19)	0.00	157	0.72*	-0.08	(-1.61)	0.02	161	0.61*	-0.02	(-0.27)	0.00
Size: Small	648	0.06*	-0.04*	(-2.05)	0.01	702	0.08*	-0.11*	(-2.58)	0.01	826	0.35*	-0.10	(-1.48)	0.00
Size: Medium	279	-0.01	-0.08	(-1.67)	0.01	326	-0.14*	-0.01	(-0.23)	0.00	295	0.09*	-0.01	(-0.15)	0.00
Size: Large	202	-0.06*	-0.10	(-1.47)	0.01	182	-0.23*	-0.01	(-0.23)	0.00	162	-0.57	-7.63*	(-2.58)	0.04

Table 5.3: Summary of Five-Year Plant Growth Persistence, continued

Sample	1972					1977					1986				
	Plants (N)	α_p	β_p	$t\beta_p$	R^2	Plants (N)	α_p	β_p	$t\beta_p$	R^2	Plants (N)	α_p	β_p	$t\beta_p$	R^2
15: Food and beverages	150	0.13	-0.04	(-0.35)	0.00	139	0.20*	-0.18	(-1.07)	0.01	136	0.18*	-0.00	(-0.00)	0.00
16: Tobacco products	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
17: Textiles	43	0.06	0.03	(0.44)	0.00	53	0.02	-0.13	(-0.96)	0.02	40	0.21*	0.00	(0.02)	0.00
18: Wearing apparel	67	-0.02	0.10	(1.10)	0.02	65	-0.00	0.00	(0.02)	0.00	52	0.05	0.07	(0.94)	0.02
19: Leather products	18	-0.03	-0.37	(-1.49)	0.12	14	-0.11	-0.27	(-0.74)	0.04	13	-0.07	-0.46	(-1.43)	0.16
20: Wood and wood products	83	0.02	-0.26	(-1.58)	0.03	75	0.02	0.05	(0.72)	0.01	65	0.14*	0.06	(1.02)	0.02
21: Pulp and paper products	25	-0.07	0.19*	(2.22)	0.18	30	-0.07	0.22*	(6.49)	0.60	41	0.08	0.31*	(6.09)	0.49
22: Publishing and printing	75	0.10	-0.09	(-0.50)	0.00	92	0.07	0.11	(1.18)	0.02	87	0.18*	-0.02	(-0.27)	0.00
23: Coke, petroleum and nuclear fuel	10	0.01	0.40	(1.37)	0.19	10	-0.04	-0.16	(-0.45)	0.03	..	..	..	..	..
24: Chemicals and chemical products	47	0.09	0.65*	(3.50)	0.21	57	-0.13*	0.02	(0.42)	0.00	78	0.17*	0.14*	(2.17)	0.06
25: Rubber and plastic products	38	0.02	0.05	(0.73)	0.01	61	0.31	-0.11	(-0.52)	0.00	95	0.27*	0.01	(0.20)	0.00
26: Non-metallic mineral products	102	-0.10*	-0.16*	(-2.09)	0.04	104	0.11	-0.34*	(-2.45)	0.06	93	0.35	-0.13	(-0.45)	0.00
27: Basic metals	83	-0.02	-0.16*	(-2.01)	0.05	85	-0.12*	0.00	(0.01)	0.00	82	0.22*	0.03	(0.48)	0.00
28: Fabricated metal products	145	0.00	0.07	(1.42)	0.01	178	0.08	0.01	(0.21)	0.00	191	0.30*	0.09*	(2.24)	0.03
29: Other machinery and equipment	114	0.18*	-0.06	(-0.96)	0.01	149	-0.05	0.01	(0.25)	0.00	157	0.50	-0.40	(-1.26)	0.01
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
31: Electrical machinery	34	0.15*	-0.01	(-0.08)	0.00	41	-0.08	0.23*	(2.20)	0.11	47	4.69	-4.75	(-1.35)	0.04
32: Radio, T.V. and communications	20	-0.08	-0.03	(-0.88)	0.04	21	-0.06	-0.15	(-0.74)	0.03	40	0.55*	0.04	(0.26)	0.00
33: Medical and optical instruments	23	0.01	0.07	(0.48)	0.01	34	0.13	-0.08	(-0.49)	0.01	44	0.16	0.17	(1.25)	0.04
34: Motor vehicles and trailers	33	0.03	0.03	(0.47)	0.01	50	0.16	-0.04	(-0.17)	0.00	51	0.11	0.26*	(2.58)	0.12
35: Other transport equipment	30	0.20	-0.99	(-1.95)	0.12	29	-0.12	-0.01	(-0.12)	0.00	24	0.54*	0.22	(0.84)	0.03
36: Furniture	62	0.10	-0.14	(-1.49)	0.04	70	0.12	-0.01	(-0.15)	0.00	81	0.16*	0.00	(0.04)	0.00

Notes: Model estimated using OLS (t-statistics in parentheses).

Plant size classified by initial year of past growth period. Micro: 0 - 8 employees, Small: 8 - 64 employees, Medium: 65 - 256 employees, Large: +256 employees.

F-statistics and t-statistics for the constant term, α_p, not reported.

.. Not reported due to sample < 10 plants. * Significantly significant at the 5 percent level.

suggest a rather mixed picture in terms of exhibiting either negative or positive persistence effects (the former consistent with above average growth rates being followed by a period of below average growth rates, and the latter with positive growth rates being followed by further positive growth rates). The industry-level analysis provided 31 samples with estimated non-negative coefficients (two being statistically significant in the first period, two in the second period, and four in the final period), with the remaining 28 industry samples exhibiting negative coefficients (two being statistically significant in the first period, one in the second period and none in the final period). At the industry-level, the past growth variable explained more than 10 percent of the variance of growth in only 10 cases.

In combination, these results at both aggregate and disaggregated levels provide support for the third proposition of Gibrat's Law that there is no persistence in plant growth rates, with past growth generally being a poor predictor of subsequent plant growth, a finding broadly consistent with, for example, Dunne and Hughes (1994), Wagner (1994), Hart and Oulton (1998b) and Lotti et al. (1999). Therefore it would be expected that the β_1 coefficient in the OLS estimated growth regression model would not be subject to significant bias from this source.

5.4. Plant Size and Growth

To examine the growth (and survival) dynamics of manufacturing plants, transition matrices (presented in Tables A5.1a-d, excluding disclosive cells) were constructed for each of the four examined time periods, 1972-1977, 1972-1997, 1977-1986 and 1986-1997. These transition matrices provide a summary of the movements of plants between size-classes and the survival of plants (by size-class) from the initial sample over the examined time periods, with the standardised probability of plant (growth) transitions between each size-class during the period given by the ratio of each cell value to the respective column total (Hart and Prais, 1956, p.162 and p.165).

Following Hart and Prais (1956) the size-classes were constructed on a geometric basis so that the upper bound of each size-class was twice the lower bound, with the plants classified by their size in each respective initial year using the employee numbers measure. Consequently, this construction permits a direct assessment of proportionate growth across plant size-classes (by examining the main top-left to bottom-right diagonal of the matrix and the distribution of observations around this diagonal). The uppermost size-class was deliberately allowed to contain zero plants in the initial year in order to permit positive growth from smaller size-classes, avoiding any artificial upper-bound censoring.

These transition matrices point to some consistency of an inverse relationship between plant size and the likelihood of failure across the samples, albeit being less strongly observed for the 1972-1977 sample. A similar inverse relationship was also identified in Mansfield (1962) and Dunne et al. (1989). The observed decline in failure rates with increased size, suggests the need to account for potential sample selection effects in the econometric analysis of the plant size-growth relationship, a point returned to later.

In addition, the transition matrices also indicate that the majority of plants are observed to remain within their initial size-class or move across only one size-class within the examined time periods, except for the very smallest size-classes where some greater degree of size-class mobility is observed (in part due to the smaller absolute size change required to move between size-classes) and for the longer time period over which larger changes in size-class would be more feasible. The observation of a concentration of plant transitions along the main diagonal axis of the transition matrices provides some broad support for Gibrat's Law, though with a slight downward "curvature" of the concentration of plants in the transition matrices in some samples indicating perhaps a modest tendency for mean reversion.

To formally examine the first proposition of Gibrat's Law, that plant growth is independent of initial size, regression methods were employed, consistent with the majority of the existing literature. Given the general rejection of any strong persistence effects on plant growth a logarithmic size-growth model of the form:

$$\Delta \ln X_{it} = \alpha + \beta_1 \ln X_{it-1} + v_{it} \qquad (5.2)$$

was employed where X_{it} is the size of plant i in time t, measured by employee numbers, v_{it} is an error term, and where $\Delta \ln X_{it} = \ln X_{it} - \ln X_{it-1}$. The null hypothesis for testing the first proposition of Gibrat's Law in the model specifications considered throughout this chapter, equation (5.2) and the relevant extensions set out latter, is therefore that H_0: $\beta_1 = 0$ (as discussed in the text relating to equation (1.10) in Chapter 1) versus the alternative hypothesis that H_1: $\beta_1 \neq 0$.

Almost all previous studies of Gibrat's Law have, at least initially, used OLS to estimate this size-growth relationship, frequently supplemented with additional maximum likelihood or Heckman two-step estimators employed in the sample selection model specifications. Subject to satisfying the assumptions of the Classical Linear Regression (CLR) model[74], the OLS estimator provides a number of advantageous properties which give rise to this being the "optimal", and therefore preferred, estimator (Kennedy, 1998, pp.43-44), including providing an unbiased estimator of β_1, with minimum variance,

---

[74] The Classical Linear Regression model requires a number of assumptions to be valid: that the mean, or expected, value of the error term v_i is zero, $E(v_i | X_i = 0)$; that there is no autocorrelation between the error terms, $cov(v_i, v_j) = 0$; that there is homoscedasticity of the error terms, $var(v_i | X_i) = \sigma^2$; that there is no covariance between v_i and X_i; and that the model is appropriately specified, including the appropriate variables and with the correct functional form, with the dependent variable being a linear function of the independent variables (Kennedy, 1998, pp.43-44; Gudjarati, 1988, pp.52-59).

being the Best Linear Unbiased Estimator (BLUE) of β_1. Further, where the disturbance terms are normally distributed the OLS estimator is the best estimator (not just the best linear estimator). The construction of the OLS estimator also means that it will be optimal in minimising the sum of squared residuals and in providing the highest R^2 (Kennedy, 1998, p.46).

As Kennedy (1998, p.28) notes, where the error terms are non-normally distributed, being heavy-tailed, robust estimators may in some cases be preferred to OLS[75] (which remains BLUE). However, Greene (2003, p.17) notes that such a normality condition is frequently considered as not necessary to produce appropriate results in the context of standard multiple regression. The specific diagnostic results from each of the regressions estimated with OLS are presented in stages throughout this chapter, but as will become apparent, in the majority of cases such an estimator provided a suitable starting point. Importantly, across the full range of samples examined in this chapter the assumption of a mean expected disturbance term of zero is appropriate, alleviating the possibility of a biased intercept problem (Kennedy, 1998, p.43).

However, before commencing such estimation, it is worth considering some other estimators which have also been considered in the context of analysing microdata with the short time dimension and large cross-section characteristics of this dataset, including pooled estimators and the fixed effects panel data estimator. As discussed in Hart and Oulton (1998, p.46), examining company accounts data, whilst pooling data can be used to increase the degrees of freedom, given the large number of degrees of freedom in their data there was no statistical need to pool the data (and they indeed follow the cross-section approach). A similar argument can also be made here. It is also of note that the study of Gibrat's Law conducted by Goddard, Wilson and Blandon

---

[75] See, for example, Bottazzi, Coad, Jacoby and Secchi (2005) who, in the context of examining firm growth, employed the Least Absolute Deviation (LAD) estimator, a robust estimator, being the maximum likelihood estimator when the error is Laplace distributed, this estimator providing the best unbiased estimator under these conditions and being more efficient than OLS (Kennedy, 1998, p.307).

(2002, p.331) also suggested that the OLS estimator of their cross-section specification exhibited similar properties to the pooled data estimator (although with lower power).

Additionally, although a small number of recent studies have conducted panel data estimates of the firm size-growth relationship, as Bell and Ritchie (1996) (discussed in Hart and Oulton, 1998b) note, cross-section estimates may be preferable to panel (or pooled) data where the estimated β_1 coefficient and firm-level (plant-level) fixed effects are not time invariant (Hart and Oulton, 1998b, p.41) even when industry fixed effects[76] are included in the model specification if these do not suitably reflect such firm-level factors. Hart and Oulton (1998b, p.38) argue that firm-level effects are highly unlikely to be constant over economic cycles making such fixed effects assumptions in the standard panel data models inappropriate, providing empirical support for such an objection in their analysis of U.K. companies.

Under such a misspecification systematic components may enter the error term giving rise to autocorrelation and inconsistent OLS parameter estimates, the estimated β_1 coefficient being excessively high where the autocorrelation coefficient is positive. Although such bias would be less significant in the case of weak autocorrelation (Hart and Oulton, 1998b, p.40) and could be potentially alleviated through the use of longer time-spans of cross-section data (Hart and Oulton, 1998b, p.40), such an issue might be of some relevance to the shorter time periods of analysis also examined here.

Furthermore, although some recent panel data studies have used panel unit root tests to test the random walk firm growth proposition, despite the increased power of these tests over the univariate tests there remains considerable controversy regarding the ability to effectively discriminate between a unit root and near unit root process. Franses (2002, p.40), for example, raises the potential "impossibility" of such a statistical testing

---

[76] Additionally, as shown by Nickell (1981), the fixed effects estimator is also downward biased in short panels (Hart and Oulton, 1998, pp.46-47; Goddard and Wilson, 2001, p.330).

procedure. A similar point is also made by Kennedy (1998, p.269).

Given these potential objections to the panel data approach, the first step conducted here is to estimate equation (5.2) using OLS cross-section methods over very short time periods to examine the variability of the regression parameters over time and economic cycles[77]. In forming such a cross-section model specification it is assumed that the intercept terms are homogeneous[78] so that $\alpha_i = \alpha$. Without such an assumption either the n + 2 parameters in the model (for example, with persistence terms) would be greater than the number of observations from the cross-sections preventing estimation or the intercept terms would need to be integrated within the disturbance terms (Goddard et al., 2002). In the latter case the induced correlation between the $\ln X_{it-1}$ terms and the disturbance terms would result in inconsistent OLS estimates of the β_1 regression coefficient (Goddard et al., 2002).

Specifications both with and without 2-digit industry-level dummy variables were estimated using White's (1980) heteroscedasticity-consistent estimator of the variance-covariance matrix of the OLS estimator, providing robust standard errors, given the results of White's test for heteroscedasticity[79], over both annual growth periods and following Hart and Oulton (1998b) across time series regimes corresponding to the contractionary and expansionary phases of the U.K. economic cycle (described in Chapter 4) covering the full time period of available data, 1966-2003. The inclusion of the industry dummy variables might be expected to absorb some of the plant-level effects (Hart and Oulton, 1998b) and therefore provides some attempt to control for

---

[77] Consistent with most studies (except Hall, 1987) no attempt is made to estimate the effects of any bias from potential measurement error.

[78] Such an assumption is employed here for consistency with previous cross-section studies, with some analysis of this feature examined using disaggregated samples. However, if plant-level heterogeneity is present, with firms exhibiting permanent, but unstable, differences in size, and convergence to different long-run sizes (Geroski, 1999) such a treatment could provide biased estimates of the β_1 coefficients.

[79] Where heteroscedasticity was detected, in the absence of such a correction the estimates would be inefficient (not minimum variance and giving rise to unnecessarily wide confidence intervals and potentially invalid inferences from t-tests and F-tests of significance), but not biased, with biased standard errors (Dunne and Hughes, 1994, p.130).

unobserved heterogeneity, which if ignored could give rise to inconsistent estimates (Hsiao, 2003, p.8).

A summary of the regression results, with test statistics at the 5 percent level of significance (given the large number of regressions only the relevant β_1 coefficient and associated t-statistics are presented) are set out in Tables A5.2 and A5.3. These suggest the rejection of Gibrat's Law in each of the time series regimes over 1966-2003 in both models specifications (with and without industry dummy variables), with mean reversion observed in all time periods. The β_1 coefficients of both specifications were generally similar, although often being slightly more negative in the model with industry dummy variables, with some variation across time periods.

Similar results also emerged for the annual cross-sections with Gibrat's Law again rejected in each sample under both model specifications. Although the inclusion of a one-year persistence term produced a non-rejection of the null hypothesis that H_0: $\beta_1 = 0$, in both 2001-2002 and 2002-2003, in all other year's the first proposition of Gibrat's Law could be rejected at the 5 percent level of significance. The R^2 values were generally small, consistent with prior expectations (Geroski, 1999).

Further analysis using the samples of plants surviving over each of the earlier described economic cycles and the long time period 1972-1997, shown in Table A5.4, produced similar results regarding the variability in the β_1 coefficient (with the first proposition of Gibrat's Law rejected in all years for the 1972-1977, 1977-1986 and 1986-1997 samples, and in all but two years in the longer 1972-1997 sample, the years 1976-1977 and 1994-1995), with coefficients, although being relatively modest for each annual estimate, being somewhat larger (more negative) than found in the annual cross-sections conducted over a broadly similar length time period by Geroski et al. (2003).

These results clearly indicate that the estimated β_1 coefficients are not time invariant exhibiting some cyclical variability, generally being most negative during recessionary periods, especially so in the early 1980s recessionary period. This variability supports the argument of Hart and Oulton (1998b) that it would be inappropriate to pool the data (Hart and Oulton, 1998b), with the cross-section approach employed in the remainder of the analysis.

Since the omission of plant-level effects correlated with the model regressors could produce inappropriately high OLS estimates of β_1 (Hart and Oulton, 1998b, p.37), following Hart and Oulton (1998b, p.17) simple correlation tests of the residuals from the cross-section models were conducted both across time series regimes and between annual cross-sections, for specifications with and without industry dummy variables. Though with some exceptions, for both the annual cross-sections and time series regimes the correlations were generally small (but exhibiting some consecutive periods of positive or negative coefficients), so that any bias from the omission of plant-level effects is likely to be modest, consistent with the finding of Hart and Oulton (1998b).

Although the above approach provides a detailed analysis of the variation of the estimated model parameters over time and minimises the effect of sample exit, as Hart and Oulton (1996, p.1247) note, over very short time periods the transitory components of firm (or plant) growth are likely to have increased influence relative to the permanent components giving rise of a downward bias in the OLS estimates of the true β_1 coefficient. To alleviate some of the noise in the very short cross-section estimates conducted above, and potentially reduce any inconsistency in the estimated regression coefficients arising from serially correlated error terms (Geroski et al., 2003), equation (5.2) was again estimated using OLS on the cross-sections of plants surviving over the three respective economic cycles and the long period covering 1972-1997.

The use of the cross-section covering the very long time period of 1972-1997 also provides the opportunity to examine the longer-run dynamics of surviving manufacturing plants, although these benefits may be somewhat countered by such a sample running the risk of neglecting the behaviour of more typical plants, and be biased towards supporting the convergence of plant sizes given the likely shakeout of non-optimal or under-performing sized plants (Geroski et al., 2003, p.51; see also Jovanovic, 1982; Ericson and Pakes, 1995; and Audretsch, 1995). However, given the analysis of a wide range of samples making use of the broader cross-section dimensions of the data such a potential issue is not considered to be highly problematic here.

Given the use of such time periods, macroeconomic variables are not included in the model specification[80]. Additionally, although many previous studies have included industry dummy variables when testing the validity of Gibrat's Law (an approach which is also employed here in some of the aggregate-level regressions[81]) to provide further detail regarding plant dynamics disaggregated estimates of the plant size-growth relationship for a number of industry, ownership and structure disaggregations, size-classes and age-classes are also produced, extending the disaggregated approach of Dunne and Hughes (1994). The estimation of separate regressions for constructed age-classes is consistent with, for example, Evans (1987a, 1987b). Some additional sensitivity analysis is also conducted through the inclusion of growth persistence terms.

The analysis by disaggregated size-class may be particularly important given that, as Hart and Oulton (1996, p.1248) reflect, concentration on a single estimated overall β_1 coefficient could provide to a misleading impression of the relationship between firm (or

---

[80] Also supporting this specification, as Geroski (1999) notes, although macroeconomic variables included in such models have frequently been found to be statistically significant (with often positive coefficients), they have not generally significantly added to the explanatory power of the models.

[81] Although previous research, such as Audretsch et al. (2002), suggests that industry-specific factors such as economies of scale, capital intensity and sunk costs, may affect the observed firm or plant size-growth relationship and the probability of survival, given the absence of suitable variables within the WRME dataset it is not feasible to examine the influence of such factors on any observed deviations from Gibrat's Law. The approach to considering the role of industry factors here is, first, to include industry-level dummy variables in the aggregate regressions and, second, to conduct regressions at the 2-digit industry-level. In the case of the former approach, for the main time periods examined the majority of these industry dummy variables were not statistically significant at the 5 percent level (significant in only 20 cases for industries with at least 10 plants). An attempt is also made to identify plants above the minimum efficient scale of production and test Gibrat's Law for these plants.

plant) size and growth if there are non-linearities in this relationship relating to the smaller firm size-classes (for which they found empirical evidence using sequential regressions for firms with less than eight employees)[82], especially where significant weight is given to such observations due to large sample sizes (Hart and Oulton, 1996, p.1248). In addition, they note the potential effect of increased employment through part-time employees examined using headcount measures giving rise to a particularly misleading impression of the rate of firm (or plant) growth for the smallest size-classes. However, such a comparison using alternative size measures cannot be considered here given the absence of alternative measures of plant size.

Following the size-classes constructed for the transition matrices, the (aggregated) size-classes are defined as 8 or fewer employees (consistent with the observed non-linearities found by Hart and Oulton, 1996), over 8 to 64 employees, over 64 to 256 employees and over 256 employees, respectively, size-classes broadly corresponding to accepted definitions of Micro, Small, Medium and Large firms (although it is recognised that the firms in which some of these plants are a part may be suitably classified in larger size-classes).

The separate regressions estimated for individual 2-digit level industries also assist in addressing the potential aggregation bias issue raised by Dunne and Hughes (1994)[83]. Further estimates are also provided for single-plant firms (which provides a link to the existing research conducted at the firm-level, although recognising that these samples are unlikely to be fully representative of such firms, in particular excluding the non-Wales located parts of multi-plant firms), for foreign owned plants (consistent with, for

---

[82] Rather than estimate sequential regressions further analysis was conducted here on highly disaggregated size-classes. While often providing small sample sizes, the increasingly negative regression coefficients found by Hart and Oulton (1996) with decreasing firm size were also detected in these samples. Such consistency with the extremely large dataset used by Hart and Oulton (1996), which they claim provides suitable coverage of the lower part of the size distribution, alleviates some concern regarding the properties of the retained samples of manufacturing plants below the official size threshold of this dataset.

[83] Such a disaggregated analysis is also supported by the differences in mean plant sizes across industries with the ranges (for samples of 2-digit industries with more than 10 plants) being 36.9 to 670.5, 26.0 to 542.6, and 38.8 to 247.2 employees for the 1972, 1977 and 1986 initial year cross-sections.

example, Blonigen and Tomlin, 1999) and for plants at or above the estimated MES, corresponding to the third version of Gibrat's Law (the first two versions of Gibrat's Law being represented by the aggregate regressions of surviving plants and the later sample selection models).

Given the observed consistent pattern of declining growth rate variance with increased size, raising the likelihood of heteroscedasticity in the model estimates, White's test was conducted on each of the initial aggregate regressions with the null hypothesis of homoscedasticity rejected in all four samples, exhibiting χ^2 values of 51.44, 6.22, 10.78 and 49.42, respectively. The Breusch-Pagan (1979) test for heteroscedasticity (with a null hypothesis of constant variance) also provided support for these conclusions (the respective tests for heteroscedasticity for all OLS estimated growth models over these time periods, except for the serial correlation models, are presented in Tables A5.7 and A5.8). Further evidence of heteroscedasticity was also detected using these tests for a number of disaggregated industry samples, size-classes and age-classes, though being less frequently statistically significant in the longer 1972-1997 samples.

Although examination of the residuals frequently suggested a broad pattern of declining heteroscedasticity with plant size, there was no simple transformation[84] (see Stanley et al, 1996, for the use of such a transformation approach) that appeared suitable across all of the very large number of aggregate and disaggregated samples considered in this analysis to alleviate the cases of observed heteroscedasticity. A weighted least squares regression approach to provide efficient parameter estimates was therefore not pursued. Given this observation, and the results of the tests for heteroscedasticity, consistent with most recent studies White's (1980) correction was applied to generate robust standard errors for each regression model, with this procedure applied in a consistent manner

---

[84] The use of model transformations to remediate heteroscedasticity also becomes more difficult when additional variables are included (Gudjarati, 1988, p.341) as conducted latter in this chapter.

across the regressions.

As an additional examination of the model, the D'Agostino et al. and Shapiro-Francia tests were applied to test for normality of the residuals. The results (presented in Tables A5.9 - A5.13 for all growth model regressions) suggest some mixed results rejecting (strict) normality with both tests for all but the Micro size-class, Wearing apparel, Leather products, Radio, T.V. and communications, and Medical and optical instruments industries in 1972-1977, but being rejected only in the Micro and Small size-classes, single-plant firms and Rubber and plastics products industry in 1972-1997[85]. Results were more variable for the 1977-1986 and 1986-1997 time periods being rejected by both tests in 16 cases in the former and 15 cases in the latter.

Closer inspection of the skewness and kurtosis indices suggested generally modest deviations from the index values associated with a normal distribution, with 58 of the 110 samples being (slightly) negative and all but 6 samples having single-digit kurtosis indices. Using the same basic rule employed earlier to examining the normal versus the Laplace distribution, a kurtosis value of 1.5 - 4.5 more coherent with the normal distribution and 4.5 - 7.5 to the Laplace distribution, the normal distribution found support in 69 cases and the Laplace in 30 samples. The use of the LAD estimator is therefore not generally supported by these results, although for those samples with residuals more appropriately described as Laplace distributed this may be an interesting area for further research.

The results of these OLS regression estimates with White's correction are presented in Tables 5.4a-c. For the regressions estimated on each of the full samples of surviving plants the β_1 coefficients were negative (being -0.11 for 1972-1977, -0.22 for 1977-1986, -0.27 for 1986-1997 and -0.41 for the period 1972-1997) and statistically

---

[85] Indeed, across the various model specifications the assumption of normally distributed residuals is most supported in the longer time-span of data.

significant at the 5 percent level, suggesting a process of mean reversion. Smaller plants therefore exhibited higher proportional growth rates than larger plants in each period, rejecting the first proposition of Gibrat's Law[86].

Briefly returning to the assumption of homogeneous plant-level effects. As Goddard et al. (2002) note, if the modelling approach fails to adequately capture individual plant-level heterogeneity the cross-section estimator would provide inconsistent estimates of β_1 being generally over-estimated in the case of $\beta_1 < 0$ (and biased towards accepting the first proposition of Gibrat's Law). Such a feature would, however, seem likely only to serve to strengthen the conclusions of this analysis of a statistically significant process of mean reversion.

As a robustness check the model was re-estimated sequentially removing extreme (outlier) growth values, excluding in stages the top and bottom 50, 100 and 200 plants, finding that it made no significant difference to the observed aggregate results. Consistent with the discussion of Geroski (1999), the models produced generally low R^2 values, indicative of significant size mobility across plants, although being higher for the longer time-spans (use of the alternative approach of regressing the logarithm of plant size in time t on the logarithm of size in time t - 1 (equation 1.6) produced much higher R^2 values consistent with those produced in studies such as Dunne and Hughes, 1994).

However, the results for size-classes and industries suggest somewhat more mixed conclusions regarding the first proposition of Gibrat's Law. In terms of the size-class estimates, the first panel covering the period 1972-1977 exhibited a pattern consistent with a number of previous disaggregated analyses (for example, Dunne and Hughes, 1994) with Gibrat's Law rejected for the Micro and Small size-classes, but not rejected for the Medium or Large size-classes. A similar result was also found for the

---

[86] As Hart and Oulton (1998b, p.28) note, although the β_1 coefficients for these time periods indicate that smaller plants grew proportionately faster than larger plants, this does not necessarily indicate higher levels of employment creation.

sample covering 1986-1997, where even though Gibrat's Law could be rejected for the three smallest size-classes, the random walk proposition again provided a reasonable description for the largest size-class. For the 1977-1986 time period Gibrat's Law was rejected for all size-classes, except for the Medium size-class. The analysis covering the period 1972-1997 suggested that Gibrat's Law could be rejected for all size-classes. For the constructed age-classes, with the exception of the 11 - 20 years age-class in the sample covering 1972-1977, all of the estimates suggested a statistically significant rejection of Gibrat's Law.

At the industry-level statistically significant rejections of Gibrat's Law (all with negative coefficients supporting a process of mean reversion) were found in 62 of the 78 2-digit industry-level samples: 13 industries in the 1972-1977 sample (with the strongest mean reversion tendencies in the Wood and wood products, Fabricated metal products, Furniture and Medical and optical instruments industries), 17 of the industries in the period 1972-1997 (exhibiting the strongest mean reversion tendencies in the Rubber and plastic products, Pulp and paper products and Chemical and chemical products industries), 15 industries during 1977-1986 (the strongest mean reversion observed in the Pulp and paper products, Fabricated metal products, Electrical machinery, Radio, television and communications and Wood and wood products industries) and 17 industry samples for 1986-1997 (with the strongest mean reversion found in the Non-metallic mineral products, Pulp and paper products and Textiles industries). In the Food and beverages, Wearing apparel, Wood and wood products, Publishing and printing, Chemical and chemical products, Rubber and plastic products, Non-metallic mineral products, Fabricated metal products, Other machinery and equipment and Furniture industries the reported β_1 coefficients were statistically significant and negative in all four time periods.

In contrast, the first proposition of Gibrat's Law could not be rejected in the

Table 5.4a: Plant Size and Growth, 1972-1977

Sample	Plants (N)	α	β_1	t_{β_1}	R^2	F
All	1,754	0.44*	-0.11*	[-11.08]	0.08	122.80*
Age: 0 - 1 years	190	1.12*	-0.27*	[-5.23]	0.17	27.34*
Age: 2 - 5 years	378	0.61*	-0.13*	[-5.56]	0.07	30.92*
Age: 6 - 10 years	754	0.15*	-0.06*	[-4.73]	0.03	22.37*
Age: 11 - 20 years	192	0.02	-0.02	[-0.71]	0.00	0.51
Age: +20 years	240	0.17	-0.05*	[-2.23]	0.04	4.98*
Size: Micro	192	1.13*	-0.41*	[-3.86]	0.09	14.88*
Size: Small	930	0.34*	-0.10*	[-3.60]	0.01	12.93*
Size: Medium	411	0.11	-0.04	[-0.66]	0.00	0.44
Size: Large	221	-0.17	0.01	[0.23]	0.00	0.05
Minimum efficient scale	181	-0.46*	0.04	[1.32]	0.01	1.73
Single-plant firm	700	0.73*	-0.23*	[-8.40]	0.15	70.58*
Foreign owned	190	0.58*	-0.09*	[-4.00]	0.10	16.01*
15: Food and beverages	183	0.27*	-0.07*	[-2.28]	0.05	5.22*
16: Tobacco products	..	..	..	..	..	..
17: Textiles	69	0.64*	-0.15*	[-2.84]	0.16	8.05*
18: Wearing apparel	90	0.45*	-0.11*	[-2.50]	0.07	6.27*
19: Leather products	25	0.42*	-0.13*	[-3.56]	0.27	12.65*
20: Wood and wood products	100	0.50*	-0.18*	[-3.00]	0.09	8.97*
21: Pulp and paper products	34	0.70	-0.15	[-1.13]	0.10	1.27
22: Publishing and printing	101	0.45*	-0.13*	[-2.56]	0.10	6.57*
23: Coke, petroleum and nuclear fuel	11	-0.03	0.00	[0.08]	0.00	0.01
24: Chemicals and chemical products	64	0.70*	-0.14*	[-3.67]	0.20	13.46*
25: Rubber and plastic products	77	0.55*	-0.10*	[-2.43]	0.06	5.90*
26: Non-metallic mineral products	133	0.33	-0.14*	[-2.47]	0.07	6.12*
27: Basic metals	116	0.10	-0.03	[-1.58]	0.01	2.51*
28: Fabricated metal products	236	0.73*	-0.18*	[-5.59]	0.18	31.28*
29: Other machinery and equipment	197	0.59*	-0.13*	[-4.16]	0.10	17.26*
30: Office machinery and computers	..	..	..	..	..	..
31: Electrical machinery	48	0.44*	-0.07	[-1.85]	0.07	3.43
32: Radio, T.V. and communications	29	0.31	-0.08	[-1.68]	0.10	2.83
33: Medical and optical instruments	44	0.72*	-0.16*	[-2.95]	0.19	8.73*
34: Motor vehicles and trailers	55	0.35	-0.05	[-1.20]	0.03	1.45
35: Other transport equipment	36	0.42	-0.10	[-1.09]	0.10	1.18
36: Furniture	94	0.69*	-0.18*	[-3.21]	0.10	10.30*

Notes: Model estimated using OLS with heteroscedasticity-consistent standard errors.
Robust t-statistics for the constant term, α, not reported.
.. Not reported given small sample < 10 plants.
* Statistically significant at the 5 percent level.

Table 5.4b: Plant Size and Growth, 1972-1997

Sample	Plants (N)	α	β_1	t_{β_1}	R^2	F
All	693	1.69*	-0.41*	[-21.58]	0.42	465.75*
Age: 0 - 1 years	67	2.21*	-0.44*	[-7.18]	0.39	51.55*
Age: 2 - 5 years	151	2.11*	-0.45*	[-10.27]	0.36	105.54*
Age: 6 - 10 years	259	1.10*	-0.33*	[-11.26]	0.36	126.77*
Age: 11 - 20 years	81	1.54*	-0.37*	[-7.32]	0.39	53.61*
Age: +20 years	135	1.26*	-0.35*	[-7.63]	0.35	58.22*
Size: Micro	58	2.81*	-0.75*	[-5.26]	0.28	27.71*
Size: Small	308	1.38*	-0.38*	[-4.85]	0.07	23.54*
Size: Medium	200	1.13	-0.29*	[-2.03]	0.02	4.13*
Size: Large	127	1.23*	-0.33*	[-4.30]	0.12	18.49*
Minimum efficient scale	92	1.43*	-0.37*	[-5.27]	0.23	27.83*
Single-plant firm	266	2.16*	-0.63*	[-15.97]	0.50	254.99*
Foreign owned	126	2.37*	-0.50*	[-9.31]	0.48	86.67*
15: Food and beverages	76	1.65*	-0.38*	[-6.36]	0.37	40.50*
16: Tobacco products	..	..	..	..	..	..
17: Textiles	22	1.62*	-0.40*	[-3.52]	0.33	12.41*
18: Wearing apparel	24	1.33*	-0.32*	[-3.85]	0.23	14.84*
19: Leather products	..	..	..	..	..	..
20: Wood and wood products	24	1.28*	-0.39*	[-2.71]	0.24	7.37*
21: Pulp and paper products	22	3.06*	-0.65*	[-4.59]	0.68	21.03*
22: Publishing and printing	43	1.18*	-0.29*	[-2.52]	0.18	6.37*
23: Coke, petroleum and nuclear fuel	..	..	..	..	..	..
24: Chemicals and chemical products	35	2.49*	-0.57*	[-6.06]	0.59	36.66*
25: Rubber and plastic products	38	3.25*	-0.71*	[-7.85]	0.63	61.63*
26: Non-metallic mineral products	43	1.19*	-0.35*	[-4.62]	0.31	21.32*
27: Basic metals	52	1.35*	-0.34*	[-6.79]	0.48	46.10*
28: Fabricated metal products	88	1.90*	-0.50*	[-8.78]	0.40	77.07*
29: Other machinery and equipment	73	1.76*	-0.51*	[-10.85]	0.60	117.80*
30: Office machinery and computers	..	..	..	..	..	..
31: Electrical machinery	24	2.35*	-0.46*	[-4.93]	0.49	24.32*
32: Radio, T.V. and communications	15	2.02*	-0.47*	[-4.46]	0.46	19.87*
33: Medical and optical instruments	16	2.01*	-0.41*	[-3.08]	0.33	9.49*
34: Motor vehicles and trailers	31	1.24*	-0.30*	[-3.86]	0.29	14.93*
35: Other transport equipment	16	1.02	-0.24	[-1.83]	0.29	3.35
36: Furniture	33	2.05*	-0.50*	[-4.01]	0.47	16.05*

Notes: Model estimated using OLS with heteroscedasticity-consistent standard errors.
Robust t-statistics for the constant term, α, not reported.
.. Not reported given small sample < 10 plants.
* Statistically significant at the 5 percent level.

Table 5.4c: Plant Size and Growth, 1977-1986

Sample	Plants (N)	α	β_1	t_{β_1}	R^2	F
All	1,444	0.82*	-0.22*	[-18.61]	0.21	346.47*
Age: 0 - 1 years	97	1.30*	-0.23*	[-4.13]	0.13	17.04*
Age: 2 - 5 years	309	0.86*	-0.18*	[-5.16]	0.09	26.64*
Age: 6 - 10 years	284	0.61*	-0.15*	[-4.95]	0.08	24.49*
Age: 11 - 20 years	512	0.45*	-0.18*	[-9.71]	0.17	94.27*
Age: +20 years	242	0.55*	-0.18*	[-6.71]	0.17	45.06*
Size: Micro	206	1.28*	-0.39*	[-3.96]	0.08	15.71*
Size: Small	735	0.38*	-0.11*	[-2.49]	0.01	6.19*
Size: Medium	333	0.20	-0.09	[-1.05]	0.00	1.11
Size: Large	170	0.26	-0.13*	[-2.01]	0.02	4.05*
Minimum efficient scale	156	0.65*	-0.17*	[-4.62]	0.09	21.37*
Single-plant firm	590	1.31*	-0.41*	[-13.74]	0.28	188.73*
Foreign owned	203	1.23*	-0.27*	[-6.99]	0.25	48.88*
15: Food and beverages	143	0.93*	-0.21*	[-5.14]	0.19	26.45*
16: Tobacco products	..	..	..	..	..	..
17: Textiles	43	0.44	-0.16	[-1.50]	0.07	2.24
18: Wearing apparel	60	0.83	-0.18*	[-2.03]	0.10	4.12*
19: Leather products	14	-0.22	0.01	[0.11]	0.00	0.01
20: Wood and wood products	67	0.91*	-0.30*	[-4.53]	0.26	20.48*
21: Pulp and paper products	37	1.70*	-0.36*	[-6.66]	0.53	44.34*
22: Publishing and printing	93	0.95*	-0.27*	[-3.87]	0.19	14.99*
23: Coke, petroleum and nuclear fuel	10	0.67	-0.16	[-1.98]	0.26	3.92
24: Chemicals and chemical products	66	0.97*	-0.29*	[-5.45]	0.29	29.72*
25: Rubber and plastic products	82	1.12*	-0.27*	[-4.90]	0.21	24.03*
26: Non-metallic mineral products	123	0.44*	-0.16*	[-4.91]	0.13	24.09*
27: Basic metals	81	0.72*	-0.22*	[-7.18]	0.31	51.57*
28: Fabricated metal products	181	1.20*	-0.35*	[-8.17]	0.33	66.69*
29: Other machinery and equipment	150	0.73*	-0.24*	[-6.85]	0.23	46.89*
30: Office machinery and computers	..	..	..	..	..	..
31: Electrical machinery	47	1.29*	-0.34*	[-3.48]	0.26	12.11*
32: Radio, T.V. and communications	33	1.44*	-0.33*	[-4.82]	0.35	23.22*
33: Medical and optical instruments	43	0.85*	-0.17	[-1.91]	0.10	3.65
34: Motor vehicles and trailers	50	0.92*	-0.23*	[-5.13]	0.37	26.29*
35: Other transport equipment	26	-0.29	-0.04	[-0.47]	0.01	0.22
36: Furniture	82	0.75*	-0.18*	[-2.51]	0.09	6.32*

Notes: Model estimated using OLS with heteroscedasticity-consistent standard errors.
Robust t-statistics for the constant term, α, not reported.
.. Not reported given small sample < 10 plants.
* Statistically significant at the 5 percent level.

Table 5.4d: Plant Size and Growth, 1986-1997

Sample	Plants (N)	α	β_1	t_{β_1}	R^2	F
All	1,417	1.27*	-0.27*	[-19.71]	0.26	388.43*
Age: 0 - 1 years	131	1.94*	-0.38*	[-5.41]	0.21	29.32*
Age: 2 - 5 years	230	1.49*	-0.28*	[-7.48]	0.20	56.01*
Age: 6 - 10 years	231	1.03*	-0.19*	[-5.21]	0.13	27.17*
Age: 11 - 20 years	540	0.77*	-0.18*	[-9.42]	0.14	88.67*
Age: +20 years	285	0.92*	-0.22*	[-6.13]	0.23	37.62*
Size: Micro	134	3.31*	-1.27*	[-5.72]	0.36	32.73*
Size: Small	816	1.21*	-0.28*	[-7.68]	0.06	59.00*
Size: Medium	321	1.01*	-0.21*	[-2.41]	0.02	5.79*
Size: Large	146	0.35	-0.11	[-1.57]	0.01	2.48
Minimum efficient scale	187	0.46	-0.12*	[-2.78]	0.04	7.71*
Single-plant firm	684	1.65*	-0.43*	[-14.45]	0.30	208.69*
Foreign owned	253	1.84*	-0.34*	[-8.01]	0.35	64.24*
15: Food and beverages	125	1.56*	-0.33*	[-7.22]	0.31	52.15*
16: Tobacco products	..	..	..	..	..	..
17: Textiles	48	1.66*	-0.37*	[-3.99]	0.30	15.94*
18: Wearing apparel	46	1.19*	-0.25*	[-3.67]	0.20	13.46*
19: Leather products	11	-0.02	-0.09	[-0.46]	0.03	0.21
20: Wood and wood products	49	1.05*	-0.27*	[-2.83]	0.15	7.99*
21: Pulp and paper products	54	2.09*	-0.43*	[-5.71]	0.43	32.63*
22: Publishing and printing	72	0.96*	-0.21*	[-2.95]	0.14	8.71*
23: Coke, petroleum and nuclear fuel	..	..	..	..	..	..
24: Chemicals and chemical products	88	1.57*	-0.31*	[-6.19]	0.35	38.29*
25: Rubber and plastic products	104	1.59*	-0.35*	[-8.77]	0.37	76.92*
26: Non-metallic mineral products	72	1.74*	-0.44*	[-6.42]	0.47	41.19*
27: Basic metals	83	1.05*	-0.21*	[-6.72]	0.31	45.10*
28: Fabricated metal products	190	1.13*	-0.26*	[-6.68]	0.18	44.58*
29: Other machinery and equipment	141	1.38*	-0.35*	[-7.74]	0.38	59.93*
30: Office machinery and computers	15	0.77	-0.07	[-0.67]	0.01	0.45
31: Electrical machinery	57	1.58*	-0.31*	[-3.02]	0.27	9.12*
32: Radio, T.V. and communications	46	1.22*	-0.22*	[-3.50]	0.20	12.22*
33: Medical and optical instruments	46	1.34*	-0.28*	[-2.99]	0.21	8.97*
34: Motor vehicles and trailers	45	0.61*	-0.10	[-1.96]	0.07	3.85
35: Other transport equipment	21	1.12*	-0.21*	[-2.86]	0.28	8.20*
36: Furniture	86	1.42*	-0.31*	[-6.00]	0.30	36.03*

Notes: Model estimated using OLS with heteroscedasticity-consistent standard errors.
Robust t-statistics for the constant term, α, not reported.
.. Not reported given small sample < 10 plants.
* Statistically significant at the 5 percent level.

Pulp and paper products, Coke, petroleum and nuclear fuel, Basic metals, Electrical machinery, Radio, television and communications, Motor vehicles and trailers and Other transport equipment industries in the period 1972-1977, in the Other transport equipment industry for the 1972-1997 period, for the Textiles, Leather products, Coke, petroleum and nuclear fuel, Medical and optical instruments and Other transport equipment industries for 1977-1986, and in the Leather products, Office machinery and computers and Motor vehicles and trailers industries in 1986-1997. In combination, such evidence implies different convergence speeds across industries, supporting the use of these disaggregated samples (although further heterogeneity may exist below this level).

For the samples of single-plant firms in each of the four time periods the β_1 coefficients were statistically significant and negative, rejecting the first proposition of Gibrat's Law. Interestingly, for this group of firms the coefficients were larger (in absolute terms) than those for the complete sample, suggesting a stronger process of mean reversion (on average) for such single-plants than for the general population of plants. Foreign owned plants were also observed to have statistically significant negative β_1 coefficients in each of the time periods, but in each case of lower magnitude than for the sample of single-plants, but only for the 1972-1977 period compared to the complete samples. For the MES samples, consistent with the third version of Gibrat's Law, the β_1 coefficient was not statistically significantly different from zero in the 1972-1977 period, although in each of the other periods 1977-1986, 1986-1997 and 1972-1997 a process of statistically significant mean reversion was detected.

As a further check on the specification of these models, Ramsey's RESET test[87], with a null hypothesis of no omitted variables was applied to each regression

---

[87] Any such omissions of relevant variables may give rise to biased and inconsistent parameter estimates if the included and omitted explanatory variables are correlated, with incorrect estimated parameter and error term variances (Gudjarati, 1988, p.403) and consequently a possibility of drawing invalid inferences (Kennedy, 1998, p.95). The inclusion of irrelevant variables would give rise to less precise estimates and potentially a reduced likelihood of finding statistically significant variables.

specification (the results of this test for all OLS estimated growth models, except for the serial correlation models, over these time periods is presented in Table A5.9). At the aggregate level for all periods the null hypothesis could be rejected at the 5 percent level of significance (even with the inclusion of industry dummy variables)[88]. However, at more disaggregated levels the null hypothesis was only rejected for the Medium plant size-class in 1977-1986 and the Micro size-class in 1986-1997, for each of the single-plant firm samples, and for the foreign owned sample in the period 1986-1997. At the industry-level there were rejections of the null hypothesis in 8 samples in the period 1972-1977, 7 in the period 1977-1986, 8 for 1986-1997 and 4 for 1972-1997, a total of 27 industry-level samples out of the 78 samples examined. In total there were only 38 rejections of the null hypothesis of no omitted variables, out of 110 samples.

However, given the possibility of omitted variables in this model for some samples, in subsequent sections additional plant specific variables, plant age and quadratic size and age terms are further considered in the model specifications. The addition of persistence terms of different time-spans is also considered below, but did not, at least at the aggregate level, address the omitted variable problem for these regressions.

To further consider the potential bias arising from persistence in plant growth, and recognising the possibility that regressing plant growth in consecutive five-year time periods may not adequately account for shorter-run persistence (the initial approach considered earlier in this chapter) the logarithmic regression model with the addition of a persistence term (also known as the serial correlation model in the literature) was estimated at the aggregate-level for various lags in plant growth spanning between one-year and five-year periods following the model specification presented in Goddard et al.

---

[88] For the standard model specification estimated across different time-spans, the test also suggested the possibility of omitted variables, at the 5 percent level, in the models both with and without industry-level dummy variables.

(2002):

$$\Delta \ln X_{it} = \alpha + \beta_1 \ln X_{it-1} + \delta(\ln X_{it-1} - \ln X_{it-2}) + \nu_{it} \qquad (5.3)$$

with the variables as defined previously, and δ being the persistence coefficient (with the difference between the $\ln X_{it-1}$ and the $\ln X_{it-2}$ terms being here referred to as the reported time-span of the persistence term). The serial correlation model results, with models again estimated using OLS and White's correction, presented in Table 5.5 (diagnostic tests[89] for the serial correlation models are presented in Table A5.14) indicate that in all of the twenty aggregate samples the hypothesis that $\beta_1 = 0$ could be rejected at the 5 percent level of significance, with a process of mean reversion observed. The first proposition of Gibrat's Law could therefore again be rejected. The analysis also further supported the earlier findings with the persistence coefficient term, δ, found to be statistically significant at the 5 percent level in only 5 of the 20 samples.

The results also point to the inclusion of increasingly lagged persistence terms leading to slightly less negative β_1 coefficients than observed in the model without persistence terms. The F-tests of the overall significance of the multiple regressions suggest that the null hypothesis of the coefficients being zero simultaneously could be rejected at the 5 percent level of significance in all of the samples.

Interestingly, neither the one-year nor two-year persistence terms were found to be statistically significant in any of the four time periods, with three-year, four-year and

---

[89] In only the 1972-1997 time period with a persistence term with two-year and three-year time-spans could neither the White nor the Breusch-Pagan tests be rejected at the 5 percent level of significance. Although the D'Agostino et al. and Shapiro-Francia tests rejected normality of the residuals in all but the 1972-1997 period with a one-year persistence term, with the Shapiro-Francia test additionally not providing rejections for the same time period with two-year and four-year time-span persistence terms, the deviations from the skewness and kurtosis benchmarks for the normal distribution were modest (the maximum absolute skewness value being -0.73) and with all kurtosis values being in single-digits. In all models the residuals were suitably described by a zero mean.

Table 5.5: Plant Size and Growth: Serial Correlation Model

Sample	Plants (N)	α	β_1	t_{β_1}	δ	t_δ	R^2	F
One-Year:								
1972-1977	1,634	0.33*	-0.08*	[-9.15]	-0.12	[-1.87]	0.06	43.45*
1972-1997	649	1.61*	-0.40*	[-19.67]	0.12	[1.13]	0.40	193.96*
1977-1986	1,397	0.75*	-0.21*	[-17.23]	-0.17	[-1.70]	0.19	151.20*
1986-1997	1,345	1.17*	-0.25*	[-17.06]	0.06	[0.32]	0.24	202.77*
Two-Year:								
1972-1977	1,535	0.27*	-0.07*	[-7.99]	-0.03	[-0.57]	0.04	32.12*
1972-1997	617	1.51*	-0.39*	[-18.83]	0.16	[1.88]	0.39	178.97*
1977-1986	1,341	0.71*	-0.21*	[-16.56]	0.04	[0.73]	0.18	138.37*
1986-1997	1,265	1.10*	-0.24*	[-16.19]	0.01	[0.11]	0.22	177.25*
Three-Year:								
1972-1977	1,419	0.22*	-0.06*	[-6.97]	-0.02	[-0.43]	0.04	24.49*
1972-1997	569	1.44*	-0.38*	[-17.53]	0.17*	[2.40]	0.39	155.65*
1977-1986	1,240	0.69*	-0.20*	[-15.51]	0.11*	[2.10]	0.18	127.48*
1986-1997	1,203	1.05*	-0.23*	[-15.70]	0.04	[0.44]	0.21	161.58*
Four-Year:								
1972-1977	1,309	0.19*	-0.06*	[-6.26]	-0.03	[-0.73]	0.04	19.62*
1972-1997	525	1.35*	-0.37*	[-16.32]	0.26*	[2.98]	0.38	136.97*
1977-1986	1,160	0.65*	-0.20*	[-14.76]	0.09	[1.75]	0.17	109.65*
1986-1997	1,161	1.02*	-0.22*	[-15.02]	0.04	[0.56]	0.21	149.93*
Five-Year:								
1972-1977	1,210	0.15*	-0.05*	[-5.50]	-0.07	[-1.73]	0.04	16.18*
1972-1997	482	1.30*	-0.36*	[-15.61]	0.29*	[3.12]	0.40	125.88*
1977-1986	1,092	0.60*	-0.19*	[-13.63]	0.11*	[2.37]	0.17	94.61*
1986-1997	1,093	0.96*	-0.21*	[-13.59]	0.01	[0.09]	0.20	129.20*

Notes: Model estimated using OLS with heteroscedasticity-consistent standard errors.
Robust t-statistics for the constant term, α, not reported.
* Statistically significant at the 5 percent level.

five-year persistence terms found to be significant in the longest time period 1972-1997 and also for the three-year and five-year terms in the 1977-1986 time period. In each case of statistically significant δ coefficients, the coefficients were positive so that positive growth in one period would be expected on average to continue in the next period. The persistence coefficients exhibited parameter values of generally consistent sign for each of the time periods, with only the δ coefficient of the one-year persistence term in 1977-1986 showing differences compared to the other lagged persistence terms for that time period.

