

The Firm Size Distribution and Inter-Industry Diversification

John Hutchinson · Jozef Konings ·
Patrick Paul Walsh

Published online: 8 October 2010
© Springer Science+Business Media, LLC. 2010

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. This paper investigates whether the time to build a portfolio of products from inter-industry diversification is a mechanism that helps explain the stylized facts of the FSD by age cohorts. We find that the degree of inter-industry diversification has a separate impact on firm size when controlling for age, amongst other factors. Using the techniques of Cabral and Mata, we find that inter-industry diversification does shift the FSD to the right, and more so in older age groups.

Keywords Firm age · Firm size distribution · Inter-industry diversification

JEL Classification L10 · L11 · L16

J. Hutchinson
External Developments Division, European Central Bank, Kaiserstrasse 29,
60311 Frankfurt am Main, Germany
e-mail: John.Hutchinson@ecb.int

J. Konings
Vives, Centre for Regional Economic Policy, K. U. Leuven, Leuven, Belgium
e-mail: Joep.Konings@econ.kuleuven.be
URL: www.econ.kuleuven.be/joep.konings

P. P. Walsh (✉)
SPIRe and Geary Institute, University College Dublin, Newman Building,
Belfield Dublin 4, Ireland
e-mail: ppwalsh@ucd.ie

1 Introduction

Investigating firm size distributions (FSD) has received considerable attention since [Gibrat \(1931\)](#) “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

Industrial Organization (IO) papers by [Klepper and Thompson \(2006\)](#) and [Sutton \(1998\)](#), using theory and evidence, explicitly link “time to build” a product portfolio to the limiting FSD. The [Klepper and Thompson \(2006\)](#) model is a dynamic model that applies to time series data for a single industry: the US laser industry. They build finite duration of product cycles into their theory. The opening and closing of product lines gives them interesting short-run dynamic relationships between growth, survival, and age. Product counts end up being an important driver of industry dynamics and 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. [Sutton \(1998\)](#) explores the possibility 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 a lower bound. Firm size in this model measures the number of equally sized sub-markets in which a firm is present.³ [Sutton \(1998\)](#) outlines theorems where

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 [Ijiri and Simon \(1964, 1977\)](#) included 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.

² Another literature using rich firm-level data, starting with [Dunne et al. \(1989\)](#), focuses on modelling firm growth and survival as a deterministic outcome of firm characteristics, including age and size. In general, these empirical studies conclude that Gibrat’s law tends to fail due to turnover in small firms. 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). The emphasis of [Cabral and Mata \(2003\)](#) is on the evolution of the distribution of survivors of a given cohort and not on noisy selection. [Luttmer \(2007\)](#) model has experimentation and selection allowing permanent shocks to preferences and technology that does affect the steady state FSD. This approach assumes that an endogenous idiosyncratic technology is linked to each single-product firm. This model could be extended to a multi-product setting, which would link product portfolios to age within the FSD. [Cabral and Mata \(2003\)](#) propose that as firms get older they encounter gradually fewer financial constraints. [Angelini and Generale \(2008\)](#), while confirming the stylized facts of FSD as proposed by Cabral and Mata, conclude that financial constraints were not the main determinant in the evolution of the size distribution of firms.

³ In the limit where the number of opportunities becomes large, older firms will have a larger product portfolio than will younger firms. 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:

certain factors will cause the FSD to reside above the lower bound. One involves firms' achieving economies of scope from product diversification, which leads to a more skewed FSD. Product diversification would have an independent role on the FSD over and above age.

Beyond the FSD studies a positive relationship between firm size and inter-industry (4-digit) product diversification has been documented for manufacturing. [Gollop and Monahan \(1991\)](#) identify this as the most important structural phenomena of the post-war period. [Bernard et al. \(2006\)](#) 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 of scope at the industry level with strong economies of scope across industries. [Nocke and Yeaple \(2008\)](#) go a step further and theoretically link 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 the trade literature predicts a positive relationship between inter-industry diversification and size, there is no explicit role for company age.

Do firms become large because they accumulate a presence in a greater number of 4-digit industries over time? Can this theory of size explain the stylized facts of the FSD observed by age cohorts, as proposed by [Cabral and Mata \(2003\)](#)? In other words; is age really just a proxy for the number of 4-digit industries that a firm accumulates over time. The [Sutton \(1998\)](#) model would predict a perfect correlation between age and product counts on his lower bound. Yet, heterogeneity in sub-markets (size and price competition) or in firms (underlying costs and productivity) would lead to partial effects from product diversification on the FSD that are over and above that created by age cohort effects.

Using a series of non-parametric density plots we show an independent role for product counts on the FSD that is more pronounced as age cohorts get older. We outline the implications of this result for research in IO and Trade in our concluding section.

2 Section I

Our data are taken from the reported unconsolidated company accounts of manufacturing firms in the UK and in 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 been officially audited. The data do not cover single-proprietorship companies with zero employees.

Footnote 3 continued

e.g., the US cement industry ([Sutton 1998](#)), the Spanish retail banking sector ([Juan 2003](#)), and the Italian motor insurance industry ([Buzzacchi and Valletti 2006](#)). Due to the richness of data, all of these studies indicate how firm size (market share) is positively related to portfolios of geographical locations. [Walsh and Whelan \(2002\)](#) address this issue in a multiproduct industry: carbonated soft drinks, where configurations are in terms of products and stores (geography).

