



**UCD GEARY INSTITUTE
DISCUSSION PAPER
SERIES**

**The Firm Size Distribution &
Inter-Industry Diversification**

August 2009

The Firm Size Distribution and Inter-Industry Diversification

John Hutchinson,
European Central Bank
(John.Hutchinson@ecb.int)

Jozef Konings,
Economics Department, K.U. Leuven
(Joep.Konings@econ.kuleuven.be)

Patrick Paul Walsh,
UCD SPIRe and Geary Institute
(ppwalsh@ucd.ie)

Abstract

We show that the stylized facts of the Firm Size Distribution (FSD) by age cohorts, as shown in Cabral and Mata (2003), bind within 4-digit manufacturing industries in the UK and Belgium. As in Klepper and Thompson (2006) and Sutton (1998), we explore whether time to build a portfolio of products is a mechanism that relates age to firm size. While inter industry diversification, to some extent, accounts for the role of age, we find that the presence of economies of scope has a separate impact on firm size when controlling for age, amongst other factors. Using the techniques in Cabral and Mata's we show that the FSD by degrees of product diversification shifts to the right, but more so in older age groups. This shows a role for inter-industry diversification over and above an age effect.

JEL Classification: L10, L11 and L16.

John Hutchinson: European Central Bank, Germany; Email, john.hutchinson@ecb.int.

Jozef Konings: K.U. Leuven, Belgium; E-mail, Joep.Konings@econ.kuleuven.be

Patrick Paul Walsh: UCD SPIRe and Geary Institute, Ireland; Email ppwalsh@ucd.ie

Introduction

Investigating firm size distributions (FSD) has received considerable attention since Gibrat's (1931) "law of proportionate effects".¹ In more recent times, Cabral and Mata (2003) have pioneered an evolutionary approach to explain the FSD and show that the distribution of the logarithms of firm size of young cohorts are skewed, but, due to an unknown economic force, the distribution gradually moves towards a log normal distribution in older cohorts. They find that most of the observed changes in the FSD results from the evolution of the distribution of survivors of a given cohort, and are not due to firm selection processes.²

Cabral and Mata (2003) propose that as firms get older they encounter gradually fewer financial constraints. Angelini and Generale (2007) empirically investigate this explanation using survey based measures of financial constraints. While confirming the stylized facts of FSD as proposed by Cabral and Mata, they conclude that financial constraints cannot be the main determinant in the evolution of the size distribution of firms.

Klepper and Thompson (2006) address the role of age on firm performance by proposing an alternative mechanism by which the age of a firm is correlated with both its growth and survival. They show that all age-size regularities can be explained by a simple framework based on the creation and destruction of product lines. They show that firm growth is negatively related, and survival is positively related, to the number of product lines a firm is operating in. Product diversification is shown to diminish the role that age has on firm growth

¹ One literature used "stochastic processes" in firm growth to explain the FSD. In this vein, Gibrat (1931) postulates that current firm size is proportional to its initial size, in the size-growth relationship for active firms, which generates approximately lognormal size distributions. Hart and Prais (1956) and Iijri and Simon (1964, 1977) built in stochastic entry processes around the Gibrat size-growth relationship for active firms. As Schmalensee (1989) concluded however, the golden age of stochastic processes failed to generate limiting size distributions observed across the general run of industries. Another literature using rich firm level data, starting with Dunne, Roberts and Samuelson (1988), focus on modeling firm growth and survival as a deterministic outcome of firm characteristics, including age and size. In general, these empirical studies seem to point to the success of idiosyncratic firm characteristics to explain firm growth and survival, as a result, conclude that Gibrat's law fails to hold.

² Theories such as Jovanovic (1982), Hopenhayn (1992) and Ericson and Pakes (1995) are often cited to motivate firm turnover in young firms (noisy selection and temporary ex-post entry learning). Luttmer (2007) has a clever model with experimentation and selection allowing permanent shocks to preferences and technology which affect the steady state FSD. This approach assumes an endogenous idiosyncratic technology is linked to each single product firm.

and survival, as previously documented in empirical studies (e.g. Dunne, Roberts and Samuelson, 1988) and is an important driver of the limiting FSD.

Sutton (1998) also puts forward the idea that firm heterogeneity in the collection of sub-markets (product lines or geographical areas) during industry evolution is a core determinant of the FSD.³ Sutton (1998) asks whether it is possible that a mechanism exists that induces some minimal degree of inequality in the FSD in industries. Product opportunities are allocated across incumbents and new entrants during industry evolution using a principle of “equal treatment” which induces age and size distribution to coincide on the lower bound. Size in this model is the number of equally sized sub-markets that a firm is in. In the limit where the number of opportunities becomes large, older firms on average will end up with larger product portfolios than younger firms.⁴ Sutton (1998) outlines many theorems where certain factors will lead the FSD will go inside the lower bound. One is where firms get economies of scope from product diversification leading to more skewed FSD allowing size and age to partially disconnect.

A positive relationship between firm size and inter-industry (4-digit) product diversification has been documented for manufacturing at regular time intervals since the 1960s (see Gollop & Monahan, 1991). Gollop and Monahan (1991) identify this as the most important structural phenomena of the post-war period. Bernard, Redding and Schott (2006) do document the evolution of multi-product firms in international trade by building on the selection and single product model of Melitz (2003). Their modeling combines diseconomies

³ The major difference between the two approaches is that the Klepper and Thompson (2006) model is a dynamic model applied to time series data for a single industry, the US laser industry. The Sutton (1998) model describes a steady state and is applied to a cross section of mature manufacturing industries. Klepper and Thompson (2006) also build in finite modelling of product lines into their theory which gives them interesting short-run dynamic relationships between growth, survival and age (product counts).

⁴ To date, most empirical validations of the Sutton (1998) theory use industry studies that focus on sub-market configurations in terms of geographic locations along one product dimension: the US Cement Industry (Sutton, 1998), the Spanish Retail Banking Sector (De Juan, 2003), and the Italian Motor Insurance Industry (Buzzacchi & Valletti, 1999). Due to the richness of data, all these studies could understand how firm size (market share) is positively related to portfolios of geographical locations. Walsh and Whelan (2002) address this issue in a multi-product industry, Carbonated Soft Drinks, where configurations are in terms of products and stores (geography).

of scope at the industry level with strong economies of scope across industries. They implicitly explain the relationship between firm size and inter-industry (4-digit) product portfolios but industry diversification is not linked to age. This supply side model is in contrast to Klepper and Thompson (2006) and Sutton (1998) who link size and product diversification explicitly to age. Nocke and Yeaple (2008) is the only trade paper that theoretical links multi-product firms directly to the FSD during globalization. Firm scope and scale are endogenous in their model. They predict that due to different inherent supply side organizational capabilities of firms, economies of scope would decline and flatten the FSD during globalization. Even though they predict a positive relationship between inter-industry diversification and size there is no explicit role for company age. Yet, the inherent differences in organizational capabilities of firms could result from experience or age issues.

A major contribution of this paper is to show that the role that inter-industry (4-digit) product diversification plays in the determination of firm size, controlling for the age of companies, amongst other factors. It has a partial effect that is independent of age. Economies of scope could be just an outcome of competition (lotteries) over the ongoing opening up of sub-markets during history but are also likely to be fueled by firm heterogeneity in underlying costs. The latter would lead to partial effects on the FSD that are over and above the age cohort effects predicted by Cabral and Mata (2003). Using a series of non-parametric density plots we replicate, to a lesser degree, the basic result of Cabral and Mata (2003) using inter-industry (4-digit) product counts rather than age cohorts but we find that the independent role of product counts is much more pronounced in older age cohorts. We outline other important implications of this result for research in our concluding section.

Section I

In this paper we ask do firms become large because they accumulate a greater number of 4-digit industries over time? We investigate whether this theory of size can explain the stylized facts of the FSD, as proposed by Cabral and Mata (2003), where we proxy age with 4-digit industry counts. If they are not interchangeable we investigate the independent effect of product counts might have overall and within different age cohorts.