To examine the possibility that these empirical results might be an artifact of the selected time-span of the cross-sections between the initial and final periods, regression estimates of the logarithmic size-growth model (with robust standard errors) using only the initial plant size as an explanatory variable were generated for each of the initial samples of plants relating to 1972, 1977 and 1986 with the time-span increased incrementally by one-year up to the end of each economic cycle. The results are presented in Table A5.5. A further specification with 2-digit level industry dummy variables was also tested, with the dummy variable for the Fabricated metal products industry omitted to avoid perfect collinearity between the industry-level dummy variables and the constant term (Blonigen and Tomlin, 1999, annex).

For these surviving plants the aggregate regressions provided a statistically significant (at the 5 percent level) rejection of Gibrat's Law for both the standard model and the regression model with industry dummy variables across all time-spans and for each of the samples with the (negative) regression β_1 coefficient decreasing steadily in magnitude with increased time-span. More specifically, in the case of the 1972-1977(1997) sample the estimated β_1 coefficient changed from -0.03 (-0.04 for the model with industry dummy variables) for a one-year growth period to -0.11 (-0.12) with growth examined over a five–year time-span, to -0.41 (-0.43) when considered over the full available time-span to 1997. For the 1977-1986 sample the corresponding coefficients changed from -0.04 (-0.04) over one-year to -0.22 (-0.24) for growth measured over the full time period, with for the final period of 1986-1997 the one-year growth coefficient being -0.06 (-0.06) declining to -0.27 (-0.29) for the full time-span. The industry-level dummy variables were predominantly found not to be statistically significant at the 5 percent level (results not reported here).

5.5. Plant Size, Age and Growth

As discussed in the literature review, there are a number of reasons why the variance of firm growth may be expected to decline with increased firm size (including diversification and negatively correlated product performance), giving rise to heteroscedasticity in the regression error term and a violation of the assumptions of the CLR model. However, the heteroscedasticity observed in the preceding section across a wide range of samples may also be a consequence of the omission of other relevant variables (or indeed an incorrect functional form, an aspect considered in Section 5.6).

A number of theoretical models have identified the role of firm (or plant) age in influencing the growth and variance of firm growth. The selection and learning models of Jovanovic (1982), Ericson and Pakes (1995) and Audretsch (1995) also imply an information acquisition process that takes time, and is thus likely to be related to plant (or firm) age. Plant performance may therefore be conceived as likely to exhibit less variance with age as well as size, and therefore with an age related component contributing to the error term structure (see Gudjarati, 1988, p.318, regarding the anticipated declining variance in such error-learning models). Such variance patterns seem unlikely to exhibit a precise relationship captured by the variance reflected in the size variable, so that the inclusion of such a variable may be important in the model specification. Of course, the model of Gibrat (1931) can also be interpreted as implying an age related process, with the assumed random sampling of firm growth opportunities giving rise to a lognormal limiting distribution through a cumulative process over a number of time periods.

Additionally, the inclusion of an age variable may be justified on the argument that age may act as a proxy for other factors which may influence growth, such as financial constraints (Elston, 2002), with such constraints moderating with the development of a track-record of business performance, the acquisition of experience over time alleviating information asymmetries between firms and financiers, and the accumulation of collateral; for a discussion, see Bank of England, 2001). Although it may be argued that such learning processes and financial constraints issues may be most relevant at the firm-level, the effects of age are examined here at the plant-level given the nature of the available data.

Given the above described expected influence of plant age, the inclusion of such a plant age variable in the logarithmic size-growth model may therefore help to alleviate the previously observed heteroscedasticity (Dunne and Hughes, 1994, p.130). Simultaneously, the inclusion of an additional plant age variable might also assist in overcoming other issues relating to the omitted variable problem identified in a number of samples using the previous model specification.

The inclusion of a plant age variable, where $\ln A_{it}$ is the logarithm of plant age[90], therefore provides the model formulation:

$$\Delta \ln X_{it} = \alpha + \beta_1 \ln X_{it-1} + \gamma_1 \ln A_{it} + \nu_{it} \tag{5.4}$$

---

[90] For plants of age < 1 in the initial years, an age value of 1 was applied (given that the logarithm of zero is undefined). Experimentation using age values of 0.5 - 1.0 did not significantly alter the reported coefficients or the t-statistics of the aggregate regressions.

Although the existing empirical literature (as found by, for example, Evans, 1987a, 1987b) suggests a potential negative relationship between firm or plant age and growth, under Gibrat's Law there is no strong relationship postulated with age so that the hypothesis H_0: $\gamma_1 = 0$, H_1: $\gamma_1 \neq 0$ is tested. In addition to the t-tests for the significance of the partial regression coefficients, F-test results are also reported as a test of the null hypothesis that the coefficients of both variables, plant size and age, are jointly zero (Gudjarati, 1988), that H_0: $\beta_1 = \gamma_1 = 0$, with the tests of statistical significance again conducted at the 5 percent level.

Although the inclusion of the age variable might be expected to assist in the alleviation of heteroscedasticity if the variance of plant growth is also related to plant age, for the equations estimated on the aggregate samples prior to White's correction for heteroscedasticity, in all except the longer 1972-1997 time period, both tests rejected homoscedasticity at the 5 percent level of significance. In the disaggregated samples mixed results were again observed, with rejections of homoscedasticity detected for a number of the industry, size-class and age-class samples for each time period. However, some improvement in the model specification arising from the inclusion of the age variable were observed in the form of slightly fewer rejections of Ramsey's RESET test in the case of the 1972-1997, 1977-1986 and 1986-1997 time periods[91]. In total, only 36 out of the 110 samples exhibited a rejection of the test for no omitted variables (38 in the model without age). The results from the analysis of the regression residuals were broadly consistent with the previous model estimates, with using the previous discussed procedure, 70 samples more closely approximating to the normal distribution, and 27 to the Laplace distribution.

---

[91] In the 1972-1997 period the null hypothesis of no omitted variables could be rejected in only 4 samples in the model specification with age (6 in the previous model specification), 9 in the model with age for 1977-1986 (10 in the model without age), and 10 samples in the 1986-1997 sample (12 in the model without age). For the shortest period, 1972-1977 the RESET test could be rejected in 13 samples, compared to 10 samples in the model specification without an age variable.

The multiple regression results for equation (5.4) estimated using OLS with White's correction, given the above reported results, are presented in Table 5.6a-d. The estimates for the aggregate samples of surviving plants provided consistent evidence of statistically significant negative coefficients for both plant size, β_1, and age, γ_1, variables. The F-tests[92] of overall significance also rejected the null hypothesis that the coefficients of these variables were jointly zero in each period. The β_1 coefficients were in each case slightly less negative than estimated in the previous specification, equation (5.2), (with coefficients of -0.08, -0.18, -0.21 and -0.38, respectively), a finding which generally, but not universally, also held across disaggregated samples. The estimated γ_1 coefficients ranged between -0.11 and -0.20 across the time periods (with the coefficient being highly similar for the 1977-1986, 1986-1997 and the 1972-1997 periods) supporting the previous research evidence suggesting that at the aggregate-level plant growth is negatively related to age.

However, some variation was observed across the size-class estimates. For the 1972-1977 sample, both the Micro and Small size-classes exhibited statistically significant negative relationships for plant size and age, which could be rejected as being jointly zero, but for the Medium and Large size-classes, neither β_1 nor γ_1 could be rejected as statistically different from zero, either individually or jointly. The 1972-1997 estimates showed a less clear picture with the β_1 coefficient statistically significant across all size-classes, except for Medium-sized plants, and with the γ_1 coefficient being statistically significant only for the Small and Medium size-classes. The F-test rejected the joint hypothesis of the β_1 and γ_1 coefficients being zero in all four samples. In the case of the

---

[92] The F-test also tests the hypothesis $\beta = 0$ in the single explanatory variable regression (Gudjarati, 1988, p.118). It should be noted that although an adjusted-R^2 may be preferred to the R^2 measure of regression fit, the R^2 potentially providing a misleading measure of model fit (Gudjarati, 1988, p.183), in some cases the adjusted-R^2 was not reported by STATA. Where the adjusted-R^2 was provided the differences were modest. However, the use of the R^2 here (consistent with Dunne and Hughes, 1994) means that comparisons between the fit of alternative regression specifications with different numbers of explanatory variables cannot be directly made (Gudjarati, 1988, pp.182-183).

Table 5.6a: Plant Size, Age and Growth, 1972-1977

Sample	Plants (N)	α	β_1	t_{β_1}	γ	t_γ	R^2	F
All	1,754	0.54*	-0.08*	[-8.57]	-0.11*	[-6.11]	0.11	78.54*
Size: Micro	192	1.23*	-0.31*	[-2.84]	-0.21*	[-3.71]	0.17	15.54*
Size: Small	930	0.49*	-0.09*	[-3.41]	-0.11*	[-4.81]	0.04	19.81*
Size: Medium	411	0.12	-0.04	[-0.61]	-0.02	[-0.45]	0.00	0.26
Size: Large	221	-0.12	0.01	[0.27]	-0.02	[-0.53]	0.00	0.25
Minimum efficient scale	181	-0.47	0.04	[1.32]	0.01	[0.21]	0.01	0.87
Single-plant firm	700	0.84*	-0.19*	[-6.71]	-0.13*	[-4.76]	0.18	47.97*
Foreign owned	190	0.69*	-0.07*	[-3.01]	-0.12*	[-3.29]	0.15	11.13*
15: Food and beverages	183	0.31*	-0.07*	[-2.25]	-0.02	[0.04]	0.05	2.70
16: Tobacco products	..	..	..	..	..	..	..	..
17: Textiles	69	0.64*	-0.15*	[-2.78]	0.00	[0.04]	0.16	4.01*
18: Wearing apparel	90	0.52*	-0.07	[-1.98]	-0.11*	[-0.11]	0.12	4.95*
19: Leather products	25	0.46*	-0.11*	[-2.18]	-0.07	[-0.85]	0.30	6.34*
20: Wood and wood products	100	0.71*	-0.16*	[-2.74]	-0.17*	[-2.01]	0.12	5.93*
21: Pulp and paper products	34	0.70	-0.15	[-1.31]	-0.00	[-0.00]	0.10	0.92
22: Publishing and printing	101	0.41	-0.14*	[-2.59]	0.03	[0.35]	0.10	3.37
23: Coke, petroleum and nuclear fuel	11	-0.21	0.00	[0.03]	0.10	[0.36]	0.01	0.08
24: Chemicals and chemical products	64	0.72*	-0.13*	[-3.28]	-0.03	[-0.67]	0.20	6.95*
25: Rubber and plastic products	77	0.63*	-0.04	[-0.95]	-0.20*	[-2.48]	0.12	5.16*
26: Non-metallic mineral products	133	0.54*	-0.11*	[-2.05]	-0.16*	[-2.44]	0.11	7.16*
27: Basic metals	116	0.17	-0.02	[-1.22]	-0.07	[-1.25]	0.03	1.57
28: Fabricated metal products	236	0.84*	-0.13*	[-3.84]	-0.18*	[-4.50]	0.25	23.51*
29: Other machinery and equipment	197	0.64*	-0.11*	[-3.40]	-0.07	[-1.29]	0.11	9.23*
30: Office machinery and computers	..	..	..	..	..	..	..	..
31: Electrical machinery	48	0.60*	-0.04	[-1.34]	-0.12	[-1.65]	0.14	2.55
32: Radio, T.V. and communications	29	0.39	-0.02	[-0.28]	-0.20	[-1.51]	0.17	2.70
33: Medical and optical instruments	44	0.76*	-0.13*	[-2.07]	-0.09	[-1.27]	0.21	5.04*
34: Motor vehicles and trailers	55	0.63*	-0.01	[-0.12]	-0.26*	[-2.79]	0.22	7.65*
35: Other transport equipment	36	0.22	-0.11	[-1.19]	0.13	[1.02]	0.14	0.71
36: Furniture	94	0.71*	-0.17*	[-2.55]	-0.03	[-0.24]	0.11	5.27*

Notes: Model estimated using OLS with heteroscedasticity-consistent standard errors.
Robust t-statistics for the constant term, α, not reported.
.. Not reported given small sample < 10 plants.
* Statistically significant at the 5 percent level.

Table 5.6b: Plant Size, Age and Growth, 1972-1997

Sample	Plants (N)	α	β_1	t_{β_1}	γ	t_γ	R^2	F
All	693	1.91*	-0.38*	[-19.59]	-0.19*	[-6.30]	0.45	255.99*
Size: Micro	58	2.86*	-0.75*	[-5.40]	-0.04	[-0.46]	0.28	16.07*
Size: Small	308	1.68*	-0.36*	[-4.81]	-0.21*	[-4.40]	0.13	19.24*
Size: Medium	200	1.40*	-0.27	[-1.93]	-0.16*	[-3.00]	0.05	6.63*
Size: Large	127	1.57*	-0.33*	[-4.35]	-0.16	[-1.94]	0.15	10.76*
Minimum efficient scale	92	1.74*	-0.34*	[-5.13]	-0.20	[-1.96]	0.26	15.83*
Single-plant firm	266	2.25*	-0.59*	[-13.41]	-0.13*	[-2.56]	0.51	133.77*
Foreign owned	126	2.48*	-0.48*	[-7.44]	-0.21*	[-3.51]	0.52	53.41*
15: Food and beverages	76	1.77*	-0.38*	[-6.36]	-0.06	[-0.64]	0.38	20.38*
16: Tobacco products	..	..	..	..	..	..	..	..
17: Textiles	22	1.87*	-0.36*	[-3.36]	-0.22	[-1.21]	0.38	11.73*
18: Wearing apparel	24	1.47*	-0.16	[-0.98]	-0.39	[-1.78]	0.36	9.32*
19: Leather products	..	..	..	..	..	..	..	..
20: Wood and wood products	24	1.54*	-0.36*	[-2.60]	-0.17	[-0.93]	0.27	3.95*
21: Pulp and paper products	22	3.00*	-0.53*	[-3.61]	-0.22	[-1.26]	0.70	10.93*
22: Publishing and printing	43	1.84*	-0.30*	[-2.53]	-0.38	[-1.74]	0.28	6.49*
23: Coke, petroleum and nuclear fuel	..	..	..	..	..	..	..	..
24: Chemicals and chemical products	35	2.50*	-0.56*	[-5.29]	-0.03	[-0.19]	0.59	18.03*
25: Rubber and plastic products	38	3.54*	-0.64*	[-6.95]	-0.32*	[-2.16]	0.69	60.03*
26: Non-metallic mineral products	43	1.49*	-0.34*	[-4.82]	-0.16	[-1.20]	0.34	13.01*
27: Basic metals	52	1.63*	-0.31*	[-6.70]	-0.22*	[-2.46]	0.54	27.00*
28: Fabricated metal products	88	2.13*	-0.44*	[-7.37]	-0.24*	[-2.76]	0.45	42.60*
29: Other machinery and equipment	73	1.82*	-0.45*	[-7.14]	-0.17	[-1.77]	0.62	73.57*
30: Office machinery and computers	..	..	..	..	..	..	..	..
31: Electrical machinery	24	2.62*	-0.44*	[-5.27]	-0.17	[-1.08]	0.51	14.55*
32: Radio, T.V. and communications	15	2.02*	-0.39*	[-2.47]	-0.22	[-0.64]	0.48	8.65*
33: Medical and optical instruments	16	2.12*	-0.39*	[-2.73]	-0.13	[-0.88]	0.35	4.90*
34: Motor vehicles and trailers	31	1.64*	-0.25*	[-3.11]	-0.30*	[-2.55]	0.37	10.56*
35: Other transport equipment	16	0.91	-0.24	[-1.74]	0.06	[0.23]	0.30	1.84
36: Furniture	33	2.07*	-0.47*	[-3.73]	-0.07	[-0.47]	0.48	8.09*

Notes: Model estimated using OLS with heteroscedasticity-consistent standard errors.
Robust t-statistics for the constant term, α, not reported.
.. Not reported given small sample < 10 plants.
* Statistically significant at the 5 percent level.

Table 5.6c: Plant Size, Age and Growth, 1977-1986

Sample	Plants (N)	α	β_1	t_{β_1}	γ	t_γ	R^2	F
All	1,444	1.05*	-0.18*	[-14.14]	-0.19*	[-8.45]	0.25	203.10*
Size: Micro	206	1.34*	-0.35*	[-3.44]	-0.08	[-1.61]	0.09	9.39*
Size: Small	735	0.67*	-0.07	[-1.74]	-0.20*	[-6.66]	0.07	25.72*
Size: Medium	333	0.47	-0.06	[-0.70]	-0.18*	[-4.20]	0.05	10.02*
Size: Large	170	0.64	-0.12	[-1.67]	-0.17	[-1.60]	0.05	6.83*
Minimum efficient scale	156	0.91*	-0.16*	[-4.16]	-0.13	[-1.95]	0.10	11.91*
Single-plant firm	590	1.50*	-0.35*	[-10.77]	-0.19*	[-5.73]	0.32	123.11*
Foreign owned	203	1.54*	-0.16*	[-4.55]	-0.36*	[-5.34]	0.36	33.39*
15: Food and beverages	143	1.26*	-0.18*	[-4.43]	-0.19*	[-2.60]	0.24	16.49*
16: Tobacco products	..	..	..	..	..	..	..	..
17: Textiles	43	0.74	-0.11	[-0.97]	-0.22	[-1.56]	0.12	2.57
18: Wearing apparel	60	1.00*	-0.12	[-1.37]	-0.18	[-1.83]	0.18	2.87
19: Leather products	14	-0.30	0.00	[0.02]	0.05	[0.17]	0.00	0.02
20: Wood and wood products	67	0.96*	-0.29*	[-3.93]	-0.04	[-0.34]	0.26	11.21*
21: Pulp and paper products	37	1.77*	-0.24*	[-3.86]	-0.26*	[-2.14]	0.60	26.45*
22: Publishing and printing	93	1.42*	-0.15*	[-2.65]	-0.40*	[-3.65]	0.33	12.85*
23: Coke, petroleum and nuclear fuel	10	0.89	-0.14	[-1.71]	-0.13	[-0.38]	0.27	1.77
24: Chemicals and chemical products	66	1.11*	-0.21*	[-5.10]	-0.22*	[-2.33]	0.32	17.29*
25: Rubber and plastic products	82	1.12*	-0.27*	[-3.79]	0.00	[0.05]	0.21	12.18*
26: Non-metallic mineral products	123	0.37*	-0.18*	[-4.14]	0.06	[0.70]	0.13	11.84*
27: Basic metals	81	1.15*	-0.18*	[-6.84]	-0.27*	[-3.34]	0.40	29.50*
28: Fabricated metal products	181	1.36*	-0.31*	[-6.74]	-0.15*	[-2.89]	0.36	41.62*
29: Other machinery and equipment	150	0.97*	-0.16*	[-4.15]	-0.25*	[-4.26]	0.30	36.92*
30: Office machinery and computers	..	..	..	..	..	..	..	..
31: Electrical machinery	47	1.51*	-0.26*	[-3.34]	-0.24	[-1.46]	0.30	6.63*
32: Radio, T.V. and communications	33	1.45*	-0.26*	[-2.30]	-0.16	[-0.82]	0.37	10.89*
33: Medical and optical instruments	43	1.20*	-0.16	[-1.71]	-0.21	[-1.91]	0.16	4.64*
34: Motor vehicles and trailers	50	1.27*	-0.20*	[-3.20]	-0.19	[-1.48]	0.41	22.24*
35: Other transport equipment	26	-0.25	-0.04	[-0.40]	-0.02	[-0.13]	0.01	0.16
36: Furniture	82	1.04*	-0.11	[-1.55]	-0.28*	[-3.52]	0.19	9.82*

Notes: Model estimated using OLS with heteroscedasticity-consistent standard errors.
Robust t-statistics for the constant term, α, not reported.
.. Not reported given small sample < 10 plants.
* Statistically significant at the 5 percent level.

Table 5.6d: Plant Size, Age and Growth, 1986-1997

Sample	Plants (N)	α	β_1	t_{β_1}	γ	t_γ	R^2	F
All	1,417	1.49*	-0.21*	[-13.89]	-0.20*	[-10.53]	0.33	291.53*
Size: Micro	134	3.51*	-1.26*	[-5.14]	-0.16*	[-3.12]	0.41	32.93*
Size: Small	816	1.41*	-0.20*	[-5.76]	-0.21*	[-10.26]	0.18	82.15*
Size: Medium	321	1.24*	-0.19*	[-2.23]	-0.12*	[-2.92]	0.05	6.37*
Size: Large	146	0.79	-0.05	[-0.81]	-0.27*	[-5.03]	0.12	13.84*
Minimum efficient scale	187	0.88*	-0.10*	[-2.28]	-0.19*	[-3.91]	0.10	11.02*
Single-plant firm	684	1.77*	-0.35*	[-11.84]	-0.17*	[-8.35]	0.37	140.98*
Foreign owned	253	1.96*	-0.27*	[-4.17]	-0.19*	[-2.83]	0.40	75.24
15: Food and beverages	125	2.14*	-0.27*	[-6.55]	-0.33*	[-4.94]	0.44	39.18*
16: Tobacco products	..	..	..	..	..	..	..	..
17: Textiles	48	1.68*	-0.28*	[-3.38]	-0.15*	[-2.27]	0.34	8.80*
18: Wearing apparel	46	1.36*	-0.20*	[-3.02]	-0.17*	[-2.70]	0.30	9.02*
19: Leather products	11	0.23	-0.07	[-0.37]	-0.10	[-0.63]	0.04	0.23
20: Wood and wood products	49	1.25*	-0.22*	[-2.19]	-0.15	[-1.58]	0.20	7.14*
21: Pulp and paper products	54	1.99*	-0.31*	[-3.29]	-0.19*	[-2.02]	0.49	19.77*
22: Publishing and printing	72	1.23*	-0.17*	[-2.34]	-0.17*	[-3.23]	0.22	10.19*
23: Coke, petroleum and nuclear fuel	..	..	..	..	..	..	..	..
24: Chemicals and chemical products	88	1.60*	-0.24*	[-3.75]	-0.15*	[-2.30]	0.39	22.46*
25: Rubber and plastic products	104	1.60*	-0.34*	[-6.04]	-0.02	[-0.21]	0.37	38.95*
26: Non-metallic mineral products	72	2.20*	-0.34*	[-6.48]	-0.33*	[-4.87]	0.60	35.54*
27: Basic metals	83	1.17*	-0.20*	[-5.30]	-0.07	[-0.75]	0.32	22.87*
28: Fabricated metal products	190	1.23*	-0.22*	[-5.28]	-0.09*	[-2.23]	0.20	25.28*
29: Other machinery and equipment	141	1.67*	-0.28*	[-6.58]	-0.23*	[-4.24]	0.47	44.22*
30: Office machinery and computers	15	0.72	-0.07	[-0.67]	0.05	[0.12]	0.01	0.26
31: Electrical machinery	57	1.49*	-0.33*	[-2.37]	0.08	[0.47]	0.28	8.86*
32: Radio, T.V. and communications	46	1.39*	-0.10	[-1.70]	-0.35*	[-4.50]	0.40	14.93*
33: Medical and optical instruments	46	1.73*	-0.18*	[-2.71]	-0.35*	[-3.25]	0.39	7.87*
34: Motor vehicles and trailers	45	1.26*	-0.05	[-1.16]	-0.32*	[-4.53]	0.31	12.14*
35: Other transport equipment	21	1.87*	-0.10	[-2.05]	-0.41*	[-3.59]	0.50	12.87*
36: Furniture	86	1.50*	-0.19*	[-2.98]	-0.26*	[-3.96]	0.41	33.62*

Notes: Model estimated using OLS with heteroscedasticity-consistent standard errors.
Robust t-statistics for the constant term, α, not reported.
.. Not reported given small sample < 10 plants.
* Statistically significant at the 5 percent level.

1977-1986 time period the β_1 coefficient was only statistically significant for the Micro plant size-class with γ_1 statistically significant for the Small and Medium size-classes. The F-test rejected the joint hypothesis of zero coefficients in each of these samples. For the 1986-1997 sample, γ_1 was statistically significant in all four samples with β_1 only not statistically significant for Large plants, with the F-test again showing rejections across all samples.

At the industry-level, for the 1972-1977 sample, the β_1 coefficient was statistically significant in 11 samples (in all cases with a negative coefficient). The largest negative β_1 coefficients were observed in the Furniture, Wood and wood products and Textiles industries. The γ_1 coefficients were statistically significant in 6 industries. The F-test rejected jointly zero coefficients in 12 of the 20 samples. The first proposition of Gibrat's Law could not be rejected in the Wearing apparel, Pulp and paper products, Coke, petroleum and nuclear fuel, Rubber and plastic products, Basic metals, Electrical machinery, Radio, television and communications, Motor vehicles and trailers and Other transport equipment industries, consistent with the earlier model specification, given by equation (5.2), with the addition of the Wearing apparel and Rubber and plastic products industries.

In the 1972-1997 sample the null hypothesis that $\beta_1 = 0$ could be rejected in all but two industries, Wearing apparel and Other transport equipment (the first proposition of Gibrat's Law being previously rejected in the equation (5.2) specification for Wearing apparel). The strongest mean reversion tendencies were exhibited in the Rubber and plastic products, Chemical and chemical products and Pulp and paper products industries, consistent with the previous regression model without age. As might be expected given the longer time-span of this cross-section, γ_1 was statistically significant in only a small number of industries, Rubber and plastic products, Basic metals, Fabricated metal products and Motor vehicles and trailers. The F-test rejected the joint hypothesis of zero coefficients in every industry sample for this period except for the Other transport equipment industry.

The β_1 coefficient was statistically significant in 13 of the 20 samples in the 1977-1986 sample (the largest negative coefficients being found in the Fabricated metal products, Wood and wood products, Rubber and plastic products, Electrical machinery and Radio, television and communications industries) with γ_1 statistically significant in 8 samples. The F-test rejected the joint null hypothesis in 15 industries. The first proposition of Gibrat's Law, however, could not be rejected in the Textiles, Wearing apparel, Leather products, Coke, petroleum and nuclear fuel, Medical and optical instruments, Other transport equipment and Furniture industries. The corresponding rejections for the 1986-1997 sample were 15 for plant size, 14 for plant age and 18 for the joint null hypothesis, respectively, with Gibrat's Law not being rejected in the Leather products, Office machinery and computers, Radio, television and communications equipment, Motor vehicles and trailers and Other transport equipment industries.

Across the full range of time periods and industry samples the F-test of jointly zero β_1 and γ_1 coefficients could not be rejected at the 5 percent level in two or more samples in only three industries, Leather products, Coke, petroleum and nuclear fuel and Other transport equipment. For the Food and beverages, Wood and wood products, Publishing and printing, Chemical and chemical products, Non-metallic mineral products, Fabricated metal products and Other machinery and equipment industries the β_1 coefficients were statistically significant and negative in each of the four time periods, indicating fairly persistent trends of mean reversion.

Broadly similar coefficient estimates (compared to the model specification without the plant age variable) were also generated for single-plant firms, foreign owned plants and plants above the MES. The first proposition of Gibrat's Law was rejected in all of these estimates with the exception of the MES sample in 1972-1977, where a not statistically significant β_1 coefficient of 0.04 was estimated.

Consistent with the previous model estimations, persistence terms were added to equation (5.4) providing the model specification:

$$\Delta \ln X_{it} = \alpha + \beta_1 \ln X_{it-1} + \gamma \ln A_{it} + \delta(\ln X_{it-1} - \ln X_{it-2}) + \nu_{it} \qquad (5.5)$$

The results, presented in Table 5.7, from the estimations using OLS[93] with White's correction[94], confirmed the above findings with β_1 again found to be negative and statistically significant in all but one of the twenty samples considered (the exception being for the 1972-1977 period with a three-year persistence term where the negative coefficient was not significant), with negative γ_1 coefficients in all but two samples (commencing in 1972 with five-year persistence terms), being statistically significant in 15 cases. The inclusion of the age term gave rise to similar or slightly less negative β_1 coefficients compared to the serial correlation model without age terms (so that at the aggregate-level the inclusion of a persistence term generally had a slight, but consistent effect, modestly reducing the magnitude of the mean reversion process). Statistically significant δ coefficients were observed in only four of the corresponding samples, two with one-year persistence terms, relating to 1972-1977 and 1977-1986 and the other two

---

[93] Although the formal tests of the residuals tended to reject normality at the 5 percent level of significance (rejected by the D'Agostino et al and Shapiro-Francia tests in all cases except for the 1972-1997 samples with both one-year and two-year persistence terms, and additionally with non-rejections for the three-year and four-year persistence terms for the same time period using the Shapiro-Francia test) the deviations from the skewness and kurtosis benchmark values associated with the normal distribution were modest (a maximum absolute skewness value of -0.74, and with a minimum kurtosis value of 3.10 and a maximum value of 8.94). Again, the residuals in all regressions had mean values very close to zero.

[94] Consistent with the serial correlation model without age, in only the 1972-1997 samples with two-year and three-year time-span persistence terms could neither the null hypothesis of the Breusch-Pagan nor White's test be rejected at the 5 percent level of significance. Ramsey's RESET test provided a rejection of the null hypothesis of no omitted variables in each case.

Table 5.7: Plant Size, Age and Growth: Serial Correlation Model

Sample	Plants (N)	α	β_1	t_{β_1}	δ	t_δ	γ	t_γ	R^2	F
One-Year:										
1972-1977	1,634	0.46*	-0.07*	[-7.33]	-0.18*	[-2.75]	-0.09*	[-4.96]	0.08	37.75*
1972-1997	649	1.88*	-0.37*	[-18.69]	-0.02	[-0.19]	-0.19*	[-4.99]	0.42	138.68*
1977-1986	1,397	1.01*	-0.17*	[-13.07]	-0.28*	[-2.87]	-0.19*	[-7.77]	0.23	116.63*
1986-1997	1,345	1.48*	-0.20*	[-15.11]	-0.15	[-0.84]	-0.20*	[-9.70]	0.30	176.26*
Two-Year:										
1972-1977	1,535	0.40*	-0.06*	[-6.67]	-0.06	[-1.37]	-0.08*	[-4.16]	0.06	27.23*
1972-1997	617	1.73*	-0.37*	[-18.13]	0.09	[1.02]	-0.13*	[-3.08]	0.40	123.01*
1977-1986	1,341	0.96*	-0.18*	[-13.77]	-0.04	[-0.63]	-0.15*	[-5.40]	0.20	97.27*
1986-1997	1,265	1.54*	-0.20*	[-14.39]	-0.16	[-1.39]	-0.23*	[-8.67]	0.28	149.96*
Three-Year:										
1972-1977	1,419	0.30*	-0.06	[-6.30]	-0.03	[-0.81]	-0.05*	[-2.21]	0.04	18.56*
1972-1997	569	1.58*	-0.37*	[-17.20]	0.14	[1.90]	-0.08	[-1.56]	0.39	104.88*
1977-1986	1,240	1.00*	-0.18*	[-13.83]	0.04	[0.71]	-0.16*	[-4.79]	0.20	90.02*
1986-1997	1,203	1.56*	-0.19*	[-14.11]	-0.10	[-1.01]	-0.25*	[-7.65]	0.26	131.15*
Four-Year:										
1972-1977	1,309	0.22*	-0.06*	[-5.91]	-0.03	[-0.83]	-0.02	[-0.76]	0.04	13.68*
1972-1997	525	1.40*	-0.36*	[-16.10]	0.26*	[2.88]	-0.03	[-0.51]	0.38	91.47*
1977-1986	1,160	0.94*	-0.18*	[-13.61]	0.04	[0.70]	-0.14*	[-3.79]	0.19	75.62*
1986-1997	1,161	1.67*	-0.18*	[-13.45]	-0.11	[-1.42]	-0.29*	[-8.45]	0.26	124.42*
Five-Year:										
1972-1977	1,210	0.12*	-0.05*	[-5.40]	-0.07	[-1.71]	0.02	[0.70]	0.04	10.83*
1972-1997	482	1.28*	-0.36*	[-15.51]	0.30*	[3.14]	0.01	[0.18]	0.40	84.05*
1977-1986	1,092	0.80*	-0.18*	[-13.22]	0.08	[1.70]	-0.09*	[-2.15]	0.17	63.22*
1986-1997	1,093	1.56*	-0.18*	[-12.42]	-0.10	[-1.38]	-0.26*	[-7.02]	0.24	102.84*

Notes: Model estimated using OLS with heteroscedasticity-consistent standard errors.
Robust t-statistics for the constant term, α, not reported.
* Statistically significant at the 5 percent level.

for the four-year and five-year persistence terms in the 1972-1997 sample. The F-test rejected the joint hypothesis of zero coefficients in all cases.

Overall, the estimation of the plant size-age-growth models provides results highly consistent with the previous economic evidence. Although with some variation at disaggregated levels, the aggregate samples exhibit a relationship of plant growth declining with both size (exhibiting statistically significant mean reversion and rejecting the first proposition of Gibrat's Law) and also with age. In the following section, some further analysis is conducted to investigate whether such negative size-growth and age-growth relationships are of a potentially non-linear functional form.

5.6. Non-Linearities in the Plant Size-Growth Relationship

To further examine the functional form of the model specifications non-linear (quadratic) size and age terms were incorporated into the regression model (consistent with, for example, Evans, 1987a, 1987b), providing a specification of following structure:

$$\Delta \ln X_{it} = \alpha + \beta_1 \ln X_{it-1} + \beta_2 \ln X_{it-1}^2 + \gamma_1 \ln A_{it} + \gamma_2 \ln A_{it}^2 + v_{it} \qquad (5.6)$$

with the variables as denoted previously. The estimation was again conducted using OLS[95] with White's correction for heteroscedasticity, given the further frequently detected rejections of both White's test and the Breusch-Pagan test (apparent at the aggregate level in all samples except for the 1972-1997 time period, and across a number

---

[95] Examination of the D'Agostino et al. and Shapiro-Francia tests for normality of the regression residuals again showed broadly similar results to the earlier regression models, although slightly less frequently rejecting the null hypothesis of normality. In all cases the assumption of a zero mean was satisfied, with generally modest deviations from the benchmark skewness and kurtosis values for normality. Only 6 kurtosis indices were not single-digit values, with 72 samples of the residuals being more suitably described, following the previously discussed criteria, by the normal distribution and 26 by the Laplace distribution, again supporting the use of OLS rather than the LAD estimator.

of size-classes (in total, 6 using White's test, 7 using the Breusch-Pagan test), industries (11 and 17 times, respectively) and other samples disaggregated (3 and 7 times, respectively) in the examined time periods.

Given the inclusion of quadratic terms within equation (5.6) a simple analysis of collinearity matrices for each regression model was conducted to further examine the relationship between the size and age variables and their respective quadratic terms. For the aggregate samples correlation coefficients between $\ln X_{it-1}$ and $\ln X_{it-1}^2$ of -0.98 and of -0.95 between the $\ln A_{it}$ and $\ln A_{it}^2$ terms were observed for the 1972-1977 regression model, with corresponding correlation coefficients of -0.97 and -0.95 for the 1972-1997 period, -0.97 and -0.95 for 1977-1986, and -0.99 and -0.94 for 1986-1997, respectively. Similarly high correlation coefficients were also detected across a range of sub-samples. However, despite these relatively high correlation coefficients, as Gudjarati (1988, p.306) notes, strictly, the inclusion of quadratic terms does not contravene the further assumption made in the multiple regression form of the CLR model (Gudjarati, 1988, p.166) that there are no exact linear relationships across the explanatory variables in the model (Kennedy, 1998, p.44)[96].

The estimates of these non-linear model specifications, presented in Tables 5.8a-d, provide consistent evidence of statistically significant coefficients for β_1, γ_1 and β_2 in all four aggregate samples, with the γ_2 coefficient statistically significant at the 5 percent in three of the four aggregate cross-section estimates, being not significant in the 1986-1997 period. The estimated β_1 coefficients were more negative in this model with the inclusion of such quadratic terms than in the previous model specifications, with the coefficients of -0.33 for the 1972-1977 period, -0.68 for the 1972-1997 period, -0.33 for 1977-1986

---

[96] Multicollinearity between explanatory variables could produce excessively large standard errors (although unbiased estimators) and wider parameter standard errors (with smaller estimated t-values) the potential effect being a Type II error and acceptance of false hypotheses (Gudjarati, 1988, p.292). It is also of note that the inclusion of such non-linear terms may impact on the interpretation of the magnitude of the effect of plant size on growth (Elston, 2002, p.4).

and -0.62 in 1986-1997, again firmly rejecting the first proposition of Gibrat's Law. In all cases, the γ_1 coefficient was also of predicted sign, being negative in all four aggregate samples. The coefficients β_2 and γ_2 terms were both positively signed, although of generally modest magnitude. The null hypothesis of the F-test was rejected in all four samples, so that it cannot be concluded that the inclusion of the previously omitted plant-level terms, $\ln X_{it-1}^2$ and $\ln A_{it}^2$, are not statistically significant at the 95 percent level of confidence.

Examination of the effects of the inclusion of the additional quadratic variables on the results of Ramsey's RESET test suggested somewhat less severe rejections of the null hypothesis in the non-linear model, but with continued rejections in all time periods at the aggregate level except for 1977-1986 period. Overall, the RESET test for the non-linear model specification provided statistically significant rejections (at the 5 percent level) of the null hypothesis in 8, 5, 4 and 7 samples, in each of the time periods respectively, in 24 samples in total out of the 110 samples examined, lower than either of the linear specifications (38 and 36 rejections, respectively). However, for these cases, since the RESET test examines whether the powers of the dependent variable have non-zero coefficients (Verbeek, 2000, p.58) this may indicate some greater complexity in the form of non-linearities of the plant size-growth relationship (or indeed the possibility of the omission of other relevant variables).

For the size-class samples the β_1 coefficient was found to be statistically significant (rejecting Gibrat's Law) in only two of the sixteen samples, for Small plants in the period 1972-1977 and for Micro plants in 1986-1997. Such a result is in sharp contrast to the findings from the previous linear model specifications where size was observed to exert a generally negative effect for the smallest size-classes. Comparison of the β_1 coefficients clearly suggests some notable differences, often being of a greater

Table 5.8a: Non-Linear Specification: Plant Size, Age and Growth, 1972-1977

Sample	Plants (N)	α	β_1	t_{β_1}	γ_1	t_{γ_1}	β_2	t_{β_2}	γ_2	t_{γ_2}	R^2	F
All	1,754	1.06*	-0.33*	[-7.53]	-0.26*	[-5.14]	0.03*	[6.47]	0.05*	[4.02]	0.15	46.92*
Size: Micro	192	1.29*	-0.43	[-0.98]	-0.28*	[-1.70]	0.05	[0.32]	0.02	[0.39]	0.17	7.85*
Size: Small	930	1.80*	-0.92*	[-2.68]	-0.19*	[-3.12]	0.13*	[2.44]	0.03	[1.69]	0.05	12.79*
Size: Medium	411	-3.14	1.34	[0.91]	-0.07	[-0.38]	-0.14	[-0.93]	0.01	[0.34]	0.00	0.37
Size: Large	221	1.10	-0.22	[-0.69]	-0.47	[-1.95]	0.02	[0.76]	0.10	[1.96]	0.03	1.16
Minimum efficient scale	181	0.53	-0.04	[-0.13]	-0.73*	[-2.33]	0.01	[0.30]	0.16*	[2.41]	0.04	1.93
Single-plant firm	700	1.67*	-0.76*	[-6.93]	-0.25*	[-3.72]	0.09*	[5.94]	0.04	[1.95]	0.27	35.43*
Foreign owned	190	1.17*	-0.20	[-1.53]	-0.37*	[-2.58]	0.01	[1.13]	0.07*	[2.05]	0.18	6.03*

Table 5.8a: Non-Linear Specification: Plant Size, Age and Growth, 1972-1977, continued

Sample	Plants (N)	α	β_1	t_{β_1}	γ_1	t_{γ_1}	β_2	t_{β_2}	γ_2	t_{γ_2}	R^2	F
15: Food and beverages	183	0.60	-0.22	[-1.15]	-0.09	[-0.78]	0.02	[0.85]	0.02	[0.51]	0.06	1.68
16: Tobacco products	..	..	..	..	..	..	..	..	..	..	..	..
17: Textiles	69	1.05	-0.38	[-1.42]	-0.02	[-0.09]	0.03	[0.94]	0.01	[0.12]	0.18	3.13*
18: Wearing apparel	90	0.29	0.16	[0.75]	-0.55*	[-3.18]	-0.03	[-0.98]	0.12*	[2.74]	0.20	4.82*
19: Leather products	25	1.05	-0.31	[-1.17]	-0.35	[-1.89]	0.02	[0.54]	0.09	[1.35]	0.38	5.00*
20: Wood and wood products	100	1.68*	-0.74*	[-2.27]	-0.35	[-1.66]	0.09	[1.91]	0.06	[1.10]	0.17	5.80*
21: Pulp and paper products	34	3.03*	-1.28*	[-3.56]	-0.19	[-0.84]	0.13*	[3.34]	0.06	[1.27]	0.37	3.69*
22: Publishing and printing	101	0.83*	-0.45*	[-2.01]	0.04	[0.16]	0.05	[1.73]	0.01	[0.08]	0.16	2.16
23: Coke, petroleum and nuclear fuel	11	-2.69	-0.05	[-0.15]	2.78	[0.89]	0.01	[0.17]	-0.68	[-0.93]	0.09	0.36
24: Chemicals and chemical products	64	0.99*	-0.28	[-1.47]	-0.01	[-0.06]	0.02	[0.92]	-0.00	[-0.05]	0.21	4.50*
25: Rubber and plastic products	77	1.02*	-0.21	[-1.16]	-0.40	[-1.23]	0.02	[1.08]	0.06	[0.77]	0.14	4.34*
26: Non-metallic mineral products	133	1.27*	-0.41*	[-2.57]	-0.58*	[-2.94]	0.04	[1.82]	0.12*	[2.39]	0.17	6.82*
27: Basic metals	116	0.59*	-0.19	[-1.90]	-0.17	[-1.15]	0.02	[1.94]	0.03	[1.11]	0.06	2.30
28: Fabricated metal products	236	1.40*	-0.40*	[-2.90]	-0.51*	[-4.69]	0.04*	[2.26]	0.11*	[3.67]	0.33	15.72*
29: Other machinery and equipment	197	1.19*	-0.42*	[-2.96]	-0.16	[-0.93]	0.04*	[2.41]	0.03	[0.68]	0.16	5.58*
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	..	..	..
31: Electrical machinery	48	1.40*	-0.37*	[-2.49]	-0.39	[-1.53]	0.04*	[2.24]	0.07	[1.09]	0.26	3.36*
32: Radio, T.V. and communications	29	-0.34	0.44*	[2.52]	-0.22	[-0.65]	-0.06*	[-2.67]	0.03	[0.38]	0.33	2.79*
33: Medical and optical instruments	44	0.80*	-0.26	[-1.23]	0.24	[0.87]	0.02	[0.60]	-0.10	[-1.29]	0.24	3.07*
34: Motor vehicles and trailers	55	0.82	-0.12	[-0.50]	-0.21	[-0.98]	0.01	[0.56]	-0.02	[-0.25]	0.22	4.07*
35: Other transport equipment	36	1.55*	-0.59*	[-2.36]	-0.36*	[-2.10]	0.06*	[2.12]	0.11*	[2.17]	0.38	3.11*
36: Furniture	94	1.88*	-0.94*	[-3.38]	0.08	[0.23]	0.10*	[3.14]	-0.03	[-0.37]	0.21	7.02*

Notes: Model estimated using OLS with heteroscedasticity-consistent standard errors.

Robust t-statistics for the constant term, α, not reported.

.. Not reported given small sample < 10 plants.

* Statistically significant at the 5 percent level.

Table 5.8b: Non-Linear Specification: Plant Size, Age and Growth, 1972-1997

Sample	Plants (N)	α	β_1	t_{β_1}	γ_1	t_{γ_1}	β_2	t_{β_2}	γ_2	t_{γ_2}	R^2	F
All	693	2.61*	-0.68*	[-9.54]	-0.44*	[-4.56]	0.03*	[4.42]	0.07*	[2.82]	0.47	152.47*
Size: Micro	58	2.71*	-0.51	[-1.18]	0.14	[0.60]	-0.11	[-0.51]	-0.06	[-0.82]	0.29	12.84*
Size: Small	308	2.32	-0.68	[-0.75]	-0.48*	[-3.53]	0.05	[0.36]	0.08*	[2.22]	0.14	11.18*
Size: Medium	200	1.08	-0.04	[-0.01]	-0.48*	[-2.67]	-0.02	[-0.06]	0.08	[1.67]	0.06	5.81*
Size: Large	127	4.06	-0.96	[-1.24]	-0.47	[-1.33]	0.05	[0.84]	0.07	[0.90]	0.16	6.15*
Minimum efficient scale	92	4.70*	-1.03	[-1.91]	-0.91*	[-2.60]	0.05	[1.31]	0.16	[1.91]	0.29	12.34*
Single-plant firm	266	2.95*	-1.04*	[-7.33]	-0.24	[-1.81]	0.07*	[2.90]	0.03	[0.91]	0.53	89.46*
Foreign owned	126	3.11*	-0.68*	[-2.07]	-0.28	[-1.41]	0.02	[0.77]	0.02	[0.32]	0.52	29.38*

Table 5.8b: Non-Linear Specification: Plant Size, Age and Growth, 1972-1997, continued

Sample (SIC)	Plants (N)	α	β_1	t_{β_1}	γ_1	t_{γ_1}	β_2	t_{β_2}	γ_2	t_{γ_2}	R^2	F
15: Food and beverages	76	2.86*	-0.83*	[-4.11]	-0.55*	[-2.16]	0.06*	[2.35]	0.13	[1.90]	0.42	14.75*
16: Tobacco products	..	..	..	..	..	..	..	..	..	..	..	..
17: Textiles	22	0.57	1.06	[1.65]	-1.83*	[-4.17]	-0.18*	[-2.16]	0.42*	[3.14]	0.55	13.71*
18: Wearing apparel	24	-1.39	1.65*	[2.34]	-1.62*	[-3.56]	-0.18*	[-2.64]	0.24*	[2.13]	0.47	40.56*
19: Leather products	..	..	..	..	..	..	..	..	..	..	..	..
20: Wood and wood products	24	2.64	-0.67	[-0.39]	-1.09*	[-3.09]	0.06	[0.22]	0.26*	[2.58]	0.38	26.44*
21: Pulp and paper products	22	5.48*	-1.90*	[-5.86]	-0.22	[-0.44]	0.16*	[4.57]	0.03	[0.21]	0.82	124.54*
22: Publishing and printing	43	3.59*	-1.28*	[-3.63]	-0.91	[-1.82]	0.14*	[3.16]	0.20	[1.47]	0.49	6.93*
23: Coke, petroleum, nuclear fuel	..	..	..	..	..	..	..	..	..	..	..	..
24: Chemicals, chemical products	35	2.74	-0.57	[-0.84]	-0.46	[-1.08]	-0.00	[-0.02]	0.14	[1.08]	0.61	8.71*
25: Rubber and plastic products	38	3.97*	-1.04*	[-2.71]	0.30	[0.80]	0.04	[1.02]	-0.17	[-1.45]	0.71	41.32*
26: Non-metallic mineral products	43	3.44*	-1.22*	[-2.66]	-0.44	[-0.84]	0.10	[1.97]	0.07	[0.57]	0.40	7.80*
27: Basic metals	52	2.41*	-0.57*	[-2.23]	-0.50	[-1.77]	0.02	[1.17]	0.08	[1.04]	0.56	16.24*
28: Fabricated metal products	88	2.37*	-0.56*	[-2.21]	-0.34	[-1.52]	0.02	[0.52]	0.03	[0.53]	0.45	21.36*
29: Other machinery, equipment	73	2.77*	-0.97*	[-5.66]	-0.31	[-1.12]	0.06*	[3.90]	0.05	[0.67]	0.66	56.62*
30: Office machinery, computers	..	..	..	..	..	..	..	..	..	..	..	..
31: Electrical machinery	24	2.40*	-0.21	[-0.79]	-0.42	[-0.67]	-0.03	[-0.72]	0.07	[0.41]	0.52	10.87*
32: Radio, T.V., communications	15	0.80	-0.66	[-0.94]	2.00	[1.34]	0.02	[0.28]	-0.49	[-1.43]	0.55	5.31*
33: Medical, optical instruments	16	2.50*	-0.88*	[-2.33]	0.49	[0.67]	0.07	[0.91]	-0.17	[-0.82]	0.41	8.50*
34: Motor vehicles and trailers	31	2.13	-0.41	[-0.91]	-0.49	[-0.79]	0.02	[0.38]	0.05	[0.29]	0.37	9.66*
35: Other transport equipment	16	2.88*	-1.09*	[-7.89]	-0.31	[-0.65]	0.10*	[5.22]	0.04	[0.32]	0.63	23.64*
36: Furniture	33	2.75*	-0.82	[-1.39]	-0.48	[-1.00]	0.05	[0.55]	0.13	[1.01]	0.52	6.71*

Notes: Model estimated using OLS with heteroscedasticity-consistent standard errors.
Robust t-statistics for the constant term, α, not reported.
.. Not reported given small sample < 10 plants.
* Statistically significant at the 5 percent level.