Firm Size

In accordance with standard accounting protocols, companies are not required to report all financial and operational items. However, full information is available on total assets of firms (intangible, tangible, financial, and current assets). For this reason, we will compute the market share of a firm, f , within its primary 4-digit industry, j , based on the firm's total assets and the industry's total assets; this is denoted as MS_{fj} (natural log of market share).⁴ We gathered full information on firms' 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.

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 over which a firm operates. We realize that the sum of the 4-digit NACE industries in which a firm is active constitutes a lower bound to the product count of a firm. In principle, 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 striking and has important implications for research areas, which we outline in the conclusion.

In Tables 1 and 2 we present co-production matrices for multi-industry firms in the UK and Belgium for 2002 at the 2-digit industry level.⁵ Each row in the matrices shows the percentage of multi-industry firms that operate in a 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 Tables 1 and 2 we see that in both countries, for most industries the diagonal contains the highest percentage, with co-production occurring within rather than across 2-digit categories.⁶ This is broadly in line with the findings of Bernard et al. (2010) where they conduct a similar exercise for the United States.

⁴ Although Sutton (1998) uses sales to measure firm size, he argues that total assets would be the preferred measure. Consequently, this is the measure that is employed in this paper. Mariuzzo et al. (2003) explore the link between firm size and estimates of market power in carbonated soft drinks. They show that inequality in market share does not reflect inequality in market power. Small companies can extract rents in niches, just as can large companies that operate across niches.

⁵ 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 in which firms operate, we restrict ourselves to focusing on firms' secondary activities that are in the manufacturing sector (NACE: 15–37). We also choose to look at the industry level (2-digit NACE). Making matrices using 4-digit level information would be impractical.

⁶ Several interesting observations emerge from these tables. For NACE 23 (manufacture of coke, refined petroleum products and nuclear fuel), 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

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

Primary Nace	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	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	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	0	1	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
18	0	19	72	5	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0					
19	0	5	20	35	0	0	0	0	5	5	0	0	0	0	10	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0					
20	1	0	1	1	49	1	0	0	1	4	2	0	15	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0					
21	1	5	0	0	3	71	9	0	1	4	0	0	2	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
22	0	1	0	0	0	3	94	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
23	0	5	0	0	0	5	0	5	53	11	0	0	0	0	5	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
24	3	7	0	0	1	1	1	2	65	11	3	0	1	2	0	0	0	0	1	1	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
25	0	3	2	2	6	5	2	0	12	43	2	0	11	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
26	0	0	0	0	2	1	1	0	2	4	81	0	13	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
27	0	0	0	0	1	0	1	0	1	3	1	25	55	8	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
28	0	0	0	0	2	0	0	0	0	2	0	7	73	11	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
29	0	1	0	0	1	0	0	0	1	1	0	1	23	65	0	4	1	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	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	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	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	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	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	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	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	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	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	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	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	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
36	0	4	2	0	10	1	2	0	2	2	2	1	7	1	0	1	1	1	1	1	0	65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
37	9	11	0	0	5	2	2	2	4	5	13	5	5	2	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

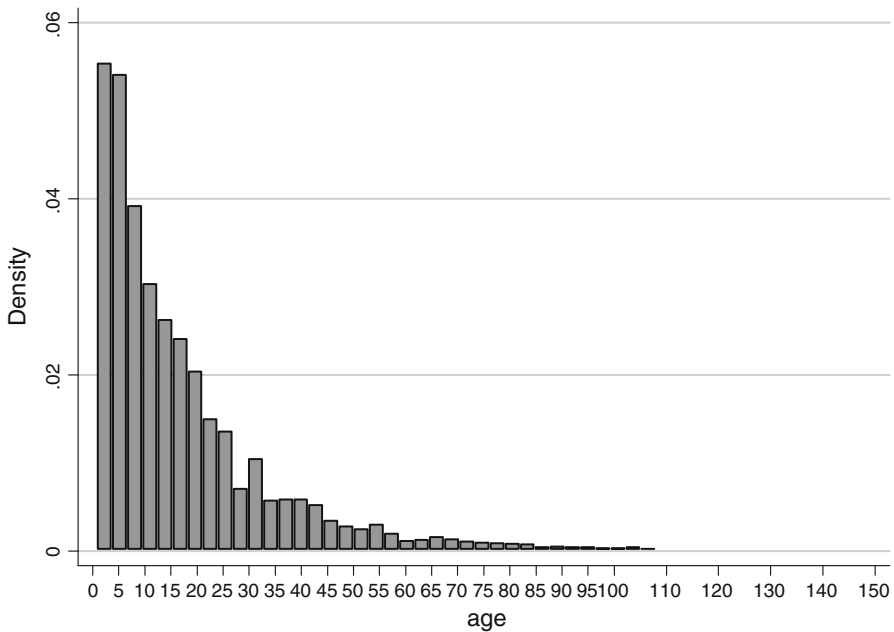


Fig. 1 Age distribution of all UK firms

Age

We measure the age of a company based on the year of incorporation in our analysis.⁷ In Fig. 1 for the UK and Fig. 2 for Belgium we graph the age distribution of firms. We note that in both countries there are many young firms, while only a small fraction of firms are older than 20 years.⁸

FSD by Age Cohorts

We follow the analysis in Cabral and Mata (2003) and plot the kernel density estimates of the FSD by age cohorts in Figs. 3 and 4, for the UK and Belgium, respectively, where size is measured as the log of the share of a firm's total assets (lms) relative to its primary 4-digit industry.