Our data is taken from the reported unconsolidated company accounts of manufacturing firms in the UK and Belgium collected by an electronic publishing company, Bureau Van Dijck (BvD), which is commercialized under the name of AMADEUS (see www.bvdep.com). These data cover virtually the entire population of businesses in manufacturing and have all been officially audited. The data do not cover the single proprietorship companies with zero employees.

Firm Size

Due to the accounting legislation, companies are not required to report all financial and operational items. However, full information is available on total assets of firms. Total assets means all fixed assets (intangible, tangible, financial and current assets). For this reason, we will compute the market share of total assets of a firm, f , within a 4-digit industry, j , and this is denoted as MS_{fj} (natural log of market share).⁵ We gathered full information on firm's total assets for active UK and Belgian manufacturing firms with our final sample consisting of 78,911 UK firms and 21,697 Belgian firms in 2002.

⁵ Although Sutton (1998) uses sales to measure firm size, he argues that total assets would be the preferred measure. Consequently, this is the measure employed in this paper. Mariuzzo, Walsh and Whelan (2003) explore the link between market share in sales and estimated market power in Carbonated Soft Drinks. They show that inequality in market share does not reflect inequality in market power. Small companies can extract rent in niches just as much as large companies that operate across niches.

Industry

We define industries using a 4-digit NACE classification. At the company level we have information on the number of 4-digit NACE industries that a firm operates over. We realize that the sum of the 4-digit NACE industries that a firm is active in constitutes a lower bound to the product count of a firm. Theoretically they could be producing many 8-digit products within 4-digit codes. The fact that our results come from product diversification across 4-digit industries is very striking and has important implications for research areas, which we outline in the conclusion.

Age

We also work with the age of a company based on the year of incorporation in our analysis⁶. Before moving to the results section we investigate the co-production of multi-industry firms in the UK and Belgium. In table 1 and table 2 we present co-production matrices for multi-industry firms in the UK and Belgium for 2002 at the industry level (2-digit level).⁷ Each row in the matrices shows the percentage of multi-industry firms operating in particular secondary industry. For example, in row 1 in table 1, for those multi-industry UK firms whose primary industry is food (nace 15), 85% of their secondary activities are also in the food industry (nace 15). This compares with 92% of Belgian firms as reported in table 2. From table 1 and table 2 we see that in both countries, for most industries that the diagonal contains the highest percentage with co-production occurring within rather than across 2-digit

⁶ The dataset does not allow for identifying whether a firm was formed as a result of a merger between two firms.

⁷ In the Amadeus database, the main activity of a company is characterised by a single NACE code, referred to as the 'primary code'. The additional activities of the company are allocated to 'secondary codes'. Although we have information on the number of secondary industries firms operate in, we restrict ourselves to focusing on firms secondary activities that are in the manufacturing sector (nace codes: 15 - 37). We also choose to look at the industry level (2-digit nace) because making matrices using 4-digit level information would be very impractical.

industries.⁸ This is broadly in line with the findings of Bernard et al. (2005) where they conduct a similar exercise for the United States.

Section II

FSD by age cohorts

In figure 1.1 for the UK and figure 1.2 for Belgium we graph the age distribution of firms in the population of firms. We can note that in both countries there are many young firms, while only a small fraction of firms are older than 20 years.⁹ We follow the analysis in Cabral and Mata (2003) and plot the kernel density estimates of the FSD by age cohorts in figure 2.1 for the UK and figure 2.2 for Belgium, where size is measured as the market share of physical assets within 4-digit industries. We see from figures 2.1 and 2.2 that for both the UK and Belgium, that as firms grow older, the size distribution shifts more to the right. Furthermore, in table 3, we report the degree of skewness for both countries and we see that the level of skewness is decreasing as firms age, indicating that the distributions are becoming more symmetrical. This finding highlights the point that as time advances some economic force is pushing the distribution of firm size within 4-digit industries towards log normality in surviving firms. We can note that the basic result of Cabral and Mata (2003) is confirmed.

Lorenz Curves of the FSD

Sutton (1998) asks whether there exists a mechanism that induces some minimal degree of inequality in firms' sizes. In his theory of limiting firm size distribution, a key structural

⁸Two industries which are an exception to this are nace 23(manufacture of coke, refined petroleum products and nuclear fuel) and nace 30(manufacture of office machinery and computers). For nace 23, the number of multi-industry firms is small, 23 in Belgium and 39 in the UK, so if a relatively small number of firms have secondary activities in another industry then the percentage appears to be high. However, the majority of secondary activities are in nace 24 (manufacture of chemicals and chemical products), a related industry. For multi-industry firms whose primary activity is within nace 30, the secondary industries for these firms are based in related industries, namely nace 31 (manufacture of electrical machinery and apparatus n.e.c), nace 32 (manufacture of radio, television and communication equipment and apparatus) and nace 33 (manufacture of medical, precision and optical instruments, watches and clocks)

⁹ Over-time tax laws in Belgium have made it worthwhile opening and closing companies, which may have lead to younger company profiles.

feature of the theory is the arrival, in the history of a defined market, of a number of discrete investment opportunities over an infinite period for firm populations. If it is assumed that opportunities are of equal size and that each of these opportunities can be taken up by one firm, then starting with a single firm of size 1, all subsequent opportunities can be taken up by either a new entrant or existing firm. If all opportunities are taken up by new entrants, then the limiting size distribution would be equally distributed with all firms being of size 1. It is differences in the opportunity portfolios of firms, referred to in Sutton (1998) as roles, which results in the differences in the size of firms. Sutton (1998) assumes two simple conditions, firstly independent of the size of that firm, the next market opportunity to be filled by any currently active firm is based on “equal treatment”, in a game theoretic sense of the term. The second condition that is required is that the probability p , the next market opportunity is filled by a new entrant, is constant over time. Using these two conditions, Sutton (1998) derives a game theoretic model of the evolution of firm size, which yields a limiting FSD. In the lower bound, unequal firm size distributions at the product level can only be driven by the heterogeneous take-up of roles across segments. FSD is based on the simple count of roles across segments and is restricted to a lower bound Lorenz curve. We refer to Sutton (1998) for the derivation which yields the following proposition, under conditions 1 and 2 for any fixed ratio k/N , an asymptotic lower bound to the k -firm concentration ratio:

$$C_{k/N} \geq \frac{k}{N} \left(1 - \ln \frac{k}{N} \right). \quad (1)$$

Equation (1) shows that the lower bound to concentration is independent of the entry parameter p . This parameter affects the average firm size, but not the shape of the size distribution, which contrasts sharply with the traditional literature in which the skewness of size distributions was parametrized by p (e.g. Hart & Prais, 1956).

The basic idea behind the theoretical lower bound to FSD is that opportunities to firms present themselves gradually overtime and hence firms need time to build up product portfolios. Age and size (product counts) become linked as firms spread themselves across product opportunities over time. If in the history of the games we allow economies of scope, this will create more inequality and disconnection between product counts (size) and age, pushing FSD inside the lower bound.

In figure 3.1 for the UK and figure 3.2 for Belgium we plot as in Sutton (1998) the k firm concentration ratio in the market, C_k , for the top 4, 8 and 20 companies in every 4-digit industry against the corresponding k/N . We can note that for both countries the mathematically predicted lower bound seems to hold for all of the 4-digit industries. Yet, we see that many observations are inside the bound indicating that there may be factors other than age driving size, such as economies of scope.

The approach in Sutton (1998) used 4-digit US 1987 and 4-digit German 1990 manufacturing data to support his predictions regarding the mathematically derived lower bound to the size distributions of firms. While the lower bound is motivated by heterogeneous operations of firms across sub-markets, his data does not allow him to discriminate between an aggregation where you are big because of a dominance in one big market or you where big because who covered multi-markets. In the next section, we examine whether company size (market share of assets) is an outcome of product counts, controlling for age and other factors.