Table 5.8c: Non-Linear Specification: Plant Size, Age and Growth, 1977-1986

Sample	Plants (N)	α	β_1	t_{β_1}	γ_1	t_{γ_1}	β_2	t_{β_2}	γ_2	t_{γ_2}	R^2	F
All	1,444	1.39*	-0.33*	[-6.38]	-0.34*	[-4.84]	0.02*	[3.28]	0.05*	[2.49]	0.26	110.19*
Size: Micro	206	1.54*	-0.66	[-1.76]	-0.27	[-1.60]	0.14	[0.89]	0.06	[1.19]	0.10	4.96*
Size: Small	735	1.22	-0.38	[-0.69]	-0.33*	[-3.46]	0.05	[0.58]	0.04	[1.57]	0.08	13.10*
Size: Medium	333	7.98	-2.99	[-1.23]	-0.68*	[-5.03]	0.30	[1.22]	0.12*	[3.86]	0.08	9.85*
Size: Large	170	-2.57	0.77	[0.87]	-0.01	[-0.03]	-0.06	[-1.06]	-0.03	[-0.39]	0.05	4.72*
Minimum efficient scale	156	-0.32	0.31	[0.98]	-0.35	[-1.30]	-0.04	[-1.54]	0.05	[0.87]	0.12	8.37*
Single-plant firm	590	2.20*	-0.86*	[-7.22]	-0.23*	[-2.40]	0.08*	[4.29]	0.01	[0.56]	0.36	76.39*
Foreign owned	203	1.71*	-0.32	[-1.77]	-0.17	[-0.82]	0.02	[0.98]	-0.05	[-0.93]	0.37	23.03*

Table 5.8c: Non-Linear Specification: Plant Size, Age and Growth, 1977-1986, continued

Sample	Plants (N)	α	β_1	t_{β_1}	γ_1	t_{γ_1}	β_2	t_{β_2}	γ_2	t_{γ_2}	R^2	F
15: Food and beverages	143	1.85*	-0.52*	[-2.32]	-0.37	[-1.83]	0.05	[1.64]	0.06	[0.98]	0.28	9.40*
16: Tobacco products	..	..	..	..	..	..	..	..	..	..	..	..
17: Textiles	43	1.09	0.16	[0.29]	-1.31*	[-3.56]	-0.03	[-0.50]	0.28*	[3.00]	0.26	9.02*
18: Wearing apparel	60	0.16	0.33	[0.76]	-0.32	[-0.86]	-0.05	[-1.13]	0.04	[0.43]	0.20	3.07*
19: Leather products	14	1.98	-0.88	[-0.90]	-0.39	[-0.21]	0.11	[0.91]	0.06	[0.17]	0.06	0.28
20: Wood and wood products	67	1.29*	-0.55	[-1.36]	-0.01	[-0.02]	0.04	[0.64]	-0.00	[-0.05]	0.27	6.31*
21: Pulp and paper products	37	2.45*	-0.58*	[-2.96]	-0.41	[-0.98]	0.04	[1.94]	0.04	[0.43]	0.65	16.69*
22: Publishing and printing	93	1.90*	-0.60*	[-2.47]	-0.37	[-0.89]	0.07*	[2.20]	0.02	[0.16]	0.38	7.72*
23: Coke, petroleum and nuclear fuel	10	0.63	0.86	[1.84]	-2.65	[-0.52]	-0.11	[-2.31]	0.72	[0.50]	0.56	16.45*
24: Chemicals and chemical products	66	1.78*	-0.55*	[-2.84]	-0.38	[-1.44]	0.04	[1.73]	0.06	[0.87]	0.35	10.86*
25: Rubber and plastic products	82	1.85*	-0.77*	[-2.33]	0.36	[1.11]	0.06	[1.72]	-0.10	[-1.28]	0.24	6.44*
26: Non-metallic mineral products	123	0.82*	-0.48*	[-3.58]	-0.26	[-0.98]	0.05*	[2.62]	0.11	[1.64]	0.19	8.97*
27: Basic metals	81	1.48*	-0.28	[-1.80]	-0.43	[-1.66]	0.01	[0.68]	0.05	[0.82]	0.41	15.56*
28: Fabricated metal products	181	1.84*	-0.59*	[-3.66]	-0.25	[-1.87]	0.04	[1.86]	0.03	[0.88]	0.37	23.07*
29: Other machinery and equipment	150	1.49*	-0.39*	[-2.83]	-0.44*	[-2.32]	0.03	[1.87]	0.06	[1.25]	0.32	20.14*
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	..	..	..
31: Electrical machinery	47	2.77*	-1.35*	[-5.02]	0.83	[1.33]	0.13*	[4.33]	-0.29	[-1.44]	0.47	16.51*
32: Radio, T.V. and communications	33	2.66*	-0.82*	[-3.16]	-0.42	[-0.74]	0.07	[1.97]	0.06	[0.40]	0.43	6.94*
33: Medical and optical instruments	43	1.80*	-0.76*	[-3.15]	0.17	[0.36]	0.09*	[2.38]	-0.10	[-0.75]	0.25	5.20*
34: Motor vehicles and trailers	50	2.61*	-0.64*	[-3.14]	-0.72*	[-2.10]	0.05*	[2.41]	0.11	[1.61]	0.53	11.89*
35: Other transport equipment	26	1.41	-0.69	[-1.22]	-0.41	[-0.88]	0.07	[1.32]	0.09	[0.76]	0.11	1.54
36: Furniture	82	1.10	-0.09	[-0.26]	-0.41	[-1.33]	-0.00	[-0.07]	0.04	[0.49]	0.19	5.94*

Notes: Model estimated using OLS with heteroscedasticity-consistent standard errors.
Robust t-statistics for the constant term, α, not reported.
.. Not reported given small sample < 10 plants.
* Statistically significant at the 5 percent level.

Table 5.8d: Non-Linear Specification: Plant Size, Age and Growth, 1986-1997

Sample	Plants (N)	α	β_1	t_{β_1}	γ_1	t_{γ_1}	β_2	t_{β_2}	γ_2	t_{γ_2}	R^2	F
All	1,417	2.18*	-0.62*	[-7.19]	-0.14*	[-2.58]	0.05*	[5.21]	-0.01	[-0.90]	0.35	165.01*
Size: Micro	134	4.39*	-2.54*	[-2.26]	-0.33*	[-2.06]	0.45	[1.29]	0.05	[1.08]	0.44	19.34*
Size: Small	816	1.86*	-0.54	[-1.28]	-0.10	[-1.47]	0.05	[0.80]	-0.03	[-1.84]	0.18	45.81*
Size: Medium	321	2.69	-0.96	[-0.44]	0.23	[1.50]	0.08	[0.37]	-0.08*	[-2.40]	0.06	5.03*
Size: Large	146	5.29	-1.29	[-1.46]	-0.63*	[-2.83]	0.09	[1.46]	0.07	[1.54]	0.14	10.07*
Minimum efficient scale	187	1.34	-0.30	[-0.85]	-0.05	[-0.33]	0.02	[0.59]	-0.03	[-0.81]	0.10	5.89*
Single-plant firm	684	2.65*	-0.99*	[-6.06]	-0.11	[-1.69]	0.10*	[3.78]	-0.02	[-1.00]	0.40	96.19*
Foreign owned	253	3.21*	-0.94*	[-2.65]	-0.05	[-0.40]	0.07*	[2.10]	-0.03	[-1.19]	0.44	38.08*

Table 5.8d: Non-Linear Specification: Plant Size, Age and Growth, 1986-1997, continued

Sample	Plants (N)	α	β_1	t_{β_1}	γ_1	t_{γ_1}	β_2	t_{β_2}	γ_2	t_{γ_2}	R^2	F
15: Food and beverages	125	2.85*	-0.66*	[-2.91]	-0.50	[-1.89]	0.05	[1.88]	0.05	[0.78]	0.46	22.68*
16: Tobacco products	..	..	..	..	..	..	..	..	..	..	..	..
17: Textiles	48	2.54*	-0.79	[-1.46]	-0.25	[-1.04]	0.07	[0.98]	0.03	[0.47]	0.36	4.78*
18: Wearing apparel	46	2.91*	-1.07*	[-3.68]	-0.04	[-0.16]	0.11*	[2.84]	-0.04	[-0.50]	0.39	10.99*
19: Leather products	11	6.05	-2.05	[-2.23]	-1.18	[-0.49]	0.23	[2.02]	0.16	[0.34]	0.31	1.38
20: Wood and wood products	49	2.22*	-1.00*	[-3.84]	0.09	[0.21]	0.11*	[3.43]	-0.05	[-0.50]	0.26	10.45*
21: Pulp and paper products	54	3.09*	-0.87*	[-2.06]	-0.29	[-1.02]	0.07	[1.47]	0.03	[0.46]	0.53	20.22*
22: Publishing and printing	72	2.76*	-1.22*	[-3.05]	0.13	[0.53]	0.15*	[2.85]	-0.08	[-1.31]	0.39	9.15*
23: Coke, petroleum and nuclear fuel	..	..	..	..	..	..	..	..	..	..	..	..
24: Chemicals and chemical products	88	0.88	0.21	[0.83]	-0.29	[-1.72]	-0.05	[-1.78]	0.03	[0.72]	0.42	13.27*
25: Rubber and plastic products	104	1.31*	-0.24	[-0.92]	0.20	[1.04]	-0.01	[-0.36]	-0.06	[-1.24]	0.39	23.43*
26: Non-metallic mineral products	72	3.31*	-0.94*	[-3.66]	-0.62*	[-2.72]	0.08*	[2.60]	0.08	[1.49]	0.66	32.77*
27: Basic metals	83	1.65*	-0.60*	[-3.78]	0.30	[0.91]	0.04*	[2.89]	-0.08	[-1.28]	0.38	21.06*
28: Fabricated metal products	190	1.60*	-0.54*	[-2.88]	0.15	[1.46]	0.04	[1.85]	-0.07*	[-2.22]	0.23	13.16*
29: Other machinery and equipment	141	2.57*	-0.77*	[-4.25]	-0.33*	[-2.07]	0.06*	[2.53]	0.03	[0.85]	0.51	30.68*
30: Office machinery and computers	15	0.99	-0.08	[-0.05]	-1.12	[-0.85]	0.00	[0.02]	0.54	[0.85]	0.15	0.89
31: Electrical machinery	57	2.84*	-1.06*	[-2.11]	-0.14	[-0.56]	0.09	[1.65]	0.06	[0.77]	0.39	5.11*
32: Radio, T.V. and communications	46	1.44*	-0.22	[-0.87]	-0.01	[-0.05]	0.02	[0.53]	-0.09	[-1.32]	0.43	8.13*
33: Medical and optical instruments	46	2.17*	-0.40	[-1.05]	-0.61	[-1.29]	0.03	[0.73]	0.08	[0.70]	0.42	3.96*
34: Motor vehicles and trailers	45	0.34	0.25	[0.84]	-0.04	[-0.14]	-0.03	[-1.09]	-0.06	[-1.06]	0.34	9.59*
35: Other transport equipment	21	2.49*	-0.59	[-1.37]	-0.08	[-0.17]	0.05	[1.21]	-0.07	[-0.57]	0.54	11.57*
36: Furniture	86	1.50*	-0.23	[-0.83]	-0.12	[-0.68]	0.01	[0.17]	-0.04	[-0.79]	0.41	19.68*

Notes: Model estimated using OLS with heteroscedasticity-consistent standard errors.
Robust t-statistics for the constant term, α, not reported.
.. Not reported given small sample < 10 plants.
* Statistically significant at the 5 percent level.

magnitude (but of the same sign), than in the previous linear regression models.

The γ_1 coefficient was found to be statistically significant in half of these samples, for Micro and Small plants in the 1972-1977 sample, for Small- and Medium-sized plants in both the 1972-1997 and 1977-1986 periods, and in both the Micro and Large plants for the 1986-1997 period. In all but two of these samples the coefficients were negative, consistent with the previous results. The squared terms were generally not statistically significant, being significant in one and three samples, respectively, for the $\ln X_{it-1}^2$ and $\ln A_{it}^2$ terms. F-tests rejected the null hypothesis at the 5 percent level of significance in all samples except for the Medium and Large plant size-classes in the 1972-1977 sample.

At the industry-level the results were mixed with the β_1 coefficient being statistically significant in 10 (1972-1977), 11 (1972-1997), 12 (1977-1986) and 10 (1986-1997) samples. The coefficient of the plant age variable, γ_1, was statistically significant in four, four, three and two samples, respectively. In all of these cases these statistically significant coefficients were negative, except for β_1 for the Radio, television and communications industry in the 1972-1977 sample, and Wearing apparel in 1972-1977. In relation to the quadratic size and age variables the former suggested 7, 7, 5 and 6 statistically significant coefficients, respectively, with the latter giving rise to statistically significant coefficients in 4, 3, 1 and 1 cases. For the quadratic size term, whilst most of the statistically significant β_2 coefficients were positive, consistent with the aggregate and size-class results, three were negative. For the quadratic age term only one of the nine statistically significant γ_2 coefficients were negative, the remainder being positive.

The first proposition of Gibrat's Law, of independence between plant size and growth, could not be rejected in the Food and beverages, Textiles, Wearing apparel, Leather products, Coke, petroleum and nuclear fuel, Chemical and chemical products, Rubber and plastic products, Basic metals, Medical and optical instruments and Motor vehicles and trailers industries in the 1972-1977 time period. For the long 1972-1997 period similar supportive findings also emerged for the Textiles, Wood and wood products, Chemical and chemical products, Electrical machinery, Radio, television and communications, Motor vehicles and trailers and Furniture industries (in several cases exhibiting relatively large coefficients).

Such non-rejections of Gibrat's Law were also found for the Textiles, Wearing apparel, Leather products, Wood and wood products, Coke, petroleum and nuclear fuel, Basic metals, Other transport equipment and Furniture industries for the 1977-1986 period. For the 1986-1997 sample support for Gibrat's Law was found in the Textiles, Leather products, Chemical and chemical products, Rubber and plastic products, Office machinery and computers, Radio, television and communications, Medical and optical instruments, Motor vehicles and trailers, Other transport equipment and Furniture industries.

Statistically significant β_1 coefficients for all four time periods were found in only five industries: Pulp and paper products, Publishing and printing, Non-metallic mineral products, the Fabricated metal products and Other machinery and equipment industries and a further two industries in three of the four time periods, Food and beverages and Electrical machinery. The remaining industries did not provided statistically significant β_1 coefficients in more than half of the samples.

For the single-plant firms the first proposition of Gibrat's Law could again be rejected, with statistically significant coefficients of greater magnitude than estimated under the previous linear model specifications for all four time periods. Again it should be noted that relatively high correlation coefficients were observed between $\ln X_{it-1}$ and $\ln X_{it-1}^2$ and also between $\ln A_{it}$ and $\ln A_{it}^2$. However, despite observing larger negative β_1 coefficients in the samples of foreign owned plants, in contrast to the previous model estimates, Gibrat's Law could only be rejected in the case of the 1986-1997 and 1972-

1997 samples. In none of the four time periods could Gibrat's Law be rejected for plants above the estimated MES, supporting the third version of Gibrat's Law.

Although the inclusion of quadratic size and age terms to examine potential non-linearities in the plant size-growth relationship did not, at least at the aggregate-level, alter the main conclusions that growth is negatively related to both size (rejecting Gibrat's Law) and age, some differences are clearly observed in the magnitude, sign and significance of the estimated coefficients across some disaggregated samples. The role of plant age was also observed to be somewhat less statistically robust than in the previous linear model specifications.

To this point the specifications of the size-growth relationship models for surviving plants, the first with size terms only, the second with size and age terms as explanatory variables and also the specification with additional quadratic size and age terms, estimated on the aggregate samples across the economic cycle time periods have not included industry dummy variables. Re-estimating each of these specifications for all four time periods with the inclusion of 2-digit industry-level dummy variables did not significantly change the sign, magnitude or significance of the estimated β_1, γ_1, β_2 and γ_2 coefficients but did provide slightly increased reported R^2 (or adjusted R^2) values.

5.7. Sample Selection Models and Estimates

An important development in the econometric analysis of Gibrat's Law has been the recognition of possible sample selection bias in the estimation of the firm (or plant) size-growth relationship. As Hall (1987, p.593) notes, even if the initial sample of plants provides a representative sample of the full population of plants, over time slower growing plants are more likely to exit the sample, so that the sample may no longer be random (Kennedy, 1998, p.251). The standard approach to such potential sample selection bias has involved the construction of a model with a two non-independent equations structure: a probit selection equation reflecting the probability of plants surviving until the final period of examination and the growth equation as previously formulated. The sample selection model essentially provides an estimate of the expected growth rate (conditional on the other covariates) that a plant (or firm) would have attained had it survived (Cosh et al., 1996, p.22). Such sample selection models essentially imply that there are two separate relationships at work, the first being a specific relationship applying to plant growth, and an independent process determining plant survival (Cosh et al., 1996, p.22)[97].

Hall (1987) examining a generalised Tobit sample selection model comprising of size-growth and probit selection equations (see Hall, 1987, p.593, for the model formulation, assumptions and covariance matrix), demonstrates that estimating such a model on only those surviving plants will in fact be estimated as:

$$\begin{aligned} E(\Delta \ln X_{it} \mid \Delta \ln X_{it} \text{ observed}) &= \beta Z_i + E(\varepsilon_{it-1} \mid \Delta \ln X_{it} \text{ observed}) \qquad (5.7) \\ &= \beta Z_i + \rho \sigma_\varepsilon \lambda \end{aligned}$$

where Z_i are industry or firm characteristics, ρ is the correlation between the error terms of the two equations, σ_ε is the standard deviation of the error term ε, and λ is the inverse Mills' ratio generated from the residuals of the probit model. The inverse Mills' ratio is included as an additional explanatory variable in the size-growth equation (Lotti et al., 2003, p.5) as a means of (asymptotically) alleviating the potential sample selection bias (Kennedy, 1998, p.259). Correlated error terms in the size-growth and survival equations

---

[97] Such a relationship between firm size and survival has been recently questioned (see, for example, Cosh et al., 1996, p.22, who argue that for some firms there is an inseparable relationship between firm growth and survival). Wagner (1992) and Audretsch et al. (2002) have also pointed to difficulties in appropriately controlling for sample selection bias in this context.

would produce biased OLS estimates, upwards in the case of positive correlation and downwards in the case of negative correlation, with the bias of greater magnitude for small plants (or firms) if size influences the likelihood of survival, given that the λ will be larger in the case of smaller plants (Hall, 1987, p.593). Where there is no correlation between error terms, so that $\rho = 0$, the estimates will be unbiased.

Such two-equation sample selection models were constructed here. However, as Hall (1987, p.598) notes, even with comprehensive data it is generally difficult to find variables that would appear in a such a growth equation that would also not be relevant to survival. In such circumstances similar variables are likely to be available for inclusion in both the growth equation and selection equation of the sample selection model, a feature which may lead to an identification problem[98]. Unlike Hall (1987) who was able to incorporate some additional information on factors influencing the type of reported firm exit (mergers and acquisition, bankruptcy and liquidation), the addition of further useful information to alleviate this issue is not possible here given the available data. The absence of specific information on the nature of plant exits also means that an assumption of a relatively homogeneous exit process is required. There may therefore be some limitations in the interpretation of this element from the pooling of potentially differing processes (Harhoff et al., 1998, p.479).

Given this available data, following, for example, Dunne and Hughes (1994), Lotti and Santarelli (1999) and Lotti, Santarelli and Vivarelli (2003) the approach employed here is to specify linear growth models with non-linear (quadratic terms) in the selection equations only, providing the required additional term(s) in the selection equation which are assumed to be not correlated with size (and age) in the growth

---

[98] Hall (1987) entered firm size variables and higher-order expansions in both the growth and selection equations (see also Cosh et al, 1996) relying on the selectivity correlation, ρ, to permit identification. As Griliches et al. (1978), Maddala (1983, p.271) and Hall (1987) note, such expanded models with quadratic terms could induce collinearity between these terms and the inverse Mills' ratio, giving potentially misleading estimates of the selection effects. However, such an issue may be alleviated through the inclusion of additional variables in the selection equation (Hall, 1987, p.598).

equation. Specifically, the first growth equation component includes only a plant size term, with the other two versions including both size and age terms. Plant size and quadratic size terms were included in the respective probit selection equations for the first two sample selection model equations with age and a quadratic age variable included in the third specification[99].

Specifically, the initial sample selection model is given by:

$$P(d_i = 1) = F[\alpha_0 + \alpha_1 \ln X_{it-1} + \alpha_2 \ln X_{it-1}^2 + \mu_{it}] \qquad (5.8)$$

$d_i = 1$ survivor; $d_i = 0$ non-survivor

$$\Delta \ln X_{it} = \beta_0 + \beta_1 \ln X_{it-1} + \nu_{it}$$

The sample selection models with age terms in the growth equation, denoted as Without Age (in the selection equation), equation (5.9), and With Age (in the selection equation), equation (5.10), following the classification of Dunne and Hughes (1994) are given by:

$$P(d_i = 1) = F[\alpha_0 + \alpha_1 \ln X_{it-1} + \alpha_2 \ln X_{it-1}^2 + \mu_{it}] \qquad (5.9)$$

$d_i = 1$ survivor; $d_i = 0$ non-survivor

$$\Delta \ln X_{it} = \beta_0 + \beta_1 \ln X_{it-1} + \gamma_1 \ln A_{it} + \nu_{it}$$

---

[99] Probit models of the probability of plant failure were examined using specifications of the form: $P(d_i = 1) = F[\alpha + \beta \boldsymbol{X}_i + \mu_{it}]$, $d_i = 1$ (survivor), $d_i = 0$ (non-survivor), $\boldsymbol{X}_i$ being a vector of explanatory variables, including plant size and age, industry dummy variables (2-digit SIC level) and dummy variables for ownership and firm structure (consistent with the findings of, for example, Geroski and Gregg, 1996), industry MES and location in an assisted-area. Various model specifications were estimated for each of the economic cycle periods and the longer time-span 1972-1997 using maximum likelihood with robust standard errors. The results supported a generally positive relationship between plant size and survival (consistent with Doms et al. 1995, and Gort and Jensen, 2002, for firms), and also between plant age and survival, consistent with for example Doms et al. (1995) and Disney et al. (1999). The other variables exhibited some sensitivity to the model specification, although foreign ownership was in many cases found to be statistically significant and positively related to plant survival. Although industry dummy variables were predominantly not statistically significant the estimated results suggested the need for some reflection of potential industry effects. However, low pseudo-R^2 values were generally observed, with the explanatory variables added to size and age providing very little additional explanatory power to the model.

$$P(d_i = 1) = F[\alpha_0 + \alpha_1 \ln X_{it-1} + \alpha_2 \ln X_{it-1}^2 + \alpha_3 \ln A_{it} + \alpha_4 \ln A_{it}^2 + \mu_{it}] \qquad (5.10)$$

$d_i = 1$ survivor; $d_i = 0$ non-survivor

$$\Delta \ln X_{it} = \beta_0 + \beta_1 \ln X_{it-1} + \gamma_1 \ln A_{it} + \nu_{it}$$

Whilst it is accepted that the inclusion of these quadratic terms in the probit selection equation but not in the size-growth equation (in the case where it should be appropriately included as an additional explanatory variable) could result in a spurious acceptance of a sample selection problem (Verbeek, 2000, p.210), such an approach has been widely applied within this element of the literature, and in the absence of other information is pursued here.

Estimation of the above sample selection models can be conducted using either the full maximum likelihood (ML) estimator (discussed by Griliches et al., 1978), an approach which selects parameter values that maximises the probability of obtaining the observed sample of data (Kennedy, 1998, p.21), where the likelihood function can be suitably produced (Kennedy, 1998, p.252), or the two-step estimator (Heckman, 1976), both estimators being available in STATA (described as Heckman ML and Heckman two-step)[100].

Although the Heckman two-step estimator has been employed in some studies of Gibrat's Law, as Kennedy (1998, p.256) reflects, such an estimator is inefficient (but consistent) if the error terms are both of constant variance and normally distributed (Hsiao, 2003) and also produces some additional measurement error through the process of estimating the error term's expected value for inclusion in the second step. As Kennedy (1998) and Lotti et al. (2003) also observe the Heckman two-step estimator is also biased as well as inefficient in small samples. Full maximum likelihood estimation

---

[100] Although Cosh et al. (1996) employed semi-parametric sample selection models to examine Gibrat's Law, given the practical difficulties of implementing such semi-parametric sample selection models (Verbeek, 2000, p.220), a feature found by Cosh et al. (1996), this approach is not used here.

may therefore be preferred (Nawata, 1994; Nawata and Nagase, 1996) and is used here following studies such as Hall (1987), Dunne and Hughes (1994) and Lotti et al. (2003, p.6), given its greater efficiency (Hsiao, 2003), including for each 2-digit industry (for samples of at least 10 plants) and for the constructed size-classes. The results from the Heckman two-step estimator are also presented for the aggregate-level samples of each time period for comparison.

Again, the models were estimated with robust standard errors, being particularly important given the sensitivity of the probit equation to heteroscedasticity (Evans, 1987a, p.573). Given the use of robust standard errors likelihood-ratio tests are not appropriate, so a Wald test for the parameters in the size-growth component of the model being zero (except for the constant term) is presented with additional Wald χ^2 tests for whether $\rho = 0$, testing whether the two-equations are independent. A significance test of the inverse Mills' ratio can only be directly conducted using a standard z-test in the Heckman two-step procedure.

The results of the sample selection models are summarised in Tables 5.9a-d[101]. Although departing from standard econometric practice and the earlier presentation, due to the large size of the tables setting out sample selection model equation estimates, and the large number of these tables, the z-statistics (testing for significant differences from zero) are not presented for all of the variables. Such information has however been used in noting the statistical significance of the variables in the estimated equations, the 5 percent level of significance being identified with an asterix.

For the maximum likelihood estimates of equation (5.8) on the aggregate samples Gibrat's Law can be rejected for all four time periods with statistically significant (at the 5

---

[101] In these tables, some results are not presented due to either the log-pseudo-likelihood (LL in the tables) being non-concave, failure to converge or since the specification reduced to a model which could be estimated by OLS (the results being presented earlier). In some cases STATA did not produce Wald χ^2 values for the overall model, a feature not suggesting any deficiency in the model specification (StataCorp LP, 2005).

percent level) negative β_1 coefficients. Although the null hypothesis that $\rho = 0$ could be rejected in all cases[102], the estimated β_1 coefficients of -0.12, -0.48, -0.19 and -0.18 were generally similar to those produced by the OLS estimation of equation (5.2) where coefficients of -0.11, -0.41, -0.22 and -0.27, respectively, were reported. In three of the four cases the probit selection equation estimates provided evidence of a positive and statistically significant relationship between plant size, as shown by the α_1 coefficients, and survival consistent with expectations based on previous empirical evidence, though deviating from the strongest form of Gibrat's Law (the reverse being the case for the long 1972-1997 time period). Although, in the case where such an inconsistent relationship is detected the estimated plant growth equation would not be biased by such a relationship where ρ is zero (Hall, 1987, p.594), for this sample this is not the case, so that some caution may be required in interpreting this result.

Although the Heckman two-step estimator provided broadly similar results for the first two time periods, 1972-1977 and 1972-1997 (in terms of the β_1 estimates), both rejecting Gibrat's Law, in the case of the 1977-1986 and 1986-1997 periods the hypothesis that $\beta_1 = 0$ could not be rejected at the 5 percent level of significance, with a positive β_1 coefficient in the case of the latter. However, the probit selection equation provided a statistically significant negative α_1 coefficient (for plant size) for the 1972-1977 period, being also negative in the 1972-1997 estimate (but not statistically significant). For the 1977-1986 and the 1986-1997 periods the α_1 coefficient followed expectations, being positive and statistically significant. The inverse Mills' ratio was statistically significant in three of the models, with the exception being the final period

---

[102] A positive sign on the correlation coefficient indicates that slower growing plants tend to be more likely to exit (Evans, 1987a, pp.576-577). However, in a number of cases a negative correlation coefficient was observed (as also found in Dunne and Hughes, 1994; and Harhoff et al, 1998), a finding inconsistent with expectations. Harhoff et al. (1998, p.481) attributed such an observation to the under-representation of firms exhibiting above average employment growth.

Table 5.9a: Sample Selection Model: Plant Size and Growth, 1972-1977

Sample	All	Survivors	α_0	α_1	α_2	β_0	β_1	z_{β_1}	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
All: Maximum likelihood	2,044	1,754	2.81*	1.08*	-0.14*	0.58*	-0.12*	[-11.78]	0.59	-0.71	-0.42	-2,206.13	64.89*	138.78*
All: Heckman two-step	2,044	1,754	2.06*	-0.66*	0.10*	1.03*	-0.15*	[-4.95]	1.74	-1.00	-1.74*	-	-	47.05*
Size: Micro	217	192	3.25*	-2.30	0.59	1.14*	-0.33*	[-3.09]	0.81	-0.75	-0.61	-287.87	7.97*	9.54*
Size: Small	1,115	930	4.93*	-2.50*	0.38	0.40*	-0.09*	[-3.19]	0.55	-0.55	-0.30	-1,202.55	9.30*	10.17*
Size: Medium	482	411	-5.48	2.52	-0.24	0.45	-0.08	[-1.30]	0.56	-0.86	-0.48	-478.92	48.25*	1.69
Size: Large	230	221	27.75	-8.41	0.67	-0.08	-0.00	[-0.18]	0.41	-0.62	-0.26	-146.57	4.70*	0.03
Minimum efficient scale	191	181	11.69	-3.83	0.35	-0.28	0.02	[0.65]	0.51	-0.83	-0.42	-158.82	13.48*	0.43
Single-plant firm	846	700	3.40*	-1.52*	0.21*	0.82*	-0.20*	[-7.69]	0.66	-0.88	-0.58	-942.19	103.81*	59.18*
Foreign owned	201	190	3.00	-0.97	0.13	0.58*	-0.10*	[-3.89]	0.43	-0.01	-0.01	-148.65	0.02	15.17*

Table 5.9a: Sample Selection Model: Plant Size and Growth, 1972-1977, continued

Sample (SIC)	All	Survivors	α_0	α_1	α_2	β_0	β_1	z_{β_1}	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
15: Food and beverages	215	183	4.03*	-1.67*	0.21*	0.36*	-0.07*	[-2.09]	0.51	-0.85	-0.43	-194.24	27.65*	4.37*
16: Tobacco products	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..
17: Textiles	84	69	1.55	-0.55	0.09	0.81*	-0.17*	[-2.65]	0.53	-0.57	-0.30	-87.56	0.61	7.02*
18: Wearing apparel	116	90	0.52	-0.25	0.07	0.48*	-0.11*	[-2.55]	0.42	-0.07	-0.03	-107.76	0.48	6.52*
19: Leather products	29	25	2.44	-0.53	0.04	0.44*	-0.12*	[-2.78]	0.28	-0.76	-0.22	-12.70	0.47	7.70*
20: Wood and wood products	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
21: Pulp and paper products	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
22: Publishing and printing	111	101	4.51*	-1.96*	0.26*	0.50*	-0.13*	-2.50	0.51	-0.89	-0.45	-93.32	30.88*	6.23*
23: Coke, petroleum and nuclear fuel	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
24: Chemicals and chemical products	70	64	2.69	-1.20	0.20	0.71*	-0.14*	[-3.64]	0.49	-0.06	-0.03	-63.39	0.09	13.27*
25: Rubber and plastic products	87	77	0.42	0.22	0.00	0.54*	-0.10*	[-2.36]	0.64	0.04	0.03	-103.96	0.07	5.55*
26: Non-metallic mineral products	158	133	0.72	0.24	-0.04	0.30	-0.14*	[-2.50]	0.62	0.16	0.10	-192.10	1.40	6.27*
27: Basic metals	127	116	2.69*	-0.63	0.07	0.13	-0.03	[-1.68]	0.46	-0.31	-0.14	-109.23	1.62	2.83
28: Fabricated metal products	278	236	2.30*	-0.82	0.11	0.86*	-0.19*	[-5.43]	0.58	-0.67	-0.39	-301.26	4.59*	29.46*
29: Other machinery and equipment	238	197	3.77*	-1.62*	0.20*	0.70*	-0.13*	[-4.06]	0.63	-0.72	-0.46	-269.67	15.45*	16.44*
30: Office machinery and computers	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..
31: Electrical machinery	55	48	5.89*	-2.66*	0.34*	0.50*	-0.07*	[-1.99]	0.39	-0.48	-0.18	-39.82	2.88	3.98*
32: Radio, T.V. and communications	34	29	4.19	-2.49	0.35	0.05	-0.05	[-0.94]	0.56	0.94	0.52	-29.63	0.69	0.88
33: Medical and optical instruments	46	44	15.73*	-8.43*	1.19*	0.73*	-0.16*	[-2.92]	0.54	-0.88	-0.47	-38.84	2.63	8.54*
34: Motor vehicles and trailers	60	55	4.79	-2.79	0.49	0.49	-0.08	[-1.76]	0.54	-0.58	-0.32	-56.21	0.65	3.10
35: Other transport equipment	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
36: Furniture	115	94	4.49*	-2.54*	0.39*	0.97*	-0.20*	[-3.86]	0.83	-0.93	-0.77	-142.54	22.76*	14.93*

Notes: Model estimated using maximum likelihood with heteroscedasticity-consistent standard errors (STATA default standard errors applied for the Heckman two-step estimate).
z-statistics only reported for the Size (β_1) coefficient in the growth equation.
The statistical significance of λ (z-test reported) is only available for the Heckman two-step estimator. Two-step $\rho = -1.53$ truncated to -1.00.
.. Not reported given small sample < 10 plants
- Not reported by STATA.
* Statistically significant at the 5 percent level.

Table 5.9b: Sample Selection Model: Plant Size and Growth, 1972-1997

Sample	All	Survivors	α_0	α_1	α_2	β_0	β_1	z_{β_1}	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
All: Maximum likelihood	2,044	693	-0.33*	-0.24*	0.05*	2.55*	-0.48*	[-21.53]	0.92	-0.60	-0.55	-2,084.24	43.28*	463.58*
All: Heckman two-step	2,044	693	-0.51*	-0.14	0.04*	5.21*	-0.70*	-5.82	2.27	-1.00	-2.27*	-	-	36.81*
Size: Micro	217	58	-0.70*	0.23	-0.11	2.94*	-0.75*	[-5.33]	0.68	-0.16	-0.11	-185.36	0.54	28.43*
Size: Small	1,115	308	-0.12	-0.28	0.04	1.46*	-0.37*	[-4.85]	0.77	-0.09	-0.07	-1,011.48	0.27	23.54*
Size: Medium	482	200	-9.43	3.74	-0.38	2.21*	-0.36*	[-2.18]	1.02	-0.78	-0.80	-561.18	7.39*	4.75*
Size: Large	230	127	-4.66	1.03	-0.04	0.65	-0.27*	[-3.52]	0.77	0.35	0.27	-292.91	1.24	12.36*
Minimum efficient scale	191	92	2.04	-1.22	0.14	0.27	-0.24*	[-2.23]	0.80	0.56	0.45	-219.31	1.30	4.97*
Single-plant firm	846	266	0.36	-0.71*	0.13*	2.90*	-0.67*	[-17.41]	0.89	-0.65	-0.58	-813.24	42.10*	303.15
Foreign owned	201	126	0.23	-0.32	0.06	2.72*	-0.54*	[-4.79]	0.77	-0.31	-0.24	-264.26	0.24	22.96*

Table 5.9b: Sample Selection Model: Plant Size and Growth, 1972-1997, continued

Sample (SIC)	All	Survivors	α_0	α_1	α_2	β_0	β_1	$z\beta_1$	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
15: Food and beverages	215	76	0.41	-0.63*	0.10*	2.50*	-0.42*	[-7.09]	0.92	-0.73	-0.67	-218.92	5.47*	50.22*
16: Tobacco products	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..
17: Textiles	84	22	-1.97	0.75	-0.09	1.98	-0.41*	[-3.63]	0.73	-0.40	-0.29	-70.23	0.12	13.18*
18: Wearing apparel	116	24	-0.21	-0.37	0.05	1.45	-0.32*	[-3.55]	0.70	-0.10	-0.07	-84.13	0.03	12.60*
19: Leather products	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
20: Wood and wood products	116	24	-2.84*	1.36*	-0.21*	0.82	-0.41*	[-2.67]	0.63	0.60	0.38	-77.13	2.33	7.14*
21: Pulp and paper products	42	22	2.32	-1.53*	0.21*	3.95*	-0.75*	[-7.32]	0.75	-0.83	-0.62	-44.58	9.56*	53.58*
22: Publishing and printing	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
23: Coke, petroleum and nuclear fuel	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
24: Chemicals and chemical products	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
25: Rubber and plastic products	87	38	-2.14*	0.79	-0.06	3.08*	-0.69*	[-8.02]	0.74	0.15	0.11	-97.19	0.55	64.26*
26: Non-metallic mineral products	158	43	-2.92*	0.97*	-0.08	-0.18	-0.21	[-1.84]	0.86	0.82	0.71	-124.78	3.83	3.38
27: Basic metals	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
28: Fabricated metal products	278	88	-0.44	-0.13	0.03	2.31*	-0.52*	[-8.40]	0.88	-0.34	-0.30	-281.50	1.77	70.62*
29: Other machinery and equipment	238	73	-0.08	-0.47*	0.08*	2.96*	-0.60*	[-11.10]	0.92	-0.83	-0.77	-211.67	33.11*	123.20*
30: Office machinery and computers	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..
31: Electrical machinery	55	24	2.49	-1.64*	0.21*	2.27*	-0.45*	[-5.86]	0.87	0.08	0.07	-63.21	0.07	34.35*
32: Radio, T.V. and communications	34	15	-4.14*	1.54*	-0.13*	3.91*	-0.66*	[-4.17]	1.25	-1.00	-1.25	-34.02	3,207.89*	-
33: Medical and optical instruments	46	16	-1.39*	0.72*	-0.11*	0.15	-0.38	[-1.74]	1.68	1.00	1.68	-46.73	33,334.70*	-
34: Motor vehicles and trailers	60	31	-0.21	-0.18	0.05	1.36*	-0.31*	[-3.62]	0.86	-0.09	-0.08	-76.78	0.22	13.11*
35: Other transport equipment	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
36: Furniture	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-

Notes: Model estimated using maximum likelihood with heteroscedasticity-consistent standard errors (STATA default standard errors applied for the Heckman two-step estimate).

z-statistics only reported for the Size (β_1) coefficient in the growth equation.

The statistical significance of λ (z-test reported) is only available for the Heckman two-step estimator. Two-step $\rho = -1.11$ truncated to -1.00.

.. Not reported given small sample < 10 plants

- Not reported by STATA.

* Statistically significant at the 5 percent level.

Table 5.9c: Sample Selection Model: Plant Size and Growth, 1977-1986

Sample	All	Survivors	α_0	α_1	α_2	β_0	β_1	z_{β_1}	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
All: Maximum likelihood	2,812	1,444	-1.09*	0.49*	-0.03*	0.55*	-0.19*	[-16.19]	0.71	0.31	0.22	-3,310.05	20.64*	261.99*
All: Heckman two-step	2,812	1,444	-1.04*	0.45*	-0.03*	-0.68	-0.06	[-1.05]	1.24	1.00	1.24*	-	-	76.22*
Size: Micro	767	206	-1.01*	0.16	0.09	1.51*	-0.43*	[-2.54]	0.78	-0.19	-0.15	-673.42	0.12	6.46*
Size: Small	1,320	735	0.85	-0.43	0.06	0.52	-0.11*	[-2.27]	0.68	-0.32	-0.22	-1,640.13	0.10	5.13*
Size: Medium	504	333	-27.29*	11.39*	-1.16*	0.07	-0.08	[-0.94]	0.67	0.21	0.14	-655.65	3.05	0.89
Size: Large	-	-	-	-	-	-	-	[NC]	-	-	-	-	-	-
Minimum efficient scale	211	156	0.44	-0.16	0.03	0.63*	-0.17*	[-4.25]	0.58	0.04	0.02	-253.44	0.04	18.02*
Single-plant firm	987	590	1.56*	-0.92*	0.14*	1.66*	-0.40*	[-13.39]	0.82	-0.76	-0.62	-1,254.44	40.57*	179.32*
Foreign owned	251	203	1.86*	-0.62*	0.08*	1.71*	-0.32*	[-7.06]	0.87	-0.85	-0.74	-341.79	10.38*	49.84*

Table 5.9c: Sample Selection Model: Plant Size and Growth, 1977-1986, continued

Sample (SIC)	All	Survivors	α_0	α_1	α_2	β_0	β_1	$z\beta_1$	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
15: Food and beverages	270	143	-1.21*	0.63*	-0.06*	0.20	-0.12*	[-2.57]	0.78	0.78	0.61	-308.22	9.19*	6.61*
16: Tobacco products	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..
17: Textiles	88	43	-0.92	0.53	-0.07	0.40	-0.16	[-1.53]	0.73	0.08	0.06	-107.56	0.03	2.35
18: Wearing apparel	127	60	-2.14*	0.72	-0.05	1.05*	-0.21*	[-2.47]	0.56	-0.24	-0.14	-131.58	0.39	6.11*
19: Leather products	29	14	-1.92	0.89	-0.09	-0.41	0.03	[0.26]	0.52	0.31	0.16	-29.26	0.88	0.07
20: Wood and wood products	161	67	-2.33*	1.34*	-0.17*	0.74*	-0.28*	[-3.54]	0.50	0.24	0.12	-142.62	0.75	12.53*
21: Pulp and paper products	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
22: Publishing and printing	184	93	-1.61*	0.96*	-0.09*	-0.12	-0.06	[-0.64]	0.84	0.86	0.72	-195.36	6.22*	0.41
23: Coke, petroleum and nuclear fuel	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
24: Chemicals and chemical products	97	66	-1.49*	0.93*	-0.08*	0.73*	-0.26*	[-5.22]	0.80	0.35	0.28	-127.65	7.89*	27.27*
25: Rubber and plastic products	135	82	-2.22*	1.35*	-0.15*	0.49	-0.21*	[-3.87]	0.83	0.85	0.71	-158.88	5.66*	14.96*
26: Non-metallic mineral products	252	123	-0.14	-0.38	0.14*	0.89*	-0.23*	[-3.12]	0.74	-0.48	-0.36	-285.72	1.11	9.73*
27: Basic metals	148	81	-1.22*	0.48*	-0.03	0.55*	-0.20*	[-7.36]	0.63	0.25	0.16	-171.59	2.17	54.10*
28: Fabricated metal products	402	181	-1.32*	0.65*	-0.07*	0.59*	-0.30*	[-5.94]	0.77	0.69	0.53	-436.48	7.61*	35.25*
29: Other machinery and equipment	316	150	-1.26*	0.54*	-0.04	0.43*	-0.21*	[-6.32]	0.67	0.36	0.24	-348.09	4.54*	39.94*
30: Office machinery and computers	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..
31: Electrical machinery	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
32: Radio, T.V. and communications	51	33	-0.25	-0.08	0.05	2.40*	-0.45*	[-3.33]	0.90	-0.79	-0.71	-66.50	0.87	11.09*
33: Medical and optical instruments	73	43	-1.70*	1.05*	-0.11*	-0.05	-0.06	[-0.56]	0.94	0.93	0.88	-87.96	1.58	0.31
34: Motor vehicles and trailers	85	50	-2.03*	1.02*	-0.09*	0.30	-0.15*	[-3.34]	0.67	0.79	0.53	-89.38	7.68*	11.16*
35: Other transport equipment	52	26	-1.91*	0.77*	-0.05	-0.66	-0.00	[-0.01]	0.77	0.36	0.28	-58.75	1.47	0.00
36: Furniture	178	82	-1.59*	0.69*	-0.06	0.69*	-0.18*	[-2.93]	0.80	0.05	0.04	-209.93	0.04	8.56*

Notes: Model estimated using maximum likelihood with heteroscedasticity-consistent standard errors (STATA default standard errors applied for the Heckman two-step estimate).
z-statistics only reported for the Size (β_1) coefficient in the growth equation.
The statistical significance of λ (z-test reported) is only available for the Heckman two-step estimator. Two-step $\rho = 1.06$ truncated to 1.00.
.. Not reported given small sample < 10 plants
- Not reported by STATA.
* Statistically significant at the 5 percent level.

Table 5.9d: Sample Selection Model: Plant Size and Growth, 1986-1997

Sample	All	Survivors	α_0	α_1	α_2	β_0	β_1	z_{β_1}	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
All: Maximum likelihood	2,505	1,417	-1.20*	0.58*	-0.04*	0.52*	-0.18*	[-12.34]	0.79	0.82	0.65	-2,940.21	107.17*	152.30*
All: Heckman two-step	2,505	1,417	-0.79*	0.31*	-0.01	-4.62	0.45	[0.72]	4.96	1.00	4.96	-	-	13.17*
Size: Micro	383	134	-1.85*	1.60*	-0.40	1.59*	-0.77*	[-2.60]	0.95	0.90	0.85	-362.31	29.54*	6.74*
Size: *Small*	1,489	*816*	*-0.05*	*0.04*	0.00	1.22*	-0.28*	[-7.50]	0.60	-0.02	-0.01	-1,769.95	0.00	56.15*
Size: Medium	469	321	10.44	-4.25	0.45	1.31*	-0.21*	[-2.19]	0.70	-0.78	-0.55	-570.78	28.66*	4.80*
Size: Large	164	146	10.90	-3.09	0.24	0.53	-0.12	[-1.69]	0.59	-0.78	-0.46	-169.82	6.73*	2.86
Minimum efficient scale	243	187	-3.22	0.91	-0.03	0.45	-0.12*	[-2.22]	0.56	0.01	0.01	-274.60	0.00	4.92*
Single-plant firm	1,226	684	-0.90*	0.57*	-0.06*	0.97*	-0.37*	[-10.47]	0.79	0.89	0.70	-1,408.45	117.13*	109.64*
Foreign owned	336	253	-0.77	0.43	-0.02	1.67*	-0.32*	[-6.01]	0.66	0.28	0.18	-422.23	0.30	36.11*

Table 5.9d: Sample Selection Model: Plant Size and Growth, 1986-1997, continued

Sample (SIC)	All	Survivors	α_0	α_1	α_2	β_0	β_1	z_{β_1}	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
15: Food and beverages	220	125	0.08	-0.25	0.07	1.94*	-0.38*	[-6.06]	0.74	-0.41	-0.30	-275.75	2.09	36.72*
16: Tobacco products	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..
17: Textiles	83	48	-2.14*	1.30*	-0.16	1.32*	-0.33*	[-3.81]	0.67	0.51	0.34	-97.49	2.00	14.55*
18: Wearing apparel	115	46	1.52	-1.26	0.19*	2.26*	-0.35*	[-4.24]	0.78	-0.88	-0.69	-109.79	17.59*	17.94*
19: Leather products	18	11	-7.72	3.78	-0.39	-0.48	-0.01	[-0.05]	0.58	0.71	0.41	-15.06	0.32	0.00
20: Wood and wood products	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
21: Pulp and paper products	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
22: Publishing and printing	138	72	-3.02*	1.71*	-0.21*	0.20	-0.14	[-1.71]	0.81	0.96	0.78	-142.43	38.36*	2.91
23: Coke, petroleum and nuclear fuel	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..
24: Chemicals and chemical products	126	88	-1.09	0.61	-0.03	1.82*	-0.35*	[-5.27]	0.59	-0.41	-0.24	-143.75	0.92	27.75*
25: Rubber and plastic products	197	104	-0.32	-0.05	0.05	1.60*	-0.35*	[-8.81]	0.54	-0.01	-0.00	-214.43	0.01	77.61*
26: Non-metallic mineral products	170	72	-1.96*	0.80*	-0.05	1.06	-0.34*	[-2.55]	0.70	0.60	0.42	-164.89	0.53	6.49*
27: Basic metals	119	83	-1.74*	0.93*	-0.08	0.69*	-0.17*	[-4.38]	0.58	0.65	0.38	-129.44	2.29	19.20*
28: Fabricated metal products	310	190	-1.30*	0.77*	-0.07	0.48*	-0.17*	[-3.66]	0.72	0.88	0.64	-351.20	33.90*	13.43*
29: Other machinery and equipment	239	141	-1.52*	0.86*	-0.08*	0.59*	-0.26*	[-4.84]	0.80	0.98	0.78	-258.94	3.37	23.38*
30: Office machinery and computers	31	15	3.35	-1.84	0.23	0.43	-0.08	[-0.91]	0.99	0.51	0.50	-40.12	0.35	0.83
31: Electrical machinery	97	57	0.81	-0.56	0.10	2.19*	-0.37*	[-2.19]	0.90	-0.67	-0.61	-129.95	0.73	4.79*
32: Radio, T.V. and communications	67	46	4.61*	-2.58*	0.36*	1.75*	-0.28*	[-3.85]	0.83	-0.78	-0.64	-84.80	4.00*	14.84*
33: Medical and optical instruments	76	46	-1.62	0.97	-0.11	0.72	-0.20*	[-2.25]	0.80	0.72	0.57	-96.40	2.87	5.05*
34: Motor vehicles and trailers	78	45	-3.51*	1.47*	-0.11	0.89*	-0.15*	[-2.72]	0.56	-0.36	-0.20	-74.09	1.02	7.42*
35: Other transport equipment	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-
36: Furniture	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-

Notes: Model estimated using maximum likelihood with heteroscedasticity-consistent standard errors (STATA default standard errors applied for the Heckman two-step estimate).
z-statistics only reported for the Size (β_1) coefficient in the growth equation.
The statistical significance of λ (z-test reported) is only available for the Heckman two-step estimator. Two-step $\rho = 1.31$ truncated to 1.00.
.. Not reported given small sample < 10 plants
- Not reported by STATA.
* Statistically significant at the 5 percent level.

1986-1997.

For plant size-classes, in the 1972-1977 period Gibrat's Law could be rejected for the Micro and Small plant size-classes, but not in the case of the Medium-sized and Large plants, supporting Gibrat's Law for larger plant sizes. In all four cases the null hypothesis that $\rho = 0$ could be rejected with the β_1 coefficient being somewhat less negative in the case of the Micro plants than provided under OLS estimation. However, in the case of the period 1972-1997 Gibrat's Law was again rejected for all size-classes. The null hypothesis that $\rho = 0$ was statistically rejected only in the case of the Medium size-class.

For the 1977-1986 period there was again some supportive evidence of Gibrat's Law being a more reasonable description for plants of larger size, with the sample selection models rejecting Gibrat's Law for the two smallest size-classes, but not in the case of Medium-sized plants. For the Largest size-class estimates could not be generated, so there remains the possibility of non-linearities in the size-growth relationship across size-classes as observed in the corresponding OLS estimates. Again, for the three reported size-classes, the null hypothesis that $\rho = 0$ could not be rejected.

The estimates for 1986-1997 again provided evidence of Gibrat's Law being more appropriate for larger plants, with Gibrat's Law rejected for the Micro-, Small- and Medium-sized plant size-classes, but not in the case of the Largest size-class. A notable upwards effect on the β_1 coefficient in the Micro size-class was detected (-0.77 compared to -1.27 under OLS). The ρ term was significant in three of the samples.

At the industry-level $\rho = 0$ could be rejected at the 5 percent level of significance in 5 samples in the period 1972-1977 (Food and beverages, Publishing and printing, Fabricated metal products, Other machinery and equipment and Furniture), 5 in the 1972-1997 period (Food and beverages, Pulp and paper products, Other machinery and equipment, Radio, television and communications and Medical and optical instruments), 7 for 1977-1986 (Food and beverages, Publishing and printing, Chemical and chemical products, Rubber and plastic products, Fabricated metal products, Other machinery and equipment and Motor vehicles and trailers) and 4 in 1986-1997 (Wearing apparel, Publishing and printing, Fabricated metal products and Radio, television and communications industries). However, it should be noted that in some cases using this estimation procedure, particularly for 1972-1997, a number of sample estimates were not produced (some of which had no sample exit and so reduced to the OLS estimated model). For a number of the disaggregated industry samples (as well as the size, ownership and structure classes) the probit selection equation did not consistently provide the expected positive and statistically significant α_1 coefficient estimates for plant size. The coefficient, α_2, of the quadratic size term also provided somewhat mixed results across estimators, time periods and aggregations.

In the 1972-1977 time period, the sample selection model estimates rejected Gibrat's Law in 13 samples (out of the 16 reported) with all sample selection model estimates also being rejected under OLS. Both methodologies generally provided very similar β_1 coefficients. The first proposition of Gibrat's Law could only not be rejected for the Basic metals, Radio, television and communications, and Motor vehicles and trailers industries. In the probit selection equation both the $\ln X_{it-1}$ and $\ln X_{it-1}^2$ terms were statistically significant at the 5 percent level (being negative in the case of the significant α_1 coefficients but positive for the α_2 coefficients) in Food and beverages, Publishing and printing, Other machinery and equipment, Electrical machinery, Medical and optical instruments and Furniture. In the remaining industries neither α_1 nor α_2 was significant.

For the 1972-1997 period Gibrat's Law could be rejected in 11 of the 13 industries reported. In only two industry samples did the sample selection correction significantly effect the validity of Gibrat's Law, with the OLS estimates rejecting Gibrat's

Law for the Non-metallic mineral products (with a β_1 coefficient of -0.35) and the Medical and optical instruments industries (with a β_1 coefficient of -0.41), but not providing such a rejection in the sample selection estimates (with corresponding coefficients of -0.21 and -0.38, respectively). In the probit selection equations, both the α_1 and α_2 coefficients, the size and quadratic size terms, were statistically significant in 7 of the reported industry samples, with α_1 being individually significant in the Non-metallic mineral products industry. However, the relationship between plant size and survival for these industries was not consistent, being negative in the case of 4 industries and positive in 4 other industries, with similar inconsistency observed in respect of the statistically significant α_2 coefficients. In the remaining 5 industries (Textiles, Wearing apparel, Rubber and plastic products, Fabricated metal products and Motor vehicles and trailers) neither variable was statistically significant.