Footnote 6 continued

primary activity is within NACE 30 (manufacture of office machinery and computers), 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)

⁷ The dataset does not allow for identifying whether a firm was formed as a result of a merger between two firms.

⁸ Over time, the tax laws in Belgium have made the opening and closing of companies worthwhile, which may have led to younger company profiles.

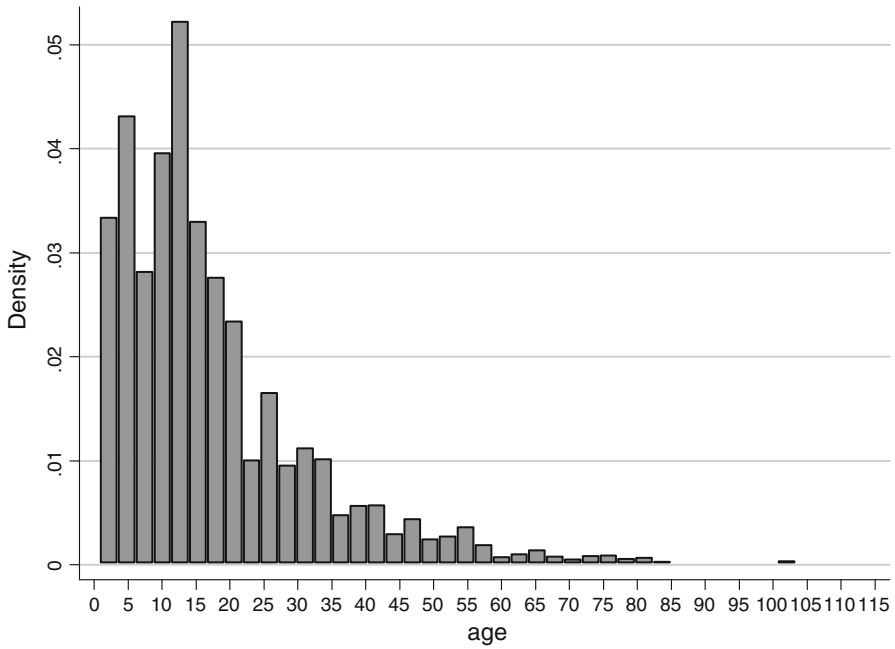


Fig. 2 Age distribution of all Belgian firms

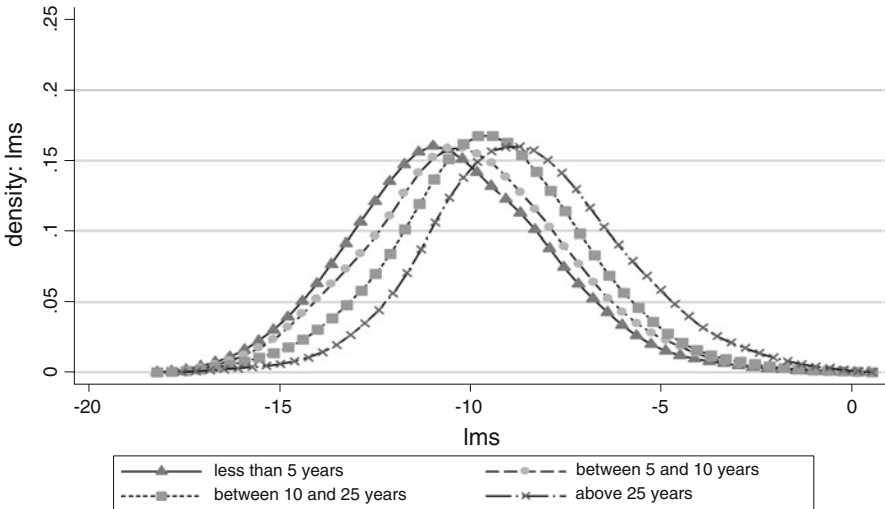


Fig. 3 Size distribution of UK firms by age cohorts

We see from Figs. 3 and 4 that for both the UK and Belgium, as firms grow older, the size distribution shifts more to the right. This finding highlights the point that as time advances some economic force is pushing the distribution of firm size within 4-digit