¹⁰ Age and product counts can proxy each other for market share observations on the bound but should have separate effects on observations inside the bound. ¹¹

¹⁰ The cross section approach used in Sutton (1998) had a few other data limitations that we do not have. A problem with official statistics in various countries is that at the 4-digit level of manufacturing there is usually no data available on top four, eight or twenty firm concentration ratios, let alone the number of firms active in each market. If the number of firms active is reported, it often refers to firms with at least 5 or 10 employees. Many countries are characterized by a large number of small firms with less than 10 employees. For this reason Sutton (1998) works with top four, eight, twenty and fifty firm concentration ratios within 4-digit industries in the US in 1987 and top three, six, ten, twenty five fifty firm concentration ratios within 4-digit industries in Germany in 1990. He has no data on the total active number of firms in each industry. He re-derives his mathematically derived lower bound to the size distributions of firms for top twenty firms conditional on knowing the top fifty

Section III

In our data we have a list of 4-digit industries which firms have activities in. In this section we provide evidence that product diversification matters for firm size in terms of the market share of total assets within 4-digit industries, controlling for age and industry specific effects. We set out to reject the null hypothesis, that counts of 4-digit products line have no independent effect on firm size, in terms of market share of assets, in a simple econometric model. The basic model of firm size f in each 4-digit industry j is as follows,

$$MS_{fj} = \alpha + \beta_1 \ln Counts_f + \beta_2 \ln Age_f + Sector_j + \varepsilon_{fj} \quad (2)$$

MS_{fj} is the log size of a firm in terms of its market share of assets within a 4-digit industry, $Counts_f$ is firm counts of 4-digit industries. Age_f is the number of years since the year of the firm's incorporation and 4-digit industry dummies control for unobserved industry effects.¹² We assume that operating over industries is something that takes a long time to do. Our product counts variable is a structural feature of the company as is age. We do not instrument industry counts for this reason. We can think of the results as associations rather than causalities. Our main empirical evidence will be based on non-parametric density plots used by Cabral and Mata (2003).

We now turn to testing our hypothesis in tables 4 - 7, where we estimate equation (2). We report the estimates of equation (2) estimating the effect of age non-parametrically in Table 4 but allow for a number of functional forms in later tables. In columns (i) and (iii) in tables 4, where $\ln(Counts_f)$ refers to the number of 4-digit NACE industries a firm operates over, we

firm concentration ratio. An important contribution of our approach is that rather than obtaining sector level information at the 4-digit level from the official statistical offices, we use individual firm level data covering virtually the entire population of businesses in the various sectors that we investigate. This has the advantage that we not only can compute the desired concentration ratios used, but also that we have a complete picture about the number of firms active within and across industries, without having the deal with the cut-off point often used by the statistical offices.

¹¹ As one would expect, for both the UK and Belgium, there is a positive correlation between age and industry diversification with a coefficient of 0.02 and 0.09 respectively.

¹² The industry dummies were constructed by the primary NACE code for each firm.

report the estimates for the UK and Belgium with no 4-digit industry dummies.¹³ In columns (ii) and (iv) we include a full set of 4-digit industry dummies.¹⁴ Interestingly, irrespective of the functional form of age or whether we include industry dummies, the coefficient β_1 is always positive and significant. Overall, it appears that after controlling for age and industry effects, the number of 4-digit NACE industries a firm is operating over is a feature of why some firms are small and others are large.

FSD by Industry Counts

In this paper we explore what age proxies for as a source of firm size. Our candidate is the time needed to build product portfolios in the history of industry. We examine whether this mechanism can generate similar basic stylized facts to FSD as proposed by Cabral and Mata (2003) by plotting the FSD of single and multi industry cohorts (4-digit NACE).

In Table 8 we report the degree of skewness for both countries and we see that the level of skewness is decreasing, becoming more symmetrical, in firms that host multi-products. Clearly age and collection of products, as theory predicts (Klepper and Thompson (2006) and Sutton (1998)) are proxies for each other to some degree. Taking Belgium first, in figure 4.2 we plot the kernel density estimates of the FSD by single industry and multi-industry firms measuring size as the market share of assets within 4-digit industries. We see that Belgian firms operating in more than one industry have size distributions further to the right. In figure 4.1 we see that for the UK shifts of FSD with industry diversification are not as clear cut.

In line with our regressions it is interesting to see whether we get an effect on the FSD from product counts over and above those generated by age cohorts. In figure 5.1 to 5.4 for the UK and Belgium we see that, once we condition on age cohorts, multi-industry firms aged

¹³ In table 4 we present the results having estimated age non-parametrically. It should be noted that when estimating the effect of age non-parametrically, a coefficient for age is not generated and therefore one is not included. The non-parametric estimations are performed in the statistical package R, where the default smoothing parameter is used.

¹⁴ The coefficients on the majority of industry dummies are statistically significant.

ten years or more have size distributions that are much further to the right in companies that operate over three or more 4-digit industries. Within age cohorts we see clear signs of independent product diversification effects on the FSD. Those that document and model economies of scope, whether due to competition (lotteries) over the opening of sub-markets during history or differences in supply side primitives such as entry costs, or a combination, have identified an important empirical factor that drives the FSD.¹⁵ Product diversification is not fully captured by age but has a clear independent partial effect.

Conclusions

We have showed the related stylized facts for the FSD for age cohorts, as proposed by Cabral and Mata (2003), bind within 4-digit manufacturing industries in the UK and Belgium. Klepper and Thompson (2006) and Sutton (1998)) suggest that the role of age is accounted for by the time needed for companies to expand their product portfolio. They open up the possibility that age and product counts could proxy for it other in the determination of firm size. Yet, they realize that cost reductions coming from economies of scope would lead to more inequality in size and an effect that would be greater than that accounted for by age.

International trade papers such as Bernard, Redding and Schott (2006) and Nocke and Yeaple (2008) do link firm size to inter-industry (4-digit) economies of scope that result from supply side heterogeneity in firm's productivity, capabilities and costs. Here product diversification is not linked directly to age.

We show that the 4-digit industry count has an independent effect on firm size (in terms of the market share of assets within 4-digit industries), controlling for the age, among other factors. We provide evidence that time to build a product portfolio across industries is a

¹⁵ Other reasons for scope economies include the need for risk diversification for shareholders, economies of scale in distribution (trade) costs and internalizing the costs of intermediates in vertical chain of production.

mechanism that does relate age to firm size to a certain degree. One can create Cabral and Mata's(2003) stylized facts for the FSD using counts of inter-industry (4-digit) product diversification rather than age cohorts. Firms operating in more than one industry have size distributions further to the right and more symmetrical. A better approach is to document an effect on the FSD from product counts that are over and above those generated by age cohorts. Once we condition on age cohorts, multi-industry firms aged ten years or more have size distributions that are much further to the right in companies that operate over three or more 4-digit industries. Clearly product diversification and age interact but there are clear independent and partial effects coming from inter industry diversification, particularly in older companies.

The implications for empirical work are important. For example, taking the productivity literature, the notion that multi-product firms can affect the estimation of total factor productivity has been shown by Levinsohn and Melitz (2000) and more recently by De Loecker (2006). If products are spread across industries one would need output, raw material and capital deflators to be a weighted sum of 4-digit deflators across several industries.

In the international trade literature, Bernard, Redding and Schott (2006) show that with increased international economic integration, the product portfolio of firms has been changing greatly. If one used company level data to evaluate the impact of trade liberalization on market structure within an industry one would have to realize that company expansions and contractions could result from inter 4-digit activity as well as reallocations intra-industry which is the focus of the new trade literature on firm heterogeneity. Can theoretical, empirical and policy makers continue to ignore the role of inter-industry diversification in companies?

References

Angelini, P. and A. Generale (2008), "On the Evolution of Firm Size Distribution: Further Empirical Evidence", *American Economic Review* 98(1), pp. 426-438.

Audretsch, D.B. (1995), "Innovation and Industry Evolution", Cambridge, MA: MIT Press.

Bernard, A.B., S. J. Redding and P. K. Schott. (2005), 'Multi-Product Firms and Product Switching', NBER Working Paper No. 12293.

Bernard, A.B., S. J. Redding and P. K. Schott. (2006), "Multi-Product Firms and Trade Liberalization", December 2006.

Berry, S., J. Levinsohn and A. Pakes (1995), "Automobile Prices in Market Equilibrium" *Econometrica*, Vol. 63, pp.841-890.

