CSEF

Centre for Studies in Economics and Finance

WORKING PAPER NO. 228

Firm Size Distribution and Returns to Scale. Non-Parametric Frontier Estimates from Italian Manufacturing

Lisa Crosato, Sergio Destefanis and Piero Ganugi

May 2009

University of Naples Federico II | University of Salerno | Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS – UNIVERSITY OF NAPLES
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unisa.it

WORKING PAPER NO. 228

Firm Size Distribution and Returns to Scale. Non-Parametric Frontier Estimates from Italian Manufacturing

Lisa Crosato [*], **Sergio Destefanis** [**] **and Piero Ganugi** [***]

Abstract
This paper explores the relationship between firm size distribution and technology. We analyse firm technology across selected manufacturing industries by means of a non-parametric production analysis, the Free Disposal Hull approach (Deprins et al., 1984; Kerstens and Vanden Eeckaut, 1999) and appraise the links between size and scale elasticity, finding a clear inverse relationship. Building on this result, we inquire whether the shape of the firm size distribution is related to a particular pattern of scale elasticities. We rely on the Zipf Plot (Stanley et al., 1995) of the Pareto IV distribution, which is concave up to a given threshold, and then approximately linear. Firms in the concave part of the plot are overwhelmingly found to experience increasing returns to scale. On the contrary, firms in the linear part are mainly characterised by constant returns to scale.

JEL classification: L11, L6, D20, C14

Acknowledgements: This paper was developed within the project: "Industry evolution: innovation, profitability and firm's growth", coordinated by Professor Maurizio Baussola (DISES - Università Cattolica del Sacro Cuore, Piacenza) in cooperation with ISTAT (Italian Statistical Office, Lombardy regional office). We thank Marco Vivarelli for helpful comments on a previous draft and Francesco Addesa for skilful research assistance.

[*] Università di Milano-Bicocca (lisa.crosato@unimib.it)

[**] Università di Salerno, CELPE and CSEF (destefanis@unisa.it)

[***] Università Cattolica del Sacro Cuore, Piacenza (piero.ganugi@unicatt.it)

Table of contents

1 Introduction

Firm size has been typically modelled in the literature by means of the Lognormal and Pareto distributions, both because of their descriptive power and of their link to Gibrat's Law (Gibrat, 1931; Kalecki, 1945; Simon, 1955; Steindl, 1965). Nonetheless, these distributions have systematically shown some shortcomings. The Lognormal over- or under-estimated the right tail of firm sizes (Stanley *et al.*, 1995; Hart and Oulton, 1997; Voit, 2001), while the Pareto I, noticeably better in the right tail, has problems in fitting the left one (Ijiri and Simon, 1977; Steindl, 1965; Okuyama *et al.*, 1999). Departures from the Pareto I and Lognormal distributions have been interpreted as deviations from Gibrat's Law and therefore as instances of different regimes of growth for differently sized firms.

A similar problem has been noticed for some time in the income distribution literature. Attempts at its solution involved the application either of Paretian tail models (Singh and Maddala, 1976; Dagum, 1977; Stoppa, 1990) or of a right-truncated Lognormal on the left tail of the distribution and a Pareto I on the rest of it (Clementi and Gallegati, 2005).

Crosato and Ganugi (2007) first carried over a similar approach to the field of firm size distribution, applying the Pareto IV distribution (Arnold, 1983) within Italian manufacturing industries. They find favourable evidence for the Pareto IV both at the aggregate and at the sectoral level, shedding again doubts on the relevance of the law of proportionate effect in this ambit. Still, the size distribution of large[1] firms can be successfully fitted through

[1]The definition of large firms here adopted is spelled out in section 2

a Pareto I distribution, as the Pareto IV's right tail matches the Pareto I's. This aspect is made particularly clear by the Zipf Plot (Stanley *et al.*, 1995), a double log scale plot providing an estimate of any firm's size under the null that the size is distributed according to a specific distribution: linearity in the right tail of the plot seems to narrow the domain of validity of Gibrat's Law to large firms.