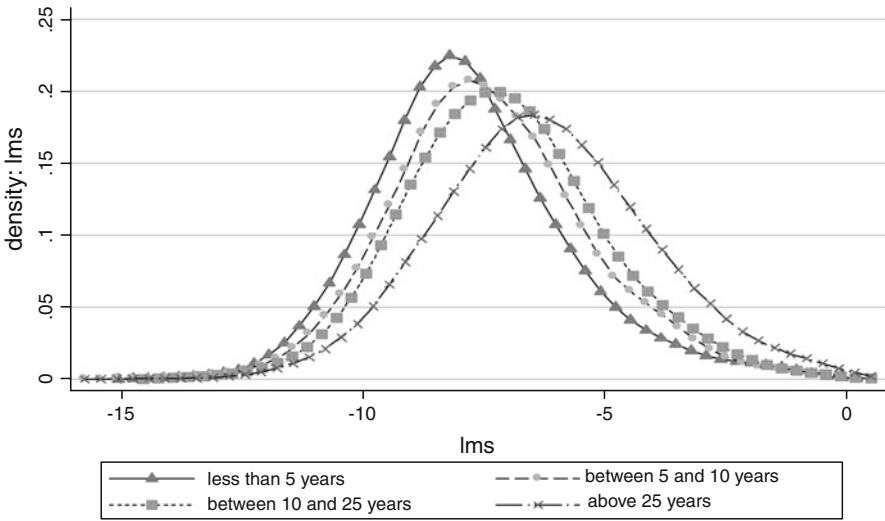


Fig. 4 Size distribution of Belgian firms by age cohorts

industries towards log normality in surviving firms. We 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 feature of the theory is the arrival, in the history of a defined market, of a number of discrete investment opportunities in sub-markets over an infinite period for firm populations. 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 sub-markets. We refer to Sutton (1998) for the derivation of the lower bound, which yields the proposition that, for any fixed ratio k/N (e.g., the top 20 percent of companies in the population), where N is the total number of firms in an industry, an asymptotic lower bound to the k -firm concentration ratio ($C_{k/N}$) exists such that:

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

The basic idea behind the theoretical lower bound to FSD is that opportunities for firms present themselves gradually over time and hence firms need time to aggregate product portfolios. Age and size (product counts) become linked as firms spread themselves across product opportunities over time. If in the history of Sutton’s games we allow heterogeneity in sub-markets (size), or in firms (scope economies), this will

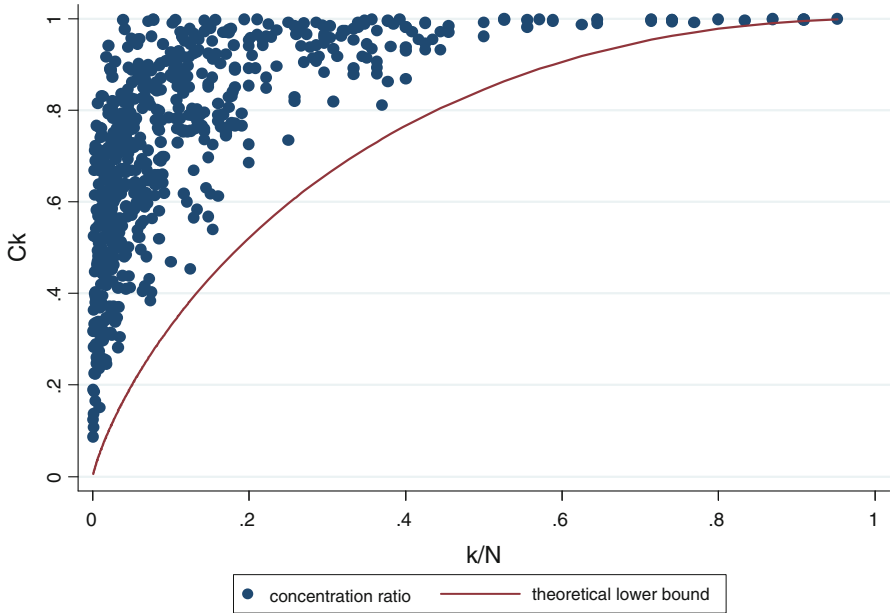


Fig. 5 Top 4, 8, and 20 UK companies’ concentration ratios in each 4-digit industry

Table 3 Correlation coefficients

	UK			Belgium		
	Market share	Age	Industry count	Market share	Age	Industry count
Market share	1.0			1.0		
Age	0.1	1.0		0.14	1.0	
Industry count	0.09	0.02	1.0	0.03	0.09	1.0

create more inequality and a disconnection between product counts (size) and age, pushing FSD inside the lower bound.

In Fig. 5 for the UK and Fig. 6 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 note that for both countries the mathematically predicted lower bound holds for almost all of the 4-digit industries. Yet, we see that many observations are well above the bound, suggesting that there may be factors other than age that drive the FSD. In Table 3 we report simple correlations between firm-level market share, age and industry (product) counts. Age and product counts are positively correlated but not especially strong. The many observations inside the lower bound suggest that product counts may have an additional role that is separate from age.

