Company Size, Industry and Product Diversification Evidence from the UK and Belgium

John Hutchinson[a], Jozef Konings[ab] and Patrick Paul Walsh[c]

a: LICOS, Centre for Transition Economics, K.U.Leuven

b: Department of Economics, K.U.Leuven

c: Trinity College Dublin, Ireland

March 2006

In the UK and Belgium, the Sutton (1998) mathematically derived lower bound to the firm size distribution is binding across all manufacturing industries. Using 8-digit product counts, we verify the Sutton theory that large firms are large as a result of operating across more sub-industries than smaller firms by showing that product count is indeed a key determinant of firm size, measured in terms of asset size, controlling for the age of companies, amongst other factors

Key Words: Firm Size, Product Counts, Manufacturing Industries.

JEL Classification: L11, L25, L66, L81

1 Introduction

Investigating firm size distributions (FSD) has received considerable attention within the growth of firms literature initiating with Gibrat (1931) "law of proportionate effect". In more recent times, Cabral and Mata (2003) pioneer an evolutionary approach to explain the FSD and show that the distribution of the logarithms of firm size of a given cohort is skewed to the right at the time of birth but gradually moves towards a normal distribution. 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 not due to firm selection processes. They propose that as firms get older they face increasingly less financial constraints and subsequently this determines the evolution of the FSD. However, Klepper and Thompson (2005) show through an alternative mechanism, namely via submarkets, that they are able to match the predicted properties of the size distribution of firms of Cabral and Mata (2003). In their study, Klepper and Thompson (2005) further emphasize the role submarkets play in determining firm size, age, growth and survival. They re-examine the firm age-growth relationship from an alternative perspective to previous approaches and show that all age-size regularities can be explained by a simple framework based on the concept of submarkets (product lines).

It is through this submarket approach that Sutton (1998) provides an explanation to an important empirical regularity that has become more pronounced over-time and which previous literature had left us with little theory to explain. This regularity relates to the positive relationship between firm size and product diversification that has been documented for manufacturing at regular time intervals since the 1960s (see Gollop and Monahan, 1991). Across the general run of industries, larger firms co-exist with many small firms where larger firms host more product lines than small firms. Gollop and Monahan (1991) see this as the most important structural phenomena of the postwar period. Sutton (1998) puts forward an evolutionary idea that focuses on heterogeneity in the collection of opportunities (product lines) during industry evolution as a determinant of the FSD. Companies grow by expanding their product lines or geographically areas. Sutton (1998), using manufacturing data for Germany in 1990 and the US in 1987, shows that the mathematically derived lower bound to the FSD is binding across 4-digit industries. The major shortcoming with his empirical evidence is that, using relatively aggregated data, one cannot discriminate whether this outcome is driven by market definition problems, large firms operating in a different set of sub-industries to small companies or product portfolio effects, where large companies are large because they have a greater coverage of subindustries than small companies[1]. The key insight of the Sutton (1998) theory is that outcomes should be driven by the latter, product portfolio effects. In this paper, using richer data, we document the nature of aggregation of companies over 8-digit products within 242 different 4-digit industries to show that product counts of companies are an important empirical reason why the bound is binding using manufacturing data for the UK and Belgium. We show similar outcomes in terms of the FSD by age cohorts, using Cabral and Mata (2003) but we also show similar outcomes in terms of the FSD are obtained by multi-industry and multi-product cohorts, for the UK and Belgium. This highlights the role of 8-digit product counts as a key determinant of firm size, measured in terms of asset size, controlling for the age of companies, amongst other factors.

In addition to verifying the key insight of the Sutton (1998) theory, the impact of product portfolio effects reinforces the notion that the modelling of competition

[1] Sutton discusses that when using relatively aggregated data that a four digit industry could be made up of five digit industries that are very different from each other. He refers to this as “independent subindustries”.

within industries, of any aggregation, should not be done with just company level data.

In section II we review the approach put forward by Sutton (1998). Section III introduces the data that we use and highlights the difference between our approach and the existing evidence to date. In section IV and we give and discuss our empirical results, while section V concludes the paper.

II. The Sutton Approach

Sutton (1998) marries a game theoretic approach on firm growth with elements of the stochastic approach. Firm growth is modeled as a collection of discrete investment opportunities (product lines), which arrive over an infinite period as an outcome of a stochastic process. While the evolution of the investment opportunities is a stochastic process, the manner in which firm collect these opportunities (firm growth) is deterministic. The limiting firm size distribution is an outcome of deterministic entry games among active firms and potential entrants across investment opportunities. Differences in firm size emerge due to firms taking up a different count of investment opportunities during market evolution. Testing whether this proposition holds across 4-digit nace industries is the focal point of our empirical work[2].

[2] One literature tried to use only “stochastic processes” in firm growth to explain the FSD. In this vein, Gibrat (1931) postulates that the size-growth relationship for active firms 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. No *stochastic* mechanism in firm growth could explain observed size distributions. Another literature, Dunne, Roberts and Samuelson (1988), focus on modeling firm growth as a deterministic outcome. Using rich firm level data, they suggest that the relationship between firm growth and firm characteristics, including size, is more complex than a stochastic process. In general, these empirical studies seem to point to the success of idiosyncratic firm and sector characteristics and conclude that no simple *deterministic* mechanism explains short run firm growth and, as a result, determines the limiting FSD. Most studies conclude that that Gibrat’s law fails to hold. The failure of Gibrat´s law is motivated by the Jovanovic (1982) theory of firm selection and industry evolution under ex-ante uncertainty concerning the ex-post performance of firms. Extensions of Jovanovic (1982) can be found in Hopenhayn (1992) and Ericson and Pakes (1995). As we demonstrate, using Cabral and Mata (2001) and Sutton (1998), it is the nature of the evolution of surviving companies that really matters for the FSD.

Assuming opportunities of equal size, Sutton (1998) imposes only one restriction, a *Symmetry Principle*, on the form of the entry game into each of these opportunities to model a lower bound on the FSD. In the limit, the FSD is restricted to a lower bound Lorenz curve, with a measure of inequality that is approximately equal to a Gini coefficient of 0.5, that graphs the fraction of top k ranking firms in the population N of firms (k/N) against their corresponding share of market sales given by the k-firm concentration ratio (Ck) that satisfies,

$$C_k \geq \frac{k}{N}\left(1 - \ln\frac{k}{N}\right) \tag{1}$$

where the size of the market is the total number of opportunities captured by all firms, and the size of each firm is total number of opportunities captured by the firm. Introducing a size advantage (scope economies) in the take up of investment opportunities or allowing for differences in the size of these opportunities (competition within opportunities) will have the effect of introducing greater heterogeneity between firms in the market and a resulting Lorenz curve that will lie inside the lower bound. The main theoretical outcome is that large firms will host more investment opportunities than small in mature industries. Firm configurations of opportunities are predicted to *at least* induce a FSD that is greater than or equal to Suttons' (1998) mathematically derived lower bound and could explain most of the observed FSD.

III. Existing Evidence and Data

Sutton (1998) provides evidence for his theory using two approaches. One is based on cross sections of 4-digit manufacturing industries and another is based on a specific industry (The US Cement Industry). To date most empirical validations of this theory use industry studies focus on investment 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 and Valletti, 1999). These studies find that firm size is mainly determined by the degree to which firms have counts of geographical locations and not by the nature of competition within geographic locations. Due to the richness of data all these studies could understand how firm (aggregate) market share is determined by the nature of aggregation over geographical locations. The nature of the aggregation can take two forms. The first, big firms target large geographical locations and small firms locate away in peripheral locations. This would be consistent with the observed aggregate distribution of firm size. The Sutton theory is based on the second outcome, within each geographical location firms have similar market share. What makes a firm big in aggregate is that it has a presence across more geographical locations than small firms. All these studies document that nature of aggregation of companies over geographical locations to be of the latter kind. Mariuzzo, Walsh and Whelan (2003) address this issue in a multi-product industry, Carbonated Soft Drinks, where investment configurations reflect outcomes of investments over product and location space. The Coca Cola company is not big because its specializes in Cola, but because it invested in brands across all flavour packaging and diet product lines, with a presence in most stores. Small companies have brands that can extract market share within segments as high as multi-nationals but their portfolio of such brands tends to be narrow. The nature of aggregation of companies over product lines is of the second form.