Again, for the 1977-1986 sample there were limited effects on the validity of Gibrat's Law from the implementation of the sample selection model across industries, with the only change in statistical significance being for the Publishing and printing industry where Gibrat's Law (rejected with a statistically significant β_1 coefficient of -0.27 using OLS regression) could no longer be rejected after the correction for sample selection bias (with a coefficient of -0.06). Gibrat's Law could be rejected at the 5 percent level of significance in 12 of the 17 industry samples reported, not being rejected in the Textiles, Leather products, Publishing and printing, Medical and optical instruments and Other transport equipment industries. The coefficient of the plant size variable in the probit selection equation, α_1, was significant in 12 industries, with the coefficient of the quadratic size variable, α_1, significant in 9 industries (both significant in 8 samples). In all statistically significant cases, the α_1 coefficient was positive (consistent with expectations), with the α_2 coefficients in all but one case being negatively signed.

Finally, for the 1986-1997 time period, for those industries with sample selection estimates, only two industries were again observed to change their results in terms of the validity of Gibrat's Law, with Gibrat's Law being not rejected in the sample selection estimates relating to the Publishing and printing industry (with a β_1 coefficient of -0.14) but being rejected under OLS regression estimates (with a statistically significant β_1 coefficient of -0.21), with the Motor vehicles and trailers industry exhibiting the opposite result, being rejected in the sample selection model (with a β_1 coefficient of -0.15 compared to -0.10 in the OLS regression). Gibrat's Law could again be rejected in 13 of the 16 industries with reported results (not being rejected in the Leather products, Publishing and printing and Office machinery and computers industries). The coefficient of the plant size variable, α_1, in the probit selection equation was significant in 8 industries, with the α_2 coefficient significant in 4 industries (both significant in 3 samples). Again, consistent with expectations, the statistically significant α_1 coefficients in the probit equation were in all but one case positively signed, with mixed results for the α_2 coefficients.

The results for the samples of single-plant firms, foreign owned plants and plants above the MES were highly consistent with the OLS estimated linear regression model with in all cases the β_1 coefficients being of the same sign and very similar magnitude. In all four samples of single-plant firms the hypothesis that $\rho = 0$ could be rejected at the 5 percent level of significance using the Wald test, with similar rejections only in the 1977-1986 sample for foreign owned plants, and the 1972-1977 period for the MES samples.

Turning to the sample selection model with age in the growth equation but not in the selection equation, with results reported in Tables 5.10a-d. The probit selection equation for the 1972-1977 time period under both the maximum likelihood and Heckman two-step estimators provided statistically significant negative α_1 coefficients for the size variable: inconsistent with prior expectations, with corresponding negative coefficients also reported for the 1972-1997 period (although not statistically significant in the case of the Heckman two-step estimator). However, in both the 1977-1986 and the 1986-1997 periods both estimators provided α_1 coefficients which were positive and statistically significant. The coefficients of the quadratic size term, α_2, were statistically significant and positive in the two time periods commencing in 1972, with the opposite results emerging for the latter two time periods (but not statistically significant for the Heckman two-step estimate in 1986-1997).

Similar results to those of the previous model specification emerged with Gibrat's Law rejected for the four aggregate samples employing full maximum likelihood estimation. Although the null hypothesis that $\rho = 0$ could be rejected at the 5 percent level of significance, the estimated sample selection models did not provide highly notable changes in the reported coefficient estimates compared to the OLS results. A similar pattern to that of the initial sample selection model was also observed for the Heckman two-step estimator with no notable changes for the time periods commencing in 1972, but with no rejection of Gibrat's Law for either the 1977-1986 or 1986-1997 time periods. The inverse Mills' ratio was statistically significant in two of the time periods (1972-1977 and 1977-1986), but only affected the significance of the β_1 coefficient in the latter. Under both estimators the γ_1 age coefficient was negative and statistically significant, with highly similar coefficients across both time periods and estimators.

The size-class estimates also supported some non-linearity in the size-growth relationship with Gibrat's Law being rejected for the smallest two size-classes, but not for the largest plants in the period 1972-1977. A similar result also emerged for the 1986-1997 sample. Gibrat's Law was rejected for all size-classes in the 1972-1997 sample but only for the Largest size-class in 1977-1986. The γ_1 plant age coefficient was always

Table 5.10a: Sample Selection Model: Plant Size, Age and Growth, Without Age, 1972-1977

Sample	All	Survivors	α_0	α_1	α_2	β_0	β_1	$z\beta_1$	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
All: Maximum likelihood	2,044	1,754	2.82*	-1.08*	0.14*	0.70*	-0.09*	[-9.26]	-0.11*	0.59	-0.75	-0.44	-2,176.74	82.30*	190.15*
All: Heckman two-step	2,044	1,754	2.06*	-0.66*	0.10*	1.10*	-0.12*	[-4.19]	-0.10*	1.66	-1.00	-1.66*	-	-	55.88*
Size: Micro	217	192	3.46*	-2.71	0.75	1.24*	-0.24*	[-2.19]	-0.22*	0.77	-0.73	-0.56	-279.64	15.95*	27.08*
Size: Small	1,115	930	4.78*	-2.41*	0.37*	0.57*	-0.08*	[-2.92]	-0.11*	0.55	-0.62	-0.34	-1,186.60	17.11*	41.09*
Size: Medium	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
Size: Large	230	221	25.78	-7.76	0.62	-0.01	-0.00	[-0.17]	-0.03	0.41	-0.67	-0.28	-146.23	7.95*	0.55
Minimum efficient scale	191	181	11.91	-3.92	0.36	-0.25	0.02	[0.74]	-0.02	0.51	-0.84	-0.43	-158.72	22.53*	0.71
Single-plant firm	846	700	3.34*	-1.49*	0.20*	0.92*	-0.16*	[-6.08]	-0.13*	0.64	-0.89	-0.57	-926.11	102.14*	91.54*
Foreign owned	201	190	3.05	-0.99	0.14	0.70*	-0.07*	[-2.94]	-0.12*	0.42	-0.06	-0.02	-142.54	0.17	21.92*

Table 5.10a: Sample Selection Model: Plant Size, Age and Growth, Without Age, 1972-1977, continued

Sample (SIC)	All	Survivors	α_0	α_1	α_2	β_0	β_1	z_{β_1}	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
15: Food and beverages	215	183	4.03*	-1.67*	0.21*	0.38*	-0.07*	[-2.07]	-0.01	0.51	-0.84	-0.43	-194.22	25.77*	4.45
16: Tobacco products	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
17: Textiles	84	69	1.56	-0.55	0.09	0.81*	-0.17*	[-2.74]	-0.00	0.53	-0.57	-0.30	-87.55	0.54	7.52*
18: Wearing apparel	116	90	0.52	-0.26	0.07	0.55*	-0.08*	[-2.10]	-0.11*	0.41	-0.09	-0.04	-105.54	0.57	10.31*
19: Leather products	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
20: Wood and wood products	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
21: Pulp and paper products	-	-	-	-	-	-	-	[-	-	-	-	-	-	-	-
22: Publishing and printing	111	101	4.54*	-1.98*	0.27*	0.54*	-0.12*	[-2.37]	-0.03	0.51	-0.91	-0.46	-93.18	19.88*	6.22*
23: Coke, petroleum and nuclear fuel	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
24: Chemicals and chemical products	70	64	2.68	-1.19	0.20	0.73*	-0.13*	[-3.30]	-0.03	0.49	-0.05	-0.03	-63.24	0.07	13.79*
25: Rubber and plastic products	87	77	0.42	0.22	0.00	0.62*	-0.04	[-0.92]	-0.20*	0.61	0.04	0.03	-101.12	0.09	10.23*
26: Non-metallic mineral products	158	133	0.72	0.24	-0.04	0.52*	-0.12*	[-2.09]	-0.16*	0.60	0.15	0.09	-189.09	1.41	14.66*
27: Basic metals	127	116	2.60*	-0.59	0.06	0.20	-0.02	[-1.32]	-0.07	0.45	-0.29	-0.13	-108.33	1.26	3.46
28: Fabricated metal products	278	236	2.55*	-0.98*	0.14*	1.00*	-0.13*	[-3.81]	-0.19*	0.58	-0.80	-0.46	-288.68	11.58*	48.09*
29: Other machinery and equipment	238	197	3.74*	-1.60*	0.20*	0.76*	-0.11*	[-3.27]	-0.07	0.63	-0.73	-0.46	-268.49	18.67*	18.00*
30: Office machinery and computers	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
31: Electrical machinery	55	48	5.69*	-2.57*	0.32*	0.66*	-0.05	[-1.50]	-0.13	0.37	-0.50	-0.19	-37.93	3.74	6.13*
32: Radio, T.V. and communications	34	29	5.29	-2.96	0.39	0.08	0.02	[0.31]	-0.19	0.55	1.00	0.55	-27.80	2,815.55*	-
33: Medical and optical instruments	46	44	14.90*	-7.90*	1.11*	0.78*	-0.13*	[-2.05]	-0.09	0.53	-0.87	-0.46	-38.28	3.45	10.15*
34: Motor vehicles and trailers	60	55	3.64	-2.16	0.41	0.63*	-0.00	[-0.09]	-0.26*	0.47	0.03	0.02	-50.46	0.02	16.64
35: Other transport equipment	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
36: Furniture	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-

Notes: Model estimated using maximum likelihood with heteroscedasticity-consistent standard errors (STATA default standard errors applied for the Heckman two-step estimate).
z-statistics only reported for the Size (β_1) coefficient in the growth equation.
The statistical significance of λ (z-test reported) is only available for the Heckman two-step estimator. Two-step ρ = -1.52 truncated to -1.00.
.. Not reported given small sample < 10 plants.
- Not reported by STATA.
* Statistically significant at the 5 percent level.

Table 5.10b: Sample Selection Model: Plant Size, Age and Growth, Without Age, 1972-1997

Sample	All	Survivors	α_0	α_1	α_2	β_0	β_1	z_{β_1}	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
All: Maximum likelihood	2,044	693	-0.34*	-0.23*	0.05*	2.75*	-0.45*	[-19.36]	-0.19*	0.90	-0.61	-0.54	-2,067.05	34.33*	493.12*
All: Heckman two-step	2,044	693	-0.51*	-0.14	0.04*	5.12*	-0.64*	[-5.77]	-0.18*	2.08	-1.00	-2.08	-	-	57.58*
Size: Micro	217	58	-0.70*	0.23	-0.10	2.96*	-0.75*	[-5.52]	-0.04	0.68	-0.13	-0.09	-185.26	0.48	33.20*
Size: Small	1,115	308	-0.14	-0.26	0.04	1.74*	-0.35*	[-4.82]	-0.21*	0.74	-0.07	-0.05	-1,001.65	0.22	38.66*
Size: Medium	482	200	-8.35	3.29	-0.33	2.52*	-0.35*	[-2.11]	-0.16*	1.03	-0.80	-0.82	-557.57	8.52*	14.50*
Size: Large	230	127	-4.68	1.04	-0.04	0.85	-0.25*	[-2.95]	-0.16	0.77	0.44	0.34	-291.00	1.02	13.29*
Minimum efficient scale	191	92	2.13	-1.25	0.14	0.70	-0.23*	[-1.98]	-0.20	0.77	0.52	0.40	-217.47	0.87	8.89*
Single-plant firm	846	266	0.37	-0.72*	0.13*	3.03*	-0.62*	[-14.75]	-0.13*	0.90	-0.68	-0.61	-809.57	43.57*	310.31*
Foreign owned	201	126	0.40	-0.39	0.07	3.06*	-0.50*	[-5.43]	-0.21*	0.78	-0.52	-0.40	-259.74	1.14	51.41*

Table 5.10b: Sample Selection Model: Plant Size, Age and Growth, Without Age, 1972-1997, continued

Sample (SIC)	All	Survivors	α_0	α_1	α_2	β_0	β_1	$z\beta_1$	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
15: Food and beverages	215	76	0.42*	-0.63*	0.10	2.61*	-0.42*	[-7.15]	-0.06	0.92	-0.72	-0.66	-218.74	6.28*	51.59*
16: Tobacco products	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
17: Textiles	84	22	-1.98*	0.75	-0.09	2.28*	-0.37*	[-3.46]	-0.22	0.72	-0.46	-0.33	-69.39	0.19	24.45*
18: Wearing apparel	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
19: Leather products	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
20: Wood and wood products	116	24	-2.75*	1.30*	-0.20	1.24*	-0.16*	[-2.74]	-0.38	0.57	0.42	0.24	-76.75	1.09	8.62*
21: Pulp and paper products	42	22	2.30*	-1.52*	0.21	3.85	-0.66*	[-5.55]	-0.16	0.73	-0.81	-0.59	-44.16	8.39*	55.18*
22: Publishing and printing	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
23: Coke, petroleum and nuclear fuel	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
24: Chemicals and chemical products	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
25: Rubber and plastic products	87	38	-2.18*	0.81	-0.07	3.27*	-0.61*	[-6.69]	-0.32*	0.69	0.26	0.18	-94.05	1.07	112.42*
26: Non-metallic mineral products	158	43	-2.89*	0.95	-0.08	0.37	-0.23*	[-2.49]	-0.14	0.77	0.73	0.56	-124.09	2.05	6.91*
27: Basic metals	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
28: Fabricated metal products	278	88	-0.48	-0.10	0.03	2.37*	-0.45*	[-7.14]	-0.24*	0.82	-0.21	-0.17	-278.05	0.72	83.04*
29: Other machinery and equipment	238	73	-0.10	-0.45*	0.08*	3.00*	-0.55*	[-7.89]	-0.16	0.91	-0.84	-0.76	-210.13	29.26*	132.15*
30: Office machinery and computers	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
31: Electrical machinery	55	24	2.49	-1.64*	0.21*	2.51*	-0.43*	[-6.82]	-0.17	0.86	0.10	0.09	-62.83	0.10	51.70*
32: Radio, T.V. and communications	34	15	-3.91*	1.42*	-0.12*	3.84*	-0.83*	[-2.78]	0.40	1.24	-1.00	-1.24	-33.72	14,944.17*	-
33: Medical and optical instruments	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
34: Motor vehicles and trailers	60	31	-0.21	-0.18	0.05	1.75*	-0.26*	[-2.97]	-0.30*	0.81	-0.09	-0.08	-75.03	0.13	18.14*
35: Other transport equipment	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
36: Furniture	115	33	0.39	-0.86*	0.14*	4.02*	-0.58*	[-4.25]	0.06*	1.57	-1.00	-1.57	-97.81	39,992.98*	-

Notes: Model estimated using maximum likelihood with heteroscedasticity-consistent standard errors (STATA default standard errors applied for the Heckman two-step estimate).
z-statistics only reported for the Size (β_1) coefficient in the growth equation.
The statistical significance of λ (z-test reported) is only available for the Heckman two-step estimator. Two-step ρ = -1.10 truncated to -1.00.
.. Not reported given small sample < 10 plants.
- Not reported by STATA.
* Statistically significant at the 5 percent level.

Table 5.10c: Sample Selection Model: Plant Size, Age and Growth, Without Age, 1977-1986

Sample	All	Survivors	α_0	α_1	α_2	β_0	β_1	z_{β_1}	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
All: Maximum likelihood	2,812	1,444	-1.07*	0.48*	-0.03*	0.84*	-0.16*	[-13.44]	-0.18*	0.68	0.25	0.17	-3,273.16	18.70*	359.63*
All: Heckman two-step	2,812	1,444	-1.04*	0.45*	-0.03*	-0.18	-0.05	[-1.03]	-0.17*	1.02	0.98	1.01*	-	-	148.37*
Size: Micro	767	206	-1.00*	0.15	0.09	1.61	-0.40	[-1.95]	-0.08	0.78	-0.22	-0.17	-672.32	0.10	8.15*
Size: Small	1,320	735	0.99	-0.51	0.07	0.97*	-0.07	[-1.54]	-0.20*	0.72	-0.59	-0.43	-1,614.42	5.63*	51.77*
Size: Medium	504	333	-26.82*	11.20*	-1.14*	0.37	-0.05	[-0.61]	-0.18*	0.66	0.17	0.11	-648.50	2.25	19.65*
Size: Large	221	170	3.26	-0.84	0.07	1.52*	-0.21*	[-3.19]	-0.16	0.83	-0.94	-0.78	-277.61	20.88*	17.67*
Minimum efficient scale	211	156	0.43	-0.16	0.03	0.89*	-0.16*	[-3.84]	-0.13*	0.57	0.04	0.02	-251.92	0.03	20.69*
Single-plant firm	987	590	1.55*	-0.92*	0.14*	1.86*	-0.34*	[-10.59]	-0.19*	0.81	-0.78	-0.63	-1,236.19	35.08*	233.75*
Foreign owned	251	203	1.88*	-0.64*	0.08*	2.02*	-0.21*	[-5.61]	-0.37*	0.81	-0.89	-0.72	-321.67	18.09*	82.85*

Table 5.10c: Sample Selection Model: Plant Size, Age and Growth, Without Age, 1977-1986, continued

Sample (SIC)	All	Survivors	α_0	α_1	α_2	β_0	β_1	$z\beta_1$	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
15: Food and beverages	270	143	-1.18*	0.61*	-0.05*	0.55*	-0.10*	[-2.16]	-0.18*	0.75	0.76	0.57	-303.99	7.94*	11.50*
16: Tobacco products	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
17: Textiles	88	43	-0.92	0.53	-0.07	0.74	-0.11	[-1.01]	-0.22	0.71	0.00	0.00	-106.24	0.00	5.60
18: Wearing apparel	127	60	-2.15*	0.73	-0.05	1.30	-0.16	[-1.32]	-0.18	0.55	-0.34	-0.18	-128.78	0.17	6.00*
19: Leather products	29	14	-1.90	0.88	-0.09	-0.47	0.02	[0.19]	0.04	0.52	0.30	0.15	-29.25	0.84	0.07
20: Wood and wood products	161	67	-2.33*	1.33*	-0.17*	0.81*	-0.27*	[-3.39]	-0.03	0.50	0.22	0.11	-142.55	0.54	13.88*
21: Pulp and paper products	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
22: Publishing and printing	184	93	-1.57*	0.86*	-0.07	0.94*	-0.08	[-1.34]	-0.37*	0.64	0.49	0.32	-188.60	3.94*	11.39*
23: Coke, petroleum and nuclear fuel	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
24: Chemicals and chemical products	97	66	-1.42*	0.89*	-0.08*	0.91*	-0.20*	[-4.62]	-0.20*	0.78	0.28	0.22	-126.28	6.61*	30.94*
25: Rubber and plastic products	135	82	-2.22*	1.34*	-0.15*	0.48	-0.21*	[-3.62]	0.02	0.84	0.86	0.72	-158.86	4.28*	15.27*
26: Non-metallic mineral products	252	123	-0.13	-0.39*	0.14*	0.96*	-0.29*	[-3.16]	0.11	0.79	-0.65	-0.51	-284.95	2.29	12.36
27: Basic metals	148	81	-1.17*	0.45*	-0.03	1.07*	-0.17*	[-6.70]	-0.27*	0.57	0.13	0.08	-165.56	0.78	58.21*
28: Fabricated metal products	402	181	-1.31*	0.64*	-0.07*	0.81*	-0.27*	[-5.17]	-0.14*	0.73	0.65	0.47	-433.17	6.10*	46.39*
29: Other machinery and equipment	316	150	-1.24*	0.53*	-0.04	0.70*	-0.13*	[-3.73]	-0.25*	0.63	0.33	0.21	-340.95	3.91*	65.79*
30: Office machinery and computers	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
31: Electrical machinery	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
32: Radio, T.V. and communications	51	33	-0.50	0.20*	-0.00*	2.81*	-0.36*	[-5.27]	-0.35*	1.01	-1.00	-1.01	-60.76	24,692.49*	-
33: Medical and optical instruments	73	43	-1.66*	0.99*	-0.10*	0.35	-0.06	[-0.56]	-0.22*	0.91	0.94	0.86	-86.00	9.86*	5.03
34: Motor vehicles and trailers	85	50	-1.90*	0.94*	-0.08	0.78*	-0.14*	[-3.11]	-0.18	0.60	0.66	0.40	-87.99	3.27	22.24*
35: Other transport equipment	52	26	-1.91*	0.77*	-0.05	-0.61	0.00	[0.04]	-0.03	0.78	0.36	0.28	-58.74	1.69	0.05
36: Furniture	178	82	-1.59*	0.69*	-0.06	1.00*	-0.11	[-1.87]	-0.28*	0.75	0.04	0.03	-205.11	0.03	22.80*

Notes: Model estimated using maximum likelihood with heteroscedasticity-consistent standard errors (STATA default standard errors applied for the Heckman two-step estimate).
z-statistics only reported for the Size (β_1) coefficient in the growth equation.
The statistical significance of λ (z-test reported) is only available for the Heckman two-step estimator.
.. Not reported given small sample < 10 plants.
- Not reported by STATA.
* Statistically significant at the 5 percent level.

Table 5.10d: Sample Selection Model: Plant Size, Age and Growth, Without Age, 1986-1997

Sample	All	Survivors	α_0	α_1	α_2	β_0	β_1	z_{β_1}	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
All: Maximum likelihood	2,505	1,417	-1.15*	0.54*	-0.04*	0.81*	-0.13*	[-9.45]	-0.18*	0.73	0.79	0.58	-2,873.60	59.51*	246.52*
All: Heckman two-step	2,505	1,417	-0.79*	0.31*	-0.01	-3.58	0.40	[0.76]	-0.18*	4.26	1.00	4.26	-	-	19.52*
Size: Micro	383	134	-1.72*	1.42*	-0.34	1.75*	-0.73*	[-2.73]	-0.16*	0.93	0.92	0.86	-355.03	47.49*	28.08*
Size: Small	1,489	816	-0.06	0.04	0.00	1.41*	-0.20*	[-5.63]	-0.21*	0.56	-0.02	-0.01	-1,715.95	0.00	163.16*
Size: Medium	469	321	10.01	-4.06	0.43	1.59*	-0.19*	[-2.00]	-0.14*	0.71	-0.82	-0.58	-565.03	38.19*	13.62*
Size: Large	164	146	9.75	-2.76	0.22	0.95*	-0.07	[-0.99]	-0.26*	0.56	-0.82	-0.46	-160.96	8.61*	26.35*
Minimum efficient scale	243	187	-3.22	0.90	-0.03	0.83*	-0.09*	[-2.16]	-0.19*	0.54	0.07	0.04	-268.08	0.26	23.65*
Single-plant firm	1,226	684	-0.91*	0.57*	-0.06*	1.14*	-0.30*	[-8.82]	-0.16*	0.74	0.87	0.64	-1,376.29	107.87*	175.65*
Foreign owned	336	253	-0.69	0.38	-0.01	1.85*	-0.26*	[-3.90]	-0.19*	0.63	0.18	0.11	-412.50	0.28	103.37*

Table 5.10d: Sample Selection Model: Plant Size, Age and Growth, Without Age, 1986-1997, continued

Sample (SIC)	All	Survivors	α_0	α_1	α_2	β_0	β_1	$z\beta_1$	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
15: Food and beverages	220	125	0.02	-0.21	0.07	2.54*	-0.32*	[-5.86]	-0.33*	0.68	-0.48	-0.33	-262.92	3.40	74.17*
16: Tobacco products	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
17: Textiles	83	48	-2.15*	1.29*	-0.15	1.37*	-0.25*	[-3.20]	-0.15*	0.64	0.48	0.31	-95.87	2.69	17.07*
18: Wearing apparel	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
19: Leather products	18	11	-14.48*	7.27*	-0.82*	-0.03	-0.01	[-0.05]	-0.17	0.65	1.00	0.65	-13.19	651.24*	-
20: Wood and wood products	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
21: Pulp and paper products	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
22: Publishing and printing	138	72	-2.99*	1.70*	-0.21*	0.43	-0.11	[-1.39]	-0.12*	0.77	0.95	0.73	-139.91	20.61*	9.21*
23: Coke, petroleum and nuclear fuel	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
24: Chemicals and chemical products	126	88	-1.12*	0.63	-0.03	1.88*	-0.28*	[-4.27]	-0.16*	0.58	-0.48	-0.28	-140.65	1.83	36.86*
25: Rubber and plastic products	197	104	-0.32	-0.05	0.05	1.60*	-0.34*	[-6.01]	-0.02	0.54	-0.00	-0.00	-214.40	0.00	79.08*
26: Non-metallic mineral products	170	72	-1.93*	0.78	-0.05	1.65*	-0.26*	[-2.74]	-0.33*	0.60	0.56	0.34	-154.60	0.54	31.15*
27: Basic metals	119	83	-1.59	0.84	-0.07	0.86*	-0.17*	[-4.68]	-0.06	0.56	0.54	0.30	-129.11	2.26	26.27*
28: Fabricated metal products	310	190	-1.36*	0.79*	-0.07*	0.55*	-0.13*	[-2.67]	-0.10*	0.73	0.92	0.67	-347.24	42.03*	22.51*
29: Other machinery and equipment	239	141	-1.60*	0.90*	-0.09*	1.01*	-0.22*	[-4.71]	-0.18*	0.68	0.91	0.62	-251.65	6.13*	41.37*
30: Office machinery and computers	31	15	3.37	-1.85	0.23	0.37	-0.08	[-0.94]	0.06	0.98	0.50	0.49	-40.10	0.51	1.00
31: Electrical machinery	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
32: Radio, T.V. and communications	67	46	4.15*	-2.31*	0.32*	1.83*	-0.15*	[-2.22]	-0.32*	0.71	-0.76	-0.54	-78.94	1.31	29.99*
33: Medical and optical instruments	76	46	-1.26	0.73	-0.07	1.56*	-0.16*	[-2.51]	-0.35*	0.61	0.25	0.16	-90.79	1.56	14.74*
34: Motor vehicles and trailers	78	45	-3.41*	1.42*	-0.10	1.48*	-0.09	[-1.60]	-0.32*	0.48	-0.32	-0.15	-67.48	0.59	27.90*
35: Other transport equipment	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
36: Furniture	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-

Notes: Model estimated using maximum likelihood with heteroscedasticity-consistent standard errors (STATA default standard errors applied for the Heckman two-step estimate).
z-statistics only reported for the Size (β_1) coefficient in the growth equation.
The statistical significance of λ (z-test reported) is only available for the Heckman two-step estimator. Two-step $\rho = 1.30$ truncated to 1.00.
.. Not reported given small sample < 10 plants.
- Not reported by STATA.
* Statistically significant at the 5 percent level.

negative but only significant in 10 of the reported estimates. Comparing the β_1 coefficients from this model specification to the OLS estimates in the model with age as an additional explanatory variable provided only three changes to the statistical significance of the estimated coefficients, with Gibrat's Law being rejected in the sample selection model for Medium-sized plants in the 1972-1997 period and Large plants in the 1977-1986 period, but being no longer rejected for Micro-sized plants in the 1977-1986 sample.

For the industry-level samples relating to 1972-1977, Gibrat's Law was rejected at the 5 percent level of significance in 9 of the 14 reported industries, in each case exhibiting mean reversion, whilst not being rejected in the Rubber and plastic products, Basic metals, Electrical machinery, Radio, television and communications and Motor vehicles and trailers industries. Only the Wearing apparel industry exhibited a change in the reported validity of Gibrat's Law being rejected in the case of the sample selection model estimates. In no cases did the γ_1 coefficient change reported statistical significance. These results emerged under the rejection of the hypothesis that $\rho = 0$ in 5 samples, with OLS and sample selection estimates of the coefficients β_1 and γ_1 in the growth equation being generally very similar. The α_1 coefficients in the probit selection equation exhibited mixed signs, although in all 6 statistically significant results, the coefficient was negative, contrary to expectations. The coefficient of the quadratic size term in the probit equation was also statistically significant in 6 industry samples, in each of these being positively signed.

Again, for the 1972-1997 time period, the effects of the sample selection correction only led to a change of significance of the γ_1 coefficient, for plant age, in the growth equation for one industry, Furniture, with no changes in the reported validity of Gibrat's Law from the estimated β_1 coefficient. Gibrat's Law could be rejected in all 12 reported industry samples, with statistically significant negative β_1 coefficients. The selection equations however returned estimated coefficients not always consistent with prior expectations with size (α_1) being statistically significant and positive in two industries (Wood and wood products and Radio, television and communications) and statistically significant and negative in five other industries, the remainder being not significant.

Modest changes were observed when comparing the sample selection estimates with the OLS estimates for the period 1977-1986. In the sample selection model Gibrat's Law could be rejected in 10 industries out of 17, all with negative coefficients, with only the Publishing and printing industry not rejecting Gibrat's Law after corrections for sample selection bias having been so using OLS procedures. The γ_1 plant age coefficient only changed reported significance (becoming statistically significant) in the case of the Radio, television and communications and the Medical and optical instruments industries. The signs of the plant size term in the probit selection equation were generally consistent with prior expectations, being positive and significant in 13 cases (negative and significant in one).

Sample selection corrections also generally had only modest effects for the final panel covering 1986-1997 with no changes in the previously reported significance of the γ_1 coefficient of the plant age variable from the OLS estimates, and only two changes in the significance of the β_1 coefficients, the Publishing and printing industry no longer rejecting Gibrat's Law after the sample selection correction but the Radio, television and communications industry showing the reverse change. Gibrat's Law could be rejected for 10 of the 14 reported industries, all exhibiting mean reversion. Probit selection equations inconsistent with prior expectations were again in some cases observed.

The results of the estimates for the single-plant firms, foreign owned plants and plants corresponding to the MES were consistent with the OLS estimated linear regression model, equation (5.4), with no changes in the significance of the β_1 coefficients. The size of these estimated coefficients were highly consistent, with such corrections again only significant once each in both the foreign owned (1977-1986) and MES (1972-1977) samples. For all three plant classifications there was some evidence of a relationship between plant size and survival probabilities not always corresponding with prior expectations, with the size, α_1, coefficients being positive in all three cases in the time period 1986-1997, but all being negative in the other periods.

Looking now at the sample selection model specification with both plant size and age variables in the growth and selection equations (with the associated quadratic terms in the selection equation), for which the results are summarised in Table 5.11a-d. In the probit selection equation, at the aggregate-level, the estimates for the time periods 1972-1977 and 1972-1997 again showed statistically significant negative α_1 coefficients corresponding to plant size (not significant for the Heckman two-step estimate in 1972-1997), with positive, statistically significant, quadratic size variable coefficients. For the shorter of these time periods the coefficient of the plant age variable followed expectations (consistent with other empirical evidence) with a positive α_3 coefficient (but being significant only in the Heckman two-step estimator). However, for the longer time period these age coefficients were both statistically significant and negative (the coefficients of the quadratic age terms were positive and statistically significant).

In the case of the two periods 1977-1986 and 1986-1997, the probit estimates were consistent with expectations, the coefficients of the size and age terms being positive and statistically significant in the 1977-1986 period using both the maximum likelihood and Heckman two-step estimators, with consistently signed coefficients in the 1986-1997 period estimates (the coefficients of the size variable being positive and statistically significant and the coefficients of the age terms also being positive, although

Table 5.11a: Sample Selection Model: Plant Size, Age and Growth, With Age, 1972-1977

Sample	All	Survivors	α_0	α_1	α_2	α_3	α_4	β_0	β_1	z_{β_1}	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
All: Maximum likelihood	2,044	1,754	2.77*	-1.09*	0.14*	0.02	0.02	0.71*	-0.09*	[-8.89]	-0.13*	0.59	-0.74	-0.43	-2,174.33	71.22*	190.51*
All: Heckman two-step	2,044	1,754	1.95*	-0.73*	0.10*	0.25*	-0.04	1.08*	-0.10*	[-4.31]	-0.16*	1.43	-1.00	-1.43*	-	-	70.08*
Size: Micro	217	192	2.38*	-0.81	0.02	-0.53	0.30*	1.18*	-0.42*	[-3.72]	-0.16*	0.80	0.90	0.72	-275.04	24.98*	29.19*
Size: Small	1,115	930	4.73*	-2.51*	0.38*	0.18	-0.01	0.58*	-0.08*	[-2.94]	-0.13*	0.54	-0.56	-0.30	-1,182.92	6.84*	49.02*
Size: Medium	482	411	-2.89	1.57	-0.14	-0.30	0.06	0.49	-0.08	[-1.23]	-0.03	0.57	-0.87	-0.49	-476.77	54.86*	2.18
Size: Large	230	221	6.10*	1.16*	-0.09*	-7.11*	1.41*	-0.00	-0.01	[-0.39]	-0.01	0.42	-1.00	-0.42	-134.80	21,784.41*	-
Minimum efficient scale	191	181	15.90	-3.95	0.38	-3.47	0.63	-0.48	0.04	[1.24]	0.01	0.48	0.03	0.01	-159.73	0.00	1.62
Single-plant firm	846	700	3.28*	-1.56*	0.21*	0.14	0.02	0.96*	-0.15*	[-5.50]	-0.18*	0.64	-0.88	-0.57	-919.95	95.71*	99.63*
Foreign owned	201	190	2.93	-0.97	0.13	0.14	-0.05	0.69*	-0.07*	[-2.89]	-0.12*	0.42	-0.02	-0.01	-142.47	0.01	21.86*

Table 5.11a: Sample Selection Model: Plant Size, Age and Growth, With Age, 1972-1977, continued

Sample (SIC)	All	Survivors	α_0	α_1	α_2	α_3	α_4	β_0	β_1	z_{β_1}	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
15: Food and beverages	215	183	3.88*	-1.75*	0.22*	0.05	0.07	0.45*	-0.06	[-2.02]	-0.06*	0.51	-0.86	-0.44	-192.17	32.39*	5.37
16: Tobacco products	..	..	..	..	..	..	..	..	..		..	..	..	..	..	..	..
17: Textiles	84	69	1.51	-0.78	0.12	0.65	-0.15	0.79	-0.16*	[-2.61]	-0.01	0.52	-0.50	-0.26	-86.62	0.33	6.82*
18: Wearing apparel	116	90	0.05	-0.20	0.05	0.58	-0.11	-0.06	-0.09	[-0.64]	0.35	0.42	0.38	0.16	-104.49	0.06	1.89
19: Leather products	29	25	5.77	-1.83	0.25	-0.59	-0.10	0.46*	-0.11*	[-2.31]	-0.07	0.25	0.04	0.01	-10.44	0.00	12.37*
20: Wood and wood products	116	100	1.91	-1.17	0.18	0.44	0.07	1.04*	-0.16*	[-2.79]	-0.28*	0.60	-0.81	-0.49	-120.03	7.70*	18.81*
21: Pulp and paper products	42	34	4.09*	-2.38*	0.37*	0.48	-0.24	1.26	-0.26*	[-2.10]	0.04	0.63	-0.85	-0.54	-44.58	3.29	4.43
22: Publishing and printing	111	101	7.58*	-2.42*	0.36*	-2.45	0.48	0.41	-0.13*	[-2.60]	0.06	0.50	-0.95	-0.47	-87.54	24.92*	6.79*
23: Coke, petroleum and nuclear fuel	-	-	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
24: Chemicals and chemical products	70	64	2.92	-1.18	0.20	-0.43	0.11	0.73*	-0.13*	[-3.34]	-0.03	0.49	-0.04	-0.02	-63.03	0.04	14.12*
25: Rubber and plastic products	87	77	0.38	0.12	-0.01	-0.02	0.24	0.62*	-0.04	[-0.96]	-0.20*	0.61	0.03	0.02	-98.50	0.11	10.06*
26: Non-metallic mineral products	158	133	2.06*	-0.18	0.02	-0.99*	0.26*	0.68*	-0.10	[-1.73]	-0.17*	0.70	-0.94	-0.67	-179.43	41.55*	13.11*
27: Basic metals	127	116	3.54*	-0.30	0.04	-1.51	0.28	0.20	-0.02	[-1.45]	-0.05	0.47	-0.58	-0.27	-106.55	3.54	3.11
28: Fabricated metal products	278	236	2.57*	-0.90*	0.12*	-0.33	0.12	1.01*	-0.13*	[-3.73]	-0.20*	0.58	-0.82	-0.48	-287.37	16.57*	48.08*
29: Other machinery and equipment	238	197	3.74*	-1.59*	0.20*	-0.02	0.01	0.76*	-0.11*	[-3.24]	-0.07	0.63	-0.73	-0.46	-268.49	19.92*	17.78*
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
31: Electrical machinery	55	48	6.06*	-3.11*	0.38*	0.92	-0.21	0.65*	-0.05	[-1.43]	-0.13	0.37	-0.41	-0.15	-37.10	2.05	5.97
32: Radio, T.V. and communications	-	-	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
33: Medical and optical instruments	46	44	33.74	-6.89	0.92	-22.32	6.03	0.77*	-0.13*	[-2.10]	-0.09	0.52	-0.53	-0.27	-37.61	0.55	10.48*
34: Motor vehicles and trailers	60	55	17.44*	-14.37*	2.51*	3.85*	-0.77*	0.70*	-0.01	[-0.31]	-0.27*	0.48	-0.45	-0.21	-44.87	0.93	19.43*
35: Other transport equipment	-	-	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
36: Furniture	115	94	2.71*	-1.04*	0.13*	-0.13	0.02*	1.16*	-0.17*	[-2.80]	-0.14	0.86	-1.00	-0.86	-138.78	42,703.17*	-

Notes: Model estimated using maximum likelihood with heteroscedasticity-consistent standard errors (STATA default standard errors applied for the Heckman two-Step estimate).

z-statistics only reported for the Size (β_1) coefficient in the growth equation.

The statistical significance of λ (z-test reported) is only available for the Heckman two-step estimator. Two-step ρ = -1.46 truncated to -1.00.

.. Not reported given small sample < 10 plants.

- Not reported by STATA.

* Statistically significant at the 5 percent level.

Table 5.11b: Sample Selection Model: Plant Size, Age and Growth, With Age, 1972-1997

Sample	All	Survivors	α_0	α_1	α_2	α_3	α_4	β_0	β_1	$z\beta_1$	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
All: Maximum likelihood	2,044	693	-0.29	-0.22*	0.05*	-0.26*	0.11*	2.82*	-0.44*	[-20.07]	-0.24*	0.90	-0.62	-0.56	-2,053.68	44.41*	516.98*
All: Heckman two-step	2,044	693	-0.47*	-0.13	0.03*	-0.23*	0.10*	4.36*	-0.54*	[-10.18]	-0.31*	1.51	-1.00	-1.51*	-	-	122.56
Size: Micro	217	58	-0.69	0.22	-0.10	0.00	-0.01	2.98*	-0.75*	[-5.48]	-0.04	0.68	-0.15	-0.10	-185.25	0.13	32.45*
Size: Small	1,115	308	-0.20	-0.25	0.03	-0.23	0.12*	2.09*	-0.35*	[-4.60]	-0.25*	0.79	-0.39	-0.31	-991.55	7.62*	44.44*
Size: Medium	482	200	-7.47	3.10	-0.31	-0.64*	0.18*	2.25*	-0.32*	[-2.08]	-0.17*	0.93	-0.68	-0.63	-552.91	3.21	14.37*
Size: Large	230	127	-1.49	0.51	-0.01	-1.50*	0.34*	2.57*	-0.43*	[-3.65]	-0.17*	0.80	-0.57	-0.45	-287.82	1.63	16.42*
Minimum efficient scale	191	92	3.41	-1.27	0.14	-0.96	0.17	3.16*	-0.52*	[-4.84]	-0.14	0.82	-0.69	-0.56	-216.01	5.72*	30.07*
Single-plant firm	846	266	0.29	-0.68*	0.12*	-0.18	0.12*	3.02*	-0.60*	[-13.94]	-0.21*	0.87	-0.63	-0.55	-799.85	26.63*	335.46*
Foreign owned	201	126	0.97	-0.44	0.08	-0.53	0.10	2.88*	-0.49*	[-5.46]	-0.19*	0.75	-0.39	-0.29	-257.82	0.96	80.95*

Table 5.11b: Sample Selection Model: Plant Size, Age and Growth, With Age, 1972-1997, continued

Sample (SIC)	All	Survivors	α_0	α_1	α_2	α_3	α_4	β_0	β_1	$z\beta_1$	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
15: Food and beverages	215	76	0.69	-0.68*	0.11*	-0.50	0.18*	2.80*	-0.43*	[-7.07]	-0.13	0.93	-0.76	-0.71	-216.06	7.78*	52.71*
16: Tobacco products	..	..	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
17: Textiles	84	22	-2.78*	1.26*	-0.17*	-0.43	0.21*	3.79*	-0.33*	[-2.00]	-0.55*	1.24	-0.98	-1.21	-66.17	9.28*	18.58*
18: Wearing apparel	116	24	0.27	-0.86	0.10	0.62	-0.11	-0.16	-0.06	[-0.42]	-0.41	1.03	0.91	0.93	-79.89	14.83*	6.31*
19: Leather products	-	-	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
20: Wood and wood products	116	24	-3.02*	0.79	-0.14	0.98	-0.16	0.55	-0.41*	[-2.75]	0.03	0.72	0.78	0.57	-74.40	2.97	7.64*
21: Pulp and paper products	42	22	2.26	-1.58*	0.21*	0.10	0.04	3.85*	-0.62*	[-4.53]	-0.26	0.72	-0.80	-0.58	-43.67	9.31*	57.57*
22: Publishing and printing	111	43	0.13	-0.49	0.11	-0.32	0.11	3.47*	-0.52*	[-3.21]	-0.40	1.02	-0.89	-0.91	-114.75	8.00*	14.36*
23: Coke, petroleum and nuclear fuel	-	-	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
24: Chemicals and chemical products	70	35	-0.70	0.26	0.00	-0.42	0.08	2.80*	-0.60*	[-4.01]	-0.00	0.85	-0.26	-0.22	-87.84	0.13	24.24*
25: Rubber and plastic products	87	38	-2.31*	0.88	-0.08	-0.14	0.09	2.92*	-0.59*	[-5.08]	-0.28*	0.75	0.53	0.40	-93.42	0.91	44.63*
26: Non-metallic mineral products	158	43	-2.95*	0.92	-0.08	-0.01	0.04	0.48	-0.25*	[-2.70]	-0.10	0.73	0.66	0.48	-123.58	1.07	7.49*
27: Basic metals	127	52	-0.44	0.15	0.00	-1.08*	0.34*	2.33*	-0.35*	[-6.59]	-0.27*	0.71	-0.63	-0.45	-127.43	2.35	71.51*
28: Fabricated metal products	278	88	-0.43	-0.12	0.02	-0.19	0.09	2.45*	-0.45*	[-7.27]	-0.26*	0.83	-0.28	-0.23	-276.46	1.27	85.25*
29: Other machinery and equipment	238	73	-0.01	-0.42	0.08*	-0.27	0.08	2.98*	-0.55*	[-7.82]	-0.16	0.89	-0.83	-0.74	-209.45	26.40*	125.23*
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
31: Electrical machinery	55	24	2.21	-1.60*	0.21*	0.14	-0.00	2.48*	-0.43*	[-6.62]	-0.17	0.86	0.13	0.11	-62.63	0.14	47.54*
32: Radio, T.V. and communications	34	15	-4.50*	1.59*	-0.19*	0.26	0.25*	3.55*	-0.48*	[-3.45]	-0.43	0.94	-1.00	-0.94	-28.29	8,081.41*	-
33: Medical and optical instruments	46	16	-0.82	0.87	-0.10	-1.33*	0.30	1.51*	-0.34*	[-2.42]	-0.22	0.97	0.63	0.61	-46.95	3.50	9.83*
34: Motor vehicles and trailers	60	31	-0.55	-0.11	0.03	-0.09	0.12	1.57*	-0.24*	[-3.36]	-0.29*	0.81	0.05	0.04	-73.24	0.10	23.40*
35: Other transport equipment	-	-	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
36: Furniture	115	33	0.64	-0.85*	0.14*	-0.51*	0.15*	3.53*	-0.53*	[-4.25]	-0.03	1.28	-0.95	-1.22	-98.95	35.78*	21.08*

Notes: Model estimated using maximum likelihood with heteroscedasticity-consistent standard errors (STATA default standard errors applied for the Heckman two-Step estimate).
z-statistics only reported for the Size (β_1) coefficient in the growth equation.
The statistical significance of λ (z-test reported) is only available for the Heckman two-step estimator. Two-step ρ = -1.03 truncated to -1.00.
.. Not reported given small sample < 10 plants.
- Not reported by STATA.
* Statistically significant at the 5 percent level.

Table 5.11c: Sample Selection Model: Plant Size, Age and Growth, With Age, 1977-1986

Sample	All	Survivors	α_0	α_1	α_2	α_3	α_4	β_0	β_1	$z\beta_1$	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
All: Maximum likelihood	2,812	1,444	-1.17*	0.30*	-0.02*	0.32*	-0.02	0.85*	-0.17*	-13.60	-0.16*	0.68	0.21	0.14	-3,232.06	15.85*	362.10*
All: Heckman two-step	2,812	1,444	-1.13*	0.28*	-0.02*	0.30*	-0.01	-0.90	-0.06	-1.06	0.06	1.43	1.00	1.43*	-	-	55.92*
Size: Micro	767	206	-1.28*	0.12	0.03	0.47*	-0.01	1.37*	-0.35*	-3.17	-0.09	0.77	-0.02	-0.02	-639.88	0.00	10.10*
Size: Small	1,320	735	1.01	-0.66	0.09	0.01	0.06	0.95*	-0.06	-1.26	-0.24*	0.68	-0.48	-0.33	-1,600.83	1.05	31.76*
Size: Medium	504	333	-15.88*	6.85*	-0.70*	-0.62*	0.16*	1.20*	-0.10	-1.06	-0.23*	0.83	-0.90	0.75	-628.65	83.25*	25.67*
Size: Large	221	170	2.13	-0.68	0.06	0.45	-0.09	1.52*	-0.20*	-3.11	-0.18	0.82	-0.94	-0.77	-277.13	21.94*	17.64*
Minimum efficient scale	211	156	-0.92	0.06	0.02	0.72	-0.17	0.82*	-0.15*	-3.69	-0.13*	0.57	0.16	0.09	-250.66	0.81	19.75*
Single-plant firm	987	590	1.41*	-0.89*	0.13*	-0.06	0.06	1.91*	-0.33*	-9.95	-0.24*	0.79	-0.75	-0.60	-1,230.03	28.84*	243.79*
Foreign owned	251	203	1.80*	-0.68*	0.08*	0.23	-0.05	2.03*	-0.20*	-5.35	-0.38*	0.82	-0.90	-0.73	-321.30	14.89*	73.16*

Table 5.11c: Sample Selection Model: Plant Size, Age and Growth, With Age, 1977-1986, continued

Sample (SIC)	All	Survivors	α_0	α_1	α_2	α_3	α_4	β_0	β_1	$z\beta_1$	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
15: Food and beverages	270	143	-0.94	0.12	0.01	0.13	0.07	1.42	-0.19*	-2.11	-0.22	0.62	-0.18	-0.11	-298.07	0.02	5.13
16: Tobacco products	..	..	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
17: Textiles	88	43	-1.00	0.29	-0.04	0.58	-0.13	0.39	-0.11	-1.00	-0.20	0.77	0.50	0.38	-105.80	0.85	4.33
18: Wearing apparel	127	60	-2.27*	0.71	-0.06	0.07	0.03	1.30	-0.15	-1.41	-0.20	0.55	-0.32	-0.17	-127.85	0.17	4.62
19: Leather products	-	-	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
20: Wood and wood products	161	67	-2.21*	1.13*	-0.15*	-0.11	0.10	0.81*	-0.28*	-3.63	-0.02	-0.02	0.20	0.10	-140.93	0.47	14.57*
21: Pulp and paper products	-	-	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
22: Publishing and printing	184	93	-1.46*	0.14	0.02	0.89*	-0.11	1.30*	-0.15*	-2.21	-0.37*	0.60	0.12	0.07	-179.78	0.13	11.43*
23: Coke, petroleum and nuclear fuel	-	-	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
24: Chemicals and chemical products	97	66	-1.44*	1.09*	-0.09*	-0.29	0.02	0.92*	-0.19*	-4.30	-0.22*	0.78	0.28	0.22	-125.62	6.54*	29.44*
25: Rubber and plastic products	135	82	-2.23*	1.41*	-0.16*	-0.22	0.08	0.43	-0.22*	-3.54	0.05	0.86	0.89	0.76	-158.46	2.13	14.21*
26: Non-metallic mineral products	252	123	-0.49	-1.00*	0.21*	1.26*	-0.21	0.71	-0.20*	-3.79	0.01	0.72	-0.38	-0.28	-268.22	0.71	14.42*
27: Basic metals	148	81	-0.71	0.27	-0.02	-0.42	0.18*	2.10*	-0.23*	-6.21	-0.40*	0.73	-0.86	-0.63	-160.00	7.37*	46.95*
28: Fabricated metal products	402	181	-1.39*	0.47*	-0.05*	0.33	-0.03	0.78*	-0.29*	-5.87	-0.07	0.72	0.63	0.46	-428.46	4.66*	40.67*
29: Other machinery and equipment	316	150	-1.24*	0.42*	-0.03	-0.01	0.07	0.74*	-0.15*	-4.00	-0.22*	0.62	0.25	0.16	-337.57	1.20	52.29*
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
31: Electrical machinery	77	47	0.09	-0.47	0.07*	0.60	-0.08	2.92*	-0.35*	-3.50	-0.45	1.21	-0.94	-1.14	-104.04	2.92	12.71*
32: Radio, T.V. and communications	51	33	0.53	0.24	0.00	-1.73	0.49	1.57	-0.28	-1.09	-0.15	0.74	-0.14	-0.10	-63.48	0.01	5.05
33: Medical and optical instruments	73	43	-1.76*	0.82*	-0.07*	0.50	-0.14	0.48	-0.06	-0.55	-0.27*	0.89	0.93	0.83	-86.56	10.22*	5.87
34: Motor vehicles and trailers	85	50	-1.93*	0.27	-0.03	0.90	0.01	0.83*	-0.19*	-3.22	-0.07	0.57	0.52	0.30	-78.21	3.30	32.92*
35: Other transport equipment	52	26	-2.18*	0.513	-0.04	0.50	-0.01	-0.75	-0.02	-0.15	0.06	0.78	0.40	0.31	-57.24	2.20	0.09
36: Furniture	178	82	-1.75*	0.58*	-0.05	0.39	-0.05	0.96*	-0.11	-1.67	-0.27*	0.76	0.08	0.06	-202.65	0.13	24.25*

Notes: Model estimated using maximum likelihood with heteroscedasticity-consistent standard errors (STATA default standard errors applied for the Heckman two-Step estimate).
z-statistics only reported for the Size (β_1) coefficient in the growth equation.
The statistical significance of λ (z-test reported) is only available for the Heckman two-step estimator. Two-step $\rho = 1.12$ truncated to 1.00.
.. Not reported given small sample < 10 plants.
- Not reported by STATA.
* Statistically significant at the 5 percent level.