In the literature, firms lying in the linear part of the Zipf Plot are expected to experience constant returns to scale, while no clear-cut opinion is provided on the nature of returns to scale characterising firms in non-linear parts of the plot (Simon and Bonini, 1958; Ijiri and Simon, 1964; Vining, 1976; Lucas, 1978).

The basic goal of the present paper is to provide novel empirical evidence on the relationship between the distribution of firm size and returns to scale, gauging the expected association between constant returns to scale and distributions compatible with Gibrat's Law, as well as the possibility that non-linear parts of the Zipf plot are systematically characterised by non-constant returns to scale. We do so by relying on the Free Disposal Hull (FDH) approach first proposed by Deprins *et al.* (1984), which imposes very little a priori structure on the pattern of the returns to scale. In particular, no hypothesis is made on the relationship between the shape of the production function and the size distribution of firms. As a by-product of the FDH approach we obtain estimates for the technical efficiency of all firms.

The paper is divided in five sections. In section 2 we briefly present our dataset. In section 3 we describe the FDH approach to the measurement of technical efficiency and scale elasticity, as well as our main results on the relationships between firm size, technical efficiency and scale elasticity. The evidence supports, in all industries, a lack of relationship between firm size and technical efficiency, while elasticities of scale are systematically lower for large firms. In section 4 we focus on the Zipf Plot, determining the threshold above which firms approximate the Pareto I distribution. We find that firms display different returns to scale regimes below and above this cut-off point. Section 5 contains some final remarks and describes further directions of research.

2 Data Description

We focus on six industries selected from the Micro1 survey, assembled by ISTAT (Italian National Institute of Statistics) through the matching of the Structural Business Statistics Survey (SCI) and the Community Innovation Survey (CIS1). The original dataset is composed of 5445 firms with at least 20 employees, followed from 1989 to 1997, and was extensively analysed in Crosato and Ganugi (2007), who also discussed the characteristics of the dataset in relation to Gibrat's Law (about Italian manufacturing in the same period, see also Cefis *et al.* (2003); Bottazzi *et al.* (2005)).

As it was constructed by ISTAT in order to analyse the dynamics of Italian industry during the 1990s, Micro1 has good data quality and is reasonably representative of its population. However, it is a closed panel and hence exposed in principle to survivor bias (Mansfield, 1962). In a given period, slow-growing small firms are less likely to survive than slow-growing large firms, which can reduce in (relative) size but still remain in the sample.

Small survivors in a closed sample should hence show higher growth than their population counterparts, leading to unwarranted conclusions against Gibrat's Law. In principle, this survivor bias should be considered and corrected in order to yield to a correct evaluation of the Law's validity (Evans, 1987; Hall, 1987; Dunne and Hughes, 1994). Our analysis, however, is not so much concerned with the Law as with the association of different parts of the Zipf Plot with different kinds of returns to scale. In section 4 we provide some evidence supporting the claim that the survivor bias does not materially affect this association.

An important, and relatively uncommon, feature of our dataset is that it includes data not only on the number of employees, but also on the total number of work hours and, alternatively, on the number of blue and white collars. Relying on either of these measures is likely to provide a better specification of the labour input than aggregate employees. On the other hand, a potential drawback of the dataset is its lack of measures for intermediate inputs. In this case, the production set can only be specified in terms of value added, and not of gross output (sales). This is not without consequence, because the gross-output specification is considered by some authors the most appropriate for the analysis of technology and productivity (Basu and Fernald, 1997). There are however, analytical derivations of the connections between value-added and gross-output specifications (Basu and Fernald, 1997; Diewert and Fox, 2004). The value-added estimate of returns to scale is greater than the corresponding gross-output estimate if there are increasing returns to scale in the gross-output model and vice versa for decreasing returns to scale in the gross-output model. For present purposes it is crucial to remark that if there are constant returns to scale in the gross-output model, then the value-added model also exhibits constant returns to scale. Furthermore, with the appropriate information about inputs and cost-shares, one could always retrieve the gross-output returns to scale from the value-added based estimates, provided no correlation exists in the data between the intermediate inputs-to-output ratio and value added. In Section 3 we provide some evidence against the existence of this correlation, at least in most industries.