In this paper we revisit the cross section approach used in Sutton (1998). 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-industries within 4-digit industries, his data does not allow him to discriminate between the nature of the aggregation, outlined in the industry studies. In this paper we examine whether company size is mainly an outcome of the size of its 8-digit product count, controlling for age and other factors[3]. In other words, is the nature of aggregation of companies over 8-digit product lines across manufacturing industries such that large firms are large because they have a greater coverage of products than smaller firms?

The cross section approach used in Sutton (1998) had a few other less important 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. However, 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

[3] We also investigate whether a firms 4-digit industry count determines firm size.

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 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 in any particular market, without having the deal with the cut-off point often used by the statistical offices.

Our data built on the reported company accounts of manufacturing firms in the UK and Belgium collected by an electronic publishing company, Bureau Van Dijck (BvD), which it commercializes under the name of AMADEUS (see www.bvdep.com)[4]. These data cover virtually the entire population of businesses and have all been officially audited. The data do not cover the single proprietorship companies with zero employees. Due to the accounting legislation companies are not required to report some financial and operational items. However, full information is available on total assets of firms. For this reason, we will compute concentration ratios based on the total assets rather than on sales for the various sectors that we investigate. Even though Sutton (1998) uses sales data, he argues that total assets would be his first choice if the data existed. By using sales it adds an additional element to the inequality in the size distribution, thus generating a greater degree of inequality. We gathered full information on firm total assets for active UK and Belgian manufacturing firms with our final sample consisting of 78,911 UK firms and

[4] The Amadeus data set has been increasingly used in academic papers (e.g. Budd et al., Desai et al.2003; Konings et al.,2004; among others).

21,697 Belgian firms in 2002[5]. Another data issue is that Sutton (1998) only considers the 4-digit industries within the 2-digit homogeneous goods industries. We do not exclude industries with generally low advertising and R&D to sales ratios. Yet the theory is a limiting outcome based on the idea that product lines arrive into infinity during industry evolution. Many 4-digit industries can be made up of a small number of 8-digit product lines. This makes it difficult for firms to have different product portfolios in a cross section. We feel the issue is not about advertising and R&D to sales ratios per se. Clearly product proliferation within 4-digit industries is related to taste and technology characteristics that may, or may not, be sensitive to endogenous sunk costs.

Before moving to the results section we investigate the co-production of multi-industry firms in the UK and Belgium. From our data we know both the number and also the type of industries all firms are operating in at the 4-digit level. As a result, we can see which secondary sectors multi-industry firms operate in. We find that 79% of UK firms operate in a single industry while 38% of Belgian firms operate in single industries. Bernard, Redding and Schott (2005) find that 71% of US manufacturing firms operate in a single industry. In table 1 and 2 we present co-production matrices for firms in the UK and Belgium for 2002 at the sector level (2-digit level)[6]. Each row in the matrices shows the percentage of firms operating in particular secondary sectors. For example, in row 1 in table 1, for those UK firms whose primary sector is

---

[5] The industry classification that we use at the 4 digit level is the nace code. For the 8-digit codes, the industry classification we follow here is the CSO activity codes, defined by the British Statistical Office. The advantage of using this classification is that it is an 8-digit CSO classification, which is a classification that gets close to the product classification. This allows us to count the number of products the firm is operating over. A drawback, however, is that the 8-digit CSO codes are only reported for the medium and large sized enterprises in our sample, the codes are only reported in 1999 and we only have the codes for the UK. Our final sample has 5052 firms reporting 8-digit CSO activity codes.

[6] Although we have information on all secondary sectors 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 sector level (2-digit nace) because making matrices using 4-digit level information would be very impractical.

food (nace 15), 85% of their secondary activities are also in the food sector (nace 15). This compares with 95% of Belgian firms as reported in table 2. From table 1 and 2 we see for most sectors that the diagonal contains the highest percentage with co-production occurring within sectors rather than across sectors. This is similar to the findings of Bernard, Redding and Schott (2005) where they do a similar exercise for the United States. For example, for UK firms whole primary sector is in publishing and printing (nace 22), 88% of their secondary industries are also in publishing and printing. This compares to 93% in Belgium. However, for the manufacture of coke, refined petroleum products and nuclear fuel (nace 23) in both countries the diagonal value is very low. Similar to Bernard, Redding and Schott (2005), we also see some intuitive co-production occurs across certain sectors. For example, manufacture of basic metals and fabricated metal products (nace 27) and manufacture of fabricated metal products, except machinery and equipment (nace 28), manufacture of electrical machinery and apparatus (nace 31) and manufacture of radio, television and communication equipment (nace 32).

IV Results

IV.I Lorenz Curves

In figure 1 for the UK and figure 2 for Belgium we plot as in Sutton (1998) the k firm concentration ratio in the market, Ck, 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 in the UK and with only 3 industries violating the bound in the UK[7]. The basic idea behind the theoretical lower bound to firm size distributions is the taking up of new

[7] These industries are the manufacture of musical instruments (nace 3630), the manufacture of other non-metallic mineral products (nace: 2682) and the dressing and dying of fur, manufacture of articles of fur (nace 1830).

opportunities in independent sub-markets. It is derived as product opportunities tend to infinity. Overall, it seems that the Sutton (1998) bound is binding in a cross section of manufacturing industries in both the UK and Belgium in 2002.

Furthermore, by having the number of active firms by four-digit sectors and by using companies' assets to measure firm size, we can overcome some data weaknesses present in Sutton (1998). The main issue with the evidence presented in Sutton (1998) was that he had no data on the way companies operated over 8-digit products. He could not test whether the nature of aggregation flawed the cross section evidence supporting his predictions. In our data we have a list of 8-digit products produced by companies but only for the right tail of the size distribution. In our econometric section we provide evidence that product counts matter for firm size, controlling for age and sector specific effects.

IV.2. Econometric Evidence

We set out to reject the null, that counts of product line tell you nothing about the firm size, in terms of assets, in an econometric model. The basic model of firm size *f* in each 4-digit sector j is as follows,

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

$Counts_f$, is firm counts of 8-digit product lines[8]. Age_f is the number of years since the year of the firms' incorporation. Sector dummies control for unobserved sector effects. We also investigate at the sector level, how the proportion of multi-product firms impacts on the GINI coefficient of the industry.

$$Gini_k = \alpha + \beta_1 pmpf_k + \beta_2 mage_k + \varepsilon_k \qquad (3)$$

[8] We also estimate equation 2 measuring counts as the number of 4-digit industries a firm operates across.

$Gini_k$ is the Gini coefficient for each 3-digit sector k, $pmpf_k$ is the number of multi-product firms in industry k divided by the total number of firms in industry k and $mage_k$ is the mean age of firms in sector k. We first graph the descriptive statistics on age and product counts for the UK and Belgium in 2002. In figure 3 for the UK and figure 4 for Belgium we graph the age distribution of firms in the population of firms. We can note that in both countries there are relatively many young firms, while only a small fraction of all firms are older than 20 years. This suggests that there may be some life cycle effect present explaining firm presence. Moreover, it is striking to note that firm age distributions just as firm size distributions are skewed. We therefore follow the analysis in Cabral and Mata (2003) and plot the kernel density estimates of the FSD by age cohorts in figure 5 for the UK and figure 6 for Belgium, where size is measured as the market share of physical assets within 4-digit sectors. We see from figures 5 and 6 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 goes on, some economic force is pushing the distribution of size on a log scale to be more symmetrical. We can note that the basic result of Cabral and Mata (2003) is confirmed or as firms grow older the size distribution shifts more to the right. This highlights the point that as time goes on, some economic force is pushing a tendency towards log normality in surviving firms.