Table 5.11d: Sample Selection Model: Plant Size, Age and Growth, With Age, 1986-1997

Sample	All	Survivors	α_0	α_1	α_2	α_3	α_4	β_0	β_1	$z\beta_1$	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
All: Maximum likelihood	2,505	1,417	-1.23*	0.52*	-0.04*	0.05	0.01	0.79*	-0.14*	[-10.00]	-0.16*	0.73	0.78	0.57	-2,866.72	49.06*	197.98*
All: Heckman two-step	2,505	1,417	-0.91*	0.29*	-0.01	0.11	-0.00	-3.10	0.26	[0.69]	0.02	3.67	1.00	3.67	-	-	15.17*
Size: Micro	383	134	-1.59*	1.40*	-0.33	0.01	-0.03	1.88*	-0.73*	[-2.65]	-0.23*	0.92	0.92	0.85	-353.15	43.30*	29.83*
Size: Small	1,489	816	-0.12	-0.09	0.02	0.28*	-0.04	1.49*	-0.20*	[-5.79]	-0.21*	0.57	-0.15	-0.09	-1,703.91	2.69	172.12*
Size: Medium	469	321	8.99	-3.95	0.41	0.54*	-0.08	1.63*	-0.17	[-1.77]	-0.20*	0.70	-0.81	-0.57	-558.55	33.74*	18.44*
Size: Large	164	146	15.92	-3.74	0.30	-2.22	0.39	0.93*	-0.07	[-1.04]	-0.25*	0.55	-0.80	-0.44	-160.01	8.84*	22.45*
Minimum efficient scale	243	187	-3.18	0.90	-0.03	0.00	-0.00	0.83*	-0.09*	[-2.08]	-0.19*	0.54	0.07	0.04	-268.07	0.20	22.59*
Single-plant firm	1,226	684	-1.03*	0.55*	-0.06*	0.08	0.01	1.11*	-0.32*	[-9.36]	-0.11*	0.73	0.86	0.63	-1,367.93	83.63*	133.03*
Foreign owned	336	253	-0.71	0.34	-0.01	0.15	-0.04	1.86*	-0.26*	[-3.97]	-0.19*	0.63	0.17	0.10	-412.26	0.27	108.49*

Table 5.11d: Sample Selection Model: Plant Size, Age and Growth, With Age, 1986-1997, continued

Sample (SIC)	All	Survivors	α_0	α_1	α_2	α_3	α_4	β_0	β_1	$z\beta_1$	γ_1	σ	ρ	λ	LL	Wald $\rho = 0$	Wald χ^2
15: Food and beverages	220	125	-0.15	-0.35	0.08	0.25	-0.01	2.53*	-0.30*	[-5.51]	-0.36*	0.67	-0.41	-0.28	-260.29	1.72	63.53*
16: Tobacco products	..	..	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
17: Textiles	83	48	-2.09*	1.29*	-0.16	-0.25	0.10	1.46*	-0.27*	[-3.43]	-0.13	0.62	0.36	0.23	-95.35	1.01	17.72*
18: Wearing apparel	-	-	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
19: Leather products	18	11	-7.93*	-4.42	1.10*	3.40*	0.06	0.01	-0.01	[-0.09]	-0.13	0.50	1.00	0.50	-9.01	265.63*	-
20: Wood and wood products	106	49	-0.78	-0.15	0.04	1.01*	-0.25*	1.64*	-0.26*	[-2.90]	-0.13	0.58	-0.62	-0.36	-106.32	1.19	18.93*
21: Pulp and paper products	78	54	-3.32*	1.99*	-0.22*	-0.18	0.00*	1.28*	-0.26*	[-2.39]	-0.11	0.76	1.00	0.76	-81.51	24,032.14*	-
22: Publishing and printing	138	72	-2.74*	1.74*	-0.21*	-0.41	0.10	0.52*	-0.10	[-1.25]	-0.17*	0.76	0.95	0.72	-138.99	21.85*	10.91*
23: Coke, petroleum, nuclear fuel	..	..	..	..	..	..	..	..	..	[..]	..	..	..	..	..	..	..
24: Chemicals, chemical products	126	88	-0.89	0.68	-0.02	-0.25	-0.01	1.84*	-0.29*	[-4.10]	-0.12*	0.58	-0.46	-0.26	-138.31	0.97	37.67*
25: Rubber and plastic products	197	104	-0.55	-0.01	0.01	0.12	0.08	1.06	-0.32*	[-5.65]	0.09	0.62	0.64	0.40	-205.06	0.83	32.59*
26: Non-metallic mineral products	170	72	-2.30*	0.93*	-0.07	0.40	-0.12	1.23*	-0.19*	[-2.50]	-0.34*	0.70	0.83	0.59	-153.83	4.44*	30.11*
27: Basic metals	119	83	-1.54	0.69	-0.05	0.10	0.01	0.92	-0.18*	[-4.24]	-0.05	0.54	0.42	0.23	-128.61	0.52	18.42*
28: Fabricated metal products	310	190	-1.43*	0.81*	-0.08*	-0.17	0.08	0.50*	-0.16*	[-3.10]	-0.05	0.74	0.93	0.69	-344.73	41.63*	14.42*
29: Other machinery, equipment	239	141	-1.73*	0.85*	-0.08*	0.27	-0.06	0.97*	-0.23*	[-4.65]	-0.16*	0.69	0.92	0.64	-250.95	2.88	33.37*
30: Office machinery, computers	-	-	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-
31: Electrical machinery	97	57	0.71	-0.88	0.12	0.99	-0.19	2.14	-0.35*	[-2.02]	-0.02	0.89	-0.68	-0.60	-123.92	0.26	4.24
32: Radio, T.V., communications	67	46	4.15*	-1.70*	0.23*	-1.22*	0.29	1.30*	-0.08	[-1.20]	-0.36*	0.63	0.23	0.14	-77.51	0.17	31.70*
33: Medical, optical instruments	76	46	-1.19	0.67	-0.06	0.24	-0.08	1.51*	-0.15*	[-2.28]	-0.35*	0.62	0.33	0.21	-90.57	2.46	14.98*
34: Motor vehicles and trailers	78	45	-3.42*	1.33*	-0.10	-0.08	0.06	1.39*	-0.07	[-1.06]	-0.33*	0.47	-0.17	-0.08	-67.03	0.08	19.72*
35: Other transport equipment	40	21	-0.70	0.01	0.01	-0.44	0.23	1.82*	-0.10*	[-2.18]	-0.40*	0.48	0.06	0.03	-38.10	0.01	30.14
36: Furniture	-	-	-	-	-	-	-	-	-	[-]	-	-	-	-	-	-	-

Notes: Model estimated using maximum likelihood with heteroscedasticity-consistent standard errors (STATA default standard errors applied for the Heckman two-Step estimate).
z-statistics only reported for the Size (β_1) coefficient in the growth equation.
The statistical significance of λ (z-test reported) is only available for the Heckman two-step estimator. Two-step $\rho = 1.30$ truncated to 1.00.
.. Not reported given small sample < 10 plants.
- Not reported by STATA.
* Statistically significant at the 5 percent level.

not statistically significant at the 5 percent level).

For the growth equations, similar results were found to those from the earlier specifications estimated using both the full maximum likelihood and Heckman two-step estimators, with the former again providing a rejection of Gibrat's Law across all time periods for the aggregate samples (and broadly similar β_1 coefficients to the OLS estimates) exhibiting a process of mean reversion. The latter estimator also provided a statistically significant rejection in the time periods commencing in 1972, but not in either of the other samples where Gibrat's Law could not be rejected. Similar outcomes to those of the OLS estimates also emerged for the plant age variable. In all four aggregate samples the null hypothesis that $\rho = 0$ could be rejected. The inverse Mills' ratio was statistically significant in all but the most recent time period.

For the 1972-1977 sample, the size-class results again supported the notion of Gibrat's Law being more appropriate for larger plants with Gibrat's Law not rejected for either the Medium or Large size-classes, and with the coefficient of the plant age term not statistically significant for these larger size-classes. For the longer time period, again commencing in 1972, the null hypothesis of the β_1 coefficient in the growth model not being significantly different from zero was rejected across all size-classes, with growth declining with age (although in all but the Micro size-class the negative coefficient was not statistically significant).

In the 1977-1986 sample, the sample selection model estimates gave rise to a rejection of Gibrat's Law only for the Micro and Large plants with the coefficient of the plant age term, γ_1, not statistically significantly different from zero, but the opposite result for the Small- and Medium-sized plant size-classes. In the period 1986-1997, growth was found to be decreasing in age across all size-classes (and statistically significant at the 5 percent level), but only decreasing with size for the Micro and Small size-classes: Gibrat's Law could not be rejected for the Medium and Large size-classes.

At the industry-level Gibrat's Law could be rejected in 10 of the 17 industries in the period 1972-1977, in each case supporting a process of mean reversion (but not being rejected in the case of the Food and beverages, Wearing apparel, Rubber and plastic products, Non-metallic mineral products, Basic metals, Electrical machinery and Motor vehicles and trailers industries). The age variable in the growth equation was only statistically significant in 6 industries, in each of these providing the expected negative relationship. This sample selection model specification gave rise to three changes in the reported significance of the β_1 coefficient in the 1972-1977 sample (compared to the corresponding OLS estimates), becoming not statistically significant in the Food and beverages and the Non-metallic mineral products industries, supporting the first proposition of Gibrat's Law, but having the reverse effect for Pulp and paper products. In only two industries did the correction for sample selection bias alter the significance of the γ_1 plant age coefficient in the growth equation, Food and beverages and Wearing apparel, with opposite effects. In all but one case the coefficient of the size variable in the probit selection equation was negatively signed (significant in 8 samples).

Despite the null hypothesis of independence between the growth and survival equations being rejected at the industry-level in 8 cases for the 1972-1997 samples, the sample selection corrections did not change the statistical significance of any of the β_1 coefficients in the growth equation. Gibrat's Law could be rejected in all industries, except for Wearing apparel, with mean reversion observed. Plant age again generally provided a negative effect on growth (but being statistically significant in only five industries) with the significance of the coefficient of the plant age variable only affected for the Textiles industry, with the coefficient becoming more negative and statistically significant.

Again, only modest differences were detected using the sample selection model compared to the earlier OLS estimates for the 1977-1986 period, with only the Radio, television and communications industry showing changes in terms of the rejection of Gibrat's Law compared to the OLS estimators, but with three changes in the statistical significance of the γ_1 plant age term in the growth equation. All of the β_1 coefficients in the growth model were again negatively signed, being statistically significant in 11 of the 17 industries (being not significant in the Textiles, Wearing apparel, Radio, television and communications, Medical and optical instruments, Other transport equipment and Furniture industries). Across the industries examined plant age was observed to exert a generally negative effect on growth (significant in 6 industry samples).

Similar modest changes were also observed for the 1986-1997 time period with the OLS and sample selection estimates only showing statistically significant differences in two of the samples for the β_1 coefficient: Publishing and printing and Other transport equipment; and three of the γ_1 plant age coefficients: Textiles, Pulp and paper products and Fabricated metal products. The β_1 coefficients in the growth equation were all negative and significant in 13 industries (not significant for the Leather products, Publishing and printing, Radio, television and communications and Motor vehicles and trailers industries). The effects of the plant age variable in the growth equation was also generally negative (being negative in all but one industry) and significant in 9 industries.

At the industry-level the three sample selection model specifications therefore suggest some consistency in those industries reporting the highest (relative) negative β_1 coefficients across each of the four time periods - indicating the strongest tendency towards mean reversion - although such a comparison is however limited by the increased number of non-estimated industry-level models.

The results of the sample selection estimates for single-plant firms, foreign owned plants and plants above the MES were highly consistent with the previous sample selection models and OLS regressions: rejecting the first proposition of Gibrat's Law in all cases for both single-plant firms and foreign owned plants, and also in all but the 1972-1977 sample of MES plants. As in the previous sample selection models, the hypothesis that $\rho = 0$ could be rejected at the 5 percent level of significance in all four samples of single-plant firms, and in the cases of foreign owned plants in the 1977-1986 period, and for the MES sample in the 1972-1997 period.

As a check on these results and to further examine the effects of the economic cycle on the validity of Gibrat's Law the initial sample selection model, equation (5.8), was estimated using both the maximum likelihood and Heckman two-step estimators over the same periods of macroeconomic expansion and contraction as previously examined using OLS regression. The maximum likelihood estimates, summarised in Table A5.2, provide broadly consistent estimates rejecting Gibrat's Law in all 10 time series regimes, with $\rho = 0$ rejected in 9 of these estimates. However, in the case of the Heckman two-step estimator Gibrat's Law could not be rejected in the time periods 1975-1979, 1982-1988, 1988-1992, 1992-2000 and 2000-2003. Clearly the choice of sample selection estimation procedure may therefore provide a significant influence on the acceptance or rejection of this particular model of firm (or plant) growth, and may have indeed done so in previous studies.

Overall, though the aggregate results for the full economic cycle periods and the 1972-1997 sample tend to suggest that employing these correction procedures tended to make little difference to the statistical significance (or non-significance), and in many cases the magnitude, of the estimated β_1 coefficients of central interest in this research. Certainly, for these samples, sample selection bias does not seem to constitute, in general, a significant effect in relation to the estimated model parameters, except for some samples under the Heckman two-step estimator. However, the inconsistent (with prior expectations) coefficient signs in some of the probit selection equations may provide grounds for caution in accepting some of the sample selection model estimates. But the fact that these estimates are generally consistent with the OLS estimates may give rise to somewhat greater confidence in these findings.

5.8. Conclusions

This chapter has examined the empirical validity of the three propositions of Gibrat's Law: that plant growth is independent of initial size, that the variance of growth is independent of initial size and that there is no persistence in growth, using both econometric and statistical methodologies across a range of time periods and constructed samples.

The assessment commenced with an examination of the empirical properties of the variance of plant growth rates testing for statistically significant differences across constructed size-classes using F-tests for homogeneity of variances supported by additional robust tests for equality of variances. Differences in mean growth rates were examined using a t-test for unequal variances following the Welch (1947) procedure. These tests suggested some statistically significant differences in both the mean (proportional) growth rate and the variance of growth rates, with both declining with increased plant size in all four time periods, rejecting the first and second propositions of Gibrat's Law.

To examine the possibility of persistence in plant growth, an analysis regressing growth in consecutive five-year time periods using OLS procedures was employed. This approach suggested generally weak persistence in plant growth rates both at the aggregate-level and for disaggregated samples, so that the OLS estimated β_1 coefficient relating to the explanatory variable of plant size in the standard logarithmic growth regression model would not be subject to significant bias from this source. At the aggregate-level the past growth variable was observed to be a poor predictor of the subsequent growth performance of manufacturing plants, a finding, although with a modest number of exceptions, which also generally held in the case of a range of sub-samples. Such findings are therefore broadly supportive of the third proposition of Gibrat's Law.

To further consider potential plant growth persistence the logarithmic size-growth model specification of Gibrat's Law with the addition of persistence terms were estimated at the aggregate-level for various persistence lags covering time-spans of between one- and five-years. The persistence term was found to be statistically significant in only 5 of the 20 samples. Estimation of this model with an additional plant age term gave rise to statistically significant persistence term coefficients in only four of the corresponding samples. The inclusion of such persistence terms tended to have a slight upwards effect on the estimated β_1 coefficients.

To provide an initial analysis of the relationship between plant size, growth and survival, transition matrices were constructed for the time periods corresponding to each economic cycle and the longer time-span 1972-1997. These exhibited some consistency pointing to an inverse relationship between plant size and the likelihood of failure across the samples, though appearing less strongly observed for the 1972-1977 sample, and suggested the need to appropriately consider potential sample selection effects in the econometric analysis of the plant size-growth relationship. These transition matrices also exhibited a concentration of plant transitions along the main diagonal axis of the transition matrices providing some broad support for Gibrat's Law. There was however some sign of a slight downward "curvature" of the concentration of plants in the transition matrices in some samples, supporting the possibility of a slight tendency towards mean reversion.

The first proposition of Gibrat's Law was formally tested using the logarithmic growth specification and OLS cross-section estimation procedures, with corrections for heteroscedasticity, examining samples covering a range of time periods broadly corresponding to U.K. economic cycles. For each of the aggregate samples Gibrat's Law was rejected at the 5 percent level of significance with a process of mean reversion observed, with smaller plants exhibiting higher proportional growth rates than larger plants. Examination of the serial correlation model with persistence terms further supported these conclusions, with additional analysis indicating that the results were not an artifact of the choice of selected cross-section time-spans. The inclusion of industry dummy variables did not change these conclusions, but did lead to very slightly higher reported R^2 values for the models.

The results for size-classes provided somewhat more mixed conclusions, with the evidence being suggestive of Gibrat's Law providing a better approximation for larger plants, with smaller plants having a somewhat greater tendency towards convergence than larger plants (a finding consistent with many previous studies). Gibrat's Law was rejected in the majority of industry-level samples in each of the time periods examined, being only a reasonable description in 16 of the 78 samples considered, with the disaggregated results implying some differences in convergence speeds across industry samples.

The extension of the model to include a plant age variable gave rise to consistent evidence of statistically significant negative coefficients for both plant size and age variables at the aggregate-level (indicating a negative relationship between both size and growth and between age and growth), but with generally, though not universally, slightly less negative β_1 coefficients than in the standard model specification. The F-tests of overall significance rejected the null hypothesis that the coefficients of both of these variables were jointly zero in each of the samples. The corresponding serial correlation model including persistence terms estimated at the aggregate-level supported these findings with the coefficient of the plant size term again found to be negative and statistically significant in all but one of the twenty samples considered, with negative plant age coefficients in all but two of the samples, being statistically significant in 15 cases. The F-test on each occasion rejected the joint hypothesis of zero coefficients. Although some variation in relation to the role of age was detected both across plant size-classes and industries the broad empirical patterns relating to the size and sign of the β_1 coefficient were consistent with the model without plant age.

The addition of quadratic size and age terms to consider the existence of possible non-linearities in the relationship between these explanatory variables and plant growth produced statistically significant coefficients for the size, age, and quadratic size variables for all four aggregate samples, with the coefficient of the quadratic age variable being statistically significant at the 5 percent level in three of the four aggregate cross-section estimates. The F-test was rejected in all four samples, so that it could not be concluded that the inclusion of the quadratic size and age terms was not statistically significant at the 95 percent level of confidence. However, these quadratic variables did not significantly alter the explanatory power of the model.

To address the potential bias arising from the non-random exit of plants from an initial sample, three sample selection models were estimated using maximum likelihood, supplemented with estimates using the Heckman two-step estimation procedure for the aggregate samples only. For the simplest sample selection model with plant size as the only explanatory variable in the growth equation, at the aggregate-level Gibrat's Law could again be rejected in all four time periods using the maximum likelihood estimator, supporting the earlier findings of a process of mean reversion. However, whilst giving similar rejections in the two periods commencing in 1972, the Heckman two-step estimator provided some contrary results for the latter two time periods, with no rejection provided at the same level of significance in either period.

Examination of Gibrat's Law for constructed size-classes did not alter the broad findings of the regression analysis with Gibrat's Law being generally more appropriate for larger manufacturing plants. In the case of the industry-level samples the sample selection model estimates did not (in the vast majority of samples) significantly alter the observed level of rejections compared to the OLS estimated regressions, with the regression and sample selection models generally providing very similar β_1 plant size coefficients in the growth equation.

In the sample selection model with plant age in the growth equation but not in the selection equation, similar results again emerged with Gibrat's Law rejected for the aggregate samples employing the maximum likelihood estimation procedure, with only modest differences in the reported coefficient estimates compared to the OLS results. A similar pattern to that of the initial sample selection model was also found for the Heckman two-step estimator with no notable changes for the time periods commencing in 1972 but with no rejection of Gibrat's Law for either the 1977-1986 or 1986-1997 time periods. The inverse Mills' ratio was only statistically significant in two of the time periods. The γ_1 plant age coefficients were negative and statistically significant, with highly similar coefficients estimated across both time periods and estimators.

For the size-classes, a comparison of the β_1 coefficients from this sample selection model with the OLS estimates in the model with plant age as an additional explanatory variable indicated only three changes to the statistical significance of the estimated β_1 coefficients. Similar results emerged at the industry-level with few changes in either the support or the rejection of Gibrat's Law. Across the time periods and samples, however, there was some evidence of a variable relationship, particularly at the industry-level, between plant size and survival and also age and survival, with some samples reporting no significant relationship whilst others indicated examples of both statistically significant negative relationships and statistically significant positive relationships. Estimation of the sample selection model with both plant age and quadratic terms in the survival equation again provided results consistent with those from the other model specifications.

Overall, a comparison of the OLS regression and sample selection model results suggest that the implementation of the latter tended to make little difference to the statistical significance (or non-significance) of the estimated β_1 coefficients, and therefore to the acceptability of Gibrat's Law. However, although the analysis of the probit selection equations at the aggregate-levels provided estimates broadly consistent with prior expectations under the alternative model specifications for the 1977-1986 and 1986-1997 time periods, some inconsistencies (with prior expectations) were observed in the periods commencing in 1972 and for some disaggregated samples. Such differing coefficient signs in the probit selection equations might provide some grounds for caution in accepting some of the sample selection model estimates. However, the strong consistency of the generated growth equation results with the OLS estimates would appear to provide some broader basis for confidence in these overall findings.

Examination of the regression models without and with plant age using OLS for plants above the estimated MES provided a rejection of Gibrat's Law for the three time periods 1972-1997, 1977-1986 and 1986-1997 for the first two linear regression model specifications but not in the case of the 1972-1977 period (some differences were however observed in the non-linear model). Sample selection model results estimated using maximum likelihood also gave rise to rejections of Gibrat's Law for the same three time periods. Repetition of these procedures on samples of single-plant organisations, which might approximate more closely to some firms (although providing a non-representative sample of the full population of firms), rejected Gibrat's Law in all estimates. Similar comprehensive rejections were found using both regression and sample selection model specifications for foreign owned plants, albeit with some not statistically significant β_1 parameters in the non-linear OLS model.

Re-estimation of the basic sample selection model using both maximum likelihood and Heckman two-step estimators across different time series regimes, corresponding broadly to periods of U.K. macroeconomic expansions and contractions, whilst finding a consistent process of mean reversion under maximum likelihood for all ten time periods, provided no rejection of Gibrat's Law using the Heckman two-step estimator in the time periods 1975-1979, 1982-1988, 1988-1992, 1992-2000 and 2000-2003. These results may suggest a potential influence on the estimated plant (or firm) size-growth relationship as a consequence of the choice of estimator – a factor which may have affected previous studies.

In summation, given these results, and consistent with many recent empirical studies of production activities, it cannot be concluded that Gibrat's Law holds in general across both aggregate samples, time periods or for the full range of manufacturing industries and ownership structures, with mean reversion the dominant process, so that smaller plants were found, on average, to exhibit proportionately greater employment growth than larger plants. However, this empirical evidence suggests that Gibrat's Law may indeed provide a more reasonable description for larger manufacturing plants, an observation again consistent with previous research.

6

CONCLUSIONS

6.1. Introduction

This chapter draws together the main findings of this research. Section 6.2 provides a brief overview of the theory and implications of Gibrat's Law, before reviewing the principle aims of this study and the methodologies employed to empirically test the propositions of this particular stochastic model of firm (or plant) growth. Section 6.3 provides a summary of some of the main empirical properties of the dataset, especially with regards to the observed plant and employment dynamics, as well as plant entry and exit, within the manufacturing sector, and for various disaggregated samples. Section 6.4 reports the conclusions from this research in relation to plant size and growth rate distributions, and the consistency of these results with both the postulated lognormal limiting distribution of sizes and the assumed normal distribution of plant growth rates. The conclusions arising from the econometric analysis of the plant size-growth relationship, and associated extensions to this model, is presented in Section 6.5. Section 6.6 discusses some of the work conducted as part of this research that could not be included here, and suggests a number of areas for potential further investigation.

6.2. Research Aims and Methodologies

At the commencement of this research two primary over-arching objectives were established: first, to provide a detailed review of the existing economic and statistical literature on Gibrat's Law; in particular considering the effects of methodological developments, such as the role of model specifications, estimation procedures, the coverage of business micro-datasets and the time periods examined on the reported empirical validity of this particular stochastic model; and second, to generate new empirical evidence regarding the applicability of Gibrat's Law, including in respect of the stated propositions regarding plant or firm dynamics and the implied size and growth rate distributions. In conducting the latter element, a need to examine the wider patterns of development and change within the manufacturing sector as context for this research was identified.

The stochastic model of Gibrat's Law is founded on a modified Gaussian process whereby a large number of small, independent components, or growth opportunities, operating multiplicatively and acting on a fixed population of firms (or plants) would generate a lognormal limiting distribution of firm sizes, a distribution skewed to the right, with a higher frequency of smaller than larger firms observed. Such a prediction is consistent with the actual distributions of firm sizes observed across a range of reasonable measures of firm sizes, time periods, industries and countries (Ijiri and Simon, 1977; Hart and Oulton, 1996), providing powerful support for the potential role of stochastic influences in shaping the development of firms and market structures (see Clarke, 1985, pp.34-35).

The model of Gibrat (1931) also provides three empirically testable propositions regarding the specific growth patterns of firms (or plants): that firm growth is independent of size; that the variance of growth is independent of size, and; that there is no serial correlation (or persistence) in growth rates across time periods. The first proposition has been frequently examined using either a standard logarithmic or growth model specification representing the firm size-growth relationship. Although such models set out a linear relationship between firm size and growth, suitable non-linear extensions have also been developed and empirically tested.

Under the assumptions of Gibrat's Law the coefficient on the explanatory firm size variable in the standard model specification is hypothesised to be of value unity (so that $\beta = 1$), implying that firm growth could be described by a non-stationary, random walk, process with possibly non-zero drift. Such a process suggests that firm growth would be observed to follow an erratic, unpredictable pattern, with firm size at any point in time being the sum of cumulative (unanticipated) shocks providing permanent effects on firm size, and as such exhibiting path-dependency.

Although sometimes criticised for a lack of economic foundations (see Sutton, 1998), such a proposition of independence between firm or plant size and growth can be relatively easily interpreted within a more deterministic framework, under the assumption of an "L-shaped" long-run average cost curve, so that average costs are constant beyond a given minimum efficient scale of production, a property frequently supported across reasonable output ranges by the empirical evidence (see Clarke, 1985).

Under the condition that $\beta = 1$, subject to the existence of some random growth variation across firms, the implication is that industrial concentration will increase continuously over time. This aspect, however, is not specifically examined within this research given the regional nature of the dataset which makes such an analysis of concentration somewhat less appropriate. The focus in testing Gibrat's Law is instead on both investigating the validity of this specific model and, by extension, the relative employment generating propensities of this sample of plants, which can be examined using the coefficient of the explanatory plant size parameter in the standard model specification of Gibrat's Law. Where the estimated size coefficient displays $\beta > 1$, large plants exhibit, on average, growth rates above that of smaller plants, with, in the case of $\beta < 1$ (mean reversion) smaller plants being subject, on average, to proportionately faster growth than larger plants (Prais, 1976, p.279; Hart and Oulton, 1996, p.1245).

In conducting a review of the theoretical and empirical evidence relating to Gibrat's Law, and associated stochastic models, the first objective of this research, a highly notable feature has been the voluminous literature which has emerged, with econometric and statistical studies spanning a wide range of different datasets, time periods, geographies and industries, and within the theoretical context a number of extensions and modifications to the theory presented by Gibrat (1931). The presentation of a detailed survey of this evidence and the developing analytical methodologies and estimators as a significant component of this thesis has aimed to draw this evidence into a consistent and coherent review, presenting both the empirical regularities and departures from Gibrat's Law.

Further, this research also set out to provide a new empirical assessment of Gibrat's Law, the second objective of this research, making use of a previously little used official U.K. regional plant-level micro-dataset, the Welsh Register of Manufacturing Employment. This micro-dataset has the attractive feature of providing a much longer time dimension than has been generally available, permitting the empirical examination of Gibrat's Law over a number of economic cycles and allowing the examination of the long-run dynamics of surviving manufacturing plants, a further relatively under-researched economic issue.

Although a number of empirical examinations of Gibrat's Law have been conducted previously using U.K. firm-level or plant-level data, none have been identified which provide such a comprehensive investigation of both the plant (or firm) growth propositions of this model and the implied lognormal distribution of plant (or firm) size distributions, or using U.K. regional data. In addition, the examination of the plant growth rate distributions using the employed techniques further augments this element of the literature.

A notable attribute of this research is also the level of disaggregation considered examining Gibrat's Law for a range of industry-level samples, across constructed plant age-classes and size-classes and for further sub-samples of foreign owned plants (of interest given U.K. and regional policies during the time-span of the dataset aimed at encouraging and supporting such investments) and single-plant organisations (which may more closely reconcile to the existing research conducted at firm-level, although such employed samples are non-representative of the complete population of firms). Of further interest, this analysis has also considered both the static plant size distributions and the evolution of these distributions over time, and specifically the effects of age and selection processes in shaping the observed plant size distributions, again an area of limited previous investigation.

Of further value is the use of a range of complementary econometric and statistical approaches within this study to test the empirical validity of Gibrat's Law as a check on the robustness of such estimation procedures. Of particular note is the application of non-parametric procedures to test the empirical dynamics of manufacturing plant size and growth rate distributions, specifically using kernel density estimation, examination of the moments of these distributions and the application of the formal statistical tests of D'Agostino et al. (1990) and Shapiro-Wilk (1965) or Shapiro-Francia (1972). Although these non-parametric techniques have been recently applied in a small number of studies covering datasets from European countries, no such studies have been identified relating to U.K. regions, or even for that matter comprehensively applied in respect of U.K. data.

Consistent with the majority of the literature, in testing the primary growth propositions of Gibrat's Law, significant use has been made of the standard logarithmic growth specification of Gibrat's Law, tested on aggregate samples of plants across different time periods. Noting the early work of Mansfield (1962), the analysis has tested the three main versions of Gibrat's Law, considering its appropriateness for samples of all plants, surviving plants, and constructed samples of plants estimated to correspond to the minimum efficient scale of production.

Again, following the mainstream of the existing literature, this study utilised cross-section estimators in the regression analysis, using an employment-based measure of plant size with the merits of this measure and plant-level analysis both having been reviewed. Whilst a very small number of recent studies have attempted to employ both panel data estimators and tests for unit roots, the cross-section approach used here can be justified on the basis of providing consistent and comparable estimates with the majority of the existing domestic and international studies, due to the potential power limitations of such unit root tests (even for more powerful panel unit root tests) in effectively discriminating between unit root and near unit root processes, and due to the particular empirical properties of the dataset which weaken the appropriateness of such standard panel data estimators here.

In estimating the plant size-growth relationship a range of econometric modelling issues were considered, such as functional form, persistence and heteroscedasticity in documenting the robustness of the estimated model coefficients. Particular attention was given to the issue of potential sample selection bias, introducing a number of sample selection models with comparisons made between the effects of estimation under both maximum likelihood and the Heckman two-step estimator, and also against OLS estimates. Further extensions to the Gibrat model were presented examining the role of plant age in the observed dynamics of manufacturing plants.

6.3. The Empirical Properties of Manufacturing Industry Dynamics

As just noted, this research has set out to test the implications and propositions of

Gibrat's Law. In so doing, the Welsh Register of Manufacturing Employment, an official business micro-database recording information on manufacturing plants with more than 10 employees (with partial information on smaller plants and some non-manufacturing establishments), has provided the examined data.

Following a number of procedures to check and clean the data, a sample of over 6,100 manufacturing plants was constructed for analysis. The properties of this established dataset were described both as context the subsequent analysis and also as a means of identifying important analytical issues. Indeed, such information both contributed to the determination of time periods for the analysis and pointed strongly, for example, to the need to appropriately address features such as heterogeneity in the econometric and statistical analysis, an aspect which emerged as a key theme in respect of plant and employment dynamics across a range of disaggregated samples.

Examination of the aggregate information across the full time-span of the dataset suggested that although there was some growth in both the stock of manufacturing plants and employment within the early years of the dataset, since about 1977 in the case of the former, when the manufacturing plant stocked peaked at 2,812 plants, and 1974 for the latter, peaking at over 332,000 employees, both recorded measures have exhibited a generally downward trend, albeit with some variation over time. Within these trends there have been some notable differences in performance observed across both industries and ownership classes, particularly in respect of the recorded employment patterns.

Although with some degree of volatility in the plant entry and exit data, the former exhibiting a particularly significant peak in 1975, perhaps the most interesting observation has been the change over time in these plant flows, which after having initially provided a positive effect of the stock of plants until around the mid-to-late 1970s, after a more mixed period during the 1980s, have from around 1990 given rise to the stock of plants declining in each subsequent year, these dynamics seemingly reflecting, at least in part, the broader structural changes within the manufacturing sector experienced over this time period. Plant entry and exit, both primarily operating into and out of the smaller size-classes, across industries indicated some consistency of those industries with the highest levels of entry and exit, again coherent with previous empirical evidence.

In conjunction with plant expansions and contractions the effect of these plant entries and exits provided significant annual gross in-flows and out-flows of employment. A number of short, intensive periods of labour shedding were observed in 1971, 1975-1976, in the early 1990s and again in the early 2000s, but especially during the early 1980s recession. The net employment effects were over 100,000 job losses in the period 1978-1983, a further net loss of almost 32,000 jobs between 1991-1993 and over 33,000 jobs lost in the period 1999-2003, with a reduction of over 41,000 manufacturing jobs in 1981 alone. However, except during the years 1981, 1992 and after 1999, in each year the number of plants reporting increased employee numbers exceeded those reporting reduced employment.

Such changes in plant numbers, arising through plant entry and exit, and employment would be expected to contribute to the evolution and shaping of the observed properties of the manufacturing plant size distributions. An initial overview of the properties of the aggregate cross-section distributions of manufacturing plant sizes, presenting the statistical data on the mean and median, the former exceeding the latter in all years, percentile measures and constructed skewness and kurtosis pointed strongly to the highly skewed nature of plant sizes across each of the examined aggregate cross-sections spanning a period of over three decades, an observation broadly consistent with a Gibrat-type stochastic process. The calculated kurtosis indices for these distributions, in particular, however, exhibited some degree of volatility over time, with a number of distinct phases of stable kurtosis indices interspersed with rapid changes.

Further analysis of the data indicated that the mean average plant size declined significantly in the early part of the dataset, until around the mid-1980s (before some degree of reversal to slightly larger mean sizes), with the standard deviation of plant sizes declining notably from the early years of the data, partly reflecting a shift from very large manufacturing operations, especially so in the early 1980s. Such declines in the sizes of the very largest plants were primarily a consequence of employment reductions within these plants, rather than the closure of such plants. However, these plant-level employment trends were also observed in the percentile measured of the size distributions, the 10th, 25th, 50th, 75th and 90th percentile measures of plant sizes declining until the mid-1980s, with the 75th and 90th percentile measures, remaining noticeably below the original values in the final year of the dataset.

At a more disaggregated level, some important differences however emerged across a range of sub-samples. Although most industries exhibited the largest mean size in the late 1960s or early 1970s, for a modest number of industries such largest mean sizes over the time-span of the dataset were observed within either 2002 or 2003. The empirical data also showed a trend towards smaller mean plant sizes across both domestic and foreign ownership classes, though with interesting differences within the sub-categories of these groups, with a significant decline in the mean sizes of Asian owned plants, especially so in the early years of the dataset.

Across industries (though not in every case), and particularly the Basic metals industry, the standard deviation of plant sizes also tended to decline, especially in the time period until the early 1980s. Such a pattern is consistent with some degree of convergence towards an "optimal" scale through size changes or the shake-out of such non-optimal sized plants. Declining standard deviations of plant size over time were also found for domestic and foreign owned plants.

6.4. Findings on Plant Size and Growth Rate Distributions

Under the assumptions of the stochastic model of Gibrat's Law, and for a fixed population of firms (or plants), the lognormal distribution emerges as the limiting distribution of firm sizes. Even under weakened assumptions the lognormal distribution may still emerge as the limiting distribution.

However, further modifications, for example, regarding firm entry and exit processes (from which Gibrat's Law abstracts), the effects of mergers and acquisitions, the allocation of growth opportunities and serial correlation processes can also give rise to a wide range of predicted statistical distributions, including various versions of Pareto distributions (including the special case of the Zipf distribution), for which perhaps a similar weight of support has been provided, the Yule distribution, and less frequently examined distributions such as the Waring, a generalisation of Waring, particular generalised hypergeometric, negative binomial, logarithmic, generalisation of logarithmic, geometric, Poisson, and extended Katz distributions. The Laplace and Subbotin distributions have also been further suggested as suitable approximations to the actual distributions of firm growth rates, rather than the Gaussian (normal) distribution assumed in Gibrat's Law. The reviewed evidence has tended to suggest that no single statistical distribution can be reasonably considered as universally appropriate across all time periods, geographies and industrial samples.

Although this range of plausible theoretical size distributions (and others) has been documented, the scope of the analysis here has been restricted to the testing of the lognormal distribution for a number of reasons: first, Gibrat's Law is the foundation from which many of these other stochastic models, and resultant distributions, have frequently been constructed; second, given the difficulties in many cases of fine discrimination between competing statistical distributions, which may also be hampered by the practical problems of data truncation (which could affect the statistical properties of the actual observed distributions) and often incomplete industry coverage; and third, given the time available for such a comprehensive analysis (such potential research is discussed in Section 6.6).

In assessing the validity of the lognormal distribution, this research applied a broader range of estimation methodologies than generally employed in previous single studies, and considered a wide range of aggregate and disaggregated samples, examining both the static and dynamic distribution of plant sizes. As discussed in Chapter 3, in conducting such an analysis use was made of an employment-based measure of size, with the plant being the unit of analysis.

In respect of the aggregate cross-section plant size distributions, the kernel density estimates, examined higher moments of the distributions and formal statistical tests provided evidence across the full time period 1966-2003 of generally stable distributions, with qualitatively similar properties frequently exhibiting modest positive skewness and leptokurtosis compared to that of the (log)normal distribution. However, across all of these examined years the statistical tests provided a strict rejection of the lognormal distribution, although the potential effects of the partial truncation of the WRME data on the statistical properties of the distribution, the incomplete sectoral coverage and the previously discussed issues regarding the statistical testing of extreme hypotheses are important to note in drawing appropriate conclusions. Given these factors, these observed plant size distributions appear to be at least strongly consistent with a stochastic process, even if perhaps under a process of a more complicated nature than proposed by Gibrat.

The analysis conducted by tracking the evolution of plant size distributions over time also did not provide evidence of any consistent convergence to the lognormal distribution, with some evidence emerging of possible bi-modal distributions for some plant age-classes. Both plant ageing and selection effects (in the form of plant exit) generally provided modest contributions in shaping the observed acceptability of the lognormal distribution. Even over durations of 25 years, these effects were not conducive to permitting a statistically acceptable fit of the lognormal distribution to the actual data on surviving manufacturing plants. Similar analyses of the evolution of cohorts of new entrant plants, suggested the frequent presence of statistical deviations from the lognormal distribution, and although the observed slightly larger initial deviations diminished relatively rapidly, no consistent convergence towards the (log)normal distribution was found across samples.

At the industry-level, and in line with previous research, evidence of notable heterogeneity in the evolution of the plant size distributions was detected, an observation consistent with Schmalensee (1989) and Sutton (1995a) regarding the lack of support for any single statistical distribution in providing a universally adequate description of the distribution of firm or plant sizes. Although for some samples there was evidence of convergence towards the lognormal distribution over time, this was not observed as a general pattern, with in many cases the plant size distributions exhibiting some degree of stability remaining either suitably described by the lognormal or not lognormal over the time periods examined. Such evidence suggests a role for industry-specific factors influencing the emergence of the observed size distributions.

However, whilst it is not clear whether over longer time periods such distributions may evolve (or not) to the lognormal distribution, the evidence from the long time period sample suggests that observed divergences may be fairly persistent, even if such an evolution towards the (log)normal distribution were to occur. In respect of the statistical acceptability of the lognormal distribution, both of the formal statistical tests employed provided a higher non-rejection than rejection rate for the lognormal distribution across the samples of selected industries.

The analysis of plant growth rate distributions provided strong evidence of highly-peaked distributions for the aggregate data, with kurtosis indices exhibiting "fatter tails" than would be expected under the normal distribution, suggesting a higher probability of extreme plant growth rates, or indeed the Laplace distribution. Considerable heterogeneity in terms of (industry-level) plant growth rates across time periods was observed, but again with consistent greater mass of the distribution in the tails. In fact, only around one half of the reported kurtosis values could be even broadly considered as approximating to either the normal or Laplace distributions, the remaining samples exhibiting heavier tails, with the normal distribution being statistically rejected in all but six industry samples.

In using these techniques, this research has therefore found evidence of plant size and growth rate distributions highly consistent with the more recently presented empirical studies. Across these plant-level samples, it is clear that, whilst in some cases the lognormal distribution implied by Gibrat's Law may provide a reasonable broad first approximation to some actual plant size distributions, more generally, it cannot be concluded that the lognormal distribution is across all of the examined time periods and samples (even within the manufacturing sector) capable of providing a strictly statistically appropriate representation of the actual manufacturing plant data. The identified heavy-tails of the plant growth rate distributions, adds to the emerging literature, with again considerable heterogeneity across disaggregated samples.

6.5. Findings on the Plant Size-Growth Relationship

As re-stated above, Gibrat's Law provides three primary empirical propositions regarding the properties of firm (or plant) growth. Whilst a significant element of previous research has suggested some support for Gibrat's Law, a number of other studies have provided evidence of deviations from all three of the model propositions, even after controlling for factors such as potential sample selection bias.

In the case of the first proposition of Gibrat's Law, that firm (or plant) growth is independent of initial size, support has been found for such a proposition, but also in other studies finding patterns of larger firms and plants growing proportionately faster than smaller firms (and plants), the opposite result of mean reversion, and indeed more complicated non-linearities in the firm (or plant) size-growth relationship. In many cases larger firms (or firms at or above the MES) appear to have been more adequately described by Gibrat's Law. More recent studies correcting for potential sample selection bias arising from the non-random exit of firms or plants have tended to suggest some prominence in finding mean reversion, although frequently the degree of mean reversion has been sufficiently modest for authors such as Geroski (1999) to conclude that the random walk process may still be appropriate as the benchmark description of firm growth.

More specifically, in respect of the three primary propositions of Gibrat's Law, at least on one point the evidence is generally, though not universally, consistent, finding that the variance of firm (or plant) growth is not independent of size but decreasing with size. Recent research has suggested that such a relationship between firm size and variance of growth may be more appropriately described by a power law relationship. Evidence relating to the persistence of growth appears to be somewhat inconclusive, with different studies providing support for the proposition of no persistence, whilst others have detected both negative and positive persistence in growth rates. Coherent with a number of more recent models of industrial dynamics, such as the learning and selection models of firm growth, incorporating stochastic elements, extensions of the basic Gibrat model have also produced fairly consistent evidence of a negative relationship between firm age and growth.

Turning to the results from this research, in agreement with most recent empirical evidence the statistical analysis of the mean (proportional) plant growth rates across constructed plant size-classes provided rejections of the first proposition of Gibrat's Law, an observation also suggested by analysis of constructed transition matrices. The econometric estimation of the standard logarithmic model of Gibrat's Law (with growth as the dependent variable) using OLS on cross-sections of surviving manufacturing plants over a range of time periods supported this finding, the regression analysis indicating a consistent process of mean reversion with smaller manufacturing plants exhibiting higher proportional growth rates than larger plants. Such a mean reversion process was also generally found across determined periods of macroeconomic expansion and contraction (although in some time periods, and under certain estimation procedures, support for Gibrat's Law was found). The statistical analysis also supported significant differences in growth rate variances across plant size-classes, declining with increased plant size, and rejecting the second proposition of Gibrat's Law. The analysis of manufacturing plant growth using regression methods suggested generally weak persistence in growth, both at the aggregate-level and for industry-level samples. Such findings are therefore not unsupportive of the third proposition of Gibrat's Law.

Again, also coherent with the existing literature, econometric analysis of the size-growth relationship across plant size-classes further supported Gibrat's Law being a better approximation for larger manufacturing plants (although not for plants being at or above the minimum efficient scale of production, which probably reflects the limitations of the measure employed here) with smaller plants generally exhibiting a somewhat greater convergence tendency than larger plants.

Gibrat's Law could be generally rejected for both samples of single-plant organisations, which might provide some consistency with previous firm-level studies (but being likely to provide a non-representative sample), and foreign owned manufacturing plants (although with some differences found in the non-linear model specification). At the industry-level Gibrat's Law could be rejected in the majority of samples in each of the time periods examined, being only a reasonable description in the case of a small minority of industries. Analysis of the industry samples further indicated some support for different convergence speeds across industries.

The inclusion of a plant age variable in the model specification did not significantly alter these findings at the aggregate-level, in agreement with the previous economic literature, supporting a negative relationship between both plant size and growth and also between plant age and growth. Some variability was detected in the effect of the age variable both across plant size-classes and industries. The further addition of quadratic plant size and age terms to examine the possibility of non-linearities in the plant size-growth relationship did not significantly alter these conclusions, with their inclusion not significantly adding to the explanatory power of the models. Further, examination of the serial correlation models with persistence terms also further supported these conclusions.

To address the potential bias arising from the non-random exit of manufacturing plants from the initial samples (the transition matrices suggesting a frequently higher probability of exit of smaller plants), a number of sample selection models were examined. Estimation using the preferred maximum likelihood procedure did not in general significantly alter the acceptability or rejections of Gibrat's Law, with the estimated size coefficients similar to those from the estimated OLS regressions. For a very small number of industry-level samples some changes in the acceptability of Gibrat's Law were detected: but these were the exception. Broadly similar results were found across all three sample selection models, although a point of note is that sometimes the coefficients of the probit selection equations were not always consistent with prior expectations, perhaps adding a degree of caution to some of these findings.

In conclusion, this section of the analysis has provided evidence highly consistent with recent econometric and statistical studies of Gibrat's Law: rejection of the first proposition that manufacturing plant growth is independent of size with a process of mean reversion generally detected, implying that smaller plants, on average, tended to have exhibit proportionately faster employment growth than larger plants, with such mean reversion implying that plant sizes would tend towards the mean size of the population (Prais, 1976, p.281; Hay and Morris, 1991, p.537); rejection of the second proposition that the variance of manufacturing plant growth rates is independent of size instead finding evidence of declining variance with increased size-classes, at least for the smallest size-classes; and non-rejection of the third proposition of no persistence in manufacturing plant growth rates. In addition, the results also add to the weight of evidence suggesting a negative relationship between plant age and growth.

6.6. An Overview of Unreported Research and Scope for Further Research

Although this research has provided the opportunity to examine in some detail the validity of Gibrat's Law - both in respect of the primary propositions regarding the growth of plants, and the implication of an (approximately) lognormal distribution of plant sizes, many relevant and associated aspects of the industrial dynamics of plants could not be covered within thesis. In some cases, these issues did form part of the research conducted, but unfortunately, due to space and time constraints had to be excluded from the final text. Other areas, which were identified at an early stage, have been beyond the scope of this work, but given the linkages to the presented analysis, would form interesting themes for further investigation. Three major aspects seem most relevant for further research: first, more detailed analysis of the size distribution of plants (or firms); second, further investigation of the propositions of Gibrat's Law, and particularly the empirical properties of plant (or firm) time series; and third, investigation of plant (or firm) entry, exit and survival durations. These areas are now briefly considered in concluding this thesis.

Although only briefly discussed here, the literature survey on firm (or plant) size distributions presented a broad range of statistical distributions which could be plausibly generated through the modification of, and additional assumptions to, the model of Gibrat (1931). Despite the limitations of attempting to statistically discriminate between such distributions (a discussion touched upon earlier within this work), it would be appealing to employ some of the techniques set out within the published literature to attempt to analyse and compare the relative acceptability of other frequently examined distributions such as the Pareto distribution (including the rarely tested Pareto Type III distribution), and to build on this to also empirically examine some of the less well known skewed statistical distributions, discussed, for example, by de Wit (2005), for which little comprehensive evidence has been published.

Similarly, in respect of the growth rate distributions of plants (or firms), a further area of interest would be to extend this current research to fully consider distributions such as the Subbotin distribution, perhaps making use of estimation programmes such as that described by Bottazzi (2004). Statistical testing for potential multi-modality of plant size and growth distributions would also be of merit, in developing a more complete description of firm or plant size and growth rate distributions. Additionally, the present analysis has considered only the growth rate distributions across selected time periods. It would be informative to examine in more detail the relationship between macroeconomic variables and the moments of such distributions.

Of course, an obvious extension to this research would be to reconsider the analysis of the plant size and growth rate distributions employed here making use of a more comprehensive dataset (allowing some analysis of non-manufacturing sectors and other geographies, and possibly across different measures of firm or plant size). Recent research linking official datasets might also permit the opportunity to examine conditional plant (or firm) size distributions using techniques such as quantile regression (for example, following Machado and Mata, 2000), which was not possible here.

In terms of the empirical growth dynamics of these plants, several issues seem worth considering in more detail. One area for additional examination would be to consider the sensitivity of the estimated regression coefficient of the size-growth model relationship implied by Gibrat's Law. Although cross-section estimates were used here, a range of other studies have increasingly made use of panel data procedures and estimators, such as that of Breitung and Meyer (1994). The application of some of the techniques employed by Geroski et al. (2003) to this data could add further to this analysis, perhaps also examining the properties of alternative unit root and panel unit root tests. Further possible econometric aspects would be to consider the use of Bayesian estimation procedures, such as employed in Cefis et al. (2001), or the semi-parametric modelling approaches used by Cosh et al. (1996), neither of these techniques being widely employed to date.

Given the deviations from Gibrat's Law observed in this research, further work might also include the empirical testing of alternative models which might more adequately capture the properties of this data. A wide range of models have been proposed in the economics literature regarding firm growth which might also lend themselves to empirical testing. Geroski (1999), for example, in addition to setting out the model of Gibrat (1931) also presented specifications consistent with, optimal-size theory models, the life-cycle (stage theory) models of firm growth, the managerial model of firm growth described by Penrose (1980) and organisation capabilities models. Hall (1987) also examined Leonard's (1984) mean-reverting (flexible accelerator) model in modelling firm time series, as an extension to her analysis of Gibrat's Law. Whilst Geroski (1999) was not overly supportive of some of these models, further empirical evidence may be informative, particularly perhaps in attempting to explain the reported deviations from Gibrat's Law, and the role of possible deterministic factors in contributing to these observations.

Further documenting and testing for persistence or serial correlation in the time series (and the form of such serial correlation) of the data would also be of merit, perhaps providing a detailed analysis of such patterns at more highly disaggregated levels than generally investigated, for example, at the plant-level. Additionally, a natural extension to the work presented here would also be to more formally test specifications of the relationship between plant (or firm) size and growth rates variances, perhaps testing power law relationships, as found in several recent studies.

Finally, there remains considerable scope for analysis relating to plant entry, exit and survival durations. Unfortunately, although significant progress was made on this theme within the scope of the work conducted, this is the area where much of this research has had to be omitted. For example, both a detailed review of firm and plant survival and survival durations and the relevant econometric methods that could be employed to examine these issues were compiled. Some preliminary empirical analysis was also conducted, for example, generating tabulations of the survival durations of new entrant plants and plant survival probabilities across different time periods, with additional tabulations investigating the relationship between plant age and survival, and plant survival by industry and ownership classes. A number of different probit model specifications examining plant survival were also estimated using standard cross-section estimators and making use of the full range of available plant-level information in the dataset. There remains considerable scope for completing and building on this analysis – perhaps again by matching-in financial and wider economic data from other sources.

The further in-depth analysis of such issues relating to plant (or firm) size and growth rate distributions and growth processes, would, it is considered, offer the opportunity to develop the evidence presented in this thesis in respect of Gibrat's Law and the lognormal distribution, contributing to the wider empirical evidence base on models of firm growth and the statistical regularities of industrial dynamics.

STATISTICAL TABLES

Table A3.1: Description of 2-Digit Standard Industrial Classifications

SIC (1992)	Description	Abbreviated Industry Descriptions in Thesis
10	Mining of coal and lignite: extraction of peat	Mining of coal and lignite
14	Other mining and quarrying	Other mining and quarrying
15	Manufacture of food products and beverages	Food and beverages
16	Manufacture of tobacco products	Tobacco products
17	Manufacture of textiles	Textiles
18	Manufacturing of wearing apparel; dressing and dyeing of fur	Wearing apparel
19	Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear	Leather products
20	Manufacture of wood and products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	Wood and wood products
21	Manufacture of pulp, paper and paper products	Pulp and paper products
22	Publishing, printing and reproduction of recorded media	Publishing and printing
23	Manufacture of coke, refined petroleum products and nuclear fuel	Coke, petroleum and nuclear fuel
24	Manufacture of chemicals and chemical products	Chemicals and chemical products
25	Manufacture of rubber and plastic products	Rubber and plastic products
26	Manufacture of other non-metallic mineral products	Non-metallic mineral products
27	Manufacture of basic metals	Basic metals
28	Manufacture of fabricated metal products, except machinery and equipment	Fabricated metal products

Table A3.1: Description of 2-Digit Standard Industrial Classifications, continued

SIC (1992)	Description	Abbreviated Industry Descriptions Employed in Thesis
29	Manufacture of machinery and equipment not elsewhere classified	Other machinery and equipment
30	Manufacture of office machinery and computers	Office machinery and computers
31	Manufacture of electrical machinery and apparatus not elsewhere classified	Electrical machinery
32	Manufacture of radio, television and communications and apparatus	Radio, T.V. (television) and communications
33	Manufacture of medical, precision and optical instruments, watches and clocks	Medical and optical instruments
34	Manufacture of motor vehicles, trailers and semi-trailers	Motor vehicles and trailers
35	Manufacture of other transport equipment	Other transport equipment
36	Manufacture of furniture; manufacturing not elsewhere classified	Furniture
37	Recycling	Recycling
51	Wholesale trade and commission trade, except of motor vehicles and motorcycles	Wholesale trade
52	Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods	Retail trade
67	Activities auxiliary to financial intermediation	Activities auxiliary to financial intermediation
72	Computer and related activities	Computer and related activities
74	Other business activities	Other business activities
99	Extra-territorial organisation and bodies	Extra-territorial organisation and bodies

Notes: Source: Office for National Statistics (www.statistics.gov.uk), HMSO.