Buzzacchi, L. and T. Valetti (2006), "Firm Size Distribution: Testing the Independent Submarkets Model in the Italian Motor Insurance Industry", *International Journal of Industrial Organisation*, Vol 24(4), pp. 809-834.

Cabral, L. and J. Mata (2003), "On the Evolution of the Firm Size Distribution: Facts and Theory", *American Economic Review*, Vol. 93(4), pp.1075-1090.

Caves, R.E. (1998), "Industrial Organization and New Findings on Turnover and Mobility of Firms", *Journal of Economic Literature*, Vol. XXXVI, pp.1947-1982.

de Juan, R. (2003), “The Independent Submarkets Model: An Application to the Spanish Retail Banking Market”, *International Journal of Industrial Organization*, Vol. 21, pp.1461-1488.

De Loecker, J. (2006), “Product Differentiation, Multi-Product Firms and structural Estimation of Productivity”, *NBER working paper series* No.13155.

Dunne, T., M.J. Roberts and L. Samuelson (1989), “The Growth and Failure of U.S. Manufacturing Plants”, *Quarterly Journal of Economics*, vol.104, pp.671-698.

Ericson, R. and A. Pakes (1995), “Markov-Perfect Industry Dynamics: a framework for Empirical Work”, *Review of Economic Studies*, Vol.62, pp.53-82.

Gibrat, R. (1931), *Les inégalités économiques; applications: aux inégalités des richesses, à la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc., d'une loi nouvelle, la loi de l'effet proportionnel*. Librairie de Recueil Sirey: Paris.

Gollop, F.M. and J.L. Monahan (1991), “A Generalized Index of Diversification: Trends in U.S. Manufacturing”, *The Review of Economics and Statistics*, Vol. 73, No. 2, pp.318-330.

Hall, B.H. (1987), “The Relationship between Firm Size and Firm Growth in the U.S. Manufacturing Sector”, *Journal of Industrial Economics*, Vol. 35, pp.583-605.

Hopenhayn, H.A. (1992), "Entry, Exit and Firm Dynamics in Long Run Equilibrium", *Econometrica*, vol. 60, pp.1127-1150.

Hart, P.E. and S.J. Prais (1956), "The Analysis of Business Concentration: A Statistical Approach", *Journal of Royal Statistical Society*, pp.119-150.

Ijiri, Y. and H. Simon (1977), *Skew Distributions and Sizes of Business Firms*, Amsterdam: North-Holland Pub. Co.

Ijiri, Y. and H. Simon (1964), "Business Firm Growth and Size", *American Economic Review*, Vol. 54, pp.77-89.

Jovanovic, B. (1982), 'Selection and the Evolution of Industry', *Econometrica*, Vol.50, pp.649-70.

Klepper, S. and P. Thompson (2006), "Submarkets and the Evolution of Market Structure", *Rand Journal of Economics*, **37**(4):861-886.

Leinson, J. and M. Melitz (2002), "Productivity in a Differentiated Products", *Harvard mimeo*.

Luttmer, G.J. (2007) "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, 2007, Vol. 122, No. 3, 1103-1144.

Melitz, M. (2003) “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71(6), 1695—1725.

Mariuzzo, F., P.P. Walsh and C. Whelan (2003), ‘Firm Size and Market Power in Carbonated Soft Drinks’, *Review of Industrial Organization*, Vol. 23, Issue 3-4, pp.283-299.

Nocke, V., and S. Yeaple (2008), “Globalizaation and the Size Distribution of Multiproduct Firms” CEPR Discussion Papers no. 6948,

Schmalensee, R. (1989), “Inter-industry Studies of Structure and Performance”, in Schmalensee R. and R. Willig (eds.), *Handbook of Industrial Organization*, Vol. 2 Amsterdam: North-Holland, 1989, pp.951-1009.

Sutton, J. (1991), *Sunk Costs and Market Structure: Price Competition, Advertising and the Evolution of Concentration*, Cambridge, MA: MIT Press.

Sutton, J. (1997), ‘Gibrat’s Legacy’, *Journal of Economic Literature*, Vol. 35, pp.40-59.

Sutton, J. (1998), *Technology and Market Structure*, Cambridge, MA: MIT Press.

Walsh, P.P. and Whelan, C. (2002). "Portfolio Effects and Firm Size Distribution - Carbonated Soft Drinks," *The Economic and Social Review*, , vol. 33(1), pages 43-54

Table 1: Distribution of secondary activities within UK manufacturing for multi-industry firms

		Secondary Nace																					
		15	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
Primary Nace	15	85%	1%	0%	0%	1%	0%	1%	1%	5%	0%	0%	0%	1%	2%	0%	0%	0%	0%	0%	0%	2%	0%
	17	0%	56%	14%	2%	1%	1%	3%	0%	3%	2%	1%	1%	2%	2%	0%	1%	0%	0%	0%	0%	12%	0%
	18	0%	10%	65%	6%	0%	1%	2%	0%	1%	2%	0%	1%	0%	2%	0%	0%	0%	1%	1%	0%	7%	0%
	19	0%	5%	14%	34%	0%	2%	5%	0%	2%	2%	7%	0%	2%	2%	0%	0%	0%	0%	0%	0%	22%	0%
	20	0%	0%	0%	0%	35%	3%	2%	0%	0%	8%	5%	1%	9%	1%	1%	2%	0%	0%	1%	1%	30%	1%
	21	1%	0%	0%	0%	2%	21%	46%	0%	3%	11%	1%	0%	1%	2%	0%	0%	0%	0%	0%	0%	9%	2%
	22	0%	0%	0%	0%	0%	5%	88%	0%	0%	1%	0%	0%	1%	1%	0%	0%	0%	0%	0%	0%	2%	0%
	23	0%	0%	0%	0%	0%	0%	0%	0%	63%	8%	0%	0%	8%	4%	0%	4%	0%	0%	0%	0%	12%	0%
	24	2%	1%	0%	0%	0%	1%	1%	1%	61%	7%	2%	2%	5%	4%	0%	1%	0%	3%	0%	0%	6%	1%
	25	0%	1%	1%	1%	1%	2%	2%	0%	2%	22%	6%	2%	26%	10%	0%	1%	1%	2%	1%	2%	15%	2%
	26	0%	0%	0%	0%	4%	0%	0%	0%	2%	3%	51%	2%	11%	2%	0%	3%	0%	2%	2%	1%	14%	0%
	27	0%	0%	0%	0%	0%	0%	2%	0%	1%	1%	0%	47%	37%	6%	0%	1%	0%	0%	0%	0%	4%	0%
	28	0%	0%	0%	0%	1%	0%	1%	0%	0%	2%	0%	2%	49%	22%	1%	5%	1%	1%	2%	3%	5%	0%
	29	0%	0%	0%	0%	0%	0%	1%	0%	1%	2%	0%	1%	19%	45%	2%	11%	2%	6%	1%	3%	4%	0%
	30	0%	0%	1%	0%	1%	1%	2%	0%	0%	0%	0%	0%	0%	3%	4%	33%	26%	21%	1%	1%	6%	1%
	31	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%	0%	12%	6%	4%	24%	25%	18%	1%	3%	6%	0%
	32	0%	0%	0%	0%	0%	0%	2%	0%	0%	1%	1%	1%	2%	3%	5%	20%	33%	32%	0%	3%	12%	0%
	33	0%	1%	0%	1%	0%	0%	1%	0%	4%	3%	1%	0%	12%	7%	1%	8%	4%	39%	2%	3%	13%	0%
	34	0%	0%	0%	0%	1%	0%	0%	0%	1%	5%	0%	0%	14%	6%	0%	2%	0%	2%	37%	17%	14%	0%
	35	0%	1%	0%	0%	0%	0%	1%	0%	2%	2%	0%	0%	22%	8%	0%	4%	1%	2%	5%	39%	14%	0%
36	0%	2%	0%	1%	5%	1%	3%	0%	2%	4%	3%	1%	5%	3%	0%	1%	0%	2%	1%	0%	61%	1%	
37	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	2%	0%	3%	0%	0%	0%	0%	0%	0%	0%	94%	