We offer an additional source driving the firm growth, that is, it is driven by the outcome of an entry game over product opportunities in the history of the industry. We do not have product counts for the entire population of firms in our data, but we do have detailed information on the number of products firms (number of 8-digit CSO codes) are producing for the medium and large sized enterprises in our sample. So, we can only study the right tail of the distribution. However, before turning to our product data, we first plot the size distributions of firms by single and multi industry cohorts (4-digit nace) and then by single and multi-product cohorts (8-digit CSO codes)[9]. In figures 7 and 10 we plot density estimates of the FSD by single industry and multi-industry firms for the UK and Belgium measuring size as the market share of physical assets within 4-digit sectors. In figures 10, 11 and 12 we see that Belgian firms operating in more than one industry have size distributions further to the right than firms operating in a single industry and firms operating in three or more industries having distributions furthest to the right. Furthermore, from columns (i), (iii) and (v) in table 5 we see that the skewness measure reduces for all groups as age increases with firms 25 years or older operating in three or more industries having the most symmetrical FSD. In figure 7 we see that for the UK when we don't control for age, shifts of FSD are not as clear as it is for Belgium. However, once we control for age, multi-industry firms aged ten years or more have size distributions that are further to the right. Similarly, from columns (i), (iii) and (v) in table 4 we see that skewness reduces as firms age with firms 25 years or older operating in three or more industries having the most symmetrical distribution. In figures 13 and 14 we summarize our previous findings by showing 3-dimensional graphs. The x, y and z axis correspond to lnage, lnsize and lnindustrycount respectively. Unsurprisingly, we see that bigger firms are both older and operate across multiple industries[10]. In figure

[9] Although four digit nace is more aggregated data than product level data we have information on all the four digit nace sectors that firms operate in. Therefore, we feel it appropriate to assume that if a firm is active in a four digit industry then it has at least one product in that industry. Therefore, we assume that the sum of four digit industries that a firm is active in constitutes a lower bound to the product count of the firm.

[10] We show a weighted graph in figure 15 and 16 of the log of the number of industries a firm operates across on the y-axis and the log of age on the x-axis. We find that a lot of the firms are single industries and that a lot of young firms are single industries. As age increases more firms are multi-industry and that multi industry firms are also older firms.

17 we plot the size distribution of UK companies for our large firm sample and we find that multi-product firms have size distributions further to the right than single product firms. We also see from table 6 that multi-product firms have a more symmetric distribution. These finding appear to suggest that being multi-product (and multi-industry) plays a role in explaining the size distribution of firms.

We now turn to testing our hypothesis more rigorously in table 7, where we estimate equation (2). We report the estimates for of equation (2) estimating the effect of age non-parametrically allowing us to avoid restrictive assumptions of the functional form of age. In columns (i) and (iii), where Ln (count) refers to the number of four digit nace industries, we report the estimates for the UK and Belgium having not included industry dummies. In columns (ii) and (iv) we include a full set of four digit sector dummies. Interestingly, irrespective of whether we include industry dummies, the coefficient β_1 is always positive and significant. In columns (v) and (vi) we estimate equation (2) using firms count of 8-digit product lines. We find that β_1 is positive and significant where a 10 % increase in the number of products that a firm is producing is associated with a 3% increase in market share. In columns (i) and (ii) in table 8 we report the estimates for equation (3) and find coefficient β_1 is negative and significant indicating that firms in sectors with a high proportion of multi-product firms are also sectors with more equality[11]. Overall it appears that after controlling for age and sector effects, the number of products a firm is operating over can explain a substantial fraction of why some firms are small and others are big.

[11] We estimate equation (3) in column (i) including mage parametrically and find mage to be statistically insignificant. Therefore, in column (ii) when we estimate the effect of age non-parametrically we unsurprisingly have the same coefficient for $\beta_{1.}$

V. Conclusions

Cabral and Mata (2001) show that the distribution of firm size of a given age cohort is very skewed to the right at the time of birth but gradually moves towards a lognormal distribution. Most of the observed changes in the firm size distribution results from the evolution of the distribution of survivors of a given cohort, and is not due to firm selection processes. We show similar outcomes across using UK and Belgian manufacturing data. Furthermore, we find comparable outcomes in terms of FSD by multi-industry and multi-product cohorts.

Their story of firm growth was based on the idea that financial constraints of companies decline monotonically with age. Our evolutionary story of firm growth is based on Sutton (1998) where companies grow by expanding their product portfolio. We show that the mathematically derived lower bound to firm size distribution is binding across 4-digit nace industries in Belgium and the UK. In addition, we show that the 8-digit product count (and 4-digit count) is empirically a key determinant of firm size, measured in terms of asset size, controlling for the age, among other factors. This finding verifies the key insight of the Sutton (1998) theory and reiterates the credence that modelling competition within industries should not be done with just company level data due to the complexity of competition between companies created by the importance of the product portfolio effect.

References

Audretsch, D.B (1995) *Innovation and Industry Evolution*, Cambridge, MIT Press.

Berry, S., Levinsohn, J. and A. Pakes, (1995) "Automobile Prices in Market Equilibrium" *Econometrica*, 63, pp841-890.

Buzzacchi, L. and T. Valetti, (1999), "Firm Size Distribution: Testing the "Independent Submarkets Model in the Italian Motor Insurance Industry", Discussion Papers Series EI, STICERD, LSE.

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

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

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

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

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

Gollop, F.M. and Monahan, J.L. (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*, 35(June), 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* (series A), 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,* 54:77-89.

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

Klepper, S. and P. Thompson (2005), "Submarkets and the Evolution of Market Structure" Florida International University, *mimeo*

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

Schmalensee, R., (1989), "Inter-industry Studies of Structure and Performance", *Handbook of Industrial Organization.* Vol 2 Eds: Richard Schmalensee and Robert Willig. 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,* 35:40-59.

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

Table1: Distribution of All Secondary Activities Within UK manufacturing for Multi-Industy firms

Secondary Nace

Primary Nace	15	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
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%	1%	0%	12%	1%
18	0%	10%	**65%**	7%	0%	1%	2%	0%	1%	2%	0%	1%	0%	2%	0%	0%	0%	1%	1%	0%	7%	0%
19	0%	4%	14%	**33%**	0%	4%	6%	0%	2%	2%	6%	0%	2%	4%	0%	0%	0%	0%	0%	0%	24%	0%
20	0%	0%	0%	0%	**35%**	3%	2%	0%	0%	7%	5%	1%	9%	1%	1%	2%	0%	0%	1%	1%	31%	1%
21	1%	0%	0%	0%	2%	**21%**	46%	0%	3%	11%	1%	0%	1%	2%	0%	0%	0%	1%	0%	0%	9%	1%
22	0%	0%	0%	0%	1%	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%	13%	0%
24	2%	1%	0%	0%	0%	1%	1%	1%	**61%**	7%	2%	2%	5%	4%	0%	1%	0%	3%	0%	0%	7%	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%	4%	0%	3%	1%	2%	2%	1%	14%	0%
27	0%	0%	0%	0%	0%	0%	2%	0%	1%	1%	0%	**42%**	37%	9%	0%	1%	0%	0%	1%	0%	4%	0%
28	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	2%	**50%**	25%	1%	4%	1%	1%	2%	3%	6%	0%
29	0%	0%	0%	0%	0%	0%	1%	0%	1%	2%	0%	1%	17%	**46%**	2%	11%	3%	6%	2%	4%	5%	0%
30	0%	0%	1%	0%	1%	1%	1%	0%	0%	0%	0%	0%	1%	2%	**3%**	27%	27%	29%	1%	1%	7%	1%
31	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%	0%	11%	6%	3%	**24%**	25%	18%	1%	3%	6%	0%
32	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%	1%	2%	3%	4%	19%	**35%**	22%	0%	2%	9%	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%	15%	6%	0%	2%	1%	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%	1%	**61%**	1%
37	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	3%	0%	3%	0%	0%	0%	0%	0%	0%	0%	**93%**

Table2: Distribution of All Secondary Activities Within Belgian manufacturing for Multi-Industy firms