Table A3.2: Plant Records Disaggregated by Standard Industrial Classification

SIC	Description	Plants (N)	%	SIC	Description	Plants (N)	%
10	Mining of coal and lignite	17	0.28	29	Other machinery and equipment	634	10.37
14	Other mining and quarrying	15	0.25	30	Office machinery and computers	56	0.92
15	Food and beverages	537	8.78	31	Electrical machinery	223	3.65
16	Tobacco products	..	..	32	Radio, T.V. and communications	149	2.44
17	Textiles	198	3.24	33	Medical and optical instruments	171	2.80
18	Wearing apparel	310	5.07	34	Motor vehicles and trailers	173	2.83
19	Leather products	55	0.90	35	Other transport equipment	98	1.60
20	Wood and wood products	293	4.79	36	Furniture	459	7.51
21	Pulp and paper products	154	2.52	37	Recycling	..	..
22	Publishing and printing	320	5.23	51	Wholesale trade	..	..
23	Coke, petroleum and nuclear fuel	15	0.25	52	Retail trade	..	..
24	Chemicals and chemical products	254	4.15	67	Activities auxiliary to financial intermediation	..	..
25	Rubber and plastic products	428	7.00	72	Computer and related activities	..	..
26	Non-metallic mineral products	424	6.93	74	Other business activities	10	0.16
27	Basic metals	252	4.12	99	Extra-territorial organisation and bodies	84	1.37
28	Fabricated metal products	769	12.58				

Notes: Total number of plants = 6,115 plants.
.. Not reported given small sample < 10 plants.

Table A3.3: Plant Records Disaggregated by Country of Ownership and Business Structures

Country of Ownership	Region/Business Structure	Plants (N)	%	Country of Ownership	Region/Business Structure	Plants (N)	%	Country of Ownership	Region/Business Structure	Plants (N)	%
Wales	Status not yet confirmed	37	0.6	Austria	E.U.	..	..	China	Asia	..	..
Wales	Single-plant	2,738	44.8	Belgium	E.U.	11	0.2	Hong Kong	Asia	..	..
Wales	More than 1 plant in Wales	61	1.0	Denmark	E.U.	15	0.2	India	Asia	..	..
Wales	At least 1 plant outside Wales	..	..	Eire	E.U.	29	0.5	Japan	Asia	47	0.8
U.K.	Single-plant in Wales	1,310	21.4	Finland	E.U.	..	..	S Korea	Asia	..	..
U.K.	More than 1 plant in Wales	452	7.4	France	E.U.	33	0.5	Taiwan	Asia	..	..
				Germany	E.U.	66	1.1				
Unknown	Unknown – No classification	51	0.8	Gibraltar	E.U.	..	..	Australia	Other	19	0.3
Unknown	Unknown Classification	853	13.9	Italy	E.U.	16	0.3	Bermuda	Other	..	..
				Luxembourg	E.U.	..	..	Dubai	Other	..	..
Guernsey	Non-E.U.	..	..	Netherlands	E.U.	17	0.3	Ghana	Other	..	..
Isle of Man	Non-E.U.	..	..	Spain	E.U.	..	..	Jamaica	Other	..	..
Lichtenstein	Non-E.U.	..	..	Sweden	E.U.	17	0.3	Kuwait	Other	..	..
Norway	Non-E.U.	..	..					Nigeria	Other	..	..
Switzerland	Non-E.U.	16	0.3	Canada	North America	19	0.3	Panama	Other	..	..
				U.S.	North America	234	3.8	Saudi Arabia	Other	..	..
								South Africa	Other	13	0.2

Notes: Total number of plants = 6,115 plants.
.. Not reported given small sample < 10 plants.

Table A3.4: Summary Statistics: Plants and Employment by Standard Industrial Classification, Selected Years

Sample (SIC)	1966			1972			1977			1986			1997			2003		
	Plants (N)	Emp.	Mean	Plants (N)	Emp.	Mean	Plants (N)	Emp.	Mean	Plants (N)	Emp.	Mean	Plants (N)	Emp.	Mean	Plants (N)	Emp.	Mean
15: Food and beverages	178	15,429	86.7	215	19,051	88.6	270	18,317	67.8	220	17,263	78.5	182	18,711	102.8	154	20,499	133.1
16: Tobacco products	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
17: Textiles	56	8,551	152.7	84	10,903	129.8	88	9,538	108.4	83	4,065	49.0	64	4,647	72.6	55	3,618	65.8
18: Wearing apparel	84	12,789	152.3	116	13,988	120.6	127	13,099	103.1	115	10,752	93.5	61	6,985	114.5	21	1,301	62.0
19: Leather products	24	2,790	116.3	29	2,867	98.9	29	2,062	71.1	18	1,442	80.1	14	1,104	78.9	..	..	..
20: Wood and wood products	92	4,221	45.9	116	4,280	36.9	161	4,182	26.0	106	2,775	26.2	67	2,740	40.9	56	2,824	50.4
21: Pulp and paper products	28	5,538	197.8	42	6,681	159.1	52	6,740	129.6	78	7,079	90.8	74	7,423	100.3	60	6,247	104.1
22: Publishing and printing	92	4,270	46.4	111	5,172	46.6	184	5,387	29.3	138	5,353	38.8	99	6,247	63.1	78	6,510	83.5
23: Coke, petroleum and nuclear fuel	..	..	..	11	4,238	385.3	14	4,469	319.2	10	2,563	256.3	..	..	..	..	..	..
24: Chemicals and chemical products	53	21,190	399.8	70	20,618	294.5	97	18,780	193.6	126	10,861	86.2	133	12,748	95.8	117	12,008	102.6
25: Rubber and plastic products	42	8,604	204.9	87	11,390	130.9	135	14,629	108.4	197	11,777	59.8	192	13,142	68.4	144	11,457	79.6
26: Non-metallic mineral products	129	9,868	76.5	158	10,569	66.9	252	9,035	35.9	170	8,070	47.5	98	5,706	58.2	76	3,598	47.3
27: Basic metals	92	92,074	1,000.8	127	85,153	670.5	148	80,299	542.6	119	29,418	247.2	100	22,603	226.0	71	12,513	176.2
28: Fabricated metal products	173	18,209	105.3	278	21,225	76.3	402	20,343	50.6	310	11,949	38.5	252	13,002	51.6	206	10,450	50.7
29: Other machinery and equipment	142	21,886	154.1	238	25,868	108.7	316	27,062	85.6	239	16,085	67.3	208	12,202	58.7	167	8,636	51.7
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	31	1,886	60.8	25	2,547	101.9	20	2,150	107.5
31: Electrical machinery	41	7,307	178.2	55	9,653	175.5	77	11,096	144.1	97	9,061	93.4	95	12,242	128.9	78	7,245	92.9
32: Radio, T.V. and communications	21	13,133	625.4	34	13,712	403.3	51	12,299	241.2	67	11,551	172.4	70	17,420	248.9	63	10,175	161.5
33: Medical and optical instruments	24	2,672	111.3	46	2,895	62.9	73	4,139	56.7	76	5,546	73.0	80	6,102	76.3	59	5,603	95.0
34: Motor vehicles and trailers	36	15,886	441.3	60	19,919	332.0	85	20,707	243.6	78	14,369	184.2	62	13,978	225.5	55	14,614	265.7
35: Other transport equipment	33	11,730	355.5	38	8,925	234.9	52	8,801	169.3	40	6,816	170.4	30	7,003	233.4	29	9,237	318.5
36: Furniture	74	10,414	140.7	115	12,686	110.3	178	15,103	84.8	163	11,370	69.8	137	10,950	79.9	113	9,403	83.2
All	1,432	292,760	204.4	2,044	312,771	153.0	2,812	308,846	109.8	2,505	203,062	81.1	2,072	200,323	96.7	1,713	165,484	96.6

Notes: Only industrial sectors SIC 15 - 36 are reported, and for selected years. The columns therefore may not sum to the All totals. .. Not reported due to sample < 10 plants. Corresponding data was compiled for each year but is not presented due to space constraints. Emp. = employee numbers.

Table A3.5: Summary Statistics: Plants and Employment by Ownership and Business Structure, Selected Years

Sample (Ownership / Business Structure Classification)	1966			1972			1977			1986			1997			2003		
	Plants (N)	Emp.	Mean	Plants (N)	Emp.	Mean	Plants (N)	Emp.	Mean	Plants (N)	Emp.	Mean	Plants (N)	Emp.	Mean	Plants (N)	Emp.	Mean
Single-plant in Wales	552	24,946	45.2	846	30,418	36.0	987	30,781	31.2	1,226	30,857	25.2	1,080	38,984	36.1	880	33,360	37.9
More than 1 plant in Wales	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
At least 1 plant outside Wales	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
U.K. single-plant in Wales	498	74,492	149.6	654	78,957	120.7	681	71,002	104.3	509	35,217	69.2	323	28,984	89.7	265	25,949	97.9
U.K. more than 1 plant in Wales	208	119,574	574.9	268	119,334	445.3	311	116,700	375.2	303	66,630	219.9	236	52,627	223.0	149	35,444	237.9
Domestic plants	*1,293*	*221,602*	*171.4*	*1,843*	*231,934*	*125.8*	*2,561*	*223,685*	*87.3*	*2,169*	*136,721*	*63.0*	*1,693*	*124,756*	*73.7*	*1,339*	*99,042*	*74.0*
E.U.	44	16,613	377.6	69	24,254	351.5	90	28,557	317.3	128	20,151	157.4	136	20,584	151.4	110	17,421	158.4
Non-E.U. Europe	..	..	..	..	..	..	..	..	..	15	1,525	101.7	20	2,112	105.6	15	1,935	129.0
North America	77	47,802	620.8	105	49,558	472.0	122	48,216	395.2	150	36,172	241.1	158	33,752	213.6	136	28,849	212.1
Asia	..	..	..	..	..	..	..	..	..	19	7,098	373.6	43	17,391	404.4	39	12,287	315.1
Other	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
Foreign owned plants	*139*	*71,158*	*511.9*	*201*	*80,837*	*402.2*	*251*	*85,161*	*339.3*	*336*	*66,341*	*197.4*	*375*	*75,540*	*201.4*	*315*	*62,400*	*198.1*
All	1,432	292,760	204.4	2,044	312,771	153.0	2,812	308,846	109.8	2,505	203,062	81.1	2,072	200,323	96.7	1,713	165,484	96.6

Notes: Plants with no ownership classification or being part of the unknown classification are not included in the table. The columns therefore may not sum to the All totals.
.. Not reported due to sample < 10 plants. Emp. = employee numbers.

Table A3.6: Plant Entry and Exit by Constructed Size-Class

Sample (Size-class: employees)	Plant Entrants	Plant Exits
$0 \leq 1$	161	143
$1 \leq 2$	267	194
$2 \leq 4$	472	311
$4 \leq 8$	783	530
$8 \leq 16$	827	621
$16 \leq 32$	572	620
$32 \leq 64$	303	441
$64 \leq 128$	147	261
+128	92	205
No Size Data	1,015	1,114
All	**4,639**	**4,440**

Notes: Plant entry and exit figures correspond to the year 1967 onwards due to precise plant entry timing identification issues relating to the initial year of the dataset, 1966.
Some larger size-classes have been aggregated to avoid disclosure.

Table A3.7: Plant Entry and Exit by Standard Industrial Classification

Sample (SIC)	Plant Entrants		Plant Exits
15: Food and beverages	353	(537)	387
16: Tobacco products	..	(..)	..
17: Textiles	140	(198)	144
18: Wearing apparel	225	(310)	290
19: Leather products	33	(58)	48
20: Wood and wood products	199	(292)	237
21: Pulp and paper products	124	(154)	94
22: Publishing and printing	226	(318)	241
23: Coke, petroleum and nuclear fuel	..	(15)	..
24: Chemicals and chemical products	201	(254)	139
25: Rubber and plastic products	381	(428)	291
26: Non-metallic mineral products	293	(424)	351
27: Basic metals	159	(252)	181
28: Fabricated metal products	588	(769)	569
29: Other machinery and equipment	484	(634)	470
30: Office machinery and computers	..	(56)	49
31: Electrical machinery	178	(223)	131
32: Radio, T.V. and communications	128	(149)	93
33: Medical and optical instruments	147	(171)	113
34: Motor vehicles and trailers	137	(173)	119
35: Other transport equipment	64	(98)	69
36: Furniture	383	(459)	349
Other Sectors	..	(140)	63
All	4,639	(6,115)	4,440

Notes: Plant entry and exit figures correspond to the year 1967 onwards due to precise plant entry timing identification issues relating to the initial year of the dataset, 1966. Plant entry figures in parentheses are for all plants recorded in the dataset of which some are prior to 1966.

.. Not reported given small sample < 10 plants, or to avoid disclosure.

Table A4.1a: Sample Sizes, 1972 Cross-Section

Sample (SIC)	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984
15: Food and beverages	215	210	201	197	188	183	171	160	153	147	139	134	126
18: Wearing apparel	116	114	109	105	99	90	85	83	77	70	65	62	56
22: Publishing and printing	111	107	105	104	103	101	98	97	96	94	92	89	87
24: Chemicals and chemical products	70	67	67	66	64	64	62	61	61	59	57	52	52
27: Basic metals	127	127	123	122	120	116	111	108	102	92	85	78	76
28: Fabricated metal products	278	275	266	256	247	236	228	223	214	196	178	161	150
32: Radio, T.V. and communications	34	33	33	31	29	29	29	27	26	22	21	20	20
33: Medical and optical instruments	46	46	46	44	44	44	42	40	39	36	34	31	29
34: Motor vehicles and trailers	60	59	58	58	57	55	54	52	52	52	50	49	47
36: Furniture	115	114	110	104	98	94	89	84	78	75	70	62	58
All	2,044	1,999	1,942	1,885	1,809	1,754	1,678	1,612	1,547	1,457	1,367	1,279	1,208

Table A4.1a: Sample Sizes, 1972 Cross-Section, continued

Sample (SIC)	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
15: Food and beverages	124	119	115	112	104	99	97	95	91	86	80	79	76
18: Wearing apparel	53	49	44	44	41	38	36	29	29	25	25	25	24
22: Publishing and printing	82	78	70	69	67	67	66	59	55	53	51	46	43
24: Chemicals and chemical products	51	49	48	48	46	45	44	43	41	39	38	36	35
27: Basic metals	74	70	67	64	62	60	58	58	56	53	53	52	52
28: Fabricated metal products	141	132	125	124	119	118	112	109	101	96	95	93	88
32: Radio, T.V. and communications	20	19	17	17	16	16	16	15	15	15	15	15	15
33: Medical and optical instruments	27	25	24	24	22	22	20	18	18	17	17	17	16
34: Motor vehicles and trailers	47	44	42	42	42	41	40	38	33	33	32	32	31
36: Furniture	57	54	47	47	45	43	43	42	40	36	35	34	33
All	1,159	1,092	1,016	993	954	931	903	859	807	765	748	721	693

Table A4.1b: Sample Sizes, 1977 Cross-section

Sample (SIC)	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986
15: Food and beverages	270	239	201	193	185	175	170	161	150	143
18: Wearing apparel	127	118	108	101	89	82	78	68	65	60
22: Publishing and printing	184	158	126	122	118	115	111	107	98	93
24: Chemicals and chemical products	97	93	86	84	80	78	71	70	68	66
27: Basic metals	148	139	129	121	109	101	93	90	87	81
28: Fabricated metal products	402	370	312	300	278	249	226	210	195	181
32: Radio, T.V. and communications	51	47	44	43	38	37	35	34	34	33
33: Medical and optical instruments	73	66	62	61	57	55	52	50	47	43
34: Motor vehicles and trailers	85	81	70	66	65	61	58	55	54	50
36: Furniture	178	160	133	125	118	106	95	88	86	82
All	2,812	2,548	2,192	2,093	1,963	1,836	1,721	1,624	1,533	1,444

Table A4.1c: Sample Sizes, 1986 Cross-section

Sample (SIC)	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
15: Food and beverages	220	205	198	186	178	171	160	153	143	132	128	125
18: Wearing apparel	115	102	95	89	83	74	62	56	49	48	47	46
22: Publishing and printing	138	121	117	111	111	108	99	92	88	82	76	72
24: Chemicals and chemical products	126	119	116	110	108	106	103	99	96	93	90	88
27: Basic metals	119	111	108	104	100	96	96	90	86	85	83	83
28: Fabricated metal products	310	289	283	268	265	252	237	222	206	202	197	190
32: Radio, T.V. and communications	67	64	64	62	61	60	55	50	48	46	46	46
33: Medical and optical instruments	76	73	70	66	65	60	56	56	53	50	48	46
34: Motor vehicles and trailers	78	74	72	70	68	62	58	50	48	47	46	45
36: Furniture	163	149	143	138	130	123	112	104	94	91	87	86
All	2,505	2,311	2,214	2,112	2,052	1,935	1,796	1,679	1,582	1,521	1,465	1,417

Table A4.2: Sample Sizes, New Entrant Cohorts

Sample (Cohort)	Year 0	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8	Year 9	Year10
1972	121	143	135	130	117	113	105	101	94	89	83
1977	195	100	92	82	75	68	67	64	60	53	47
1986	145	168	145	129	117	111	99	90	85	79	74
1973	128	169	162	151	143	132	124	118	109	104	100
1979	119	153	136	128	120	108	105	94	91	90	89
1988	..	138	130	128	110	99	89	88	85	81	76
1975	472	428	395	351	153	144	139	133	124	118	97
1981	108	149	142	127	112	90	84	79	78	78	72
1992	61	83	81	76	76	71	68	65	62	59	57

Notes: Plant numbers in the initial years may be lower than the reported total new entrants due to the delayed recording of employee numbers from which to generate summary statistics (the 1988 Year 1 sample size is not reported given the small sample size).

Table A5.1a: Plant Size, Growth and Survival: Transition Matrix, 1972-1977

Plants Open in Initial Year			Size of Surviving Plants in Final Year																	
Size-class (Employees)	Plants	Surviving Plants	0 ≤ 1	1 ≤ 2	2 ≤ 4	4 ≤ 8	8 ≤ 16	16 ≤ 32	32 ≤ 64	64 ≤ 128	128 ≤ 256	256 ≤ 512	512 ≤ 1,024	1,024 ≤ 2,048	2,048 ≤ 4,096	4,096 ≤ 8,192	8,192 ≤ 16,384	16,384 ≤ 32,768	Non-Surviving Plants	%
0 ≤ 1	13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1 ≤ 2	27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2 ≤ 4	46	-	-	-	10	11	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4 ≤ 8	131	-	-	-	-	45	38	17	-	-	-	-	-	-	-	-	-	-	-	-
8 ≤ 16	391	337	-	-	-	35	184	94	-	-	-	-	-	-	-	-	-	-	54	13.8
16 ≤ 32	402	327	-	-	-	-	62	180	58	-	-	-	-	-	-	-	-	-	75	18.7
32 ≤ 64	322	266	-	-	-	-	-	49	154	52	-	-	-	-	-	-	-	-	56	17.4
64 ≤ 128	256	211	-	-	-	-	-	11	39	122	36	-	-	-	-	-	-	-	45	17.6
128 ≤ 256	226	200	-	-	-	-	-	-	-	43	124	24	-	-	-	-	-	-	26	11.5
256 ≤ 512	-	-	-	-	-	-	-	-	-	-	31	67	11	-	-	-	-	-	-	-
512 ≤ 1,024	65	-	-	-	-	-	-	-	-	-	-	14	37	-	-	-	-	-	-	-
1,024 ≤ 2,048	30	-	-	-	-	-	-	-	-	-	-	-	-	18	-	-	-	-	-	-
2,048 ≤ 4,096	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4,096 ≤ 8,192	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8,192 ≤ 16,384	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
16,384 ≤ 32,768	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
All	2,044	1,754	-	-	34	109	303	391	276	230	-	-	-	28	13	-	-	-	290	14.2

Notes: Only cells with n ≥ 10 plants reported. Some larger numbers are also suppressed to prevent disclosure.

Table A5.1b: Plant Size, Growth and Survival: Transition Matrix, 1972-1997

Plants Open in Initial Year			Size of Surviving Plants in Final Year																	
Size-class (Employees)	Plants	Surviving Plants	0 ≤ 1	1 ≤ 2	2 ≤ 4	4 ≤ 8	8 ≤ 16	16 ≤ 32	32 ≤ 64	64 ≤ 128	128 ≤ 256	256 ≤ 512	512 ≤ 1,024	1,024 ≤ 2,048	2,048 ≤ 4,096	4,096 ≤ 8,192	8,192 ≤ 16,384	16,384 ≤ 32,768	Non-Surviving Plants	%
0 ≤ 1	13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1 ≤ 2	27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2 ≤ 4	46	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4 ≤ 8	131	35	-	-	-	-	12	-	-	-	-	-	-	-	-	-	-	-	96	73.3
8 ≤ 16	391	116	-	-	-	-	50	33	15	-	-	-	-	-	-	-	-	-	275	70.3
16 ≤ 32	402	108	-	-	-	-	23	42	25	-	-	-	-	-	-	-	-	-	294	73.1
32 ≤ 64	322	84	-	-	-	-	-	19	29	21	-	-	-	-	-	-	-	-	238	73.9
64 ≤ 128	256	106	-	-	-	-	-	15	29	29	23	-	-	-	-	-	-	-	150	58.6
128 ≤ 256	226	94	-	-	-	-	-	-	-	28	38	15	-	-	-	-	-	-	132	58.4
256 ≤ 512	116	56	-	-	-	-	-	-	-	13	21	12	-	-	-	-	-	-	60	51.7
512 ≤ 1,024	65	36	-	-	-	-	-	-	-	-	10	10	-	-	-	-	-	-	29	44.6
1,024 ≤ 2,048	30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2,048 ≤ 4,096	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4,096 ≤ 8,192	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8,192 ≤ 16,384	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
16,384 ≤ 32,768	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
All	2,044	693	-	-	-	-	106	131	125	114	113	56	24	-	-	-	-	-	1,351	66.1

Notes: Only cells with n ≥ 10 plants reported. Some larger numbers are also suppressed to prevent disclosure.

Table A5.1c: Plant Size, Growth and Survival: Transition Matrix, 1977-1986

Plants Open in Initial Year Size-class (Employees)	Plants	Surviving Plants	Size of Surviving Plants in Final Year 0 ≤ 1	1 ≤ 2	2 ≤ 4	4 ≤ 8	8 ≤ 16	16 ≤ 32	32 ≤ 64	64 ≤ 128	128 ≤ 256	256 ≤ 512	512 ≤ 1,024	1,024 ≤ 2,048	2,048 ≤ 4,096	4,096 ≤ 8,192	8,192 ≤ 16,384	16,384 ≤ 32,768	Non-Surviving Plants	%
0 ≤ 1	102	14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	88	86.3
1 ≤ 2	148	33	-	-	12	-	-	-	-	-	-	-	-	-	-	-	-	-	115	77.7
2 ≤ 4	181	45	-	-	-	10	17	-	-	-	-	-	-	-	-	-	-	-	136	75.1
4 ≤ 8	336	114	-	-	-	25	46	26	-	-	-	-	-	-	-	-	-	-	222	66.1
8 ≤ 16	478	272	-	-	-	34	133	65	18	-	-	-	-	-	-	-	-	-	206	43.1
16 ≤ 32	477	264	-	-	-	-	60	118	47	20	-	-	-	-	-	-	-	-	213	44.7
32 ≤ 64	365	199	-	-	-	-	-	47	97	41	-	-	-	-	-	-	-	-	166	45.5
64 ≤ 128	280	182	-	-	-	-	-	13	51	82	23	-	-	-	-	-	-	-	98	35.0
128 ≤ 256	224	151	-	-	-	-	-	-	-	42	71	17	-	-	-	-	-	-	73	32.6
256 ≤ 512	115	84	-	-	-	-	-	-	-	10	35	29	-	-	-	-	-	-	31	27.0
512 ≤ 1,024	59	46	-	-	-	-	-	-	-	-	-	16	18	-	-	-	-	-	13	22.0
1,024 ≤ 2,048	28	-	-	-	-	-	-	-	-	-	-	-	14	-	-	-	-	-	-	-
2,048 ≤ 4,096	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4,096 ≤ 8,192	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8,192 ≤ 16,384	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
16,384 ≤ 32,768	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
All	2,812	1,444	-	14	44	89	276	282	232	207	147	76	50	-	-	-	-	-	1,368	48.6

Notes: Only cells with n ≥ 10 plants reported. Some larger numbers are also suppressed to prevent disclosure.

Table A5.1d: Plant Size, Growth and Survival: Transition Matrix, 1986-1997

Plants Open in Initial Year			Size of Surviving Plants in Final Year																	
Size-class (Employees)	Plants	Surviving Plants	0 ≤ 1	1 ≤ 2	2 ≤ 4	4 ≤ 8	8 ≤ 16	16 ≤ 32	32 ≤ 64	64 ≤ 128	128 ≤ 256	256 ≤ 512	512 ≤ 1,024	1,024 ≤ 2,048	2,048 ≤ 4,096	4,096 ≤ 8,192	8,192 ≤ 16,384	16,384 ≤ 32,768	Non-Surviving Plants	%
0 ≤ 1	16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1 ≤ 2	26	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2 ≤ 4	101	-	-	-	-	-	13	-	-	-	-	-	-	-	-	-	-	-	-	-
4 ≤ 8	240	99	-	-	-	-	50	35	-	-	-	-	-	-	-	-	-	-	141	58.8
8 ≤ 16	585	311	-	-	-	-	129	101	57	-	-	-	-	-	-	-	-	-	274	46.8
16 ≤ 32	528	293	-	-	-	-	-	117	93	37	-	-	-	-	-	-	-	-	235	44.5
32 ≤ 64	376	212	-	-	-	-	-	39	101	49	-	-	-	-	-	-	-	-	164	43.6
64 ≤ 128	284	192	-	-	-	-	-	-	30	91	54	-	-	-	-	-	-	-	92	32.4
128 ≤ 256	185	129	-	-	-	-	-	-	-	35	56	23	-	-	-	-	-	-	56	30.3
256 ≤ 512	-	-	-	-	-	-	-	-	-	-	32	41	-	-	-	-	-	-	-	-
512 ≤ 1,024	-	-	-	-	-	-	-	-	-	-	-	19	16	-	-	-	-	-	-	-
1,024 ≤ 2,048	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2,048 ≤ 4,096	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4,096 ≤ 8,192	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8,192 ≤ 16,384	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
16,384 ≤ 32,768	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
All	2,505	1,417	-	-	-	-	236	312	304	232	170	96	30	11	-	-	-	-	1,088	43.4

Notes: Only cells with n ≥ 10 plants reported. Some larger numbers are also suppressed to prevent disclosure.

Table A5.2: Summary of β_1 Regression Coefficients, Time Series Regimes

Time Period	OLS Regression							Sample Selection Models			
	Sample Plants (N: 1,2)	Model (1)		Model (2)		Correlation Coefficient Model (1)	Correlation Coefficient Model (2)	Plants (N: All)	Plants (N: Survivors)	Maximum Likelihood	Heckman Two-Step
		β_1	t_{β_1}	β_1	t_{β_1}					β_1	β_1
1966 - 1968	1,392	-0.04*	[-7.15]	-0.04*	[-7.01]	-	-	1,432	1,392	-0.04*	-0.04*
1968 - 1971	1,528	-0.06*	[-8.63]	-0.08*	[-8.88]	0.0832	0.0785	1,601	1,528	-0.06*	-0.07*
1971 - 1973	1,848	-0.06*	[-7.73]	-0.07*	[-8.10]	0.0363	0.0306	1,935	1,848	-0.06*	-0.07*
1973 - 1975	2,027	-0.06*	[-9.06]	-0.07*	[-9.40]	-0.0492	-0.0498	2,150	2,027	-0.07*	-0.08*
1975 - 1979	2,002	-0.11*	[-13.48]	-0.12*	[-13.06]	-0.0020	0.0008	2,680	2,002	-0.09*	0.02
1979 - 1982	2,045	-0.12*	[-17.04]	-0.12*	[-16.72]	0.0438	0.0343	2,460	2,045	-0.12*	-0.13*
1982 - 1988	1,741	-0.18*	[-16.20]	-0.20*	[-16.75]	0.0798	0.0840	2,516	1,741	-0.17*	0.19
1988 - 1992	1,925	-0.12*	[-13.56]	-0.12*	[-13.04]	0.0461	0.0489	2,395	1,925	-0.14*	-0.54
1992 - 2000	1,682	-0.19*	[-16.08]	-0.20*	[-15.88]	0.0216	0.0141	2,366	1,682	-0.16*	0.01
2000 - 2003	1,607	-0.06*	[-5.01]	-0.06*	[-5.09]	-0.0112	-0.0092	1,975	1,607	-0.07*	-0.13

Notes: Model (1) is equation (5.2); Model (2) is Model (1) plus industry dummy variables (not reported).
Models (1) and (2) estimated using OLS. All models estimated with heteroscedasticity-consistent standard errors.
Statistics for the constant term, α, and the R^2 and F-statistics relating to the OLS regressions are not reported.
z-statistics not reported for the sample selection models.
Heckman two-step estimates of ρ truncated to -1.00 in all periods except 1975-1979, 1982-1988 and 1992-2000, where ρ is truncated to 1.00.
* Statistically significant at the 5 percent level.

Table A5.3: Summary of β_1 Regression Coefficients, Annual Cross-Sections

Time Period	Sample Plants (N: 1,2)	Sample Plants (N: 3)	Model (1)		Model (2)		Model (3)		Correlation Coefficient Model (1)	Correlation Coefficient Model (2)
			β_1	$t\beta_1$	β_1	$t\beta_1$	β_1	$t\beta_1$		
1966 - 1967	1,412	-	-0.02*	[-6.41]	-0.02*	[-5.85]	-	-	-	-
1967 - 1968	1,467	1,392	-0.02*	[-5.25]	-0.03*	[-5.61]	-0.01*	[-3.98]	0.0249	0.0268
1968 - 1969	1,587	1,453	-0.02*	[-5.21]	-0.03*	[-5.41]	-0.01*	[-3.25]	0.0981	0.0892
1969 - 1970	1,707	1,559	-0.04*	[-5.99]	-0.05*	[-6.39]	-0.02*	[-4.96]	0.0360	0.0301
1970 - 1971	1,815	1,670	-0.04*	[-7.04]	-0.04*	[-7.40]	-0.03*	[-6.61]	0.0493	0.0542
1971 - 1972	1,891	1,773	-0.04*	[-6.68]	-0.04*	[-6.69]	-0.03*	[-5.42]	-0.0467	-0.0427
1972 - 1973	1,999	1,848	-0.03*	[-6.23]	-0.04*	[-6.52]	-0.02*	[-4.50]	-0.0627	-0.0730
1973 - 1974	2,089	1,942	-0.04*	[-6.31]	-0.04*	[-6.87]	-0.02*	[-4.46]	-0.0597	-0.0653
1974 - 1975	2,177	2,027	-0.04*	[-8.61]	-0.05*	[-8.48]	-0.03*	[-7.29]	-0.0508	-0.0526
1975 - 1976	2,525	2,086	-0.05*	[-9.56]	-0.05*	[-9.46]	-0.05*	[-8.72]	-0.0587	-0.0593
1976 - 1977	2,604	2,418	-0.01*	[-3.38]	-0.02*	[-4.55]	-0.01*	[-2.47]	-0.2945	-0.2930
1977 - 1978	2,548	2,452	-0.04*	[-9.87]	-0.04*	[-9.36]	-0.03*	[-7.14]	-0.4066	-0.4106
1978 - 1979	2,304	2,192	-0.04*	[-8.14]	-0.05*	[-8.39]	-0.03*	[-7.16]	-0.0509	-0.0461
1979 - 1980	2,347	2,194	-0.04*	[-9.95]	-0.04*	[-9.34]	-0.04*	[-9.33]	0.0452	0.0438
1980 - 1981	2,338	2,187	-0.07*	[-13.10]	-0.07*	[-13.46]	-0.05*	[-12.40]	0.0451	0.0426
1981 - 1982	2,323	2,187	-0.05*	[-10.58]	-0.05*	[-11.01]	-0.04*	[-9.14]	-0.0024	-0.0040
1982 - 1983	2,367	2,180	-0.05*	[-9.84]	-0.06*	[-9.90]	-0.04*	[-7.70]	0.0260	0.0241
1983 - 1984	2,346	2,208	-0.04*	[-9.46]	-0.05*	[-9.50]	-0.03*	[-7.44]	-0.0127	-0.0186
1984 - 1985	2,340	2,182	-0.04*	[-8.03]	-0.04*	[-8.04]	-0.03*	[-5.75]	0.0384	0.0390
1985 - 1986	2,308	2,142	-0.05*	[-9.21]	-0.06*	[-9.57]	-0.03*	[-6.78]	0.0239	0.0228

Table A5.3: Summary of β_1 Regression Coefficients, Annual Cross-Sections, continued

Time Period	Sample Plants (N: 1,2)	Sample Plants (N: 3)	Model (1) β_1	Model (1) t_{β_1}	Model (2) β_1	Model (2) t_{β_1}	Model (3) β_1	Model (3) t_{β_1}	Correlation Coefficient Model (1)	Correlation Coefficient Model (2)
1986 - 1987	2,311	2,143	-0.06*	[-9.20]	-0.06*	[-9.41]	-0.04*	[-7.65]	-0.0142	-0.0237
1987 - 1988	2,357	2,214	-0.02*	[-5.18]	-0.02*	[-5.80]	-0.01*	[-3.70]	0.1581	0.1515
1988 - 1989	2,280	2,244	-0.03*	[-6.40]	-0.03*	[-6.42]	-0.03*	[-6.60]	0.0506	0.0454
1989 - 1990	2,427	2,211	-0.04*	[-7.62]	-0.04*	[-7.99]	-0.02*	[-5.81]	0.0490	0.0463
1990 - 1991	2,419	2,294	-0.05*	[-9.61]	-0.05*	[-9.61]	-0.04*	[-8.58]	0.1177	0.1155
1991 - 1992	2,299	2,216	-0.05*	[-8.46]	-0.05*	[-7.77]	-0.05*	[-8.03]	0.0017	-0.0021
1992 - 1993	2,199	2,136	-0.05*	[-11.31]	-0.06*	[-11.16]	-0.05*	[-10.86]	-0.0513	-0.0529
1993 - 1994	2,135	2,070	-0.05*	[-8.20]	-0.05*	[-8.35]	-0.05*	[-8.01]	-0.1170	-0.1167
1994 - 1995	2,103	2,047	-0.03*	[-7.11]	-0.03*	[-7.08]	-0.03*	[-6.94]	-0.0389	-0.0365
1995 - 1996	2,067	2,020	-0.02*	[-5.49]	-0.03*	[-5.67]	-0.02*	[-4.77]	-0.0260	-0.0300
1996 - 1997	2,039	1,987	-0.03*	[-6.93]	-0.03*	[-6.83]	-0.02*	[-5.85]	-0.0777	-0.0723
1997 - 1998	1,987	1,958	-0.03*	[-4.44]	-0.03*	[-4.11]	-0.02*	[-3.69]	-0.1019	-0.1024
1998 - 1999	1,962	1,948	-0.03*	[-5.64]	-0.03*	[-5.20]	-0.03*	[-5.91]	-0.0522	-0.0506
1999 - 2000	1,911	1,893	-0.01*	[-2.93]	-0.01*	[-2.61]	-0.01*	[-3.51]	0.2014	0.2018
2000 - 2001	1,834	1,781	-0.05*	[-6.50]	-0.05*	[-6.17]	-0.04*	[-6.01]	-0.0367	-0.0357
2001 - 2002	1,737	1,697	-0.01*	[-2.63]	-0.02*	[-3.04]	-0.01	[-1.90]	-0.0374	-0.0481
2002 - 2003	1,692	1,646	-0.01*	[-2.13]	-0.02*	[-2.53]	-0.01	[-1.52]	-0.0755	-0.0731

Notes: Model (1) is equation (5.2); Model (2) is Model (1) plus industry dummy variables (not reported), Model (3) is Model (1) with additional one-year persistence term.
Models estimated using OLS with heteroscedasticity-consistent standard errors.
Model (3) is not estimated for 1966-1967 due to the absence of data for 1965 (persistence term not available).
Statistics for the constant term, α, and the R^2 and F-statistics not reported.
* Statistically significant at the 5 percent level.

Table A5.4: Summary of β_1 Regression Coefficients, Surviving Plants, Annual Cross-Sections

Time Period	1972-1997				1972-1997			
	Model (1)		Model (2)		Model (1)		Model (2)	
	β_1	t_{β_1}	β_1	t_{β_1}	β_1	t_{β_1}	β_1	t_{β_1}
1972 - 1973	-0.03*	[-5.66]	-0.03*	[-5.89]	-0.02*	[-2.83]	-0.03*	[-3.11]
1973 - 1974	-0.02*	[-4.33]	-0.03*	[-4.67]	-0.02*	[-3.14]	-0.02*	[-3.05]
1974 - 1975	-0.03*	[-6.48]	-0.03*	[-6.55]	-0.02*	[-4.43]	-0.03*	[-4.42]
1975 - 1976	-0.04*	[-7.07]	-0.04*	[-6.85]	-0.03*	[-5.07]	-0.03*	[-4.83]
1976 - 1977	-0.01*	[-3.04]	-0.02*	[-3.38]	-0.01	[-1.46]	-0.01	[-1.72]
1977 - 1978					-0.03*	[-3.93]	-0.02*	[-3.30]
1978 - 1979					-0.03*	[-3.71]	-0.03*	[-3.86]
1979 - 1980					-0.03*	[-4.77]	-0.03*	[-4.49]
1980 - 1981					-0.05*	[-8.41]	-0.05*	[-7.80]
1981 - 1982					-0.03*	[-5.52]	-0.03*	[-5.42]
1982 - 1983					-0.02*	[-3.80]	-0.02*	[-3.54]
1983 - 1984					-0.02*	[-3.07]	-0.02*	[-3.12]
1984 - 1985					-0.01*	[-2.48]	-0.02*	[-2.76]
1985 - 1986					-0.03*	[-6.37]	-0.03*	[-5.72]
1986 - 1987					-0.03*	[-3.17]	-0.03*	[-2.77]
1987 - 1988					-0.01*	[-3.23]	-0.02*	[-3.24]
1988 - 1989					-0.02*	[-2.45]	-0.02*	[-2.44]
1989 - 1990					-0.01*	[-2.85]	-0.02*	[-2.70]
1990 - 1991					-0.03*	[-7.11]	-0.03*	[-6.80]
1991 - 1992					-0.03*	[-4.18]	-0.03*	[-3.27]
1992 - 1993					-0.03*	[-6.07]	-0.03*	[-5.68]
1993 - 1994					-0.03*	[-4.28]	-0.03*	[-3.84]
1994 - 1995					-0.01	[-1.93]	-0.02*	[-2.21]
1995 - 1996					-0.01*	[-2.58]	-0.02*	[-2.69]
1996 - 1997					-0.01*	[-2.38]	-0.01*	[-2.37]

Table A5.4: Summary of β_1 Regression Coefficients, Surviving Plants, Annual Cross-Sections, continued

Time Period	1977-1986				1986-1997			
	Model (1)		Model (2)		Model (1)		Model (2)	
	β_1	t_{β_1}	β_1	t_{β_1}	β_1	t_{β_1}	β_1	t_{β_1}
1977 - 1978	-0.03*	[-6.78]	-0.04*	[-6.43]				
1978 - 1979	-0.03*	[-6.35]	-0.03*	[-6.62]				
1979 - 1980	-0.03*	[-7.05]	-0.03*	[-6.62]				
1980 - 1981	-0.04*	[-11.10]	-0.05*	[-11.40]				
1981 - 1982	-0.03*	[-8.15]	-0.03*	[-7.65]				
1982 - 1983	-0.02*	[-5.26]	-0.02*	[-4.97]				
1983 - 1984	-0.02*	[-4.31]	-0.02*	[-4.18]				
1984 - 1985	-0.01*	[-2.66]	-0.02*	[-3.35]				
1985 - 1986	-0.02*	[-4.01]	-0.02*	[-4.58]				
1986 - 1987					-0.05*	[-8.07]	-0.05*	[-7.93]
1987 - 1988					-0.02*	[-6.00]	-0.02*	[-5.85]
1988 - 1989					-0.02*	[-4.62]	-0.03*	[-4.34]
1989 - 1990					-0.02*	[-5.58]	-0.03*	[-5.79]
1990 - 1991					-0.03*	[-8.66]	-0.03*	[-8.69]
1991 - 1992					-0.04*	[-6.79]	-0.04*	[-6.13]
1992 - 1993					-0.05*	[-8.54]	-0.05*	[-8.32]
1993 - 1994					-0.04*	[-6.03]	-0.04*	[-6.07]
1994 - 1995					-0.02*	[-5.07]	-0.03*	[-4.95]
1995 - 1996					-0.01*	[-3.98]	-0.02*	[-4.22]
1996 - 1997					-0.01*	[-3.63]	-0.01*	[-3.27]

Notes: Model (1) is equation (5.2); Model (2) is Model (1) plus industry dummy variables (not reported).
Models estimated using OLS with heteroscedasticity-consistent standard errors.
Samples sizes were: 1,754 (1972-1997), 693 (1972-1997), 1,444 (1977-1986), and 1,417 (1986-1997) plants.
Statistics for the constant term, α, and the R^2 and F-statistics not reported.
* Statistically significant at the 5 percent level.

Table A5.5: Summary of β_1 Regression Coefficients for Different Time-Spans, Surviving Plants

Time-Span	1972-1997				1977-1986				1986-1997			
	Model (1)		Model (2)		Model (1)		Model (2)		Model (1)		Model (2)	
	β_1	t_{β_1}	β_1	t_{β_1}	β_1	t_{β_1}	β_1	t_{β_1}	β_1	t_{β_1}	β_1	t_{β_1}
$T_1 - t_0$	-0.03*	[-6.23]	-0.04*	[-6.52]	-0.04*	[-9.87]	-0.04*	[-9.36]	-0.06*	[-9.20]	-0.06*	[-11.89]
$T_2 - t_0$	-0.05*	[-7.20]	-0.06*	[-7.55]	-0.06*	[-11.16]	-0.07*	[-10.74]	-0.07*	[-9.54]	-0.08*	[-11.74]
$T_3 - t_0$	-0.07*	[-9.03]	-0.08*	[-9.15]	-0.09*	[-13.91]	-0.09*	[-13.32]	-0.10*	[-10.91]	-0.11*	[-13.38]
$T_4 - t_0$	-0.10*	[-11.16]	-0.11*	[-11.19]	-0.13*	[-17.44]	-0.14*	[-16.93]	-0.12*	[-11.69]	-0.13*	[-14.83]
$T_5 - t_0$	-0.11*	[-11.08]	-0.12*	[-11.16]	-0.16*	[-18.40]	-0.16*	[-17.89]	-0.15*	[-13.47]	-0.16*	[-16.92]
$T_6 - t_0$	-0.12*	[-11.70]	-0.13*	[-11.47]	-0.17*	[-18.47]	-0.17*	[-16.99]	-0.18*	[-14.70]	-0.18*	[-16.94]
$T_7 - t_0$	-0.14*	[-12.56]	-0.15*	[-12.34]	-0.18*	[-17.97]	-0.19*	[-16.75]	-0.20*	[-16.20]	-0.21*	[-18.65]
$T_8 - t_0$	-0.15*	[-13.68]	-0.16*	[-13.20]	-0.20*	[-18.52]	-0.20*	[-17.36]	-0.24*	[-17.76]	-0.25*	[-20.48]
$T_9 - t_0$	-0.19*	[-15.54]	-0.20*	[-15.00]	-0.22*	[-18.61]	-0.24*	[-18.15]	-0.24*	[-17.99]	-0.26*	[-20.65]
$T_{10} - t_0$	-0.22*	[-16.35]	-0.22*	[-15.88]					-0.26*	[-18.46]	-0.27*	[-21.40]
$T_{11} - t_0$	-0.21*	[-16.16]	-0.22*	[-15.02]					-0.27*	[-19.71]	-0.29*	[-22.49]
$T_{12} - t_0$	-0.23*	[-16.74]	-0.24*	[-15.46]								
$T_{13} - t_0$	-0.25*	[-16.63]	-0.26*	[-15.79]								
$T_{14} - t_0$	-0.27*	[-16.41]	-0.29*	[-15.68]								
$T_{15} - t_0$	-0.30*	[-17.46]	-0.32*	[-17.01]								
$T_{16} - t_0$	-0.30*	[-17.29]	-0.32*	[-16.84]								
$T_{17} - t_0$	-0.32*	[-17.03]	-0.34*	[-16.78]								
$T_{18} - t_0$	-0.33*	[-17.23]	-0.35*	[-17.20]								
$T_{19} - t_0$	-0.34*	[-18.12]	-0.36*	[-17.89]								
$T_{20} - t_0$	-0.34*	[-17.93]	-0.36*	[-17.13]								
$T_{21} - t_0$	-0.36*	[-18.57]	-0.37*	[-17.63]								
$T_{22} - t_0$	-0.38*	[-20.19]	-0.40*	[-19.07]								
$T_{23} - t_0$	-0.39*	[-20.05]	-0.40*	[-19.22]								
$T_{24} - t_0$	-0.39*	[-20.30]	-0.41*	[-19.66]								
$T_{25} - t_0$	-0.41*	[-21.58]	-0.43*	[-20.15]								

Notes: Model (1) is equation (5.2); Model (2) is Model (1) plus industry dummy variables (dummy variables not reported).
All models estimated using OLS with heteroscedasticity-consistent standard errors. Statistics for the constant term, α, and the R^2 and F-statistics not reported.
Model estimates for the time period 1972-1977 are provided by the corresponding 1972-1997 estimates. * Statistically significant at the 5 percent level.

Table A5.6: Linear Regression Diagnostics, Five-Year Plant Growth Persistence Models

Sample	1972				1977				1986			
	White	RESET	D'Agostino	S-Francia	White	RESET	D'Agostino	S-Francia	White	RESET	D'Agostino	S-Francia
All	0.95	11.31*	1,431.68*	0.61*	2.28	7.30*	2,386.84*	0.42*	5.65	11.15*	3,991.07*	0.06*
Age: 0 - 1 years	0.10	4.38*	807.76*	0.58*	0.11	0.24	167.67*	0.57*	3.20	3.52*	194.85*	0.51*
Age: 2 - 5 years	1.41	4.34*	91.44*	0.83*	2.64	1.59	552.11*	0.29*	1.08	0.66	138.20*	0.80*
Age: 6 - 10 years	0.49	4.60*	76.33*	0.80*	7.53*	9.32*	631.83*	0.66*	0.05	0.41	364.84*	0.51*
Age: 11 - 20 years	12.10*	23.71*	313.49*	0.51*	6.99*	9.49*	196.77*	0.58*	9.55*	14.50*	1,341.44*	0.16*
Age: +20 years	0.44	2.93*	33.05*	0.88*	12.03*	18.30*	231.30*	0.63*	26.19*	99.36*	560.12*	0.11*
Size: Micro	1.09	0.94	86.08*	0.62*	2.44	0.95	131.86*	0.61*	2.07	1.38	233.47*	0.43*
Size: Small	3.93	10.58*	699.80*	0.68*	3.23	11.88*	1,397.26*	0.34*	7.16*	14.65*	1,928.54*	0.23*
Size: Medium	13.53*	10.63*	114.89*	0.89*	21.43*	10.44*	96.61*	0.92*	1.12	1.64	187.56*	0.83*
Size: Large	8.94*	4.97*	51.14*	0.92*	15.82*	0.44	6.86*	0.98*	26.23*	126.12*	343.42*	0.16*

Table A5.6: Regression Diagnostics, Five-Year Plant Growth Persistence Models, continued

Sample	1972				1977				1986			
	White	RESET	D'Agostino	S-Francia	White	RESET	D'Agostino	S-Francia	White	RESET	D'Agostino	S-Francia
15: Food and beverages	0.13	0.39	258.61*	0.41*	2.51	3.47*	169.05*	0.58*	0.18	0.38	91.13*	0.78*
16: Tobacco products	..	..	..	..	..	..	..	..	..	..	..	..
17: Textiles	0.21	1.74	1.84	0.97	4.98	5.78*	98.21*	0.41*	0.43	1.21	15.24*	0.89*
18: Wearing apparel	25.22*	5.66*	59.49*	0.75*	13.19*	9.03*	58.00*	0.73*	0.73	0.46	18.23*	0.92*
19: Leather products	1.82	0.38	5.71	0.91	2.07	1.15	4.80	0.90	4.92	4.58*	1.47	0.92
20: Wood and wood products	1.97	2.76*	55.32*	0.81*	6.92*	3.02*	23.69*	0.91*	4.47	1.46	57.38*	0.80*
21: Pulp and paper products	0.79	0.22	0.50	0.97	8.64*	4.65*	5.71	0.93*	3.86	0.95	7.59*	0.95
22: Publishing and printing	0.06	0.21	141.72*	0.38*	0.03	0.88	162.98*	0.42*	3.07	1.82	76.75*	0.74*
23: Coke, petroleum and nuclear fuel	4.33	4.46	0.31	0.94	0.55	0.22	13.72*	0.79*	..	..	..	..
24: Chemicals and chemical products	11.79*	0.50	34.33*	0.75*	0.39	0.89	12.73*	0.91*	0.05	0.77	22.23*	0.88*
25: Rubber and plastic products	0.17	0.52	52.09*	0.65*	1.53	2.53	130.93*	0.21*	2.85	4.70*	127.27*	0.65*
26: Non-metallic mineral products	7.65*	8.24*	25.64*	0.93*	10.49*	4.04*	87.84*	0.69*	0.17	0.02	144.29*	0.34*
27: Basic metals	0.14	0.58	53.66*	0.86*	0.07	14.49*	38.65*	0.87*	2.31	0.72	61.00*	0.77*
28: Fabricated metal products	2.86	2.52	158.81*	0.67*	0.07	0.40	208.05*	0.66*	1.89	3.04*	140.23*	0.77*
29: Other machinery and equipment	1.81	4.90*	58.51*	0.82*	0.01	1.03	147.06*	0.70*	6.83*	10.74*	331.81*	0.15*
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	..	..	..
31: Electrical machinery	6.66*	4.77*	23.27*	0.85*	8.21*	1.78	12.32*	0.89*	9.26*	25.46*	102.02*	0.23*
32: Radio, T.V. and communications	1.26	0.43	13.53*	0.87*	1.50	0.34	0.76	0.95	0.13	0.63	23.47*	0.76*
33: Medical and optical instruments	0.98	0.01	19.55*	0.79*	0.26	0.22	49.86*	0.67*	0.54	1.00	73.93*	0.57*
34: Motor vehicles and trailers	5.01	1.75	16.72*	0.84*	1.98	2.93*	79.62*	0.41*	1.04	0.15	22.08*	0.88*
35: Other transport equipment	13.36*	21.74*	55.24*	0.55*	5.64	0.50	5.39	0.91*	1.36	1.67	16.92*	0.76*
36: Furniture	3.84	2.30	98.93*	0.55*	2.12	0.88	136.80*	0.34*	0.59	1.63	51.05*	0.81*

Notes: D'Agostino et al. test and Shapiro-Francia tests applied to test for normality of the regression residuals.
* Statistically significant at the 5 percent level.