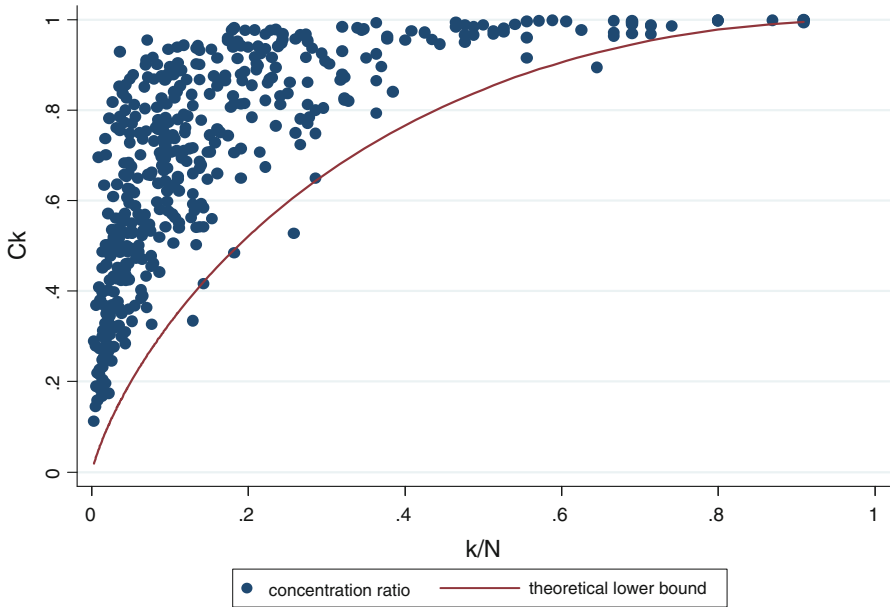


Fig. 6 Top 4, 8, and 20 Belgian companies’ concentration ratios in each 4-digit industry

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 do not allow him to discriminate between situations where the firm is big because of dominance in one big sub-market versus a presence over multiple sub-markets. In the next section, we link company size (market share of assets) to product counts to show that the aggregation assumed in his theory is empirically verified. In addition, while age and product counts can proxy for each other on the bound, we will explore whether they have separate effects due the many observations inside the bound.

3 Section II

In this section we provide evidence that *inter-industry* 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 test whether the null hypothesis—that that counts of 4-digit industries have no independent effect on firm size, in terms of the log market share of assets—can be rejected in a simple econometric model. The basic model of the size of firm f in each 4-digit industry j is as follows:

$$MS_{fj} = \alpha + \beta_1 \ln \text{Counts}_f + \beta_2 \ln \text{Age}_f + \text{Sector}_j + \varepsilon_{fj}. \tag{2}$$

Table 4 Selected summary statistics

(Mean values)	(i)	(iii)
	UK	Belgium
Total assets (value in Euros)	9878	7075
Age	17.4	18
Product counts	1.3	1.87
Number of firms	78,911	21,697

Table 5 Firm-level size regression; dependent variable is $\ln(\text{market share of a firm's total assets relative to the firm's primary 4-digit industry})$

	(i)	(ii)	(iii)	(iv)
	UK	UK	Belgium	Belgium
$\ln(\text{Counts})$	0.43*** (0.02)	0.20*** (0.02)	0.68*** (0.03)	0.42*** (0.03)
$\ln(\text{age})$	0.73*** (0.01)	0.81*** (0.01)	0.55*** (0.02)	0.54*** (0.01)
Constant	-11.42*** (0.02)	-8.61*** (0.20)	-8.80*** (0.05)	-8.19*** (0.13)
4-digit dummy variables	No	Yes	No	Yes
R2	0.09	0.43	0.07	0.37
Observations	78,911	78,911	21,696	21,696

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

MS_{fj} is the log size of a firm in terms of the share of its overall size (the total assets of the firm) relative to its primary 4-digit industry; Counts_f is firm counts of the number of 4-digit industries in which a company operates; Age_f is the number of years since the year of the firm's incorporation; and 4-digit industry dummy variables control for unobserved industry effects.⁹ We assume that operating across industries is something that takes a long time to accomplish. 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 at the mean, rather than causalities at the mean. Summary statistics are provided for the variables in Table 4.

We now turn to testing our hypothesis in Table 5, where we estimate Eq. 2. In columns (i) and (iii) in Table 5, we report the estimates for the UK and Belgium with no 4-digit industry dummy variables. In columns (ii) and (iv) we include a full set of 4-digit industry dummy variables.¹⁰ Interestingly, whether we include industry dummy variables or not, 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 in which a firm operates is a feature of why some firms are small and

⁹ The industry dummy variables are constructed to reflect the primary NACE code for each firm.

¹⁰ The coefficients on the majority of industry dummies are statistically significant.

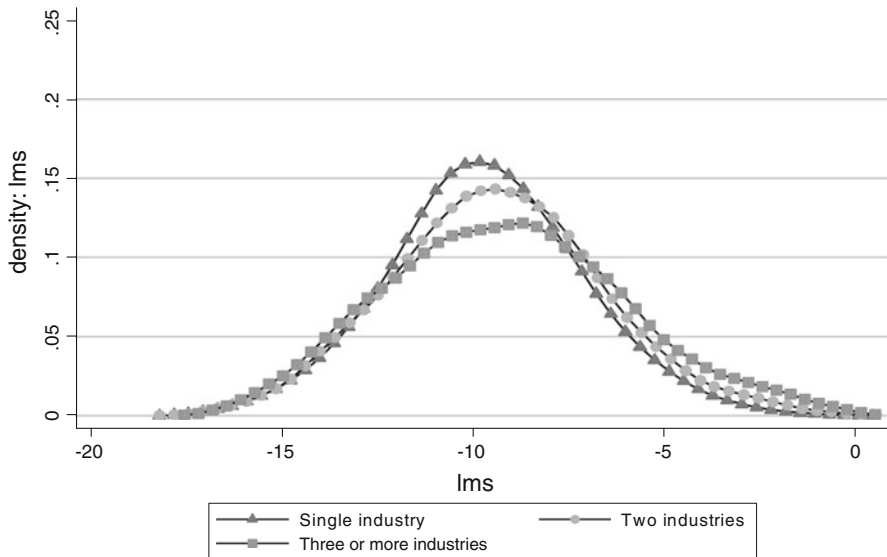


Fig. 7 Size distribution of UK companies by single and multi industry cohorts

others are large.¹¹ Our main empirical evidence will be based on the non-parametric density plots that are used in [Cabral and Mata \(2003\)](#).