Table 2: Distribution of all secondary activities within Belgian manufacturing for multi-industry firms

	Secondary Nace																						
	15	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	
15	92%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%
17	0%	82%	8%	0%	1%	0%	3%	0%	2%	2%	1%	0%	2%	0%	0%	0%	0%	0%	1%	1%	2%	2%	0%
18	0%	19%	72%	5%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	6%	0%
19	0%	5%	20%	35%	0%	0%	0%	0%	5%	5%	0%	0%	0%	10%	0%	0%	0%	10%	0%	0%	10%	10%	0%
20	1%	0%	1%	1%	49%	1%	0%	0%	1%	4%	2%	0%	15%	2%	0%	0%	0%	0%	0%	0%	20%	2%	0%
21	1%	5%	0%	0%	3%	71%	9%	0%	1%	4%	0%	0%	2%	1%	0%	0%	1%	0%	0%	0%	2%	1%	0%
22	0%	1%	0%	0%	0%	3%	94%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%
23	0%	5%	0%	0%	0%	5%	0%	5%	53%	11%	0%	0%	0%	5%	0%	5%	0%	0%	0%	0%	0%	0%	11%
24	3%	7%	0%	0%	1%	1%	1%	2%	65%	11%	3%	0%	1%	2%	0%	0%	0%	1%	1%	0%	2%	1%	0%
25	0%	3%	2%	2%	6%	5%	2%	0%	12%	43%	2%	0%	11%	3%	0%	0%	0%	0%	3%	0%	3%	3%	0%
26	0%	0%	0%	0%	2%	1%	1%	0%	2%	4%	81%	0%	13%	2%	0%	0%	0%	0%	0%	0%	1%	1%	0%
27	0%	0%	0%	0%	1%	0%	1%	0%	1%	3%	1%	25%	55%	8%	0%	1%	0%	0%	0%	0%	1%	1%	0%
28	0%	0%	0%	0%	2%	0%	0%	0%	0%	2%	0%	7%	73%	11%	0%	1%	0%	0%	1%	0%	1%	0%	0%
29	0%	1%	0%	0%	1%	0%	0%	0%	1%	1%	0%	1%	23%	65%	0%	4%	1%	2%	2%	0%	1%	0%	0%
30	0%	0%	0%	0%	0%	0%	6%	0%	0%	3%	0%	0%	10%	6%	16%	10%	23%	23%	0%	3%	0%	0%	0%
31	0%	1%	0%	0%	0%	0%	2%	0%	0%	2%	1%	0%	13%	10%	1%	42%	13%	6%	3%	0%	6%	0%	0%
32	0%	1%	0%	0%	1%	0%	5%	0%	1%	1%	0%	0%	4%	5%	10%	24%	33%	17%	0%	0%	0%	0%	0%
33	1%	0%	3%	11%	1%	0%	1%	0%	4%	0%	1%	0%	5%	4%	1%	5%	6%	57%	0%	1%	0%	0%	0%
34	0%	4%	0%	0%	3%	0%	1%	0%	1%	3%	1%	1%	19%	7%	0%	2%	0%	1%	54%	0%	2%	0%	0%
35	2%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	28%	16%	0%	4%	0%	2%	6%	24%	8%	0%	0%
36	0%	4%	2%	0%	10%	1%	2%	0%	2%	2%	2%	1%	7%	1%	0%	1%	1%	1%	1%	0%	65%	0%	0%
37	9%	11%	0%	0%	5%	2%	2%	2%	4%	5%	13%	5%	5%	2%	0%	0%	0%	2%	0%	0%	0%	34%	0%

Table 3: Measure of skewness by age cohorts

	(i)	(iii)
	UK	Belgium
Full sample	0.16	0.29
Age < 5	0.29	0.56
Age ≥ 5 & Age < 10	0.22	0.31
Age ≥ 10 & Age < 25	0.08	0.26
Age ≥ 25	0.15	0.10

Table 4: Firm level size regression dependent variable: $\ln(\text{market share of assets within 4-digit industries})$

	(i)	(ii)	(iii)	(iv)
	UK	UK	Belgium	Belgium
Ln(Counts)	0.41* (0.02)	0.18* (0.01)	0.68* (0.03)	0.41* (0.02)
Ln(Age)	-	-	-	-
Constant	- 9.7* (0.009)	- 6.68* (0.20)	-7.38* (0.02)	-6.80* (0.12)
4-digit dummies	No	Yes	No	Yes
R2	0.09	0.43	0.08	0.37
Observations	78911	78911	21696	21696

Standard errors in parenthesis.* indicates statistically significant at the 1% level, ** indicates statistically significant at the 5% level and *** indicates statistically significant at the 10% level.

Table 5: Firm level size regression dependent variable: $\ln(\text{market share of assets within 4-digit industries})$

	(i)	(ii)	(iii)	(iv)
	UK	UK	Belgium	Belgium
ln(Counts)	0.41*	0.19*	0.69*	0.44*
	(0.02)	(0.02)	(0.03)	(0.03)
Age	0.04*	0.04*	0.03*	0.03
	(0.001)	(0.0004)	(0.001)	(0.001)
Constant	-10.41*	-7.58*	-7.98*	-7.37*
	(0.01)	(0.20)	(0.03)	(0.12)
4-digit dummies	No	Yes	No	Yes
R2	0.09	0.41	0.08	0.37
Observations	78911	78911	21696	21696

Standard errors in parenthesis.* indicates statistically significant at the 1% level, ** indicates statistically significant at the 5% level and *** indicates statistically significant at the 10% level.

Table 6: Firm level size regression dependent variable: $\ln(\text{market share of assets within 4-digit industries})$

	(i)	(ii)	(iii)	(iv)
	UK	UK	Belgium	Belgium
ln(Counts)	0.43*	0.20*	0.68*	0.42*
	(0.02)	(0.02)	(0.03)	(0.03)
ln(Age)	0.73*	0.81*	0.55*	0.54*
	(0.01)	(0.01)	(0.02)	(0.01)
Constant	-11.42*	-8.61*	-8.80*	-8.19*
	(0.02)	(0.20)	(0.05)	(0.13)
4-digit dummies	No	Yes	No	Yes
R2	0.09	0.43	0.07	0.37
Observations	78911	78911	21696	21696

Standard errors in parenthesis.* indicates statistically significant at the 1% level, ** indicates statistically significant at the 5% level and *** indicates statistically significant at the 10% level.

Table 7: Firm level size regression dependent variable: $\ln(\text{market share of assets within 4-digit industries})$

	(i)	(ii)	(iii)	(iv)
	UK	UK	Belgium	Belgium
ln(Count)	0.42* (0.02)	0.20* (0.02)	0.73* (0.03)	0.48* (0.03)
1/Age	-2.94* (0.05)	-3.42* (0.04)	-2.56* (0.1)	-2.5* (0.09)
Constant	-9.2* (0.01)	-6.24* (0.21)	-7.1* (0.02)	-6.58* (0.13)
4-digit dummies	No	Yes	No	Yes
R2	0.05	0.39	0.06	0.36
Observations	78911	78911	21696	21696

Standard errors in parenthesis.* indicates statistically significant at the 1% level, ** indicates statistically significant at the 5% level and *** indicates statistically significant at the 10% level.