We selected six industries for our empirical analysis, which will be often denoted according to their 2-digit ATECO '91 (ISIC Rev.3) code. They are:

- Food products and beverages (DA 15);
- Textile (DB 17);
- Chemicals and chemical products (DG 24);
- Non-metallic mineral products (DI 26);
- Fabricated metal products except machinery and equipment (DJ 28);
- Machinery and equipment (DK 29).

These industries have been chosen through two different steps. First,as shown in Table 1 we have found eleven industries with at least 200 firms in each year to get reliable non-parametric estimates (see section 3 for details). Second, we selected six out of these eleven industries according to their importance in terms of share on total manufacturing value added. As can be seen from Table 2, industries DA15, DB17, DG24, DI 26 and DK29 are on

industry				years					
	1989	1990	1991	1992	1993	1994	1995	1996	1997
DA15	389	390	389	391	391	391	391	391	391
DB17	537	538	532	522	523	523	525	526	526
DB18	221	220	223	228	223	225	224	224	226
DG24	275	275	275	259	256	256	259	262	263
DG25	337	344	341	317	326	321	318	314	317
DI26	429	430	428	429	429	428	427	428	429
DJ27	224	219	223	215	219	215	216	215	215
DJ28	632	631	635	615	613	611	603	605	601
DK29	727	729	731	742	741	749	754	754	752
DL31	218	216	215	214	214	214	211	212	213
DN36	413	410	414	421	425	425	427	426	426

Table 1: **Number of firms by industry**

industry				years					
	1989	1990	1991	1992	1993	1994	1995	1996	1997
DA15	8.45	8.9	9.02	9.73	10.09	9.11	8.63	9.39	8.97
DB17	5.52	5.7	5.35	4.99	5.06	5.22	5.22	5.06	5.08
DB18	1.49	1.45	1.52	1.91	1.94	1.82	1.82	1.82	1.73
DG24	9.15	10.13	10.22	10.54	10.91	10.96	10.98	11.45	10.74
DG25	4.38	4.29	4.31	4.09	4.36	4.36	4.48	4.66	4.48
DI26	5.6	5.65	5.67	5.9	5.88	5.79	5.69	5.69	5.58
DJ27	5.23	4.76	4.26	4.08	4.28	4.58	5.16	4.68	4.64
DJ28	4.08	4.21	4.16	4.14	4.28	4.3	4.61	4.88	4.69
DK29	12.17	11.87	11.6	12.29	13.42	13.76	13.73	14.44	14.36
DL31	4.69	5.06	5.31	5.21	5.16	4.67	4.48	4.23	4.58
DN36	2.23	2.28	2.33	2.4	2.58	2.5	2.44	2.49	2.43

Table 2: **Percentage share by industry with respect to total manufacturing value added**

top of the value added ranking over the whole period. The three industries which occupy alternatively the sixth place are DJ27, DJ28 and DL31. We retain DJ28 because it contains almost three times as many firms as the other two industries.

Finally, Table 3 reports the standards we relied upon in order to classify firms in different size ranges. We follow the 2003 European Commission Recommendation (2003/361/EC). Among the available size proxies, we consider Total Assets (like in Crosato and Ganugi (2007)) because of the lack of volatility of this variable through time.

size proxy	micro	small	medium	large
employees (number)	< 10	$10 \leq$ and < 50	$50 \leq$ and < 250	≥ 250
sales (million Euro)	< 2	$2 \leq$ and < 10	$10 \leq$ and < 50	≥ 50
total assets (million Euro)	< 2	$2 \leq$ and < 10	$10 \leq$ and < 43	≥ 43

Table 3: **Definition of firm dimension according to the European Commission (2003/361/EC)**

3 Efficiency and Elasticity

In this section we first describe our non-parametric approach to the measurement of technical efficiency and scale elasticity, based on the concept of the Free Disposal Hull (Deprins *et al.*, 1984; Kerstens and Vanden Eeckaut, 1999). We proceed then to present our main results on the relationships between firm size, technical efficiency and scale elasticity.