Table A5.7: Linear Regression Diagnostics, Heteroscedasticity (White's Test)

Sample	1972-1977			1972-1997			1977-1986			1986-1997		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
All	51.44*	68.17*	71.90*	6.22*	8.77	5.82	10.78*	13.14*	22.98*	49.42*	64.46*	226.78*
Size: Micro	13.92*	18.69*	30.40*	3.60	5.14	13.75	4.89	8.59	9.84	24.81*	42.98*	24.92*
Size: Small	4.34	19.97*	31.94*	0.54	5.60	16.83	12.05*	15.11*	21.90*	1.24	7.52	9.62
Size: Medium	0.24	14.81*	25.23*	1.73	5.90	14.99	1.43	5.07	8.67	2.50	6.66	15.37
Size: Large	2.19	10.77	11.17	0.12	2.23	3.58	4.46	15.38*	50.16*	0.69	2.48	6.61
Minimum efficient scale	2.41	6.37	7.65	1.06	2.21	6.80	2.80	4.30	9.36	0.37	3.67	11.69
Single-plant firm	46.35*	43.47*	34.16*	1.76	5.78	16.84	7.25*	11.65*	7.67	19.36*	32.80*	61.27*
Foreign owned	10.86*	18.41*	20.31	5.53	9.98	12.00	2.26	7.33	11.98	49.73*	113.13*	61.43*

Table A5.7: Linear Regression Diagnostics, Heteroscedasticity (White's Test), continued

Sample	1972-1977			1972-1997			1977-1986			1986-1997		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
15: Food and beverages	12.64*	14.41*	17.35	0.76	2.41	10.33	5.26	9.37	22.16*	1.31	2.23	22.42*
16: Tobacco products	..	..	..	..	..	..	..	..	..	..	..	..
17: Textiles	7.62*	8.16	7.53	2.31	6.73	13.47	8.41*	8.89	17.10	4.33	14.57*	14.93
18: Wearing apparel	4.65	6.98	9.39	1.31	4.55	11.34	3.91	17.80*	19.43	0.31	1.86	9.31
19: Leather products	1.25	6.25	18.14	..	..	..	1.48	10.78	13.65	0.36	1.61	11.00
20: Wood and wood products	0.91	1.08	5.28	2.24	1.56	6.87	0.76	6.82	9.60	0.05	1.90	7.44
21: Pulp and paper products	13.87*	16.83*	17.55	10.46*	11.02	14.27	2.12	9.06	11.35	3.73	14.09*	18.69
22: Publishing and printing	13.69*	34.16*	43.10*	2.30	16.80*	19.96	11.84*	11.07	32.26*	1.96	4.11	22.90*
23: Coke, petroleum and nuclear fuel	1.62	10.51	10.90	..	..	..	0.83	6.92	10.00	..	..	..
24: Chemicals and chemical products	5.62	8.98	11.88	5.23	7.65	16.87	0.42	2.40	5.04	1.31	4.31	14.44
25: Rubber and plastic products	4.74	6.86	11.28	3.01	5.32	13.31	2.62	3.82	8.69	1.60	3.19	5.84
26: Non-metallic mineral products	4.53	6.26	12.38	0.34	4.86	14.67	2.82	3.22	7.99	15.55*	16.24*	17.59
27: Basic metals	3.49	7.94	25.28*	3.30	1.54	13.70	1.14	2.63	5.97	1.63	17.06*	40.17*
28: Fabricated metal products	28.79*	33.76*	30.87*	1.72	1.47	5.28	4.61	6.51	18.58	0.14	0.60	15.02
29: Other machinery and equipment	8.78*	14.68*	18.76	2.93	4.81	12.28	1.01	1.81	6.77	5.99	8.83	16.30
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	2.52	6.34	14.73
31: Electrical machinery	3.79	9.34	5.66	0.60	5.40	17.28	0.65	6.61	33.06*	21.99*	36.57*	26.55*
32: Radio, T.V. and communications	0.02	0.53	6.80	1.94	7.77	14.66	1.22	4.97	12.17	0.98	2.55	8.91
33: Medical and optical instruments	2.06	5.01	12.75	1.36	7.41	10.10	1.30	5.93	10.62	4.00	5.80	10.92
34: Motor vehicles and trailers	5.54	10.03	22.56*	2.76	7.95	13.13	1.40	12.06*	9.48	3.52	3.97	11.17
35: Other transport equipment	24.63*	28.60*	11.99	6.75*	9.64	14.68	0.36	1.70	6.28	3.46	5.54	9.61
36: Furniture	0.12	11.25*	17.22	5.14	7.04	14.66	5.71	6.09	16.55	0.06	3.37	17.59

Notes: Model (1) is equation (5.2); Model (2) is equation (5.4); Model (3) is equation (5.6).
* Statistically significant at the 5 percent level.

Table A5.8: Linear Regression Diagnostics, Heteroscedasticity (Breusch-Pagan Test)

Sample	1972-1977			1972-1997			1977-1986			1986-1997		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
All	108.42*	169.53*	176.19*	3.87*	693	0.23	23.26*	24.66*	27.02*	26.21*	27.26*	21.89*
Size: Micro	12.25*	7.16*	7.63*	3.51	58	3.54	5.77*	7.87*	3.73	32.00*	30.25*	16.09*
Size: Small	0.28	37.23*	21.91*	0.66	308	3.56	21.60*	20.75*	22.97*	0.09	8.42*	7.05*
Size: Medium	1.19	33.50*	29.60*	1.47	200	0.05	0.12	2.67	1.09	3.00	0.24	0.03
Size: Large	5.97*	6.05*	1.51	0.01	127	0.45	27.63*	0.08	5.41*	0.01	2.15	2.37
Minimum efficient scale	11.57*	16.39*	9.73*	0.56	92	1.21	3.77	4.29*	4.27*	0.23	0.63	0.90
Single-plant firm	53.46*	45.41*	63.12*	2.24	266	6.06*	5.16*	5.04*	4.13*	0.56	2.10	1.79
Foreign owned	25.25*	31.52*	40.63*	3.31	126	1.01	14.09*	3.17	1.59	34.34*	29.93*	35.97*

Table A5.8: Linear Regression Diagnostics, Heteroscedasticity (Breusch-Pagan Test), continued

Sample	1972-1977			1972-1997			1977-1986			1986-1997		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
15: Food and beverages	7.79*	5.88*	16.46*	0.23	76	1.74	7.85*	7.57*	5.10*	1.34	0.60	1.02
16: Tobacco products	..	..	..	..	..	..	..	..	..	..	..	..
17: Textiles	9.55*	9.62*	6.56*	0.39	22	0.19	9.86*	5.45*	0.04	1.47	1.50	0.32
18: Wearing apparel	7.72*	5.15*	0.21	0.06	24	0.00	4.10*	8.57*	7.49*	0.00	0.56	2.79
19: Leather products	1.15	0.16	1.02	..	..	..	1.08	0.86	0.09	0.00	0.12	1.48
20: Wood and wood products	1.54	1.20	0.03	1.31	24	0.07	0.96	1.43	2.22	0.02	0.01	0.22
21: Pulp and paper products	25.66*	25.68*	2.08	3.50	22	0.33	2.03	2.94	5.65*	2.97	8.03*	3.72
22: Publishing and printing	24.14*	16.11*	25.32*	0.71	43	2.57	8.44*	12.19*	13.19*	0.73	0.77	1.01
23: Coke, petroleum and nuclear fuel	2.42	1.95	2.46	..	..	..	0.80	0.87	0.98	..	..	..
24: Chemicals and chemical products	13.15*	13.03*	15.82*	1.26	35	1.16	1.27	1.44	0.48	0.05	0.85	0.00
25: Rubber and plastic products	11.98*	15.94*	11.04*	5.23*	38	0.21	7.54*	7.56*	7.96*	3.37	3.24	3.81
26: Non-metallic mineral products	2.57	0.15	1.67	0.25	43	0.02	0.07	0.09	0.45	11.95*	9.62*	3.47
27: Basic metals	14.05*	26.21*	6.91*	2.76	52	0.89	2.47	2.12	2.21	1.15	2.33	0.21
28: Fabricated metal products	15.81*	16.16*	26.42*	0.02	88	0.28	0.00	0.01	0.13	0.00	0.20	0.89
29: Other machinery and equipment	7.35*	11.19*	19.37*	2.83	73	3.57	1.58	0.14	0.37	0.48	1.92	0.17
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	0.08	0.73	0.06
31: Electrical machinery	4.36*	7.24*	1.89	0.36	24	0.27	3.97*	5.36*	24.29*	14.18*	15.70*	2.46
32: Radio, T.V. and communications	0.00	0.22	1.48	0.00	15	1.47	0.07	0.14	0.06	0.75	0.14	0.15
33: Medical and optical instruments	2.18	3.35	4.20*	0.13	16	0.06	0.10	0.04	0.75	4.74*	8.13*	8.34*
34: Motor vehicles and trailers	10.13*	0.01	0.04	1.77	31	1.65	1.76	2.65	0.89	2.70	0.21	0.16
35: Other transport equipment	18.96*	21.33*	2.69	1.34	16	0.20	0.40	0.20	0.03	1.08	0.21	0.02
36: Furniture	0.63	2.10	1.70	0.61	33	0.00	7.53*	4.43*	4.39*	0.04	0.44	0.62

Notes: Model (1) is equation (5.2); Model (2) is equation (5.4); Model (3) is equation (5.6).
* Statistically significant at the 5 percent level.

Table A5.9: Linear Regression Diagnostics, Omitted Variables (Ramsey's RESET Test)

Sample	1972-1977			1972-1997			1977-1986			1986-1997		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
All	33.11*	35.87*	6.97*	15.08*	11.29*	5.67*	12.24*	4.90*	0.26	44.14*	11.93*	13.15*
Size: Micro	0.22	0.19	0.38	0.27	0.40	0.21	1.99	2.12	5.02*	4.22*	3.38*	6.24*
Size: Small	2.17	2.64*	0.63	0.75	1.07	0.51	1.26	3.72*	5.90*	1.44	2.89*	2.27
Size: Medium	0.37	2.42	1.28	0.18	0.57	0.88	3.66*	3.84*	2.41	2.25	3.98*	1.18
Size: Large	0.22	2.47	2.68*	0.92	0.80	0.38	0.29	0.72	1.26	0.41	1.18	1.12
Minimum efficient scale	0.61	0.63	1.16	0.73	0.44	0.06	1.14	0.39	1.01	0.60	1.41	3.95*
Single-plant firm	29.05*	31.87*	5.78*	9.75*	9.52*	5.16*	14.08*	10.09*	2.16	16.83*	13.14*	5.58*
Foreign owned	1.04	2.86*	1.68	0.57	0.78	1.33	0.75	1.00	1.15	12.37*	1.29	10.88*

Table A5.9: Linear Regression Diagnostics, Omitted Variables (Ramsey's RESET Test), continued

Sample	1972-1977			1972-1997			1977-1986			1986-1997		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
15: Food and beverages	1.63	1.80	0.91	2.39	2.34	3.42*	8.82*	1.56	3.38*	3.36*	2.62	0.81
16: Tobacco products	..	..	..	..	..	..	..	..	..	..	..	..
17: Textiles	4.32*	4.40*	3.97*	0.72	0.82	0.47	5.63*	2.43	0.47	2.87*	2.47	2.24
18: Wearing apparel	1.74	6.99*	0.89	0.29	0.14	0.50	0.56	0.10	0.72	2.08	0.92	2.62
19: Leather products	0.24	1.21	1.36	..	..	..	0.99	1.56	2.60	1.31	1.10	1.11
20: Wood and wood products	3.32*	1.58	1.69	0.77	1.25	0.26	1.19	0.93	1.41	1.42	0.83	0.02
21: Pulp and paper products	6.97*	6.80*	2.39	4.19*	3.56*	0.10	3.54*	3.06*	1.50	2.43	2.52	1.46
22: Publishing and printing	2.13	2.18	2.80*	4.25*	1.81	0.90	6.39*	2.77*	1.20	8.97*	5.99*	1.42
23: Coke, petroleum and nuclear fuel	2.65	12.18*	9.37*	..	..	..	1.60	0.82	9.84	..	..	..
24: Chemicals and chemical products	0.90	1.02	1.14	2.82	2.43	1.64	1.55	0.79	0.35	3.53*	3.94*	2.60
25: Rubber and plastic products	0.54	0.89	0.83	0.89	1.24	3.70*	0.81	0.81	1.35	1.31	1.45	1.49
26: Non-metallic mineral products	2.50	3.59*	1.67	2.80	2.29	1.14	1.34	3.14*	0.27	4.95*	5.37*	2.71
27: Basic metals	2.44	1.95	1.50	1.62	1.56	1.39	0.53	0.95	2.37	2.22	1.78	0.20
28: Fabricated metal products	7.39*	10.34*	3.58*	2.49	1.15	0.81	2.52	2.87*	2.44	1.30	1.09	2.54
29: Other machinery and equipment	3.47*	3.79*	1.75	3.23*	2.30	0.51	1.70	2.96*	1.41	8.26*	5.26*	3.21*
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	0.15	0.37	2.04
31: Electrical machinery	1.64	1.83	1.44	0.32	0.20	0.36	3.84*	2.33	8.50*	13.42*	12.86*	9.33*
32: Radio, T.V. and communications	4.52*	1.21	1.20	0.31	0.40	0.99	1.12	0.76	1.07	0.62	0.35	0.16
33: Medical and optical instruments	0.26	0.62	2.00	0.47	0.27	1.43	2.81	1.18	1.99	3.51*	0.58	0.23
34: Motor vehicles and trailers	0.44	0.46	1.64	0.88	0.27	0.42	2.88*	1.80	0.70	0.64	0.91	1.74
35: Other transport equipment	8.38*	10.53*	4.61*	3.01	2.80	0.76	1.16	0.88	1.54	1.16	0.19	1.31
36: Furniture	5.27*	5.10*	1.52	8.02*	7.99*	3.44*	4.11*	1.17	0.55	0.75	0.40	1.02

Notes: Model (1) is equation (5.2); Model (2) is equation (5.4); Model (3) is equation (5.6).
* Statistically significant at the 5 percent level.

Table A5.10: Linear Regression Diagnostics, Normality of Residuals (D'Agostino et al. Test)

Sample	1972-1977			1972-1997			1977-1986			1986-1997		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
All.	239.56*	299.82*	333.87*	4.77	4.32	5.59	178.37*	197.49*	215.34*	99.97*	140.27*	86.83*
Size: Micro	0.63	0.75	0.60	8.00*	8.23*	9.02*	1.49	1.74	1.68	7.41*	14.36*	12.38*
Size: Small	150.30*	186.38*	182.04*	20.95*	18.04*	19.86*	45.61*	46.12*	47.02*	51.04*	47.40*	50.04*
Size: Medium	213.41*	218.15*	221.78*	1.97	3.81	2.62	80.60*	96.08*	90.20*	7.79*	10.25*	9.89*
Size: Large	58.47*	60.29*	63.60*	0.26	0.26	0.20	140.93*	133.50*	127.36*	13.87*	14.29*	15.10*
Minimum efficient scale	95.18*	94.38*	96.32*	1.42	1.71	0.96	12.86*	14.61*	15.73*	16.77*	15.08*	14.65*
Single-plant firm	88.94*	108.38*	166.89*	12.84*	10.12*	12.29*	26.76*	28.63*	37.80*	31.09*	18.26*	9.38*
Foreign owned	20.09*	20.22*	24.32*	1.02	0.91	1.03	116.36*	126.60*	123.30*	29.82*	65.75*	21.01*

Table A5.10: Linear Regression Diagnostics, Normality of Residuals (D'Agostino et al. Test), continued

Sample	1972-1977			1972-1997			1977-1986			1986-1997		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
15: Food and beverages	26.53*	26.32*	26.70*	3.27	3.13	2.30	16.38*	13.48*	11.28*	6.56*	2.44	3.37
16: Tobacco products	..	..	..	..	..	..	..	..	..	..	..	..
17: Textiles	6.98*	7.06*	6.64*	4.86	0.71	0.27	6.67*	8.60*	5.37	0.86	1.37	2.87
18: Wearing apparel	5.84	5.94	7.20*	1.60	0.69	1.39	4.72	2.10	2.58	9.47*	7.03*	6.97*
19: Leather products	2.10	0.36	3.77	..	..	..	2.43	2.18	5.36	2.20	1.89	2.09
20: Wood and wood products	14.16*	17.25*	16.74*	0.18	0.72	0.34	11.55*	11.98*	11.84*	0.07	0.10	0.61
21: Pulp and paper products	11.58*	11.58*	25.56*	0.93	1.14	4.10	5.26	1.99	6.12*	6.05*	2.12	2.56
22: Publishing and printing	14.41*	13.55*	11.41*	3.00	3.03	1.37	5.32	1.61	0.87	15.74*	22.65*	14.51*
23: Coke, petroleum and nuclear fuel	16.26*	14.22*	12.17*	..	..	..	4.89	6.15*	1.22	..	..	..
24: Chemicals and chemical products	12.32*	12.35*	10.99*	1.04	0.96	0.71	33.99*	39.57*	38.65*	0.79	2.35	3.31
25: Rubber and plastic products	12.06*	13.00*	11.92*	9.95*	3.56	3.32	23.47*	23.36*	19.40*	16.56*	16.72*	17.56*
26: Non-metallic mineral products	35.07*	44.14*	47.39*	0.41	0.48	1.32	3.05	3.05	0.85	8.46*	4.08	2.94
27: Basic metals	27.05*	24.88*	26.50*	1.06	0.67	0.20	20.11*	31.20*	30.43*	0.54	2.44	1.24
28: Fabricated metal products	15.06*	10.58*	8.05*	2.26	3.48	3.97	1.82	3.27	3.78	2.92	3.90	4.35
29: Other machinery and equipment	12.35*	14.18*	22.09*	1.69	3.75	2.51	13.34*	12.85*	12.44*	9.40*	1.97	1.02
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	2.58	2.06	0.87
31: Electrical machinery	8.99*	4.27	6.65*	1.30	1.70	1.89	51.28*	48.43*	42.58*	26.19*	19.73*	2.11
32: Radio, T.V. and communications	1.50	1.36	1.99	1.01	0.63	0.98	0.41	3.67	3.56	0.28	0.28	1.08
33: Medical and optical instruments	6.51*	6.50*	3.27	2.16	1.47	0.80	0.64	0.96	0.85	5.53	5.59	5.04
34: Motor vehicles and trailers	13.33*	5.89	6.38*	0.36	0.42	0.41	2.58	5.27	3.18	0.39	0.34	1.23
35: Other transport equipment	11.25*	5.72	1.73	0.93	1.58	1.10	6.77*	6.63*	8.68*	3.25	0.61	0.24
36: Furniture	64.44*	67.58*	70.24*	0.19	0.23	2.98	8.79*	12.05*	11.74*	0.28	0.46	0.33

Notes: Model (1) is equation (5.2); Model (2) is equation (5.4); Model (3) is equation (5.6).
* Statistically significant at the 5 percent level.

Table A5.11: Linear Regression Diagnostics, Normality of Residuals (Shapiro-Francia Test)

Sample	1972-1977			1972-1997			1977-1986			1986-1997		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
All	0.94*	0.93*	0.93*	1.00	1.00	1.00	0.96*	0.96*	0.96*	0.98*	0.98*	0.98*
Size: Micro	0.99	0.99	0.99	0.93*	0.93*	0.93*	0.99	0.99	0.99	0.96*	0.96*	0.97*
Size: Small	0.93*	0.92*	0.92*	0.97*	0.98*	0.98*	0.97*	0.97*	0.97*	0.98*	0.98*	0.98*
Size: Medium	0.86*	0.86*	0.85*	0.99	0.99	0.99	0.93*	0.92*	0.92*	0.99*	0.99*	0.99*
Size: Large	0.90*	0.90*	0.90*	1.00	0.99	0.99	0.81*	0.82*	0.83*	0.97*	0.97*	0.97*
Minimum efficient scale	0.83*	0.83*	0.84*	0.99	0.99	0.99	0.98*	0.97*	0.97*	0.97*	0.97*	0.97*
Single-plant firm	0.95*	0.94*	0.94*	0.98*	0.98*	0.98*	0.98*	0.98*	0.97*	0.98*	0.99*	0.99*
Foreign owned	0.95*	0.95*	0.94*	0.98	0.98	0.98	0.87*	0.85*	0.86*	0.96*	0.92*	0.96*

Table A5.11: Linear Regression Diagnostics, Normality of Residuals (Shapiro-Francia Test), continued

Sample	1972-1977			1972-1997			1977-1986			1986-1997		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
15: Food and beverages	0.92*	0.91*	0.92*	0.98	0.98	0.99	0.95*	0.95*	0.95*	0.98	0.99	0.98
16: Tobacco products	..	..	..	..	..	..	..	..	..	..	..	..
17: Textiles	0.94*	0.94*	0.94*	0.96	0.99	0.97	0.95	0.95*	0.97	0.99	0.98	0.97
18: Wearing apparel	0.96*	0.96*	0.96*	0.97	0.97	0.95	0.95*	0.97	0.97	0.95*	0.95*	0.95*
19: Leather products	0.97	0.98	0.92	..	..	..	0.92	0.91	0.89	0.91	0.92	0.87
20: Wood and wood products	0.95*	0.95*	0.95*	0.98	0.97	0.98	0.95*	0.94*	0.95*	0.99	0.99	0.98
21: Pulp and paper products	0.83*	0.83*	0.81*	0.96	0.97	0.94	0.95	0.97	0.95	0.93*	0.98	0.97
22: Publishing and printing	0.91*	0.91*	0.92*	0.96	0.97	0.98	0.97	0.98	0.98	0.94*	0.93*	0.95*
23: Coke, petroleum and nuclear fuel	0.75*	0.79*	0.80*	..	..	..	0.81*	0.83*	0.92	..	..	..
24: Chemicals and chemical products	0.91*	0.91*	0.92*	0.98	0.98	0.99	0.87*	0.84*	0.84*	0.99	0.97	0.98
25: Rubber and plastic products	0.91*	0.91*	0.91*	0.94*	0.96	0.96	0.93*	0.93*	0.95*	0.95*	0.95*	0.94*
26: Non-metallic mineral products	0.91*	0.90*	0.91*	0.99	0.99	0.98	0.98	0.98	0.99	0.97*	0.98	0.98
27: Basic metals	0.87*	0.86*	0.86*	0.97	0.98	0.98	0.93*	0.91*	0.91*	0.99	0.98	0.99
28: Fabricated metal products	0.98*	0.98*	0.98*	0.98	0.98	0.98	0.99	0.99	0.99	0.99	0.99*	0.98*
29: Other machinery and equipment	0.97*	0.97*	0.95*	0.99	0.98	0.97	0.97*	0.98*	0.97*	0.97*	0.99	0.99
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	0.93	0.93	0.96
31: Electrical machinery	0.93*	0.97	0.96	0.97	0.96	0.96	0.74*	0.77*	0.80*	0.91*	0.93*	0.98
32: Radio, T.V. and communications	0.96	0.95	0.96	0.97	0.98	0.97	0.98	0.96	0.97	0.98	0.99	0.98
33: Medical and optical instruments	0.96	0.96	0.97	0.93	0.95	0.97	0.99	0.98	0.99	0.94*	0.96	0.97
34: Motor vehicles and trailers	0.88*	0.94*	0.95*	0.98	0.98	0.98	0.97	0.96	0.97	0.99	0.99	0.98
35: Other transport equipment	0.91*	0.93*	0.97	0.95	0.94	0.97	0.88*	0.89*	0.90*	0.96	0.99	0.98
36: Furniture	0.81*	0.80*	0.80*	0.97	0.97	0.96	0.95*	0.94*	0.94*	0.99	0.99	0.99

Notes: Model (1) is equation (5.2); Model (2) is equation (5.4); Model (3) is equation (5.6).
* Statistically significant at the 5 percent level.

Table A5.12: Linear Regression Diagnostics, Normality of Residuals (Skewness Indices)

Sample	1972-1977			1972-1997			1977-1986			1986-1997		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
All	-0.37	-0.56	-0.65	0.20	0.18	0.20	-0.45	-0.54	-0.59	0.47	0.48	0.37
Size: Micro	0.10	0.15	0.13	0.89	0.90	0.93	-0.04	-0.06	-0.03	0.42	0.63	0.55
Size: Small	-0.47	-0.68	-0.65	0.63	0.55	0.56	0.15	0.01	0.00	0.53	0.45	0.46
Size: Medium	-1.97	-2.02	-2.05	-0.22	-0.26	-0.20	-1.07	-1.27	-1.19	-0.30	-0.37	-0.38
Size: Large	-1.08	-1.12	-1.23	-0.09	-0.06	-0.05	-2.82	-2.66	-2.52	-0.59	-0.63	-0.59
Minimum efficient scale	-1.92	-1.91	-1.99	-0.10	-0.12	-0.08	-0.50	-0.56	-0.58	-0.62	-0.56	-0.55
Single-plant firm	-0.43	-0.53	-0.99	0.48	0.41	0.45	-0.05	-0.08	-0.17	0.48	0.37	0.25
Foreign owned	0.31	0.08	-0.05	-0.20	-0.05	-0.12	-1.90	-2.11	-2.08	0.47	0.87	0.17

Table A5.12: Linear Regression Diagnostics, Normality of Residuals (Skewness Indices), continued

Sample	1972-1977			1972-1997			1977-1986			1986-1997		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
15: Food and beverages	-0.26	-0.23	-0.43	-0.07	-0.06	0.02	0.41	0.41	0.25	0.41	0.18	0.20
16: Tobacco products	..	..	..	..	..	..	..	..	..	..	..	..
17: Textiles	-0.08	-0.08	-0.22	0.09	0.21	0.01	-0.68	-0.87	-0.62	-0.04	-0.20	-0.26
18: Wearing apparel	0.05	-0.05	0.10	0.43	0.34	0.01	0.13	-0.27	-0.20	-0.85	-0.79	-0.71
19: Leather products	-0.31	-0.22	-0.54	..	..	..	-0.73	-0.72	-1.00	-0.81	-0.74	-0.78
20: Wood and wood products	-0.68	-0.79	-0.79	-0.17	-0.36	-0.01	-0.66	-0.71	-0.73	-0.06	0.00	0.12
21: Pulp and paper products	-0.15	-0.15	-1.50	0.31	0.32	0.64	0.80	0.51	0.81	0.81	0.35	0.42
22: Publishing and printing	-0.05	0.07	-0.25	0.57	0.34	-0.11	0.55	0.21	0.19	0.76	0.94	0.82
23: Coke, petroleum and nuclear fuel	-1.79	-1.66	-1.53	..	..	..	1.15	1.22	0.60	..	..	..
24: Chemicals and chemical products	0.40	0.41	0.31	0.04	0.05	0.08	-1.59	-1.81	-1.78	0.03	0.13	0.20
25: Rubber and plastic products	0.04	-0.27	-0.22	-0.96	-0.58	-0.56	0.90	0.90	0.83	0.49	0.50	0.52
26: Non-metallic mineral products	-1.00	-1.28	-1.40	0.03	-0.23	-0.12	-0.30	-0.33	-0.18	0.61	0.38	0.32
27: Basic metals	0.40	0.21	0.30	0.19	0.13	0.10	-0.73	-1.22	-1.22	-0.17	-0.30	0.00
28: Fabricated metal products	0.38	0.19	-0.20	0.33	0.45	0.48	0.11	0.04	-0.01	0.30	0.35	0.37
29: Other machinery and equipment	-0.10	-0.20	-0.41	-0.27	-0.40	-0.11	-0.50	-0.53	-0.51	0.56	0.27	0.20
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	0.73	0.67	0.47
31: Electrical machinery	0.96	0.66	0.69	0.33	0.34	0.35	-2.57	-2.45	-2.19	1.34	1.11	0.42
32: Radio, T.V. and communications	-0.45	-0.40	-0.54	-0.37	-0.12	0.05	-0.10	-0.32	-0.27	-0.14	-0.16	-0.23
33: Medical and optical instruments	0.79	0.70	0.50	0.20	0.24	-0.06	0.27	0.31	0.27	0.27	-0.48	-0.58
34: Motor vehicles and trailers	-0.49	-0.17	-0.21	0.19	-0.06	0.05	-0.21	-0.24	-0.53	-0.15	-0.18	-0.32
35: Other transport equipment	0.86	0.56	-0.19	-0.03	-0.04	-0.35	-1.09	-1.08	-1.14	0.02	-0.21	-0.06
36: Furniture	-2.09	-2.20	-2.37	0.11	0.01	-0.44	-0.33	-0.47	-0.44	-0.11	0.15	0.12

Notes: Model (1) is equation (5.2); Model (2) is equation (5.4); Model (3) is equation (5.6).
* Statistically significant at the 5 percent level.

Table A5.13: Linear Regression Diagnostics, Normality of Residuals (Kurtosis Indices)

Sample	1972-1977			1972-1997			1977-1986			1986-1997		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
All	8.05	8.66	8.98	2.97	3.11	3.15	6.85	6.90	7.11	4.59	5.56	4.65
Size: Micro	2.77	2.90	2.93	3.21	3.26	3.37	3.36	3.39	3.40	3.81	4.32	4.28
Size: Small	8.52	9.04	8.97	3.50	3.58	3.73	5.26	5.43	5.47	3.97	4.17	4.25
Size: Medium	13.72	14.01	14.33	3.11	3.36	3.33	6.65	7.03	6.93	3.48	3.50	3.43
Size: Large	6.97	6.97	6.66	2.80	2.74	2.76	21.83	20.02	18.79	4.20	4.12	4.41
Minimum efficient scale	10.64	10.60	10.32	2.47	2.45	2.53	4.30	4.33	4.43	4.06	4.07	4.06
Single-plant firm	6.89	7.57	7.86	3.57	3.53	3.64	4.75	4.84	5.25	3.56	3.34	3.30
Foreign owned	5.73	6.26	7.06	3.03	3.29	3.27	15.14	16.08	15.18	5.69	8.75	5.56

Table A5.13: Linear Regression Diagnostics, Normality of Residuals (Kurtosis Indices), continued

Sample	1972-1977			1972-1997			1977-1986			1986-1997		
	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)	Model (1)	Model (2)	Model (3)
15: Food and beverages	7.23	7.26	6.58	2.25	2.26	2.32	5.41	4.88	4.94	3.77	3.50	3.67
16: Tobacco products	..	..	..	..	..	..	..	..	..	..	..	..
17: Textiles	5.08	5.10	4.82	1.70	2.21	2.30	4.11	4.11	3.87	2.37	2.34	2.19
18: Wearing apparel	4.56	4.59	4.86	3.16	2.75	3.57	4.49	3.47	3.76	4.34	3.75	4.03
19: Leather products	3.64	2.79	3.90	..	..	..	2.96	2.85	3.68	2.56	2.59	2.14
20: Wood and wood products	5.47	6.00	5.82	2.59	2.58	3.02	4.97	4.91	4.78	2.86	2.93	3.16
21: Pulp and paper products	7.86	7.87	9.50	2.24	2.18	3.85	3.33	2.73	3.65	2.91	3.31	3.30
22: Publishing and printing	6.59	6.34	5.51	3.05	3.80	3.51	3.17	3.31	3.06	5.62	6.93	5.00
23: Coke, petroleum and nuclear fuel	6.00	5.53	5.09	..	..	..	3.13	3.51	2.11	..	..	..
24: Chemicals and chemical products	6.30	6.28	6.08	2.24	2.25	2.31	7.70	8.56	8.34	3.29	3.66	3.80
25: Rubber and plastic products	6.49	6.42	6.20	4.49	3.41	3.39	6.79	6.76	5.92	5.96	5.99	6.11
26: Non-metallic mineral products	6.56	6.87	6.81	3.12	2.73	2.28	3.34	3.24	3.01	4.20	3.73	3.55
27: Basic metals	9.12	9.14	9.37	3.27	3.18	2.93	6.73	7.44	7.22	2.97	3.37	3.44
28: Fabricated metal products	4.38	4.34	4.03	2.58	3.11	3.14	3.39	3.69	3.78	2.93	2.97	2.90
29: Other machinery and equipment	4.92	5.03	5.61	3.26	3.58	3.78	4.43	4.25	4.25	3.51	3.01	2.90
30: Office machinery and computers	..	..	..	..	..	..	..	..	..	3.03	2.89	2.45
31: Electrical machinery	3.66	3.14	3.93	3.22	3.41	3.49	15.36	14.09	11.98	7.25	6.05	2.59
32: Radio, T.V. and communications	2.40	2.36	2.43	2.09	2.05	1.94	2.40	1.99	1.98	2.93	2.84	2.41
33: Medical and optical instruments	3.68	3.98	3.44	1.80	1.92	2.01	2.75	2.59	3.00	4.73	4.30	3.80
34: Motor vehicles and trailers	6.73	4.86	4.97	2.87	3.09	3.09	3.80	4.63	3.21	2.55	2.86	3.07
35: Other transport equipment	5.45	4.34	2.18	1.98	1.86	2.07	3.36	3.35	4.01	1.79	2.23	2.31
36: Furniture	12.38	12.91	12.50	2.87	2.99	3.67	4.98	5.42	5.46	2.97	2.97	2.98

Notes: Model (1) is equation (5.2); Model (2) is equation (5.4); Model (3) is equation (5.6).
* Statistically significant at the 5 percent level.

Table A5.14: Linear Regression Diagnostics, Serial Correlation Models

Sample		Serial Correlation Model					Serial Correlation Model With Age				
		White	Breusch-Pagan	RESET	D'Agostino	S-Francia	White	Breusch-Pagan	RESET	D'Agostino	S-Francia
One-Year:	1972-1977	61.06*	33.23*	26.99*	285.59*	0.93*	68.21*	60.28*	37.65*	331.32*	0.93*
	1972-1997	13.87*	4.72*	14.79*	5.59	1.00	15.07	4.04*	14.49*	3.75	1.00
	1977-1986	34.88*	14.22*	10.86*	193.13*	0.96*	31.62*	16.96*	4.18*	209.02*	0.96*
	1986-1997	441.16*	13.28*	25.88*	111.21*	0.98*	353.33*	24.27*	12.14*	84.09*	0.98*
Two-Year:	1972-1977	40.05*	55.09*	12.44*	245.32*	0.93*	50.85*	79.43*	20.30*	269.94*	0.93*
	1972-1997	7.79	3.43	11.92*	6.33*	1.00	8.98	3.42	12.17*	5.39	1.00
	1977-1986	12.74*	20.12*	9.00*	200.28*	0.95*	18.04*	23.88*	4.01*	209.47*	0.95*
	1986-1997	391.81*	13.74*	29.75*	103.23*	0.98*	331.39*	24.61*	19.18*	81.83*	0.98*
Three-Year:	1972-1977	51.78*	45.97*	11.72*	186.12*	0.94*	54.67*	53.27*	15.08*	191.16*	0.94*
	1972-1997	8.46	2.58	11.17*	6.76*	0.99*	10.33	2.71	11.58*	6.35*	0.99
	1977-1986	18.43*	11.44*	13.36*	209.81*	0.95*	23.41*	11.73*	11.52*	217.58*	0.95*
	1986-1997	450.69*	12.24*	20.82*	122.47*	0.98*	397.18*	24.54*	17.00*	98.16*	0.98*
Four-Year:	1972-1977	71.79*	34.57*	13.43*	192.57*	0.93*	75.08*	36.67*	14.54*	193.90*	0.93*
	1972-1997	16.57*	3.15	7.50*	7.04*	0.99	18.09*	3.26	7.50*	6.97*	0.99
	1977-1986	16.05*	16.00*	13.09*	227.07*	0.94*	19.74*	12.93*	13.16*	227.12*	0.94*
	1986-1997	418.68*	12.70*	20.39*	132.83*	0.97*	392.61*	20.08*	17.35*	105.97*	0.98*
Five-Year:	1972-1977	68.37*	23.55*	11.26*	191.96*	0.93*	75.83*	22.54*	10.29*	191.31*	0.93*
	1972-1997	21.06*	2.27	7.78*	9.69*	0.99*	23.21*	2.25	7.77*	9.67*	0.99*
	1977-1986	19.86*	14.80*	11.47*	224.98*	0.94*	27.15*	14.44*	11.86*	224.67*	0.94*
	1986-1997	407.44*	10.93*	31.48*	116.76*	0.97*	376.23*	15.84*	18.20*	102.56*	0.98*

Notes: Serial Correlation Model (1) is equation (5.3); Serial Correlation Model With Age is equation (5.5).
D'Agostino et al. test and Shapiro-Francia tests applied to test for normality of the regression residuals.
* Statistically significant at the 5 percent level.

BIBLIOGRAPHY

Aaronovitch, S. and Sawyer, M. (1975), "Mergers, Growth and Concentration", Oxford Economic Papers, 26, pp.136-155.

Acs, Z.J. and Audretsch, D.B. (1987), "Innovation, market structure and firm size", Review of Economics and Statistics, 69, Nov., pp.567-575.

Acs, Z.J. and Audretsch, D.B. (1988), "Innovation in large and small firms: An empirical analysis", American Economic Review, 78, Sept., pp.678-690.

Acs, Z.J. and Audretsch, D.B. (1990), *Innovation and Small Firms*, Cambridge, MA: The MIT Press.

Adams, M. and Hardwick, P. (2001), "Firm Size and Growth in the United Kingdom Life Insurance Industry", European Business Management School Working Paper EBMS/2001/5, University of Wales, Swansea.

Adelman, I.G. (1958), "A Stochastic Analysis of the Size Distribution of Firms", Journal of the American Statistical Association, 53, Issue 284, pp.893-904.

Adelman, I. and Preston, L.E. (1960), "A Note on Changes in Industry Structure", Review of Economics and Statistics, 42, 1, pp.105-108.

Aislabie, C.J. (1971), "Further Evidence on the Size and Growth of Firms", Economic Record, 47, pp.230-244.

Aitchison, J. and Brown, J.A.C. (1957), *The Lognormal Distribution*, Cambridge University Press, Cambridge.

Akaike, H. (1973), "Information Theory and an Extension of the Maximum Likelihood Principal", in Petrov, B. and Csake, F. (eds.) *Second International Symposium on Information Theory*, Budapest: Akademiai Kiado, pp.267-281.

Alexander, S. (1949), "The Effect of Size of Manufacturing Corporations on the Distribution of the Rate of Return", Review of Economics and Statistics, 31, pp.229-235.

Almus, M. and Nerlinger, E.A. (2000), "Testing Gibrat's Law for Young Firms – Empirical Results for West Germany", Small Business Economics, 15, 1, pp.1-12.

Altman, E. (1968), "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", Journal of Finance, 23, 4, pp.589-609.

Amemiya, T. (1984), "Tobit Models: A Survey", Journal of Econometrics, 24, (1/2), pp.3-61.

Amaral, L.A.N., Buldyrev, S.V., Havlin, S., Leschhorn, H., Maass, P., Salinger, M.A., Stanley, H.E., Stanley, M.H. (1997), "Scaling Behaviour in Economics: I. Empirical Results for Company Growth", Journal of Physics, 7, pp.621-633.

Anderson, T.W. and Hsiao, C. (1981), "Estimation of Dynamic Models with Error Components", Journal of the American Statistical Association, 76, pp.598-606.

Aspin, A.A. (1948), "An Examination and Further Development of a Formula Arising in the Problem of Comparing Two Mean Values", Biometrika, 35, pp.88-96.

Arellano, M. and Bond, S. (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", Review of Economic Studies, 55, pp.277-294.

Artis, M. (2002), "Dating the Business Cycle in Britain", Centre for Growth & Business Cycle Research Discussion Paper No. 017, School of Economic Studies, University of Manchester.

Artis, M. (2003), "Analysis of European and U.K. business cycles and shocks", HM Treasury EMU Study, HM Treasury, London.

Amaral, L.A.N., Gopikrishnan, P., Plerou, V. and Stanley, H.E. (2001), "A model for the growth dynamics of economic organizations", Physica A, 299, pp.127-136.

Audretsch, D. (1991), "New Firm Survival and the Technological Regime", Review of Economics and Statistics, 73, 3, pp.441-449.

Audretsch, D.B. (1995a), *Innovation and Industry Evolution*, Cambridge, MIT Press.

Audretsch, D.B. (1995b), "Innovation, growth and survival", International Journal of Industrial Organisation, 13, 4, pp.441-450.

Audretsch, D.B., Klomp, L. and Thurik, A.R. (2002), "Gibrat's Law: Are the Services Different?" Erasmus Research Institute of Management Report Series, No. ERS-2-2-04-STR, Rotterdam School of Economics, Erasmus University, Rotterdam.

Audretsch, D.B. and Mahmood, T. (1991), "The hazard rate of new plants: A first report", Economics Letters, 36, pp.409-412.

Audretsch, D. and Mahmood, T. (1994), "The Rate of Hazard Confronting New Firms and Plants in U.S. Manufacturing", Review of Industrial Organisation, 9, pp.41-56.

Audretsch, D. and Mahmood, T. (1995), "New Firm Survival: New Results Using a Hazard Function", Review of Economics and Statistics, 77, 1, pp.97-103.

Audretsch, D. and Mata, J. (1995), (eds.), "The post-entry performance of firms: Introduction", International Journal of Industrial Organisation, 14, pp.413-419.

Audretsch, D.B., Santarelli, E. and Vivarelli, M. (1999), "Start-up size and industrial dynamics: some evidence from Italian manufacturing", International Journal of Industrial Organisation, 17, 7, pp.965-983.

Axtell, R. (1999), "The Emergence of Firms in a Population of Agents: Local Increasing Returns, Unstable Nash Equilibria, and Power Law Size Distributions", Centre on Social and Economic Dynamics, Working Paper No. 3, June, The Brookings Institution, Washington, D.C.

Axtell, R.L. (2001), "Zipf Distribution of U.S. Firm Sizes", Science, 293, pp.1818-1820.

Axtell, R.L. (2004), "Firm Sizes: Facts, Formulae, Fables and Fantasies", Working Paper, Centre of Social and Economic Dynamics, The Brookings Institution, Washington, D.C.

Baldwin, J, and Gorecki, P. (1991), "Firm entry and exit in the Canadian manufacturing sector", Canadian Journal of Economics, 24, pp.300-323.

Baldwin, J.R. and Rafiquzzaman, M. (1995), "Selection versus evolutionary adaptation and post-entry performance", International Journal of Industrial Organisation, 13, 4, pp.501-522.

Baltagi, B.H. (2001), *Econometric Analysis of Panel Data*, 2nd Edition, John Wiley & Sons, Ltd., Chichester, England.

Banerjee, A. (1999), "Panel Data, Unit Roots and Cointegration: An Overview", Oxford Bulletin of Economics and Statistics, Special Issue, 61, November, pp.607-629.

Bank of England (2001), *Financing of Technology-Based Small Firms*, Domestic Finance Division, Bank of England, London, February 2001.

Barro, R.J. (1991), "Economic Growth in a Cross-Section of Countries", Quarterly Journal of Economics, 106, pp.407-443.

Barro, R.J. and Sala-i-Martin, X. (1991), "Convergence Across States and Regions", Brookings Papers on Economic Activity, 1, pp.107-182.

Barro, R.J. and Sala-i-Martin, X. (1992), "Convergence", Journal of Political Economy, 106, pp.407-443.

Barro, R.J. and Sala-i-Martin, X. (1995), *Economic Growth*, McGraw-Hill: New York.

Bassett, G. and Köenker, R. (1982), "An empirical quantile function for linear models with i.i.d. errors", Journal of the American Statistical Association, 77, pp.407-415.

Baxter, M. and King, R.G. (1999), "Measuring business cycles: approximate band-pass filters for economic time series", Review of Economics and Statistics, 81, pp.575-593.

Bell, D. and Ritchie, F. (1996), "Time-varying parameters in panel models", Working Paper No. 96/11, Department of Economics, University of Stirling.

Berkson, J. (1938), "Some difficulties of interpretation encountered in the application of the Chi-square test", Journal of the American Statistical Association, Vol. 33, pp.526-542.

Bernard, A. and Durlauf, S. (1995), "Convergence in International Output", Journal of Applied Econometrics, 10, pp.97-108.

Bernard, A.B. and Durlauf, S.N. (1996), "Interpreting tests of the convergence hypothesis", Journal of Econometrics, 71, pp.161-173.

Bernard, A.B. and Jensen, J.B. (2002), "The Deaths of Manufacturing Plants", Working Paper CES-WP-02-15, Centre for Economic Studies, U.S. Bureau of the Census.

Bernard, A. and Jones, C. (1996), "Productivity Across Industries and Countries: Time Series Theory and Evidence", The Review of Economics and Statistics, 78, pp.135-146.

Bhattacharjee, A., Higson, C., Holly, S. and Kattuman, P. (2002), "Macro Economic Instability and Business Exit: Determinant of Failures and Acquisitions of Large U.K. Firms", Working Paper WP-ACCT034, London Business School.

Birch, D.L. (1979), *Job Creation in America: How Our Smallest Companies Put Most People to Work*, The Free Press, New York.

Blair, J.M. (1972), Economic Concentration: Structure, Behaviour, and Public Policy, New York: Harcourt, Brace, Jovanovich.

Bliss, C. (1999), "Galton's fallacy and economic convergence", Oxford Economic Papers, 51, pp.4-14.

Blonigen, B.A. and Tomlin, K. (1999), "Size and Growth of Japanese Plants in the United States", NBER Working Paper No. W7275, National Bureau of Economic Research.

Boeri, T. and Bellman, L. (1995), "Post-entry behaviour and the Cycle: Evidence from Germany", International Journal of Industrial Organisation, 13, 4, pp.483-500.

Bond, S., Nauges, C. and Windmeijer, F. (2002), "Unit Roots and Identification in Autoregressive Panel Data Models: A Comparison of Alternative Tests", Discussion Paper, Institute for Fiscal Studies, London.

Bottazzi, G. (2001), "Firm Diversification and the Law of Proportionate Effect", LEM Working Paper Series No. 2001/01, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa, January 2001.

Bottazzi, G. (2004), "SUBBOTOOLS User's Manual", LEM Working Paper Series No. 2004/14, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa, September 2004.

Bottazzi, G. (2007a), "On the Pareto Type III Distribution", LEM Working Paper Series No. 2007/08, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa, March 2007.

Bottazzi, G. (2007b), "On the Irreconcilability of Pareto and Gibrat Laws", LEM Working Paper Series No. 2007/10, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa, March 2007.

Bottazzi, G., Coad, A., Jacoby, N. and Secchi, A. (2005), "Corporate Growth and Industrial Dynamics: Evidence from French Manufacturing", Working Paper 2005/21, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa.

Bottazzi, G., Cefis, E. and Dosi, G. (2002), "Corporate Growth and Industrial Structures: Some Evidence from the Italian Manufacturing Industry", LEM Working Paper Series No. 2001/08, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa, April 2002.

Bottazzi, G., Cefis, E., Dosi, G. and Secchi, A. (2006), "Invariances and Diversities in the Patterns of Industrial Evolution: Some Evidence from Italian Manufacturing Industries", unpublished paper, February 2006.

Bottazzi, G., Dosi, G., Lippi, N., Pammolli, F. and Riccaboni, M. (2001), "Innovation and corporate growth in the evolution of the drug industry", International Journal of Industrial Organisation, 19, pp.1161-1187.

Bottazzi, G. and Secchi, A. (2002), "On the Laplace Distribution of Firms Growth Rates", Working Paper 2002/20, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa.

Bottazzi, G. and Secchi, A. (2003a), "Sectoral Specificities in the Dynamics of U.S. Manufacturing Plants", Working Paper 2003/18, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa.

Bottazzi, G. and Secchi, A. (2003b), "Why are distributions of firm growth rates tent-shaped?", Economic Letters, 80, pp.415-420.

Bottazzi, G. and Secchi, A. (2005), "Explaining the Distribution of Firms Growth Rates", LEM Working Paper Series No. 2005/16, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa, June 2005.

Box, G. and Cox, G. (1964), "An analysis of transformations revisited", Journal of the Royal Statistical Association, Series B, pp.211-252.

Breitung, J, and Meyer, W. (1994), "Testing for Unit Roots in Panel Data: Are Wages on Different Bargaining Levels Cointegrated?", Applied Economics, 26, pp.353-361.

Breusch, T. and Pagan, A. (1979), "A Simple Test for Heteroscedasticity and Random Coefficient Variation", Econometrica, 47, pp.1287-1294.

Buchinsky, M. (1994), "Changes in the U.S. wage structure 1963-1987: an application of quantile regression", Econometrica, 62, pp.405-458.

Buchinsky, M. (1995), "Quantile regression, Box-Cox transformation model, and the U.S. wage structure, 1963-1987", Journal of Econometrics, 65, pp.109-154.

Buldyrev, S.V., Amaral, L.A.N., Havlin, S., Leschhorn, H., Maass, P., Salinger, M.A., Stanley, H.E. and Stanley, M.H.R. (1997), "Scaling Behaviour in Economics II. Modelling of Company Growth", Journal of Physics I (France), 7, April, pp.635-650.

Bunn, P. and Redwood, V. (2003), "Company-accounts-based modelling of business failures and the implications for financial stability", Working Paper No. 210, Bank of England, London.

Buzzacchi, L. and Valletti, T. (1999), "Firm Size Distribution: Testing the "Independent Submarkets Model" in the Italian Motor Insurance Industry", Suntory and Toyota International Centres for Economics and Related Disciplines Discussion Paper No. EI/24, September 1999, London School of Economics and Political Science, London.

Cabral, L. (1995), "Sunk Costs, Firm Size and Firm Growth", Journal of Industrial Economics, 43, 2, pp.161-172.

Cabral, L. (1997), "Entry Mistakes", Centre for Economic Policy Research Discussion Paper No. 1729, CEPR, London, November 1997.

Cabral, L., and Mata, J. (2001), "On the Evolution of the Firm Size Distribution: Facts and Theory", Centre for Economic Policy Research, Discussion Paper No. 3045.

Caves, R.E. (1997), "Industrial Organisation and New Findings on the Turnover and Mobility of Firms", Harvard Institute of Economic Research, Discussion Paper No. 1808, Harvard University.

Cefis, E., Ciccarelli, M. and Orsenigo, L. (2001), "The growth of firms: from Gibrat's legacy to Gibrat's fallacy", EPRIS working paper.

Chadha, J.S., Janssen, N. and Nolan, C. (2000a), "An Examination of U.K. Business Cycle Fluctuations: 1871-1997", Department of Applied Economics Working Paper No. 0024, University of Cambridge.

Chadha, J.S., Janssen, N. and Nolan, C. (2000b), "An Examination of U.K. Business Cycle Fluctuations: 1871-1997: Supplementary Paper", Department of Applied Economics Working Paper No. 0024, University of Cambridge.

Chamberlain, G. (1994), Quantile regression, censoring and the structure of wages, in *Advances in Econometrics*, Sims, C. (Ed.), Cambridge University Press: New York; pp.171-209.

Champerknowne, D.G. (1956), "Discussion of Hart and Prais (1956), Journal of the Royal Statistical Society, Series A, 119, pp.181-182.

Chatterji, M. (1992), "Convergence Clubs and Endogenous Growth", Oxford Review of Economic Policy, 8, pp.57-69.

Chesher, A. (1979), "Testing the Law of Proportionate Effect", Journal of Industrial Economics, 27, 4, pp.403-411.

Choi, I. (2001), "Unit root tests for panel data", Journal of International Money and Finance, 20, pp.249-272.

Clarke, R. (1979), "On the lognormality of firm and plant size distributions: some U.K. evidence", Applied Economics, 11, pp.415-433.

Clarke, R. (1985), *Industrial Economics*, Blackwell Publishers Ltd, Oxford.