Secondary Nace

Primary Nace	15	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
15	**95%**	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%
17	0%	**78%**	8%	0%	1%	0%	2%	0%	2%	2%	1%	0%	2%	0%	0%	0%	0%	0%	1%	0%	3%	0%
18	0%	18%	**71%**	3%	0%	0%	1%	0%	0%	1%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	4%	0%
19	0%	4%	17%	**30%**	0%	0%	0%	0%	4%	4%	0%	0%	0%	17%	0%	0%	0%	13%	0%	0%	9%	0%
20	1%	1%	1%	1%	**46%**	2%	0%	0%	1%	5%	2%	0%	16%	3%	0%	0%	0%	0%	0%	0%	20%	2%
21	1%	4%	0%	0%	3%	**65%**	11%	0%	1%	6%	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	4%	0%
22	0%	1%	0%	0%	0%	3%	**93%**	0%	0%	1%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	1%	0%
23	0%	9%	0%	0%	0%	9%	0%	**9%**	18%	18%	0%	0%	0%	9%	0%	9%	0%	0%	0%	0%	0%	18%
24	2%	7%	0%	0%	1%	1%	1%	2%	**59%**	15%	2%	1%	1%	1%	0%	1%	0%	1%	1%	0%	3%	1%
25	0%	2%	1%	1%	4%	4%	1%	0%	8%	**51%**	2%	0%	9%	3%	0%	2%	0%	0%	3%	0%	5%	2%
26	0%	0%	0%	0%	2%	1%	1%	0%	3%	4%	**78%**	0%	5%	2%	0%	0%	0%	0%	1%	0%	2%	2%
27	0%	0%	0%	0%	1%	0%	1%	0%	1%	2%	1%	**26%**	52%	11%	0%	2%	0%	0%	2%	1%	1%	1%
28	0%	0%	0%	0%	1%	0%	0%	0%	0%	1%	0%	5%	**66%**	19%	0%	1%	0%	0%	1%	0%	2%	0%
29	0%	1%	0%	0%	1%	0%	0%	0%	1%	1%	0%	2%	43%	**34%**	1%	7%	1%	3%	3%	0%	1%	0%
30	0%	0%	0%	0%	0%	0%	5%	0%	0%	2%	0%	0%	7%	16%	**12%**	7%	19%	28%	0%	2%	2%	0%
31	0%	1%	0%	0%	0%	0%	1%	0%	0%	1%	1%	0%	11%	11%	1%	**42%**	15%	6%	3%	0%	6%	0%
32	0%	1%	0%	0%	1%	0%	4%	0%	1%	1%	0%	0%	4%	6%	7%	22%	**32%**	20%	0%	1%	1%	0%
33	1%	0%	3%	10%	1%	0%	1%	0%	4%	0%	1%	0%	5%	5%	1%	5%	7%	**56%**	0%	2%	1%	0%
34	0%	5%	0%	0%	2%	0%	1%	0%	1%	3%	1%	1%	16%	10%	0%	2%	0%	1%	**56%**	1%	3%	0%
35	2%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%	29%	19%	0%	3%	0%	2%	5%	**28%**	7%	2%
36	0%	3%	2%	0%	9%	1%	2%	0%	1%	1%	2%	1%	7%	1%	0%	1%	1%	0%	1%	0%	**67%**	0%
37	10%	10%	0%	0%	5%	2%	2%	2%	3%	5%	13%	6%	5%	5%	0%	0%	0%	2%	2%	0%	0%	**31%**

Table 3. Measure of Skewness by Age Cohorts

	(i)	(iii)
	UK	Belgium
Full sample	0.16	.29
Age <5	0.29	.56
Age>=5&age<10	0.22	.31
Age =10&age<25	0.08	.26
Age>=25	0.15	.10

Table 4. Measure of Skewness for Single and Multi-Industry Firms in the UK

	(i)	(iii)	(v)
	Single industry Firms UK	Two industry Firms UK	Three or more industry Firms UK
Full sample	0.11	0.15	0.21
Age <5	0.27	0.26	0.49
Age>=5&age<10	0.16	0.28	0.41
Age>=10&age<25	0.03	0.11	0.13
Age>=25	0.10	0.11	0.07

Table 5. Measure of Skewness for Single and Multi-Industry Firms in Belgium

	(i)	(iii)	(v)
	Single industry firms Belgium	Two industry Firms Belgium	Three of more industry firms Belgium
Full sample	0.44	0.24	0.17
Age <5	0.60	0.51	0.52
Age>=5&age<10	0.35	0.26	0.36
Age =10&age<25	0.44	0.22	0.15
Age>=25	0.27	0.06	0.003

Table 6. Measure of Skewness for Single and Multi-Product Firms in the UK

Full sample	Single Product	Two Product	Three Products	Four or more Product
0.02	0.05	-0.15	0.16	0.04

Table 7:Firm Level Size Regression Dependent variable: ln(market share)

	(i) UK	(ii) UK	(iii) Belgium	(iv) Belgium	(v) UK 8-digit product	(vi) UK 8-digit product
Ln(count)	0.41 (0.02)	0.18 (0.01)	0.68 (0.03)	0.41 (0.02)	0.27 (0.04)	0.32 (0.04)
Ln(age)	-	-	-	-	-	-
Constant	- 9.7 (0.009)	- 6.68 (0.20)	-7.38 (0.02)	-6.80 (0.12)	-5.34 (0.03)	-4.98 (0.17)
Sector dummies	No	Yes	No	Yes	No	Yes
R2	0.094	0.43	0.08	0.37	0.02	0.18
Number of Observations	78911	78911	21696	21696	5052	5052

Table 8:Sector Level Size Regression Dependent variable: Gini Coefficient

	(i)	(ii)
pmpf	-3.19** (1.27)	-3.19** (1.27)
mage	0.02 (0.02)	-
constant	1.89** (0.87)	2.48** (0.85)
R2	0.063	0.043

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.

Figure 1: Top 4,8 and 20 UK Companies Concentration Ratios in each 4 digit industries

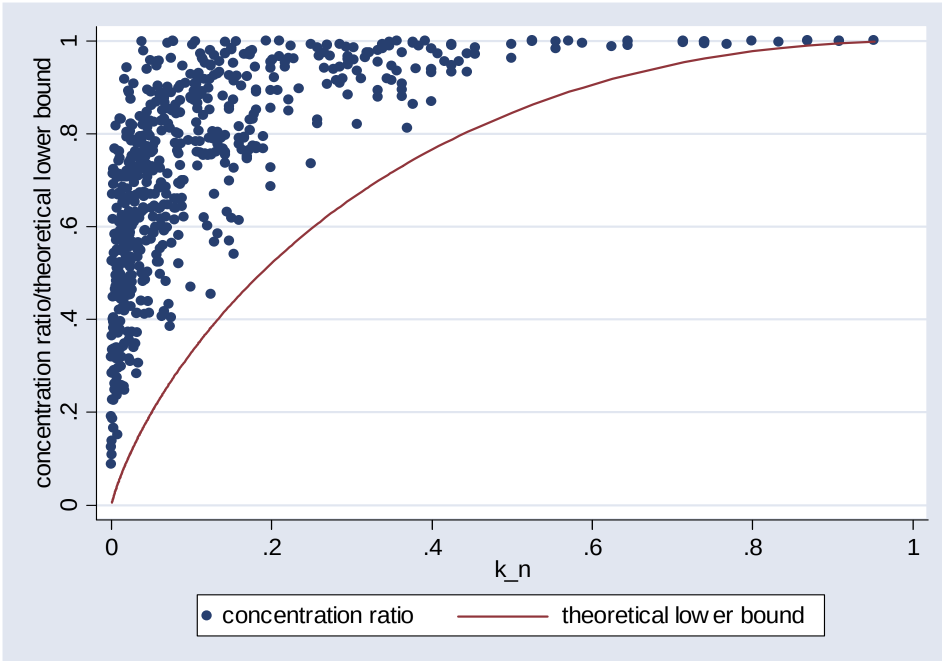

Figure 2: Top 4, 8 and 20 Belgian Companies Concentration Ratios in each 4 digit industries