Table 8: Measure of skewness by Product Count cohorts

	(i)	(iii)
	UK	Belgium
Full sample	0.16	0.29
Single Product	0.19	0.44
Two Products	0.15	0.24
Mutli-Products	0.11	0.17

Figure 1.1:
Age distribution of all UK firms

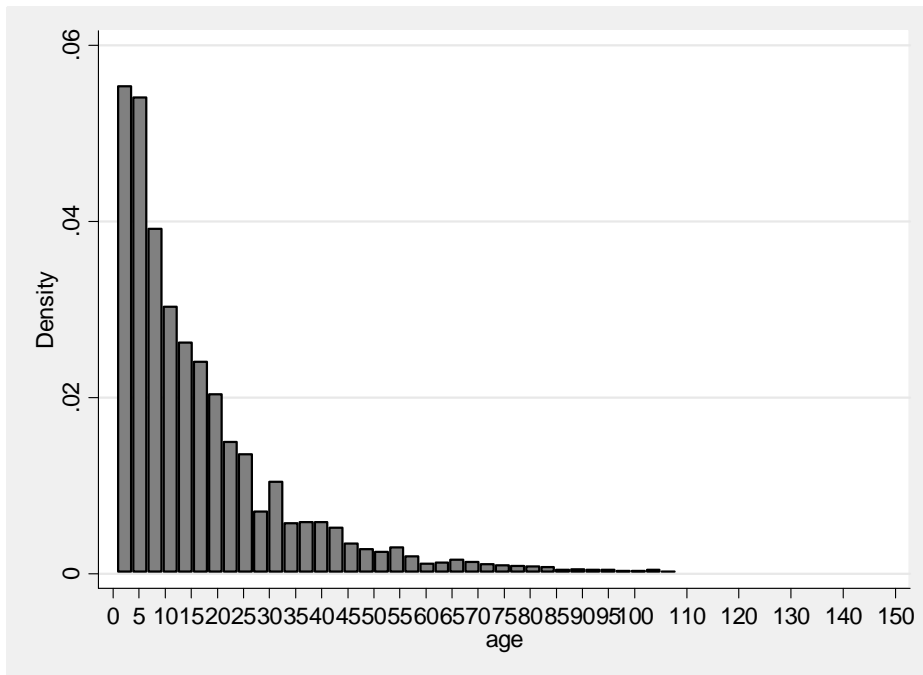


Figure 1.2:
Age distribution of all Belgian firms

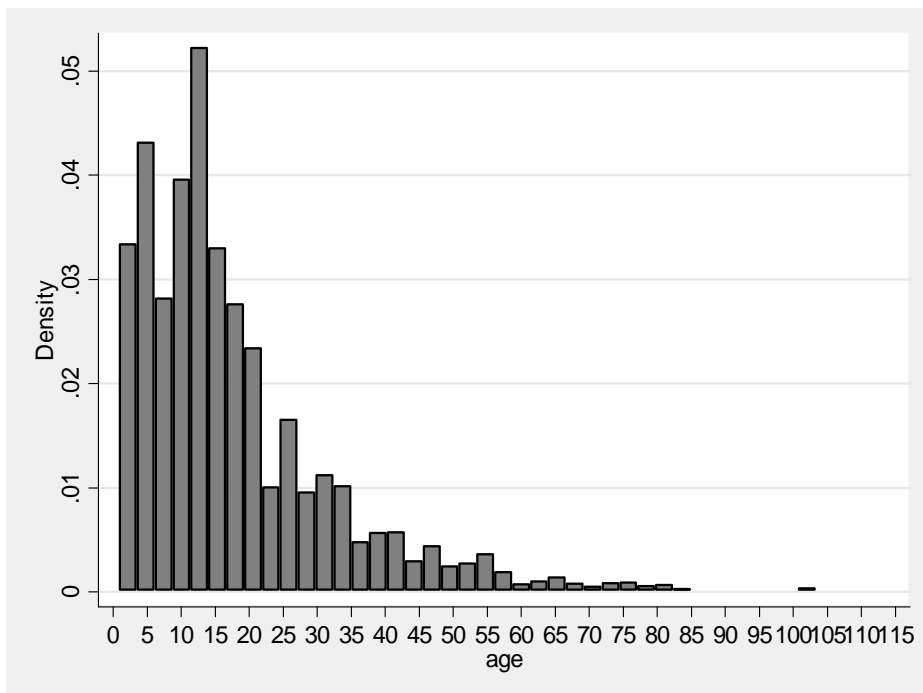


Figure 2.1:
Size distribution of UK firms by age cohorts

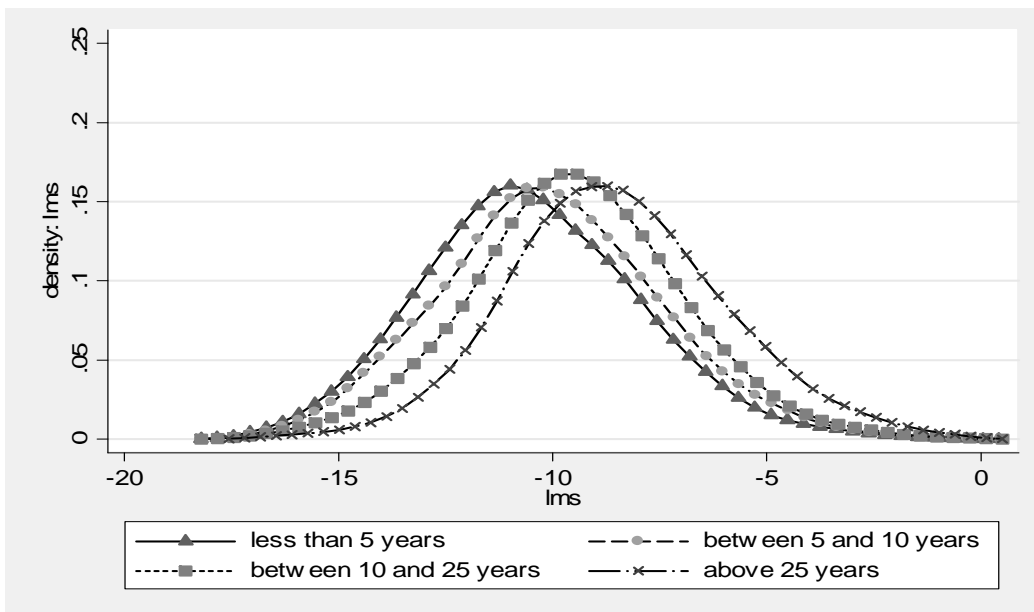


Figure 2.2:
Size distribution of Belgian firms by age cohorts

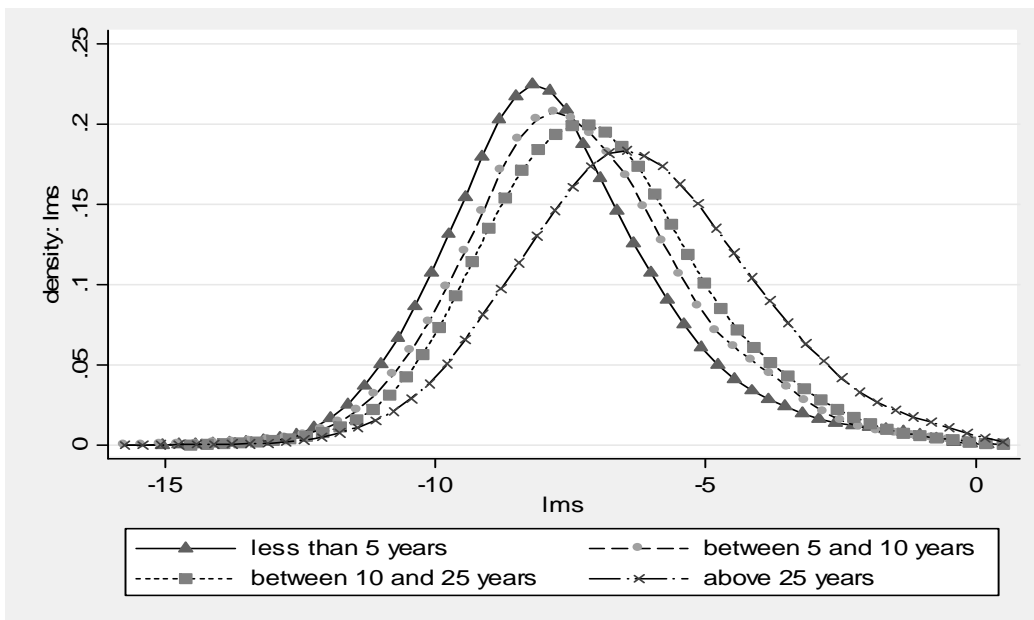


Figure 3.1: Top 4,8 and 20 UK companies concentration ratios in each 4 digit industries