Coad, A. (2006), "A Closer Look at Serial Growth Rate Correlation", LEM Working Paper Series No. 2006/29, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa, November 2006.

Cochrane, J.H. (1988), "How Big is the Random Walk in GNP?", Journal of Political Economy, 96, pp.893-920.

Cochrane, J.H. (1991), "A Critique of the Application of Unit Root Tests", Journal of Economic Dynamics and Control, 15, pp.275-284.

Comanor, W.S. and Wilson, T.A. (1967), "Advertising, Market Structure, and Performance", Review of Economics and Statistics, 49, pp.423-440.

Contini, B. and Revelli, R. (1989), "The Relationship Between Firm Growth and Labor Demand", Small Business Economics, 1, 4, pp.309-314.

Cooley, T. and Quadrini, V. (2001), "Financial Markets and Firm Dynamics", American Economic Review, 91, 5, pp.1286-1310.

Cootner, P.A. (ed.) (1964), *The Random Character of Stock Market Prices*, Cambridge, Mass.: MIT Press.

Cosh, A.D., Hughes, A. and Kambhampati, U. (1993), "Size, Growth and Failure: An Analysis of the U.K. Quoted and Unquoted Company Sectors", Small Business Centre Working Paper No. 32, University of Cambridge.

Cosh, A., Hughes, A., Lee, H. and Pudney, S. (1996), "Semi-Parametric Estimation of the Company Size-Growth Relation", Centre for Business Research Working Paper No. 32, University of Cambridge, March 1996.

Cox, D. (1972), "Regression Models and Life Tables", Journal of the Royal Statistical Society, Series B, 34, pp.187-220.

Cox, D.R. (1975), "Partial likelihood", Biometrika, 62, pp.269-275.

Curry, B. and George, K.D. (1983), "Industrial Concentration: A Survey", Journal of Industrial Economics, 31, 3, pp.203-255.

D'Agostino, R.B. (1971), "An Omnibus Test for Normality for Moderate and Large Sample Size", Biometrika, 58, pp.341-348.

D'Agostino, R.B., Belanger, A. and D'Agostino, R.B., Jr. (1990), "A Suggestion for Using Powerful and Informative Tests for Normality", The American Statistician, 44, 4, pp.316-321.

D'Agostino, R.B. and Pearson, E.S. (1973), "Testing for Departures from Normality. Fuller Empirical Results for the Distribution of b_2 and $\sqrt{b_1}$", Biometrika, 60, pp.613-622.

D'Agostino, R.B. and Stephens, M.A. (1986), *Goodness-of-fit Techniques*, Marcel Dekker: New York.

Davidson, R. and MacKinnon, J.G. (1984), "Convenient Specification Tests for Logit and Probit Models", Journal of Econometrics, 25, pp.241-262.

Davies, S., Lyons, B., Dixon, H. and Geroski, P. (1991), *Economics of Industrial Organisation*, Longman Group U.K. Limited, London.

Davis, S.J., Haltiwanger, J. and Schuh, S, (1993), "Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts", National Bureau of Economic Research, Working Paper 4492.

Davis, S.J., Haltiwanger, J. and Schuh, S, (1993), "Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts", Small Business Economics, 8, pp.297-315.

de la Fuente, A. (1997), "The Empirics of Growth and Convergence: A Selective Review", Journal of Economic Dynamics and Control, 21, pp.23-73.

de Wit, G. (2005), "Firm Size Distributions: An overview of steady-state distributions resulting from dynamics models", SCALES paper N200418, EIM Business and Policy Research working paper, January 2005.

Department of Trade and Industry (1988), *Instructions Relating to the Record of Openings and Closures of Manufacturing Plants*, unpublished paper, March 1988.

Dickey, D.A. and Fuller, W.A. (1979), "Distribution of the Estimators for Autoregressive Time Series With a Unit Root", Journal of the American Statistical Association, 74, pp.427-431.

Dickey, D.A. and Fuller, W.A. (1981), "Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root", Econometrica, 49, pp.1057-1072.

Diebold, F.X. and Rudebusch, G.D. (1991), "On the power of Dickey-Fuller tests against fractional alternatives", Economics Letters, 35, pp.155-160.

Disney, R., Haskel, J. and Heden, Y. (1999), "Entry, Exit and Plant Survival in U.K. Manufacturing", Centre for Research on Globalisation and Labour Markets Research Paper 99/9, University of Nottingham.

Doms, M., Dunne, T. and Roberts, M.J. (1995), "The role of technology use in the survival and growth of manufacturing plants", International Journal of Industrial Organisation, 13, pp.523-542.

Dong, L.B. (2004), "The Behrens-Fisher Problem: An Empirical Likelihood Approach", Econometrics Working Paper EWP0404, Department of Economics, University of Victoria, Canada.

Dosi, G. (2005), "Statistical Regularities in the Evolution of Industries. A Guide Through Some Evidence and Challenges for the Theory", Working Paper 2005/17, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa.

Dunne, P. and Hughes, A. (1994), "Age, Size, Growth and Survival: U.K. Companies in the 1980's", Journal of Industrial Economics, 42, 2, pp.115-140.

Dunne, T., Roberts, M.J. and Samuelson, L. (1988), "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries", Rand Journal of Economics, 19, 4, pp.495-515.

Dunne, T., Roberts, M.J. and Samuelson, L. (1989), "The Growth and Failure of U.S. Manufacturing Plants", Quarterly Journal of Economics, 104, 4, pp.671-698.

Durlauf, S.N. (1996), "On the Convergence and Divergence of Growth Rates", Economic Journal, 106, pp.1016-1018.

Dyckman, T.R. and Stekler, H.O. (1965), "Firm Size and Variability", Journal of Industrial Economics, 13, 3, pp.214-218.

Elston, J.A. (2002), "An Examination of the Relationship Between Firm Size, Growth and Liquidity in Neuer Markt, Discussion Paper 15/02, Economic Research Centre, Deutsche Bundesbank.

Engle, R.F. and Granger, C.W.J. (1991), Long-Run Economic Relationships: Readings in Cointegration", Oxford University Press, Oxford.

Engwall, L. (1973), *Models of Industrial Structure*, D.C. Heath, Lexington.

Ericson, R. and Pakes, A. (1995), "Markov-Perfect Industry Dynamics: A Framework for Empirical Work", Review of Economic Studies", 62, pp.53-82.

Evans, D.S. (1987a), "The Relationship Between Firm Growth, Size and Age: Estimates for 100 Manufacturing Industries", Journal of Industrial Economics, 35, 4, pp.567-581.

Evans, D.S. (1987b), "Tests of Alternative Theories of Firm Growth", Journal of Political Economy, 95, 4, pp.657-674.

Evans, P. and Karras, G. (1996), "Convergence revisited", Journal of Monetary Economics, 37, pp.249-265.

Fama, E.F. (1965), "The Behaviour of Stock-Market Prices", Journal of Business, 38, pp.34-105.

Fillipi, E. and Zanetti, G. "Exogenous and Endogenous Factors in the Growth of Firms" in Marris, R. and Wood, A. (eds.), "The Corporate Economy", (1971), Harvard University Press.

Fisher, R.A. (1932), *Statistical Methods for Research Workers*, 4th Edition Oliver & Boyd, Edinburgh.

Fisher, R.A. (1941), *Statistical Methods for Research Workers*, 8th Edition Oliver & Boyd, Edinburgh.

Fujiwara, Y., Di Guilmi, C., Aoyama, H, Gallegati, M. and Souma, W., (2004), "Do Pareto-Zipf and Gibrat Laws Hold True? An Analysis with European Firms", Physica A, 335, pp.197-216.

Frankel, A.J. and Rose, A.K. (1996), "A panel project on purchasing power parity: Mean reversion within and between countries", Journal of International Economics, 40, pp.209-224.

Franses, P.H. (2002), *A Concise Introduction to Econometrics: An Intuitive Guide*, Cambridge University Press, Cambridge.

Freeman, J.M. (1986), "A Piecewise Regression Analysis for Firm-Size Distributions", Statistician, 35, 5, pp.547-559.

Gabaix, X. (1999), "Zipf's Law for Cities: An Explanation", Quarterly Journal of Economics, 114, pp.739-767.

Gabaix, X. (2006), "Power Laws", unpublished paper, October 2006.

Gabaix, X. and Ibragimov, R. (2006), "Rank – 1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents", unpublished paper, March 2006.

Galton, F. (1889), *Natural Inheritance*, MacMillan, London.

Galton, F. (1892), *Hereditary Genius*, MacMillan, London.

Geary, R.C. and Pearson, E.S. (undated), "Tests of Normality", Biometrika, New Statistical Tables, No.1, Cambridge University Press.

Geroski, P. (1991) *Market Dynamics and Entry*, Basil Blackwell, Oxford.

Geroski, P. (1995), "What do we know about entry?", International Journal of Industrial Organisation, 13, 4, pp.421-440.

Geroski, P. (1999), "The Growth of Firms in Theory and in Practice", Discussion Paper 2092, Centre for Economic Policy Research, London.

Geroski, P. and Gregg, P. (1997), *Coping with recession: company performance in adversity*, Oxford University Press.

Geroski, P., Lazarova, S., Urga, G., Walters, C.F. (2003), "Are Differences in Firm Size Transitory or Permanent?", Journal of Applied Econometrics, 18, 1, pp.47-59.

Gibrat, R. (1931), *Les inégalitiés économiques; applications: aux inégalitiés des richesses, `a la concentration des enterprises, aux populations des villes, aux statistiques des familles, etc., d'une loi nouvelle, la loi de l'effet proportionnel.* Paris: Librairie du Recueil Sirey.

Goddard, J. and Wilson, J. (2001), "Cross-sectional and panel estimation of convergence", Economics Letters, 70, pp.327-333.

Goddard, J., Wilson, J. and Blandon, P. (2002), "Panel tests of Gibrat's Law for Japanese manufacturing", International Journal of Industrial Organisation, 20, 3, pp.415-433.

Gorman, W.M. (1956), "Discussion of Hart and Prais (1956), Journal of the Royal Statistical Society, Series A, 119, p.183.

Gort, M. and Jensen, J.B. (2002), "The Survival of Industrial Plants", Working Paper CES-WP-02-25, Centre for Economic Studies, U.S. Bureau of the Census.

Gort, M. and Klepper, S. (1982), "Time Paths in the Diffusion of Product Innovations", Economic Journal, 92, pp.630-653.

Goudie, A.W. (1987), "Forecasting Corporate Failure: The Use of Discriminant Analysis with a Disaggregated Model of the Corporate Sector", Journal of the Royal Statistical Society, A, 150, 1, pp.69-81.

Granger, C.W.J. and Morgenstern, O. (1970), *Predictability of Stock Market Prices*, Heath: Lexington.

Greene, W.H. (1996), "Marginal Effects in the Bivariate Probit Model", Working Paper No. 96-11, Department of Economics, Stern School of Business, New York University.

Greene, W.H., (2000), *Econometric Analysis*, 4th edition, Prentice-Hall International (U.K.) Limited, London.

Greene, W.H., Knapp Greene, L. and Seaks, T.G. (1995), "Estimating the Functional Form of the Independent Variables in Probit Models", Applied Economics, 27, pp.193-196.

Griliches, Z., Hall, B.H. and Hausman, J.A. (1978), "Missing Data and Self Selection in Large Panels", Annales de l'INSEE, 30-31, pp.137-176.

Griliches, Z. and Hausman, J.A. (1985), "Errors in Variables in Panel Data", Journal of Econometrics, 31, pp.93-118.

Griliches, Z and Mairesse, J. (1995), "Production functions: the search for identification", Harvard University Discussion Paper No. 1719, April.

Gudjarati, D. (1988), *Basic Econometrics*, McGraw-Hill Book Co., Singapore.

Hadri, K. (2001), "Testing for stationarity in heterogeneous panel data", Econometrics Journal, 1, pp.1-14.

Hall, B.H. (1987), "The Relationship Between Firm Size and Growth in the U.S. Manufacturing Sector", Journal of Industrial Economics, 35, 4, pp.583-606.

Hall, B.H. and Mairesse, J. (2002), "Testing for Unit Roots in Panel Data: An Exploration Using Real and Simulated Data, University of California, Berkeley, unpublished paper.

Hall, M. and Tideman, N. (1967), "Measures of Concentration", Journal of the American Statistical Association, 62, pp.162-168.

Hall, R.E. (1978), "Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence", Journal of Political Economy, 86, pp.971-987.

Han, A. and Hausman, J. (1990), "Flexible Parametric Estimation of Duration and Competing Risk Models", Journal of Applied Econometrics, 5, 1, pp.1-28.

Hannah, L. and Kay, J.A. (1977), *Concentration in Modern Industry: Theory, measurement and the U.K experience*, MacMillan Press Ltd., London.

Härdle, W. (1990), *Applied Nonparametric Regression*, Cambridge University Press: Cambridge.

Harhoff, D., Stahl, K. and Woywode, M. (1998), "Legal Form, Growth and Exit of West German Firms – Empirical Results for Manufacturing, Construction, Trade and Service Industries", Journal of Industrial Economics, 46, 4, pp.453-488.

Harris, R.I.D. (1992), "Testing for Unit Roots Using the Augmented Dickey-Fuller Test: Some Issues Relating to the Size, Power and Lag Structure of the Test", Economics Letters, 38, pp.381-386.

Harris, R.D.F. and Tzavalis, E. (1999), "Inference for Unit Roots in Dynamic Panels Where the Time Dimension is Fixed", Journal of Econometrics, 91, pp.201-226.

Hart, P.E. (1960), "Business Concentration in the United Kingdom", Journal of the Royal Statistical Society, Series A (General), 123, 1, pp.50-58.

Hart, P.E. (1962), "The Size and Growth of Firms", Economica, 29, pp.29-39.

Hart, P.E. (1965), Studies in Profit, Business Saving, and Investment in the United Kingdom 1920-62: Vol. 1, George Allen and Unwin, London.

Hart, P.E. (1973), "The Comparative Statics and Dynamics of Income Distributions", University of Reading Discussion Papers in Economics No.49.

Hart, P.E. (1975), "Moment Distributions in Economics: an Exposition", Journal of the Royal Statistical Society, Series A (General), 138, 3, pp.423-434.

Hart, P.E. (1976), "The Comparative Statics and Dynamics of Income Distributions", Journal of the Royal Statistical Society, Series A, 139, 1, pp.108-125.

Hart, P.E. (1980), "Lognormality and the Principal of Transfers", Oxford Bulletin of Economics and Statistics, 42, 3, pp.263-267.

Hart, P.E. (1995), "Galtonian Regression Across Countries and the Convergence of Productivity", Oxford Bulletin of Economics and Statistics, 57, 3, pp.287-293.

Hart, P.E. and Oulton, N. (1996), "Growth and Size of Firms", The Economic Journal, 106, pp.1242-1252.

Hart, P.E. and Oulton, N. (1997), "Zipf and the size distribution of firms", Applied Economics Letters, 4, pp.205-206.

Hart, P.E. and Oulton, N. (1998a), "Job creation and destruction in the corporate sector: the relative importance of births, deaths and survivors", NIESR Discussion Paper No. 134, National Institute for Economic and Social Research, London.

Hart, P.E. and Oulton, N. (1998b), "Job Generation in the U.K. Corporate Sector, 1986-95", NIESR Discussion Paper No. 139, National Institute for Economic and Social Research, London.

Hart, P.E. and Oulton, N. (1999), "Gibrat, Galton and Job Generation", International Journal of the Economics of Business, 6, 2, pp.149-164.

Hart, P.E. and Oulton, N. (2001), "Variance of Growth and Size of Firm", Discussion Papers in Economics and Management, Series A, 13, No. 429, University of Reading.

Hart, P.E. and Prais, S.J. (1956), "The Analysis of Business Concentration: A Statistical Approach", Journal of the Royal Statistical Society, Series A (General), 119, 2, pp.150-191.

Hassler, U. and Wolters, J. (1994), "On the power of unit root tests against fractional alternatives", Economics Letters, 45, pp.1-5.

Hausman, J.A. (1978), "Specification Tests in Econometrics", Econometrica, 46, 6, pp.1251-1271.

Hausman, J. and Taylor, W. (1981), "Panel Data and Unobservable Individual Effects", Econometrica, 49, pp.1377-1398.

Hausman, J.A. and Wise, D.A. (1979), "Attrition bias in experimental and panel data: the Gary income maintenance experiment", Econometrica, 47, pp.455-473.

Hay, D.A. and Morris, D.J. (1991), *Industrial Economics and Organisation: Theory and Evidence*, 2nd edition, Oxford University Press, New York.

Heckman, J.J. (1976), "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models", Annals of Economic and Social Measurement, 5, pp.475-492.

Heckman, J.J. (1979), "Sample Selection Bias as a Specification Error", Econometrica, 47, 2, pp.556-569.

Higson, C., Holly, S. and Kattuman, P. (2002), "The Cross-Sectional Dynamics of the U.S. Business Cycle: 1950-1999", Journal of Economic Dynamics and Control, 26, pp.1529-1555.

Higson, C., Holly, S., Kattuman, P. and Platis, S. (2002), "The Business Cycle, Macroeconomic Shocks and the Cross-section: The Growth of U.K. Quoted Companies", Department of Applied Economics Working Paper, December, University of Cambridge.

Hjarlmasson, L. (1974), "The Size Distribution of Plants and Firms Derived from an Optimal Process of Capacity Expansion", European Economic Review, 5, 2, pp.123-140.

H.M. Treasury (2005), Evidence on the U.K. economic cycle, HMSO, July 2005.

Hodrick, R.J. and Prescott, E.C. (1997), "Post-war U.S. Business Cycles: an Empirical Investigation", Journal of Money, Credit and Banking, 29, pp.1-16.

Hopenhayn, H. (1992), "Entry, Exit and Firm Dynamics in Long Run Equilibrium", Econometrica, 60, pp.1127-1150.

Horowitz, J.L. (1992), "A smoothed maximum score estimator for the binary response model", Econometrica, 60, pp.505-531.

Horowitz, J.L. (1993), "Semiparametric estimation of a work-trip model choice model", Journal of Econometrics, 58, pp.49-70.

Hsiao, C. (1996), "Logit and Probit Models", in Mátyás, L. and Sevestre, P. (1996), *The Econometrics of Panel Data: Handbook of Theory and Applications*, 2nd Edition, Boston: Kluwer Academic Publishers, pp.410-428.

Hsiao, C. (2003), *Analysis of Panel Data*, 2nd Edition, Cambridge University Press, Cambridge.

Hunter, J. and Isachenkova, N. (2002), "A Panel Analysis of U.K. Industrial Company Failure", Working Paper No. 228, ESRC Centre for Business Research, University of Cambridge.

Hymer, S. and Pashigan, P. (1962), "Firm Size and the Rate of Growth", Journal of Political Economy, 70, pp.556-569.

Ichimura, H. (1993), "Semiparametric least squares (SLS) and weighted SLS estimation of single-index models", Journal of Econometrics, 58, pp.71-120.

Ijiri, Y. and Simon, H.A. (1964), "Business Firm Growth and Size", American Economic Review, 54, pp.77-89.

Ijiri, Y. and Simon, H.A. (1967), "A Model of Business Firm Growth", Econometrica, 35, 2, pp.348-355.

Ijiri, Y. and Simon, H.A. (1971), "Effects of Mergers and Acquisitions on Business Firm Concentration, Journal of Political Economy, 79, 2, pp.314-322.

Ijiri, Y. and Simon, H.A. (1974), "Interpretation of Departures from the Pareto Curve Firm Size Distribution", Journal of Political Economy, 82, 2, pp.315-331.

Ijiri, Y. and Simon, H.A. (1977), *Skew Distributions and the Sizes of Business Firms*, Amsterdam: North-Holland.

Im, K.S, Pesaran, M.H. and Shin, Y. (1997), "Testing for Unit Roots in Heterogeneous Panels", Journal of Econometrics, 115, pp.53-74.

Ioannidis, C., Peel, D.A. and Venetis, I.A. (2003), "On the Persistence of Profit: Some New Results, Cardiff Business School Working Paper, Cardiff University.

Islam, N. (1995a), "Convergence: Variations in Concept and Empirical Results", Department of Economics Working Paper 98-10, December 1998, Emory University.

Islam, N. (1995b), "Growth Empirics: A Panel Data Approach", Quarterly Journal of Economics, 110, pp.1127-1170.

Jarque, C.M and Bera, A.K. (1980), "Efficient tests for Normality, Homoskedasticity and Serial Independence of Regression Residuals", Economic Letters, 6, 3, pp.255-259.

Jovanovic, B. (1982), "Selection and the Evolution of Industry", Econometrica, 50, 3, pp.649-670.

Jovanovic, B. and Rob, R. (1987), "Demand-driven Innovation and Spatial Competition Over Time", Review of Economic Studies, 54, pp.63-72.

Kalecki, M. (1945), "On the Gibrat distribution", Econometrica, 50, pp.161-170.

Kaplan, E.L. and Meier, P. (1958), "Nonparametric Estimation from Incomplete Observations", Journal of the American Statistical Association, 53, 3, pp.457-481.

Karlsson, S. and Löthgren, M. (2000), "On the power and interpretation of panel unit root tests", Economics Letters, 66, pp.249-255.

Katz, M.L. and Rosen, H.S. (1994), *Microeconomics*, (2nd edition), Richard D. Irwin, Inc., Burr Ridge, Illinois.

Kendall, M.G. (1956), "Discussion on Paper by Mr. Hart and Dr. Prais". Journal of the Royal Statistical Society, Series A, 119, pp.184-185.

Kendall, M.G. and Stuart, A. (1973), *The Advanced Theory of Statistics*, 3 Volumes, Charles Griffen and Company, London.

Kennedy, P. (1998), *A Guide to Econometrics*, 4th edition, Blackwell Publishers Ltd., Oxford.

Kesten, H. (1973), "Random Difference Equations and Renewal Theory for Products of Random Matrices", Acta Mathematica, CXXXI: pp.207-248.

Kiefer, N.M. (ed.) (1985), "Econometric Analysis of Duration Data", Journal of Econometrics, 28, 1, pp.1-169.

Kiefer, N.M. (1988), Economic Duration Data and Hazard Functions", Journal of Economic Literature, 6, pp.646-679.

Kiviet, J.F. (1995), "On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models, Journal of Econometrics, 68, pp.53-78.

Kleiber, C. and Kotz, S. (2003), *Statistical Size Distributions in Economics and Actuarial Sciences*, John Wiley & Sons, Inc., Hoboken, New Jersey.

Klein, R.W. and Spady, R.H. (1993), "An Efficient Semiparametric Estimators for Binary Response Models", Econometrica, 61, pp.387-421.

Klepper, S. (1996), "Entry, Exit, Growth, and Innovation over the Product Life Cycle", American Economic Review, 86, 3, pp.562-583.

Klepper, S. and Miller, J.H. (1995), "Entry, Exit, and Shakeouts in the United States in New Manufactured Products", International Journal of Industrial Organisation, 13, 4, pp.567-591.

Klette, J. and Griliches, Z., (1999), "Empirical Patterns of Firm Growth and R&D Investment: A Quality Ladder Model Interpretation", IFS Working Paper WP 25/99, Institute for Fiscal Studies, London.

Köenker, R. (1981), "A Note on Studentizing a Test for Heteroscedasticity", Journal of Econometrics, 17, pp.107-112.

Köenker, R. and Bassett, Jr., G. (1978), "Regression Quantiles", Econometrica, 46, 1, pp.33-50.

Köenker, R. and Hallock, K.F. (2001), "Quantile Regression", Journal of Economic Perspectives, 15, 4, pp.143-156.

Kruiniger, H. (1998), "Conditional Maximum Likelihood Estimation of Dynamic Panel Data Models", Working Paper, Department of Economics, Queen Mary and Westfield College, University of London.

Kumar, M.S. (1984), *Growth, Acquisition, and Investment*, Cambridge University Press, Cambridge.

Kumar, M.S. (1985), "Growth, Acquisition Activity and Firm Size: Evidence from the United Kingdom", Journal of Industrial Economics, 33, 3, pp.327-338.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992), "Testing the Null Hypothesis of Stationary Against the Alternative of a Unit Root", Journal of Econometrics, 54, pp.159-178.

Kyriazidou, E. (1997), "Estimation of a Panel Data Sample Selection Model", Econometrica, 65, pp.1335-1364.

Kyriazidou, E. (2000), "Estimation of Dynamic Panel Data Sample Selection Models", Review of Economic Studies, 68, pp.543-572.

Lancaster, T. (1990), *The Econometric Analysis of Transition Data*, Cambridge: Cambridge University Press.

Lancaster, T. (2000), "The incidental parameter problem since 1948", Journal of Econometrics, 95, pp.391-413.

Leamer, E.E. (1983), "Let's take the con out of econometrics", American Economic Review, Vol. 73, pp.31-43.

Lee, D. and Schmidt, P. (1996), "On the Power of the KPSS test of Stationarity Against Fractionally Integrated Alternatives", Journal of Econometrics, 73, pp.285-302.

Lee, K., Pesaran, M.H. and Smith, R. (1997), "Growth and Convergence in a Multi-country Empirical Stochastic Solow Model", Journal of Applied Econometrics, 12, 4, pp.357-392.

Lennox, C. (1999), "Identifying failing companies: a re-evaluation of the logit, probit and DA approaches", Journal of Economics and Business, 51, 4, pp.347-364.

Leonard, J.S. (1984), "On the Size Distribution of Employment and Plants", Institute of Industrial Relations, University of California at Berkley, mimeo.

Levin, A. and Lin, C.F. (1992), "Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties", Discussion Paper No.92-93, University of California at San Diego.

Levin, A. and Lin, C.F. (1993), "Unit Root Tests in Panel Data: New Results", Discussion Paper No. 93-56, University of California at San Diego.

Levy, M. and Solomon, S. (1996), "Power Laws are Logarithmic Boltzmann Laws", International Journal of Modern Physics, C 7, p.595.

Lippman, S.A. and Rumelt, R.P. (1982), "Uncertain Imitability: An Analysis of Interfirm Differences in Efficiency Under Competition", Bell Journal of Economics, 13, pp.418-438.

Litchenberg, F.R. (1994), "Testing the Convergence Hypothesis", Review of Economics and Statistics, 76, pp.576-579.

Ljung, L. (1987), *System identification. Theory for the user*, Prentice Hall, New Jersey.

Lo, A.W. and MacKinlay, A.C. (1988), "Stock Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test", Review of Financial Studies, 1, pp.41-66.

Lotti, F. and Santarelli, E. (2001a), "Is firm growth proportional? An appraisal of firm size distribution", Economics Bulletin, 12, 6, pp.1-7.

Lotti, F. and Santarelli, E. (2001b), "Industry Dynamics and the Distribution of Firm Sizes: A Non-Parametric Approach, Working Paper 2001/14, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa.

Lotti, F., Santarelli, E. and Vivarelli, M., (1999) "Does Gibrat's Law Hold in the Case of Young, Small Firms?", Paper presented at the 40th Annual SIE Conference, Ancona, October 1999, Working Paper No. 361, Dipartimento di Scienze Economiche, Universita` di Bologna.

Lotti, F., Santarelli, E. and Vivarelli, M., (2003), "Gibrat's Law and Market Selection in the Radio, TV & Telecommunications Equipment Industry", paper presented at the International Workshop on The Post-entry Performance of Firms: Technology, Growth and Survival, Bologna, November 2002.

Lucas, R.E., Jr. (1967), "Adjustment Costs and the Theory of Supply", Journal of Political Economy, 75, 4, pp.321-334.

Lucas, R.E., Jr. (1978), "On the size distributions of business firms", Bell Journal of Economics, 9, pp.508-523.

Lucas, R.E., Jr. and Prescott, E.C. (1971), "Investment Under Uncertainty", Econometrica, 39, pp.659-681.

Machado, J. and Mata, J. (2000), "Box-Cox Quantile Regression and the Distribution of Firm Sizes", Journal of Applied Econometrics, 15, 3, pp.253-274.

Maddala, G.S. (1977), *Econometrics*, McGraw-Hill, New York.

Maddala, G.S. (1983), *Limited Dependent and Qualitative Variables in Econometrics*, New York, Cambridge University Press.

Maddala, G.S. (1987), "Limited Dependent Variable Models Using Panel Data", Journal of Human Resources, 22, 3, pp.307-338.

Maddala, G.S. and Kim, I.M. (1998), *Unit Roots, Cointegration and Structural Change*, Cambridge University Press, Cambridge.

Maddala, G.S. and Wu, S. (1999), "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test", Oxford Bulletin of Economics and Statistics, Special Issue, 61, pp.631-652.

Maddala, G.S. and Wu, S. (2000), "Cross-country growth regressions: problems of heterogeneity, stability and interpretation", Applied Economics, 32, pp.635-642.

Malinvaud, E. (1980), *Statistical Methods of Econometrics*, 3rd Edition, North Holland: Amsterdam.

Mansfield, E. (1962), "Entry, Gibrat's Law, Innovation and the Growth of Firms", American Economic Review, 52, pp.1023-1051.

Manski, C.F. (1975), "Maximum score estimation of the stochastic utility model of choice", Journal of Econometrics, 27, pp.313-334.

Manski, C.F. (1985), "Semiparametric Analysis of Discrete Response", Journal of Econometrics, 27, pp.313-333.

Manski, C.F. (1989), "Anatomy of the Selection Problem", Journal of Human Resources, 24, pp.243-260.

Marcus, M. (1969), "A Note on the Determinants of the Growth of Firms and Gibrat's Law", Canadian Journal of Economics, 3, pp.580-589.

Marquardt, D.W. (1963), "An Algorithm for Least Squares Estimation of Non-Linear Parameters", Journal of the Society for Industrial and Applied Mathematics, 11, pp.421-441.

Marshall, A. (1890), *Principles of Economics*, MacMillan, London.

Marshall, A. (1919), *Industry and Trade*, MacMillan, London.

Marsili, O., Reichstein, T. and Salter, A.J. (2004), "Searching for Structure: The Distribution of Firm Growth Rates in the Netherlands", paper presented at the DRUID Summer Conference on Industrial Dynamics, Innovation and Development, Elsinore, Denmark, June 2004.

Mata, J. (1994), "Firm Growth During Infancy", Small Business Economics, 6, 1, pp.27-39.

Mata, J. and Machado, J.A.F. (1996), "Firm start-up size: A conditional quantile approach", European Economic Review, 40, pp.1305-1323.

Mata, J. and Portugal, P. (1995), "Life Duration of New Firms", Journal of Industrial Economics, 42, 3, pp.227-245.

Mata, J., Portugal, P. and Guimarães, P. (1995), "The survival of new plants: Start-up conditions and post-entry evolution", International Journal of Industrial Organisation, 13, 4, pp.459-482.

Mátyás, L. and Sevestre, P. (1996), *The Econometrics of Panel Data: Handbook of Theory and Applications*, 2nd Edition, Boston: Kluwer Academic Publishers.

McCloskey, D. (1986), *The Rhetoric of Economics*, Brighton: Wheatsheaf.

McCloughan, P. (1995), "Simulation of Concentration Development From Modified Gibrat Growth-Entry-Exit Processes", Journal of Industrial Economics, 43, 4, pp.405-433.

Merton, R.C. (1974), "The pricing of corporate debt: The risk structure of interest rates", Journal of Finance, 29, 2, pp.449-470.

Mills, T.C. (1999), *The Econometric Modelling of Financial Time Series*, Cambridge University Press, Cambridge.

Moon, H.R. and Phillips, P.C.B. (2003), "GMM Estimation of Autoregressive Roots Near Unity with Panel Data", Department of Economics, Yale University, unpublished paper, June 2003.

Moyes, T. and Thomas, D. (1996), "Manufacturing Change in Wales Since 1966: Evidence From the Welsh Register of Manufacturing Employment", Welsh Economic Review, Vol. 9.1, Summer 1996.

Nawata, K. (1994), "Estimation of Sample Selection Bias Models by the Maximum Likelihood Estimator and Heckman's Two Step Estimator, Economics Letters, 45, pp.33-40.

Nawata, K, and Nagase, N. (1996), "Estimation of Sample Selection Models", Econometric Reviews, 15, pp.387-400.

Neave, H.R. (1981), *Elementary Statistical Tables*, Routledge, London.

Nelson, C.R. and Plosser, C. (1982), "Trends and Random Walks in Macroeconomic Time Series", Journal of Monetary Economics, 10, pp.139-162.

Nelson, R.R. and Winter, S.G. (1982), *An Evolutionary Theory of Economic Change*, Harvard University Press: Cambridge, Mass.

Nerlove, M. (1998), "Properties of Alternative Estimators of Dynamic Panel Models: An Empirical Analysis of Cross-Country Data for the Study of Economic Growth", in Nerlove, M. (2002), *Essays in Panel Data Econometrics*, Cambridge University Press, Cambridge, pp.273-304.

Nerlove, M. (1999), "Likelihood Inference for Dynamic Panel Data Models", Annales d'Economie et de Statistique de l'INSEE, 55-56, in Nerlove, M. (2002), *Essays in Panel Data Econometrics*, Cambridge University Press, Cambridge, pp.305-348.

Nerlove, M. (2000), "Growth Rate Convergence, Fact or Artifact? An Essay on Panel Data Econometrics", in Nerlove, M. (2002), *Essays in Panel Data Econometrics*, Cambridge University Press, Cambridge, pp.237-272.

Nerlove, M. (2002), *Essays in Panel Data Econometrics*, Cambridge University Press, Cambridge.

Nickell, S. (1981), "Biases in Dynamic Models with Fixed Effects", Econometrica, 49, 6, pp.1417-1426.

Okuyama, K., Takayasu, M. and Takyasu, H. (1999), "Zipf's law in income distribution of companies", Physica A, 269, pp.125-131.

Oliveira, B. and Fortunato, A. (2003), "Does the firm growth follow a random walk? An application with panel data unit root tests, unpublished paper, July 2003.

Oulton, N. (1998), "A Tale of Two Cycles: Closure, Downsizing and Productivity Growth in U.K. manufacturing, 1973-89", NIESR Discussion Paper No. 140, National Institute of Economic and Social Research, London, August 1998.

Pagan, A. and Ullah, A. (1999), *Nonparametric Econometrics*, Cambridge University Press, Cambridge.

Pakes, A. and Ericson, R. (1998), "Empirical Implications of Alternative Models of Firm Dynamics", Journal of Economic Theory, 79, pp.1-45.

Pareto, V. (1895), "La Legge Della Domanda", Giornale degli Economisti, 10, pp.59-68, translated Rivista di Politica Economica, 87, pp.691-700, 1997.

Pareto, V. (1896), "La Courbe de la Répartition de la Richesse", reprinted in Busoni, G. (1965) (ed.) *Œeuvres Complètes de Vilfredo Pareto, Tome 3: Écrits Sur La Courbe de la Répartition de la Richesse*, Geneva: Librarie Droz, translated in Rivista di Politica Economica, 87, pp.647-700, 1997.

Pareto, V. (1897), *Cours d'Economie Politique*, Lausanne: Ed. Rouge.

Park, H.M. (2002-06), "Univariate Analysis and Normality Test Using SAS, STATA and SPSS", unpublished paper, Trustees of Indiana University, www.indiana.edu/~statmath.

Pavitt, K., Robson, M. and Townsend, J. (1987), "The Size Distribution of Innovating Firms in the U.K.: 1945-1983", Journal of Industrial Economics, 35, 3, pp.297-316.

Penrose, E.T., (1980), *The Theory of the Growth of the Firm*, Blackwell Publisher Ltd, 2nd edition, Oxford.

Perline, R. (1996), "Zipf's Law, the Central Limit Theorem, and the Random Division of the Unit Interval", Physical Review E, 54, 1, pp.220-223.

Pesaran, M.H., Shin, Y. and Smith, R.P. (1999), "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels", Journal of the American Statistical Association, 94, pp.621-634.

Pesaran, M.H. and Smith, R.P. (1995), "Estimating long-run relationships from dynamic heterogeneous panels, Journal of Econometrics, 68, pp.79-113.

Piergiovanni, R., Santarelli, E., Klomp, L. and Thurik, A.R. (2002), "Gibrat's Law and the Firm Size/Firm Growth Relationship in Italian Services", Tinbergen Institute Discussion Paper, TI 2002-080/3.

Phillips, P.C.B. and Ouliaris, S. (1986), "Testing for Cointegration", Cowles Foundation for Research in Economics Discussion Paper No. 809, Yale University.

Phillips, P.C.B. and Ouliaris, S. (1990), "Asymptotic Properties of Residual Based Tests for Cointegration", Econometrica, 58, 1, pp.165-193.

Phillips, P.C.B and Perron, P. (1988), "Testing for a Unit Root in Time Series Regression", Biometrika, 75, 2, pp.335-346.

Phillips, P.C.B. and Xiao, Z. (1999), "A Primer of Unit Root testing", Cowles Foundation for Research in Economics Discussion Paper No. 972, Cowles Foundation for Research in Economics, Yale University.

Poterba, J.M. and Summers, L.H. (1988), "Mean Reversion in Stock Prices: Evidence and Implications", Journal of Financial Economics", 22, pp.27-59.

Powell, J. (1991), Estimation of monotonic regression models under quantile restrictions, in *Nonparametric and Semiparametric Methods in Econometrics and Statistics: Proceedings of the Fifth International Symposium on Economic Theory and Econometrics*, Barnett, W., Powell, J. and Tauchen, G. (eds.), Cambridge University Press: New York; pp.357-384.

Prais, S.J. (1958), "The Statistical Conditions for a Change in Business Concentration", Review of Economics and Statistics, 40, pp.268-272.

Prais, S.J. (1974), "A New Look at the Growth of Industrial Concentration", Oxford Economic Papers", 26, pp.273-288.

Prais, S.J. (1976), *The Evolution of Giant Firms in Britain*, Cambridge University Press, Cambridge.

Prais, S.J. (1980), "Industrial Concentration: The Role of Statistical Theory", Oxford Bulletin of Economics and Statistics, 42, 3, pp.269-272.

Prescott, E.C. and Visscher, M. (1980), "Organisational Capital", Journal of Political Economy, 88, 3, pp.446-461.

Pryor, F.L. (1972), "The Size of Production Plants in Manufacturing", Economic Journal, 82, 326, pp.547-566.

Quah, D. (1993a), "Galton's fallacy and tests of the convergence hypothesis", Scandinavian Journal of Economics, 95, pp.427-443.

Quah, D. (1993b), "Empirical cross-section dynamics in economic growth", European Economic Review, 37, pp.426-434.

Quah, D. (1994), "Exploiting cross-section variations for unit root inference in dynamic data", Economics Letters, 44, pp.9-19.

Quah, D. (1996a), "Empirics for economic growth and convergence", European Economic Review, 40, pp.1353-1375.

Quah, D. (1996b), "Twin Peaks: Growth and Convergence in Models of Distribution Dynamics", Economic Journal, 106, pp.1045-1055.

Quandt, R.E. (1966), "On the Size Distribution of Firms", American Economic Review, 56, pp.416-432.

Quandt, R.E. (1968), "Lognormality and Concentration: Comment", American Economic Review, 58, pp.1367-1370.

Ramsay, J.B. (1969), "Tests for Specification Errors in Classical Linear Squares Regression Analysis", Journal of the Royal Statistical Society, Series B, 31, pp.350-371.

Reichstein, T. and Jensen, M.B. (2003), "Analysing the Distributions of the Stochastic Firm Growth Approach", paper presented at the DRUID Summer Conference 2003 paper, Copenhagen, June 2003.

Rossi-Hansberg, E. and Wright, L.J. (2005), "Firm Size Dynamics in the Aggregate Economy", NBER Working Paper No. 11261, National Bureau for Economic Research, April 2005.

Royston, J.P. (1982), "An Extension of Shapiro and Wilk's W Test for Normality to Large Samples", Applied Statistics, 31, 2, pp.115-124.

Royston, J.P. (1983), "A Simple Method for Evaluating the Shapiro-Francia W′ Test of Non-Normality", Statistician, 32, 3, pp.297-300.

Royston, J.P. (1991), "Comment on sg3.4 and an Improved D'Agostino test", Stata Technical Bulletin, 3, pp.13-24.

Royston, J.P. (1992), "Approximating the Shapiro-Wilk W-Test for Non-normality", Statistics and Computing, 2, pp.117-119.

Sala-i-Martin, X. (1996), "The Classical Approach to Convergence Analysis", Economic Journal, 106, pp.1019-1036.

Samuels, J.M. (1965), "Size and Growth of Firms", Review of Economic Studies, 32, pp.105-112.

Samuels, J.M. and Chesher, A.M. (1975), "Growth, survival and the size of companies 1960-69", in Cowling, K. (ed.), *Market Structure and Corporate Behaviour*, Gray-Mills: London.

Sapio, S. and Thoma, G. (2006), "The Growth of Industrial Sectors: Theoretical Insights and Empirical Evidence from U.S. Manufacturing", LEM Working Paper Series No. 2006/09, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Pisa, March 2006.

Sargan, J.D. (1958), "The Estimation of Economic Relationships Using Instrumental Variables", Econometrica, 26, pp.329-338.

Savage, L.J. (1954), *The Foundations of Statistics*, New-York: Wiley.

Saving, T.R. (1965), "The four-parameter lognormal, diseconomies of scale and the size distribution of manufacturing plants", International Economic Review, 6, pp.105-114.

Sawyer, M.C. (1980), "The Variance of Logarithms and Industrial Concentration: A Reply", Oxford Bulletin of Economics and Statistics, 42, 3, pp.273-278.

Schary, M.A. (1991), "The probability of exit", Rand Journal of Economics, 22, 3, pp.339-353.

Scherer, F.M. (1980), *Industrial Market Structure and Economic Performance*, 2nd edition, Rand McNally College Publishing Company, USA.

Schmalensee, R. (1989), "Inter-Industry Studies of Structure and Performance", in Schmalensee, R. and Willig, R. (eds.), *Handbook of Industrial Organisation*, Amsterdam: North-Holland.

Schreyer, P. (1996), "SMEs and Employment Creation: Overview of Selected Quantitative Studies in OECD Member Countries", Science, Technology and Industry Working Paper 1996/4, Organisation for Economic Cooperation and Development (OECD), Paris.

Schreyer, P. (2000), "High-Growth Firms and Employment", Science, Technology and Industry Working Paper 2000/3, Organisation for Economic Cooperation and Development (OECD), Paris.

Schwartz, G. (1978), "Estimating the Dimension of a Model", Annals of Statistics, 6, pp.461-464.

Shalit, S.S. and Sankar, U. (1977), "The Measurement of Firm Size", The Review of Economics and Statistics, 59, pp.290-298.

Shapiro, S.S. and Francia, R.S. (1972), "An Approximate Analysis of Variance Test for Normality", Journal of the American Statistical Association, 67, pp.215-216.

Shapiro, S.S. and Wilk, M.B. (1965), "An Analysis of Variance Test for Normality (Complete Samples)", Biometrika, 52, pp.591-611.

Shen, T.Y. (1970), "Economics of Scale, Penrose-Effect, Growth of Plants and Their Size Distribution", Journal of Political Economy, 78, pp.701-716.

Shiller, R.J. and Perron, P. (1985), "Testing the Random Walk Hypothesis: Power Versus Frequency of Observation", Economics Letters, 18, pp.381-386.

Shorrocks, A.F. (1975), "On Stochastic Models of Size Distributions", Review of Economic Studies, 42, 5, pp.631-641.

Silberman, I.H. (1967), "On Lognormality as a Summary Measure of Concentration", American Economic Review, 57, pp.807-831.

Silberman, I.H. (1968), "Lognormality and Concentration: Reply", American Economic Review, 58, pp.1370-1371.

Silverman, B.W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London.

Simon, H.A. (1955), "On a Class of Skew Distribution Functions", Biometrika, 42, pp.425-440.

Simon, H.A. (1964), "Comment: Firm Size and Rate of Growth", Journal of Political Economy, 72, 1, pp.81-82.

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

Singh, A. and Whittington, G. (1968), *Growth, Profitability and Valuation*, London: Cambridge University Press.

Singh, A. and Whittington, G. (1975), "The Size and Growth of Firms", Review of Economic Studies, 42, pp.15-26.

Smith, V., Leybourne, S., Kim, T-H. and Newbold, P. (2003), "More Powerful Panel Data Unit Root Tests with an Application to Mean Reversion in Real Exchange Rates", Journal of Applied Econometrics, 19, pp.147-170.

Smyth, D.J, Boyes, W.J. and Peseau, D.E. (1975), "The Measurement of Firm Size: Theory and Evidence for the United States and the United Kingdom", The Review of Economics and Statistics, 57, pp.111-115.

Stanley, M.H.R., Buldyrev, S.V., Havlin, S., Mantegna, R.N., Salinger, M.A., Stanley, H.E. (1995), "Zipf plots and the size distribution of firms", Economics Letters, 49, pp.453-457.

Stanley, M.H.R., Buldyrev, S.V., Mantegna, R., Havlin, S., Salinger, M.A. and Stanley, H.E. (1996), "Scaling Behaviour in the Growth of Companies", Nature, 379, pp.804-806.

StataCorp LP (2005), *STATA Windows Software*, version SE 9.0, College Station, Texas, U.S.

Stata Press (2005), *STATA Reference Manual Release 9*, College Station, Texas, U.S.: STATA Press.

Steindl, J. (1965), *Random processes and the growth of firms: A study of the Pareto law*, Griffin, London.

Steindl, J. (1968), "Size Distributions in Economics", in Sills, D.L. (ed.), *International encyclopaedia of social sciences*, London: Collier Macmillan, pp.295-300.

Subbotin, M.T. (1923), "On the Law of Frequency of Errors", Matematicheskii Sbornik, 31, pp.296-301.

Sutton, J. (1995a), "The Size Distribution of Businesses, Part 1: A Benchmark Case", Suntory and Toyota International centres for Economics and Related Disciplines Discussion Paper No. EI/9, December 1995, London School of Economics and Political Science, London.

Sutton, J. (1995b), "The Size Distribution of Businesses, Part II: A Game Theoretic Model", Suntory and Toyota International centres for Economics and Related Disciplines Discussion Paper No. EI/10, London School of Economics and Political Science, London.

Sutton, J. (1997), "Gibrat's Legacy", Journal of Economic Literature, 35, pp.40-59.

Sutton, J, (1998), *Technology and Market Structure: Theory and History*, MIT Press, Cambridge (MA) and London.

Sutton, J. (2001), "The Variance of Firm Growth Rates: The 'Scaling' Puzzle", Suntory and Toyota International centres for Economics and Related Disciplines Discussion Paper No. EI/2, London School of Economics and Political Science, London.

Swamy, P. (1970), "Efficient Inference in a Random Coefficient Regression Model", Econometrica, 38, pp.311-323.

Swamy, P. (1971), *Statistical Inference in Random Coefficient Regression Models*, Springer-Verlag: New York.

Taffler, R.J. (1982), "Forecasting Company Failure in the U.K. Using Discriminant Analysis and Financial Ratio Data", Journal of the Royal Statistical Society, A, 145, 3, pp.342-358.

Teitelbaum, D. and Axtell, R, (2005), "Firm Size Dynamics of Industries: Stochastic Growth Processes, Large Fluctuations, and the Population of Firms as a Complex System", U.S. Small Business Administration Office of Advocacy, Small Business Research Summary No. 247, January 2005.

Tobin, J. (1958), "Estimation of Relationships for Limited Dependent Variables", Econometrica, 26, pp.24-36.

Trapani, L. (2004), "Testing for Unit Roots in Heterogeneous Panels under Cross-sectional Dependence", Working Paper WP-CEA-06-2004, Centre for Econometric Analysis, Cass Business School, City University, London.

Troske, K.R. (1992), "The Time Series Pattern of Growth in Two Industries", Centre for Economic Studies, U.S. Bureau of the Census.

Tschoegl, A.E. (1983), "Size, Growth, and Transnationality among the World's Largest Banks", Journal of Business, 56, 2, pp.187-201.

Tschoegl, A.E. (1996), "Managerial (Dis)economies of Scale: The Case of Regional Banks in Japan", Working Paper 96-04, Regional H. Jones Centre for Management and Policy, Strategy and Organisation, Wharton Business School, University of Pennsylvania.

Tudela, M. and Young, G. (2003a), "A Merton-model approach to assessing the default risk of U.K. public companies", Working Paper No. 194, Bank of England, London.

Tudela, M. and Young, G. (2003b), "Predicting default among U.K. companies: a Merton approach", Financial Stability Review, Bank of England, pp.104-114, June 2003.

Upton, G and Cook, I (2002), _Dictionary of Statistics_, Oxford University Press, Oxford.

Utton, M.A. (1971), "The Effect of Mergers on Concentration: U.K. Manufacturing Industry, 1954-65", Journal of Industrial Economics, 20, 1, pp.42-58.

Utton, M.A. (1972), ""Mergers and the Growth of Large Firms", Bulletin of Oxford University Institute of Economics and Statistics, 34, 2, pp.189-197.

van Ark, B. and Monnikhof, E. (1996), "Size Distribution of Output and Employment: A Data Set for Manufacturing Industries in Five OECD Countries, 1960s-1990", Economics Department Working Paper No. 166, OECD, Paris.

Variyam, J.N. and Kraybill, D.S. (1992), "Empirical evidence on determinants of firm growth", Economics Letters, 38, pp.31-36.

Veall, M. and Zimmerman, K.F. (1996), "Pseudo-R^2 Measures for Some Common Limited Dependent Variable Models", Journal of Economic Surveys, 10, pp.241-259.

Verbeek, M. (2000), _A Guide to Modern Econometrics_, John Wiley & Sons Ltd, England.

Vining, Jr., D.R. (1976), "Autocorrelated Growth Rates and the Pareto Law: A Further Analysis", Journal of Political Economy, 84, 2, pp.369-380.

Wagner, J. (1992), "Firm Size, Firm Growth, and Persistence of Chance: Testing Gibrat's Law with Plant Data from Lower Saxony", Small Business Economics, 4, 2, pp.125-131.

Wagner, J. (1994), "The Post-Entry Performance of New Small Firms in German Manufacturing Industries", Journal of Industrial Economics, 42, 2, pp.141-154.

Walsh, P.P. and Whelen, C. (2002), "Product Differentiation and Firm Size Distribution: An Application to Carbonated Soft Drinks", Suntory and Toyota International Centres for Economics and Related Disciplines Discussion Paper No. EI/31, London School of Economics and Political Science, London.

Wand, M.P. and Jones, M.C. (1995), _Kernel Smoothing_, Chapman & Hall, London.

Wang, P. (2003), _Financial Econometrics: Methods and models_, Routledge, London.

Wedervang, F. (1965), _The Development of a Population of Industrial Firms_, Scandinavian University Books.

Weiss, C. (1998), "Size, Growth, and Survival in the Upper Austrian Farm Sector", Small Business Economics, 10, pp.305-312.

Welch, B.L. (1947), "The Generalization of 'Student's' Problem When Several Different Population Variances Are Involved", Biometrika, 34, pp.28-35.

White, H. (1980), "A Heteroscedasticity Consistent Co-Variance Matrix Estimator and Direct Test for Heteroscedasticity", Econometrica, 48, 4, pp.817-838.

White, H. and MacDonald, G.M. (1980), "Some Large Sample Tests for Non-Normality in the Linear Regression Model", Journal of the American Statistical Association, 75, pp.16-28.

Wikipedia (2007), available at www.en.wikipedia.org, accessed March 2007.

Wooldridge, J.M. (2002), *Econometric Analysis of Cross-section and Panel Data*, MIT Press, Cambridge, MA.

Wu, Y. and Zhang, H. (1996), "Mean Reversion in Interest Rates: New Evidence from a Panel of OECD Countries", Journal of Money, Credit and Banking, 28, pp.604-621.

Zipf, G.K. (1949), *Human Behaviour and the Principal of Least Effort*, Addison-Wesley, Reading, MA.

Zipf, G.K. (1972), *Human Behaviour and the Principle of Least Effort*, Hafner: New York.