FSD by Using Firm Cohorts Defined by Industry Counts

In Figs. 7 and 8 we graph the non-parametric density plots of the FSD of single- and multi-industry cohorts (4-digit NACE) in our two countries. To the degree that age and the collection of products are positively correlated, as theory would suggest ([Sutton 1998](#); [Klepper and Thompson 2006](#)), we might expect similar patterns in the FSD to that produced by age cohorts. Our analysis of the [Sutton \(1998\)](#) lower bound and simple regression suggests that there is more to the evolution of inter-industry product diversification than simply an age effect. In Fig. 8 we see that Belgian firms operating in more than one industry have size distributions farther to the right. In Fig. 7 we see such shifts of FSD with industry diversification are less pronounced in the UK data.

In line with our regressions it is interesting to see whether we get an effect on the FSD from product counts beyond those generated by age cohorts. In Figs. 9, 10, 11, 12 for the UK and Belgium we see that, once we condition on age cohorts, multi-industry firms that are ten years or older have size distributions that are much farther 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

¹¹ We did estimate a number of different functional forms other than the double log. The results in Table 5 are very robust. Basically, there is a partial correlation between firm size and industry counts that exists even when we control for age.

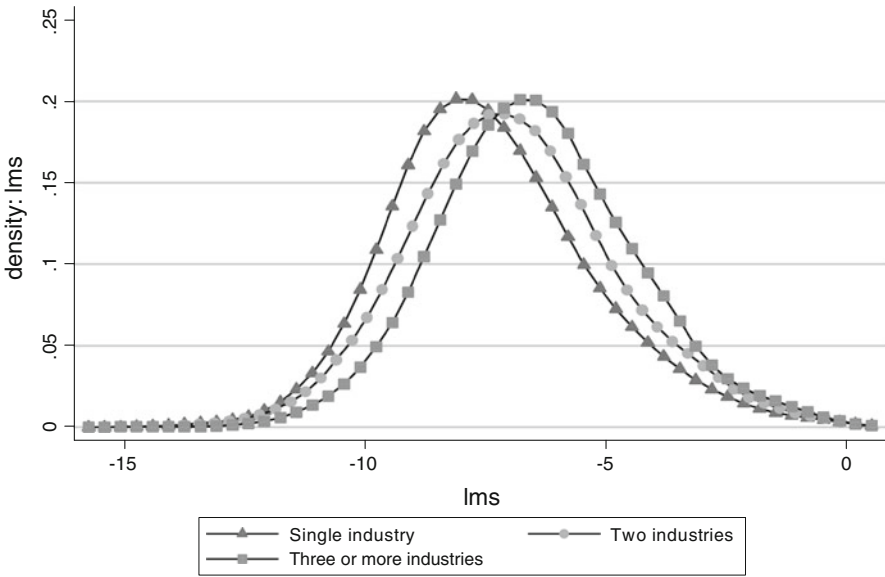


Fig. 8 Size distribution of Belgian companies by single and multi industry cohorts

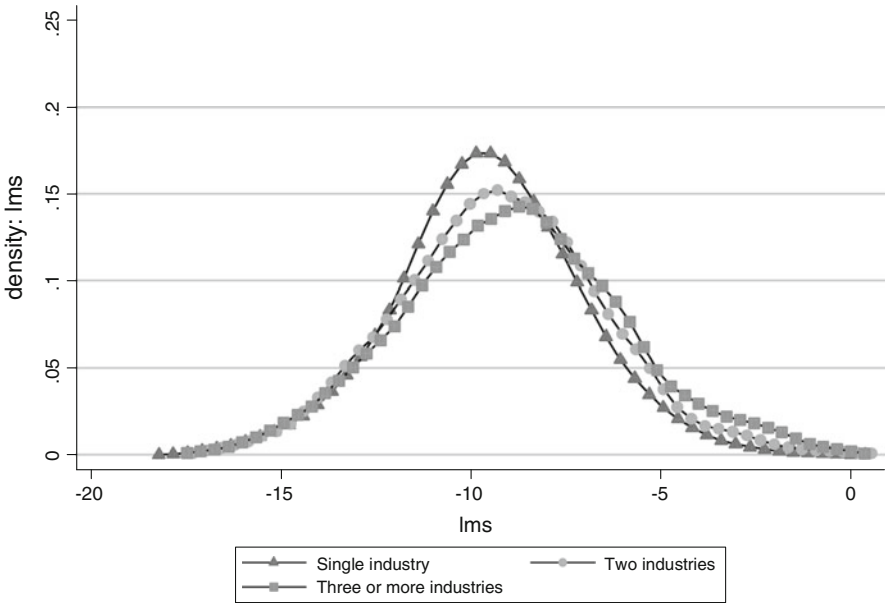


Fig. 9 Size distribution of UK companies by single and multi industry cohorts, $10 \leq \text{age} < 25$