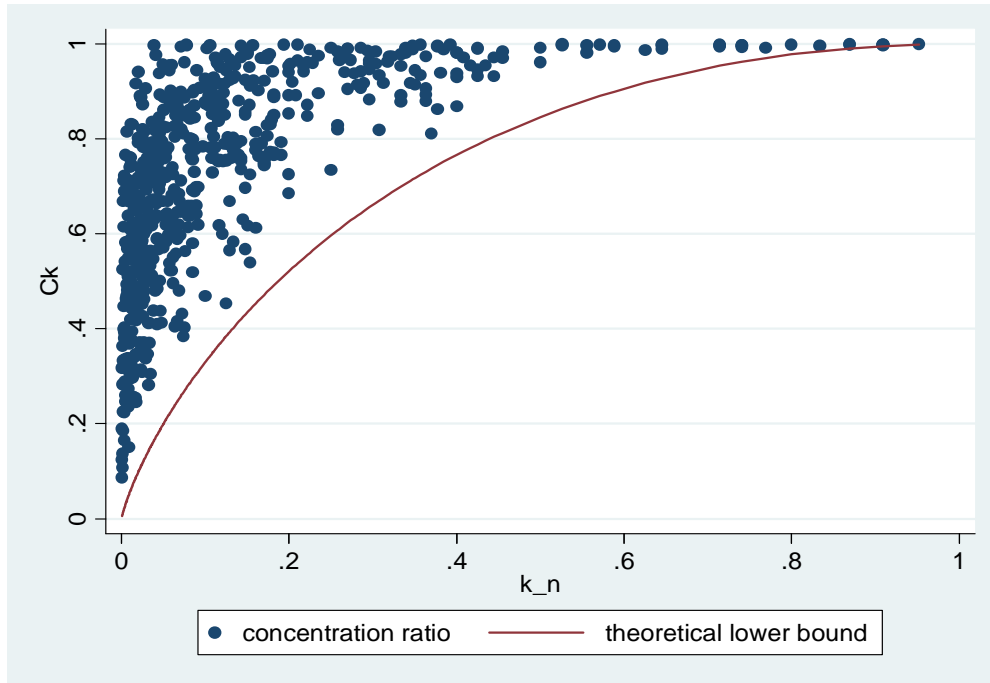


Figure 3.2: Top 4, 8 and 20 Belgian companies concentration ratios in each 4 digit industries

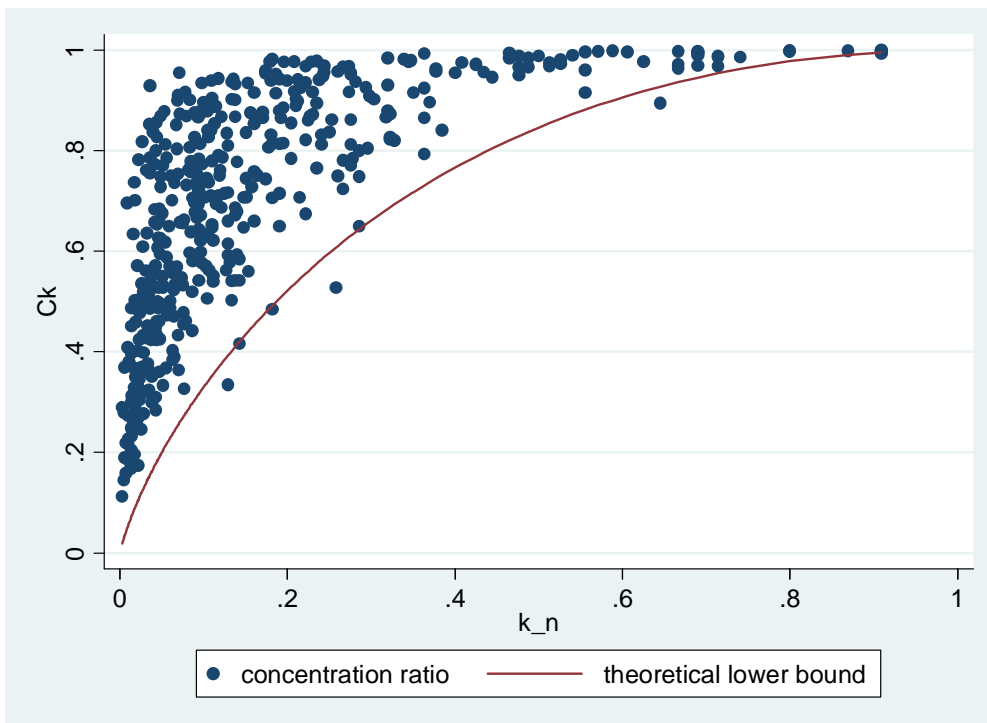


Figure 4.1: Size distribution of UK companies by single and multi industry cohorts

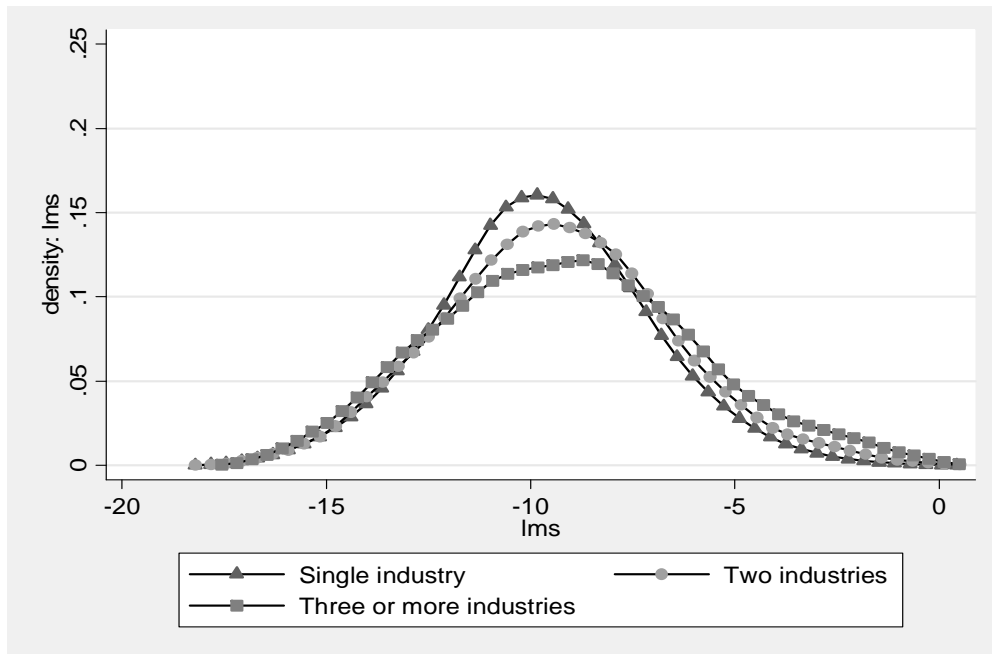


Figure 4.2: Size distribution of Belgian companies by single and multi industry cohorts

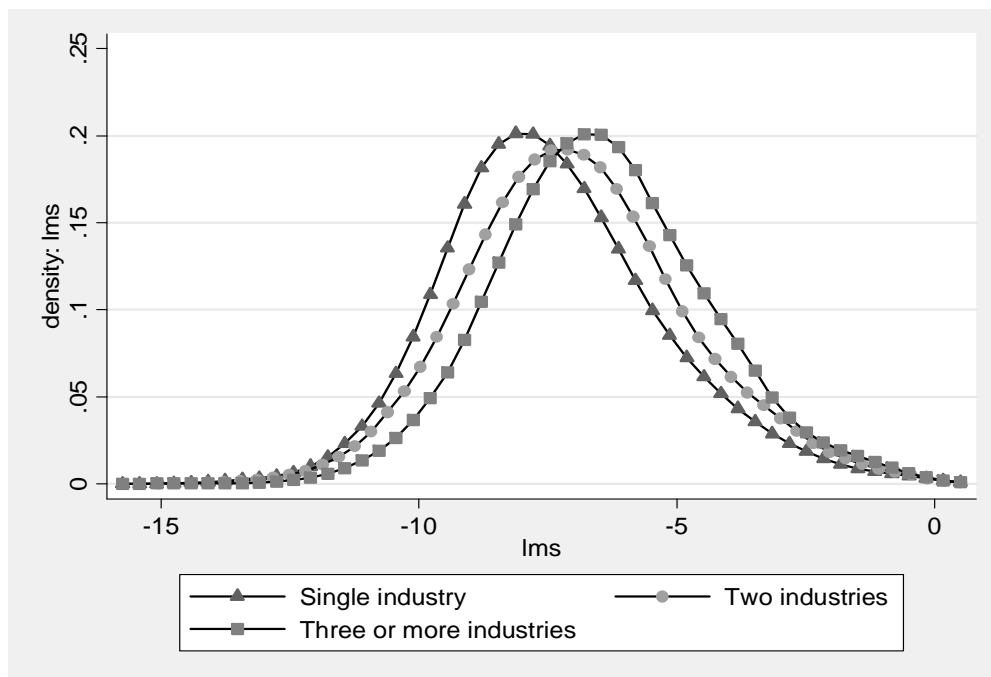


Figure 5.1: Size distribution of UK companies by single and multi industry cohorts age ≥ 10 &age < 25