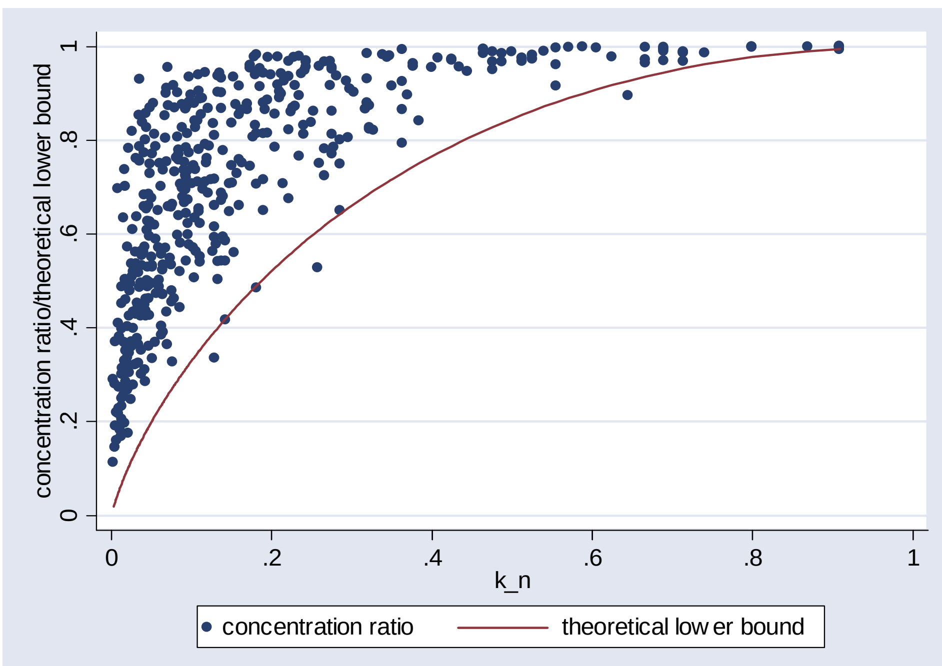

Figure 3:
Age Distribution of all UK firms

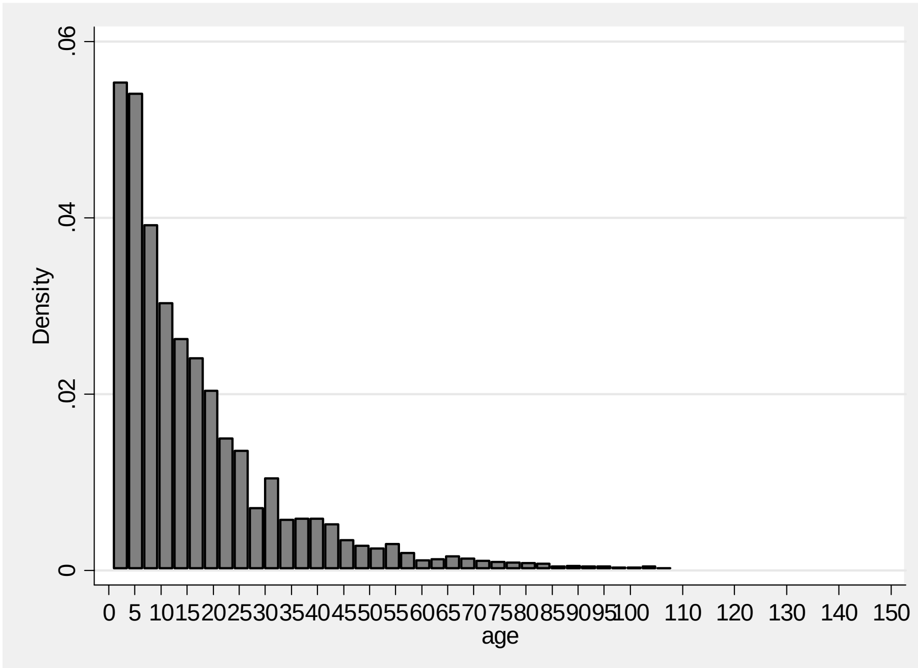

Figure 4:
Age distribution of all Belgian Firms

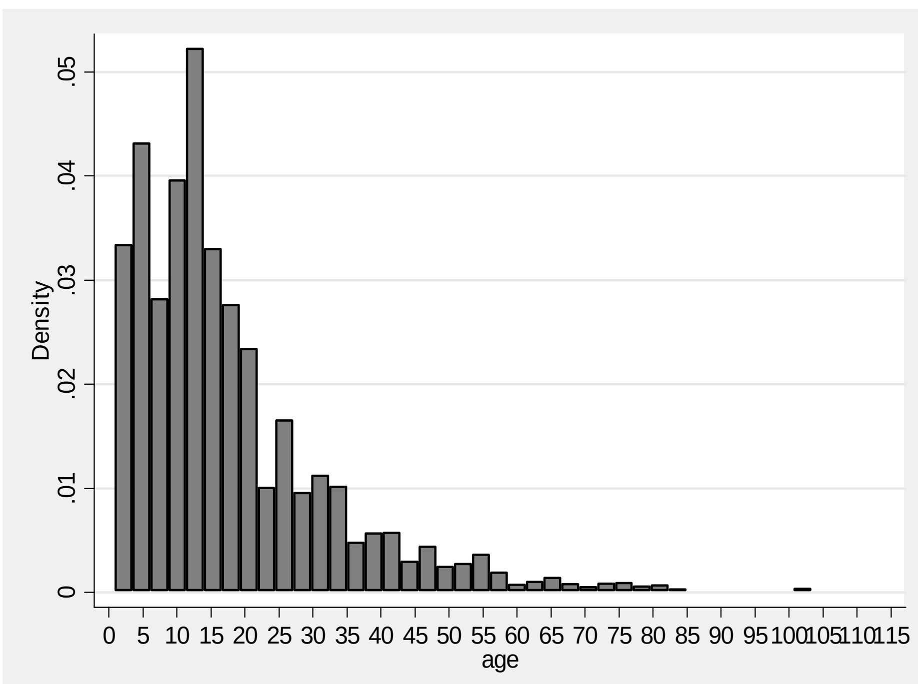

Figure 5:
Size Distribution of UK Firms by age Cohorts

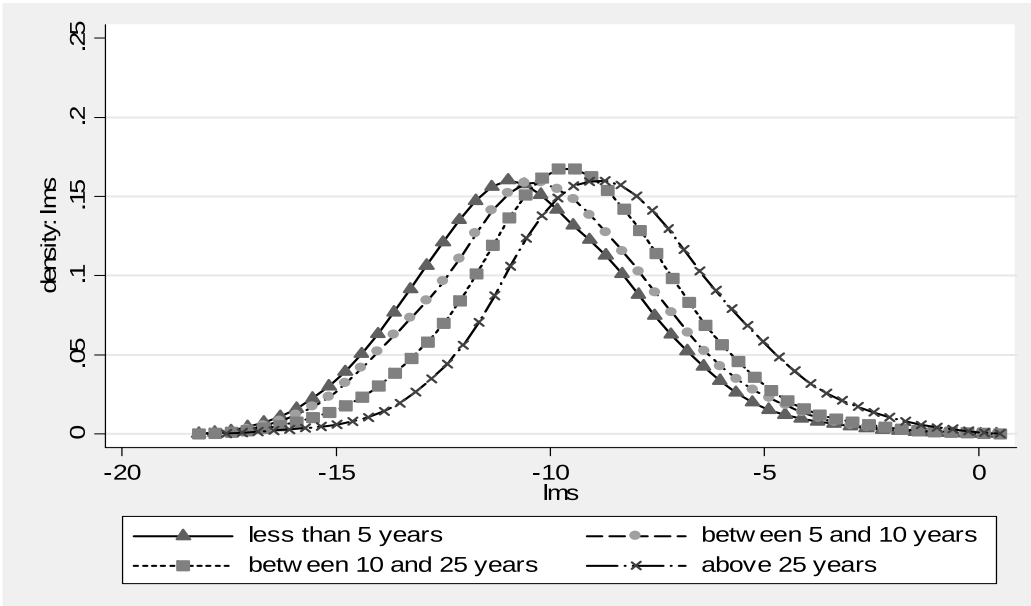

Figure 6:
Size Distribution of Belgian firms by Age Cohorts

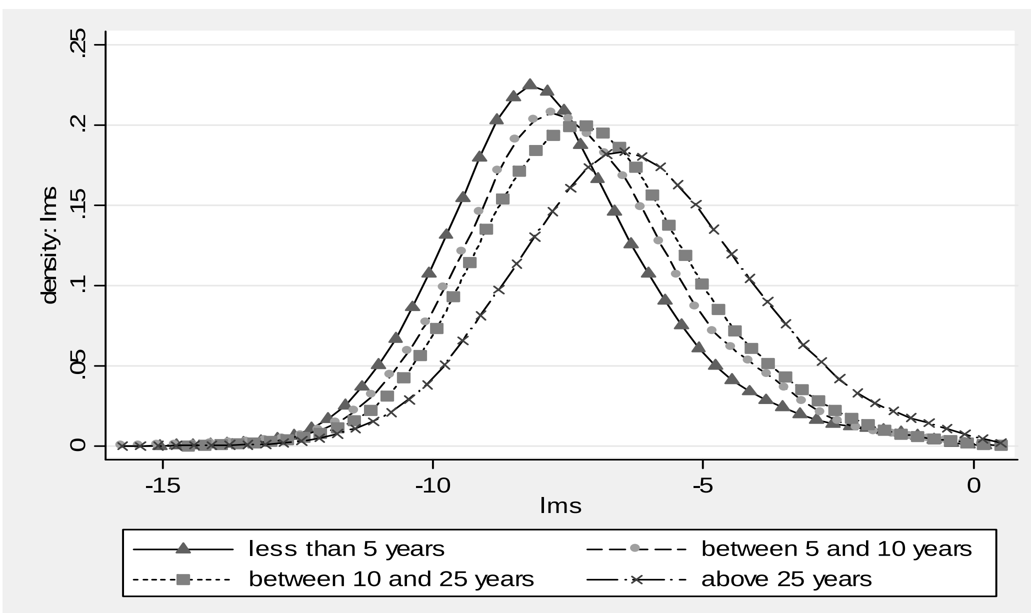

Figure 7:
Size distribution of UK companies by single and multi industry cohorts

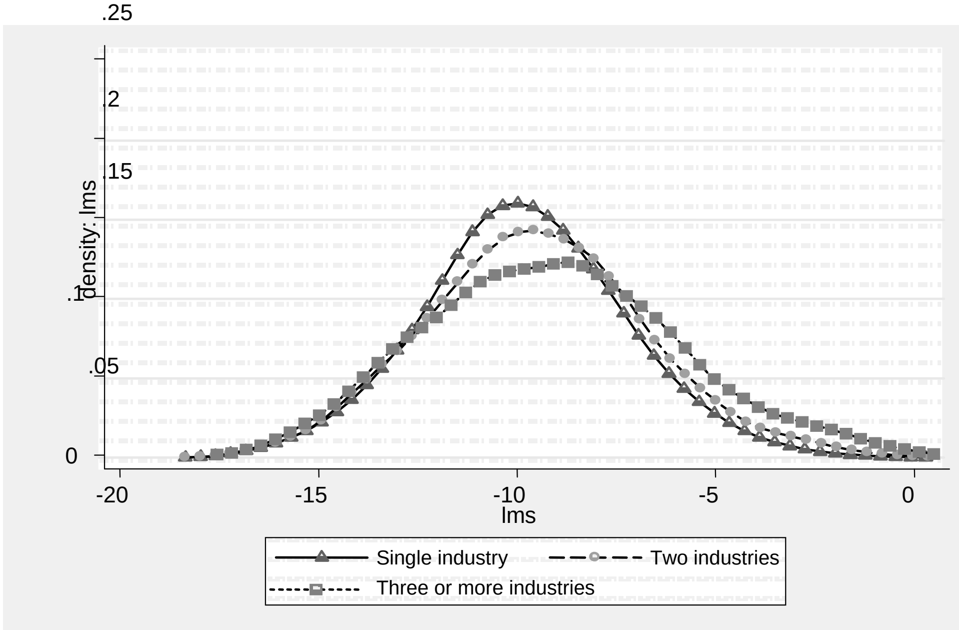

Figure 8:
Size distribution of UK companies by single and multi industry cohorts
age>=10&age<25

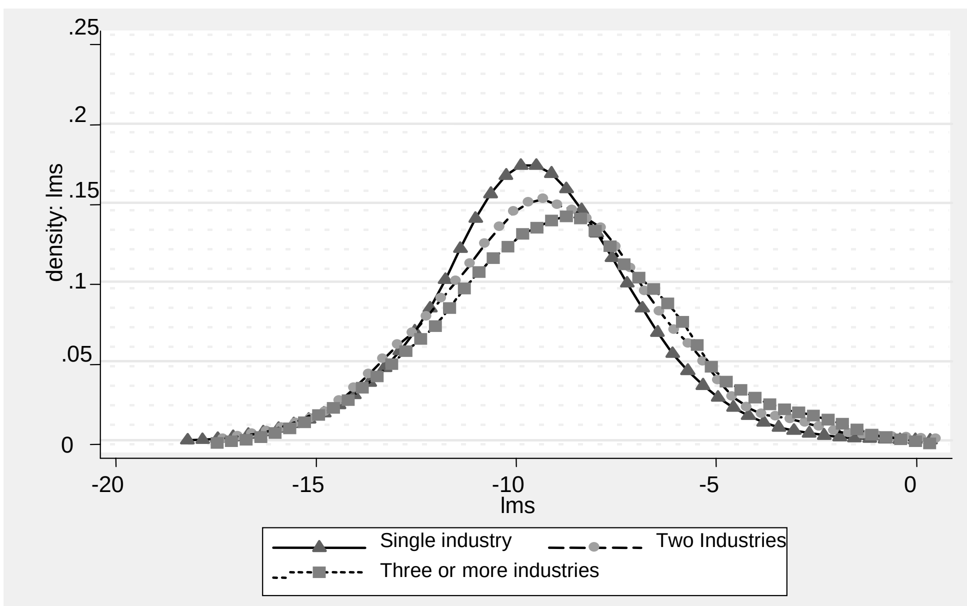

Figure 9:
Size distribution of UK companies by single and multi industry cohorts age>=25

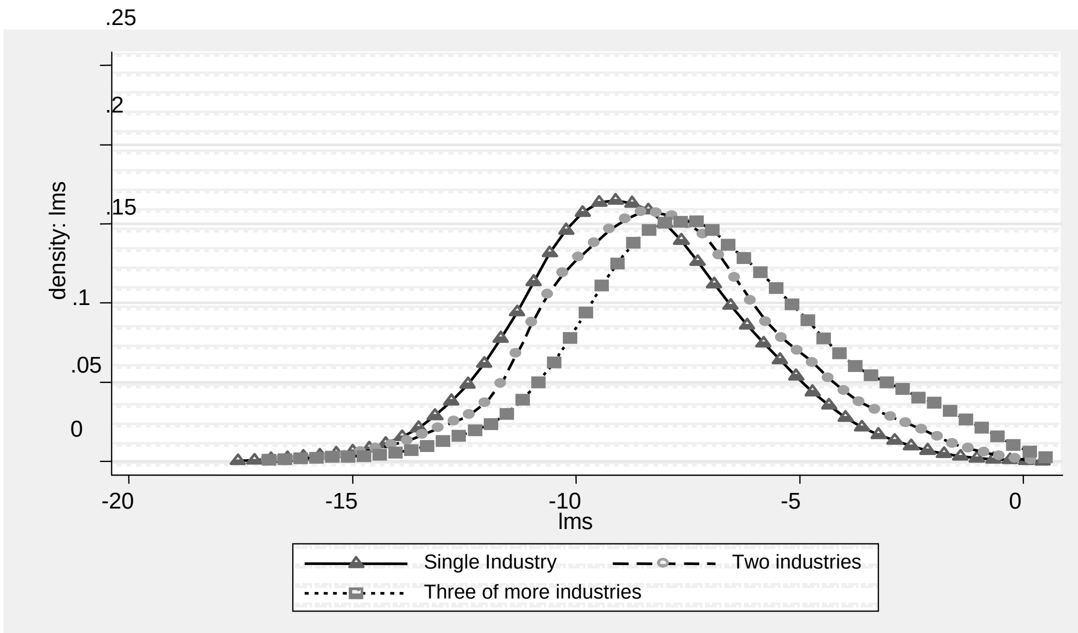

Figure 10:
Size distribution of Belgian companies by single and multi industry cohorts