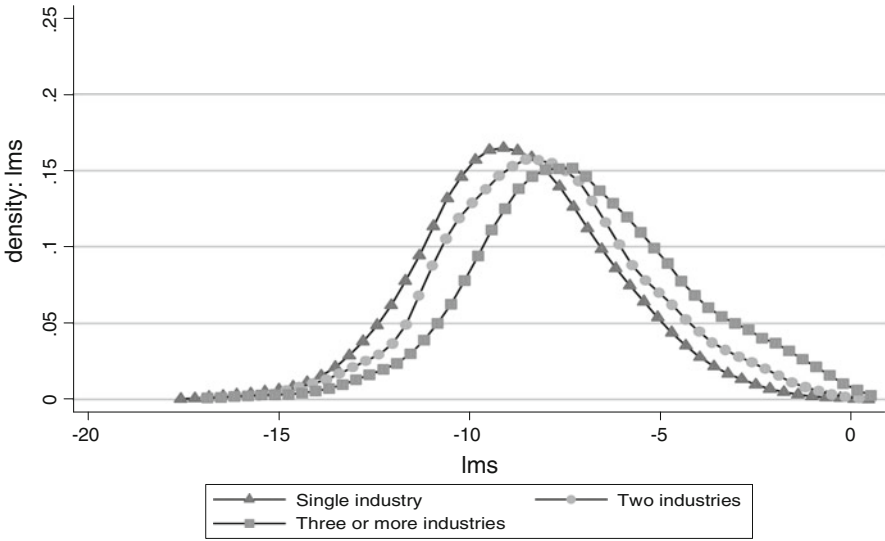


Fig. 10 Size distribution of UK companies by single and multi industry cohorts, age ≥ 25

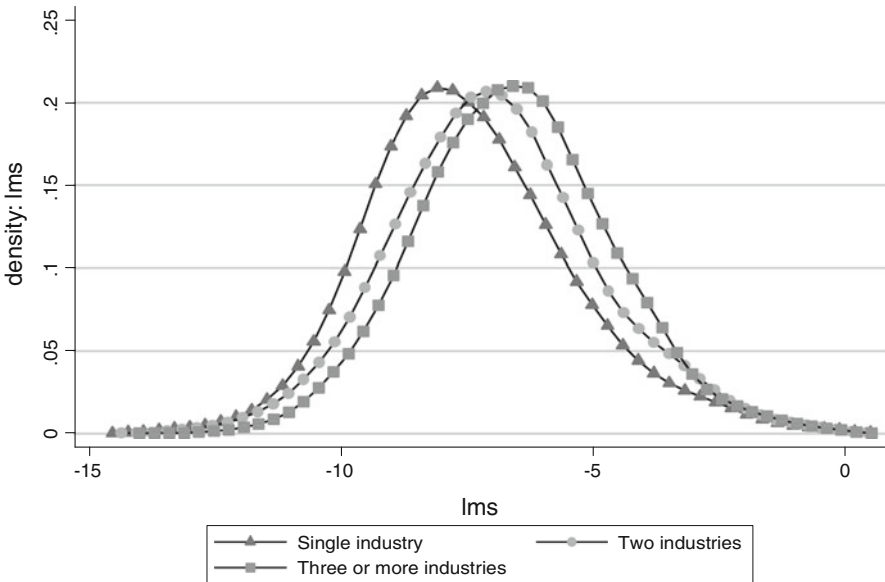


Fig. 11 Size distribution of Belgian companies by single and multi industry cohorts, $10 \leq \text{age} < 25$

researchers who document and model economies of scope, via differences in supply side primitives such as entry costs, have identified an important empirical factor that drives the FSD.¹²

¹² 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.

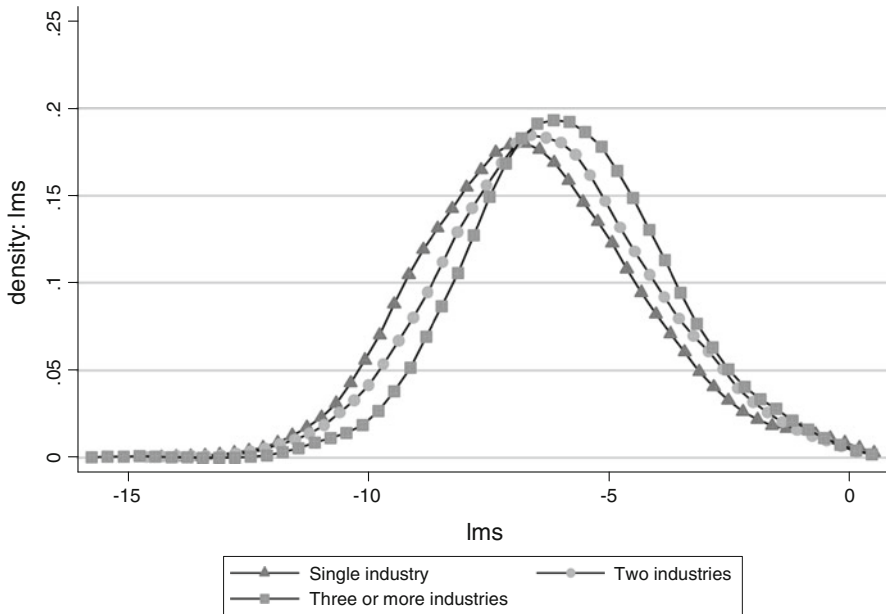


Fig. 12 Size distribution of Belgian companies by single and multi industry cohorts, age ≥ 25