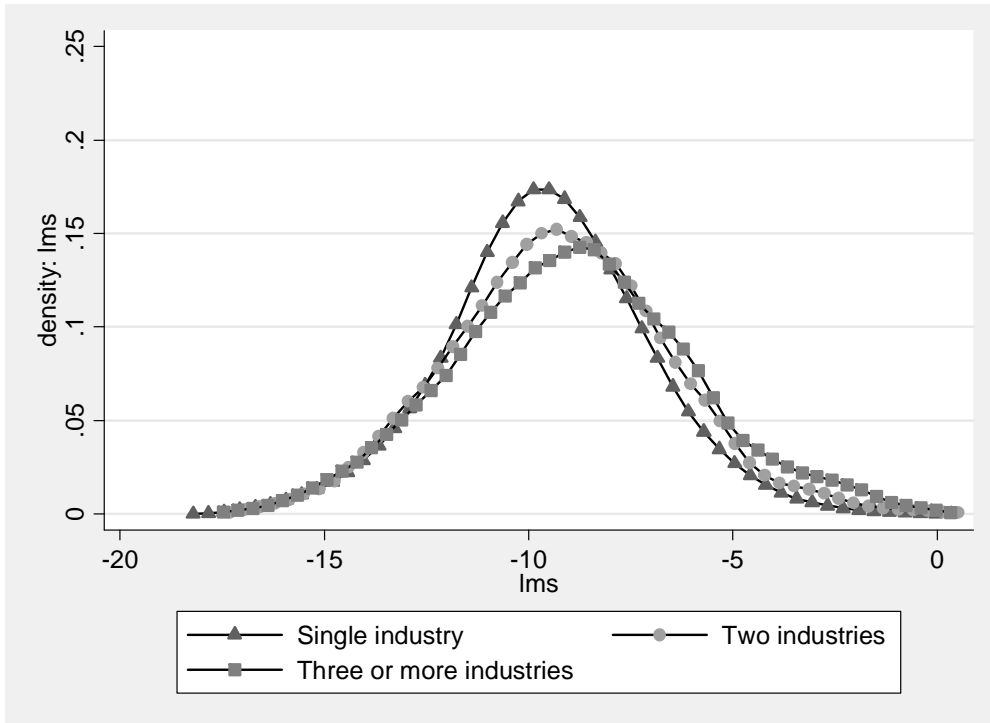


Figure 5.2: Size distribution of UK companies by single and multi industry cohorts age ≥ 25

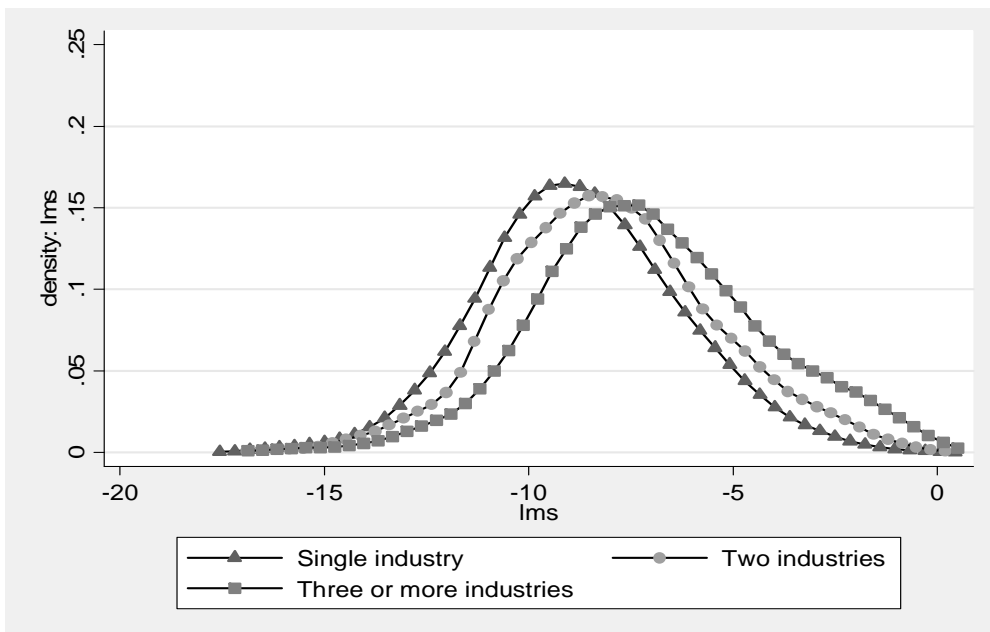


Figure 5.3: Size distribution of Belgian companies by single and multi industry cohorts age ≥ 10 &age < 25

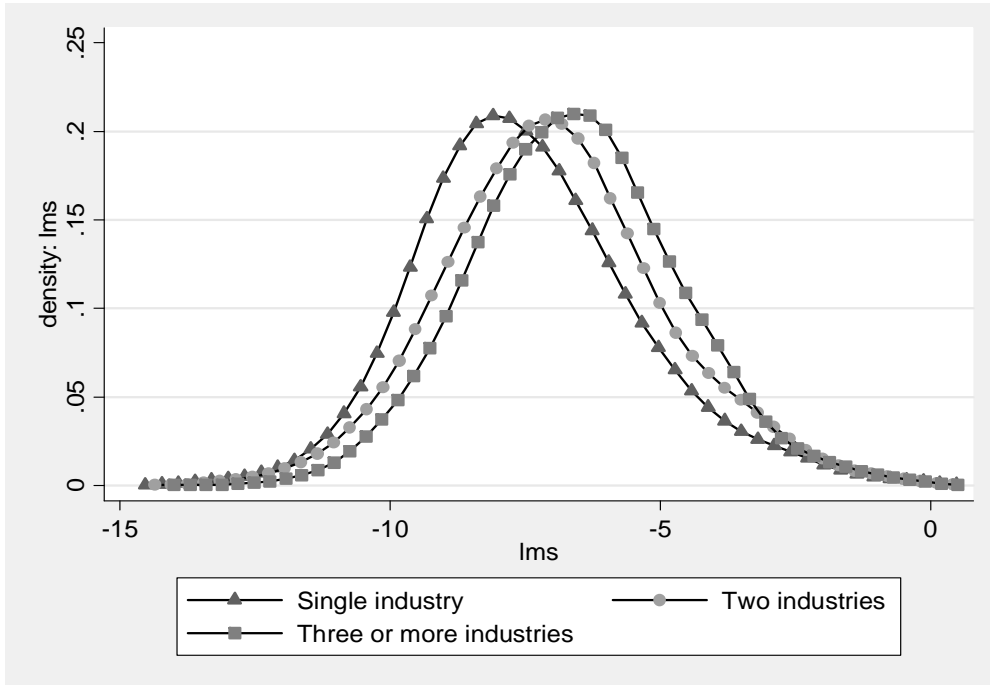


Figure 5.4: Size distribution of Belgian companies by single and multi industry cohorts age ≥ 25