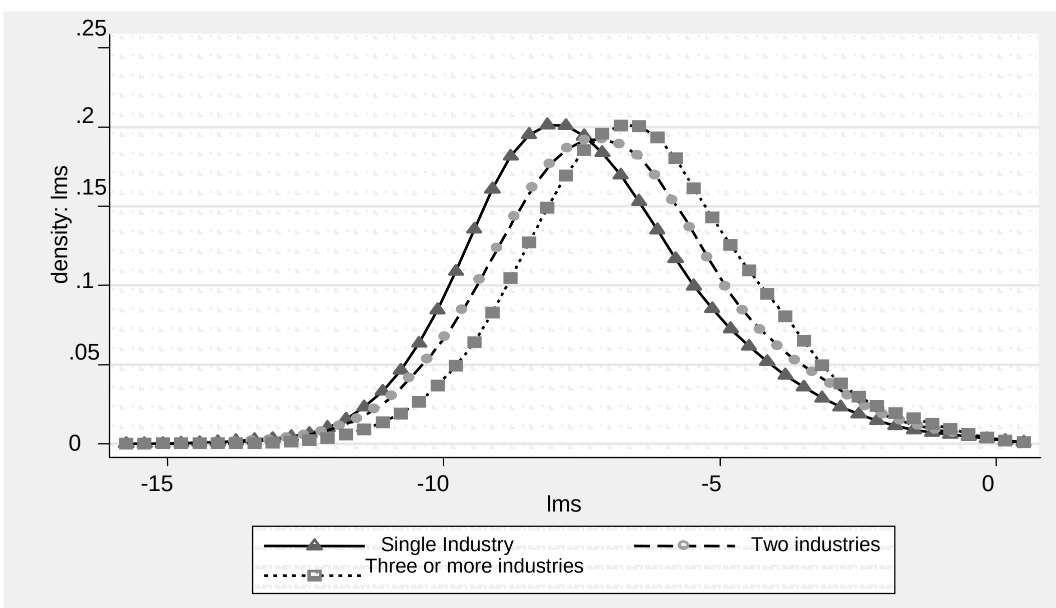

Figure11:
Size distribution of Belgian companies by single and multi industry cohorts
age>=10&age<25

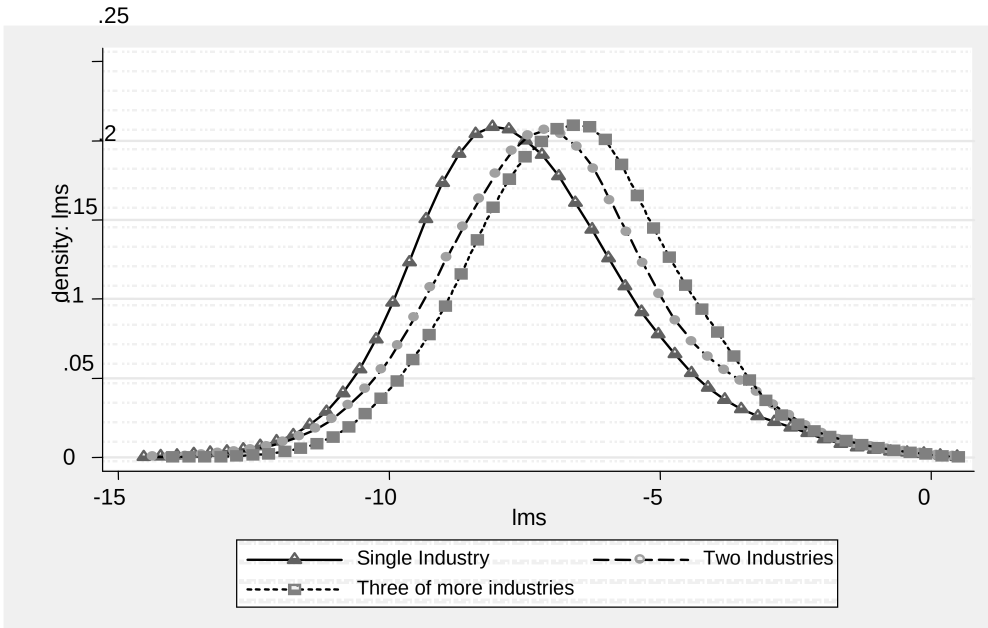

Figure12:
Size distribution of Belgian companies by single and multi industry cohorts
age>=25

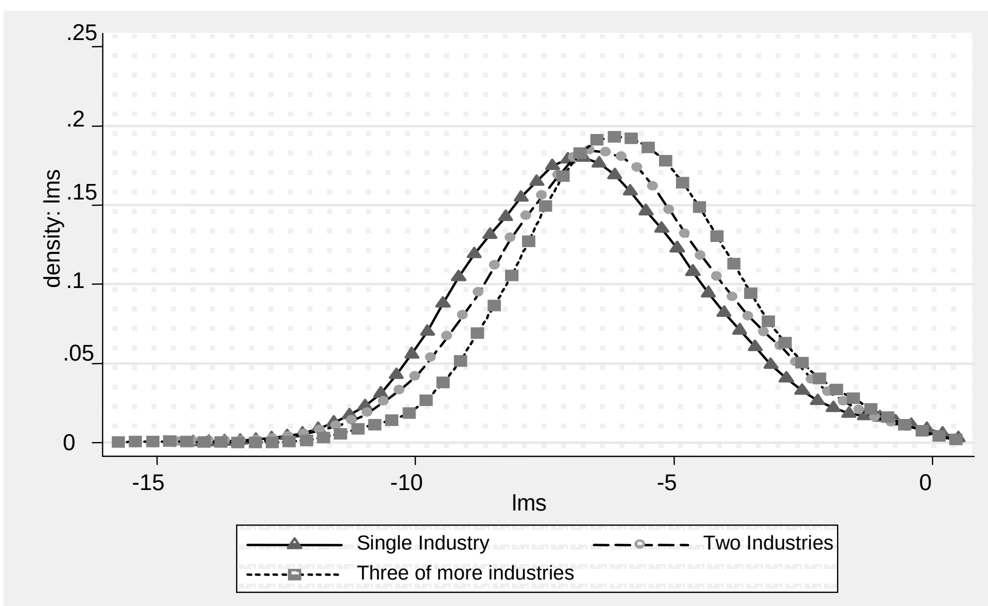

Figure 13:

Graph for UK companies of lnsize, lnAge and lnindustrycount

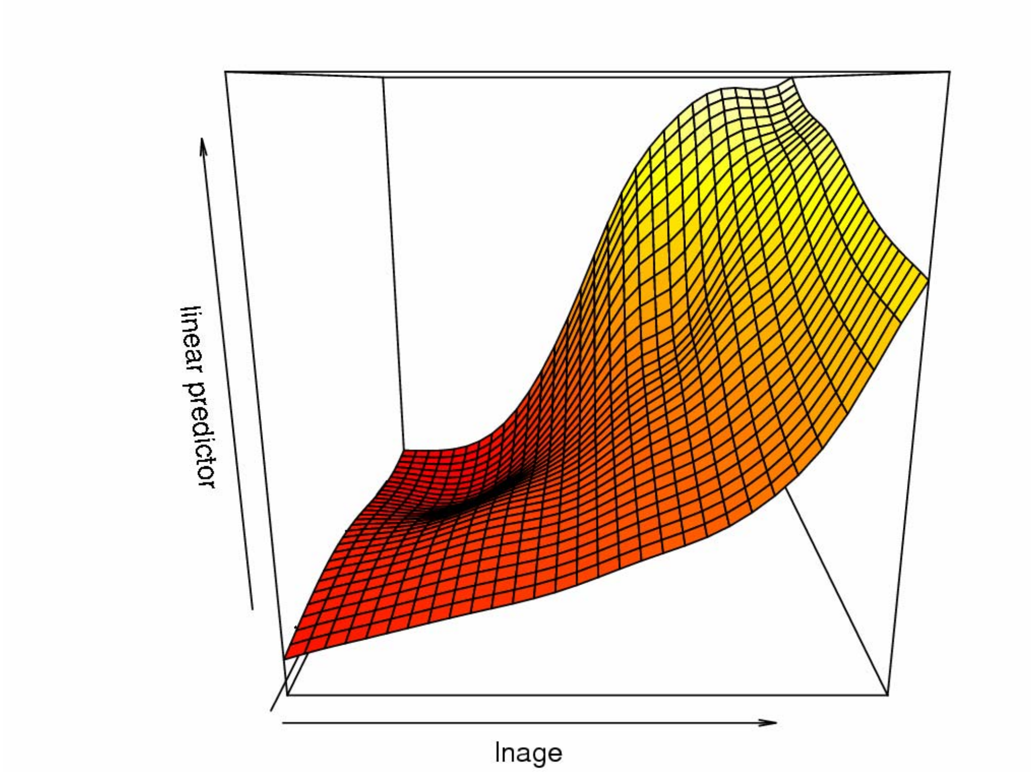

Figure 14:
Graph for Belgian companies of lnsize, lnAge and lnindustrycount

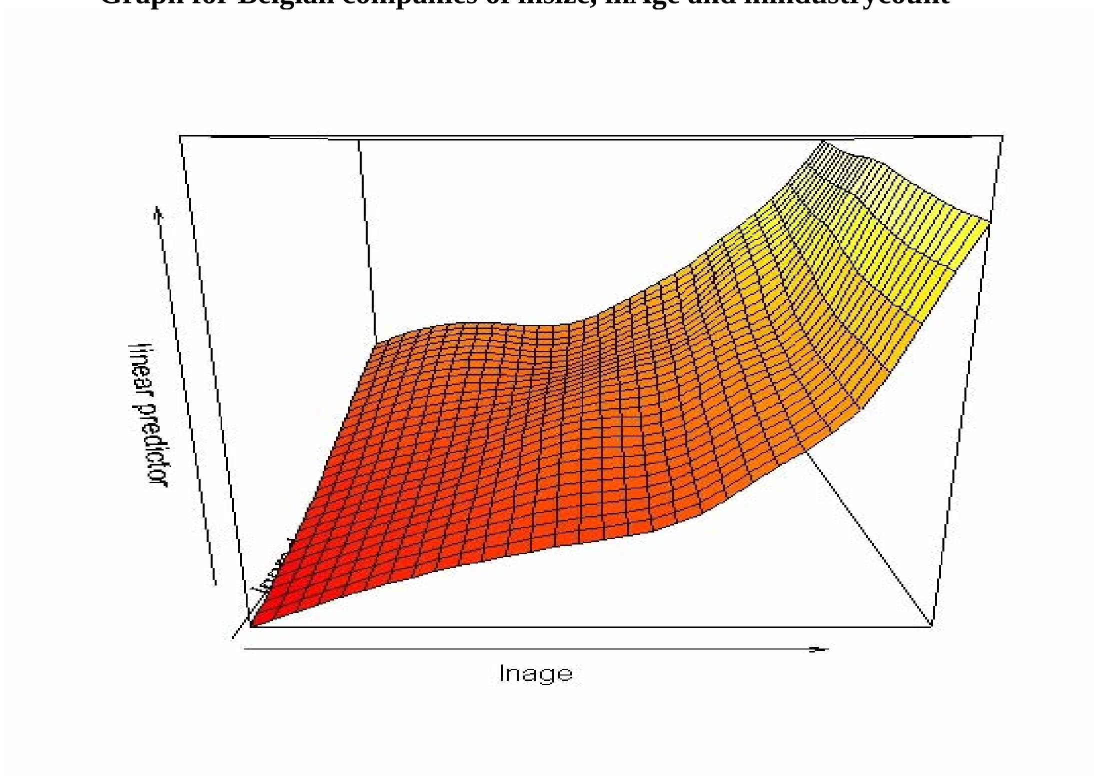

Figure 15:
Weighted graph for UK companies of ln(industry count) and ln(age)

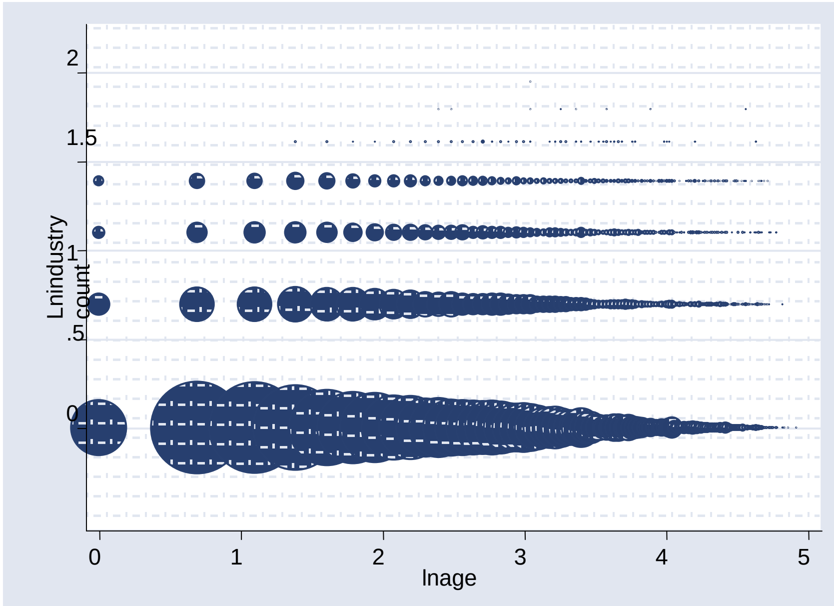

Figure 16:
Weighted graph for Belgian companies of ln(industry count) and ln(age)

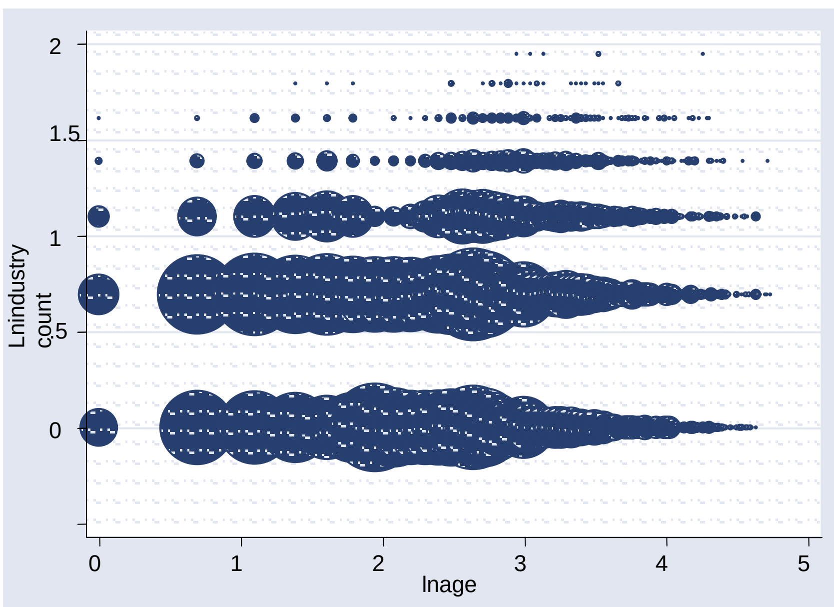

Figure 17:
Size Distribution of UK Firms by Single and Multi-product Cohorts

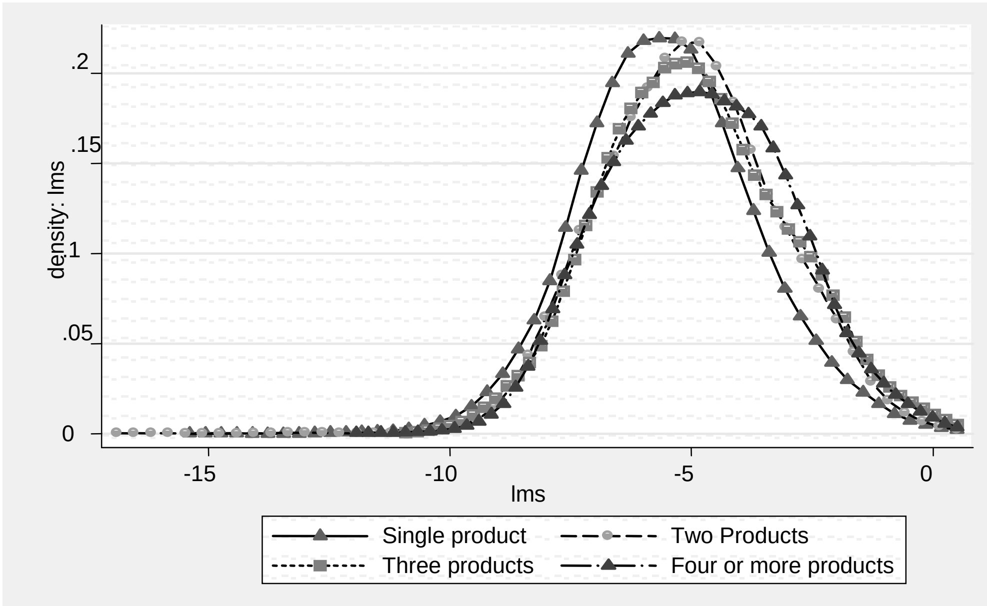

Figure 18:
Product Count Distribution of UK firms