4 Conclusions

We have showed the related stylized facts for the FSD for age cohorts, as proposed by Cabral and Mata (2003), hold within 4-digit manufacturing industries in the UK and Belgium. Studies by 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 each other in the determination of firm size. Yet, they realize that differences in supply-side primitives and 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 et al. (2006) and Nocke and Yeaple (2008) do link firm size to economies of scope from inter-industry (4-digit) diversification. This results in their theories from supply side heterogeneity in firms' 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 firm's age, among other factors. The stylized facts for the FSD that are generated by Cabral and Mata (2003) cannot be fully replicated by using counts of inter-industry (4-digit) product diversification rather than age cohorts. Yet within age cohorts, particular firms that are ten years or older 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. Economies of scope are identified as an important empirical factor that drives the FSD.

The presence of inter-industry product diversification has implications for empirical work in IO and Trade. For example, in the IO productivity literature, the presence of multi-product firms is shown to distort the estimation of total factor productivity in De Loecker (2006). If firms are spread across industries, one would need sales, raw material and capital deflators to be constructed as a weighted sum of deflators across several industries to avoid problems in interpretation and estimation.

In the international trade literature, Bernard et al. (2010) have shown that the product portfolio of firms has changed greatly. If one uses company-level data to evaluate the impact of trade liberalization on firm performance by trade orientation, one would have to realize that company expansions and contractions could result from between, as well as within, 4-digit industry reallocations. We urge theoretical and empirical researchers to pay more attention to the role of inter-industry diversification in companies.

Acknowledgments We are extremely grateful for the helpful comments provided by John Sutton, Patrick Van Cayseele, Ciara Whelan, Damiaan Persyn and anonymous referees. In addition, the editor has undertaken corrections and given us helpful comments that were over and above the call of duty.

References

- Angelini, P., & Generale, A. (2008). On the evolution of firm size distribution: Further empirical evidence. *American Economic Review*, 98(1), 426–438.
- Bernard, A. B., Redding, S. J., & Schott, P. K. (2006). Multi-product firms and trade liberalization. *NBER Working Papers* No. 12782, National Bureau of Economic Research.
- Bernard, A. B., Redding, S. J., & Schott, P. K. (2010). Multiple-product firms and product switching. *American Economic Review*, 100(1), 70–97.
- Buzzacchi, L., & Valletti, T. (2006). Firm size distribution: Testing the independent submarkets model in the Italian motor insurance industry. *International Journal of Industrial Organisation*, 24(4), 809–834.
- Cabral, L., & Mata, J. (2003). On the evolution of the firm size distribution: Facts and theory. *American Economic Review*, 93(4), 1075–1090.
- de Juan, R. (2003). The independent submarkets model: An application to the Spanish retail banking market. *International Journal of Industrial Organization*, 21, 1461–1488.
- De Loecker, J. (2006). Product differentiation, multi-product firms and structural estimation of productivity. *NBER working paper series* No. 13155, National Bureau of Economic Research.
- Dunne, T., Roberts, M. J., & Samuelson, L. (1989). The growth and failure of U.S. manufacturing plants. *Quarterly Journal of Economics*, 104, 671–698.
- Ericson, R., & Pakes, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies*, 62, 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*. Paris: Librairie de Recueil Sirey.
- Gollop, F. M., & Monahan, J. L. (1991). A generalized index of diversification: Trends in U.S. manufacturing. *The Review of Economics and Statistics*, 73(2), 318–330.
- Hopenhayn, H. A. (1992). Entry, exit and firm dynamics in long run equilibrium. *Econometrica*, 60, 1127–1150.
- Hart, P. E., & Prais, S. J. (1956). The analysis of business concentration: A statistical approach. *Journal of Royal Statistical Society, Series A*, 119–150.
- Ijiri, Y., & Simon, H. (1964). Business firm growth and size. *American Economic Review*, 54, 77–89.
- Ijiri, Y., & Simon, H. (1977). *Skew distributions and sizes of business firms*. Amsterdam: North-Holland Pub Co.
- Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica*, 50, 649–670.

- Klepper, S., & Thompson, P. (2006). Submarkets and the evolution of market structure. *Rand Journal of Economics*, 37(4), 861–886.
- Luttmer, G. J. (2007). Selection, growth, and the size distribution of firms. *Quarterly Journal of Economics*, 122(3), 1103–1144.
- Mariuzzo, F., Walsh, P. P., & Whelan, C. (2003). Firm size and market power in carbonated soft drinks. *Review of Industrial Organization*, 23(3–4), 283–299.
- Melitz, M. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6), 1695–1725.
- Nocke, V., & Yeaple, S. (2008). Globalization and the size distribution of multiproduct firms. *CEPR Discussion Papers* No. 6948, Centre for Economic Policy Research.
- Schmalensee, R. (1989). Inter-industry studies of structure and performance. In R. Schmalensee & R. Willig (Eds.), *Handbook of industrial organization*, vol. 2 (pp. 951–1009). Amsterdam: North-Holland.
- Sutton, J. (1998). *Technology and market structure*. Cambridge, MA: MIT Press.
- Walsh, P. P., & Whelan, C. (2002). Portfolio effects and firm size distribution—Carbonated soft drinks. *Economic and Social Review*, 33(1), 43–54.