3.1 The FDH approach

Non-parametric methods provide estimates of the upper boundary of a production set (the so-called production frontier) without supposing the existence of a functional relationship between inputs and outputs (Farrell, 1957; Fried *et al.*, 1993). The frontier is supported by some of the observed producers. Non-parametric methods are divided between those that impose upon the production set the hypothesis of convexity (usually gathered under

the label of Data Envelopment Analysis, or DEA) and those that do not need this assumption (the Free Disposal Hull - FDH - approach proposed in Deprins *et al.* (1984); Tulkens (1993)). In the latter case, the only property imposed on the production set is strong input and output disposability, while in DEA the additional hypothesis of convexity is made. More formally, in FDH, for a given set of producers Y_0, the reference set $Y(Y_0)$ is characterised, in terms of an observation i, by the following property: (X^i, Y^i) observed, $(X^i + a, Y^i - b) \in Y(Y_0), a, b \geq 0$
where a and b are vectors of free disposal of input and output, respectively.

In other words, due to the possibility of strong input and output disposal, the reference set includes all the producers which are using the same or more inputs and which are producing the same or less output in relation to observation i.

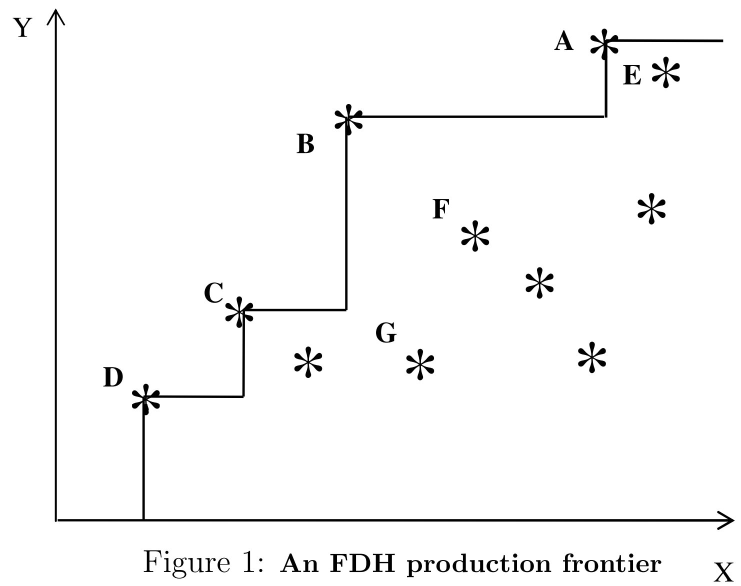

Figure 1: **An FDH production frontier**

Let us take as an example Figure 1, where we consider a technology with one input (X) and one output (Y). The input-output pairs correspond to producers examined at a given point in time. Beginning with observation B, we define every observation located at its right and/or below it (that is with more input and same output, or with less output and same input; or else with more input and less output, as F) as dominated by B.

In FDH, this comparison is carried out for every observation, and observations not dominated by any other observation are considered efficient producers, belonging to the frontier of the reference set: on the other hand, the observations that are dominated are considered inefficient. In DEA, on the other hand, the frontier of the overall reference set is found by constructing a convex envelope around the production set; this implies the assumption not only of free input and output disposal, but also of convexity. Hence, the DEA frontier must exhibit by construction non-decreasing returns to scale for relatively smaller observations, and non-increasing returns for relatively larger observations. This is not true in FDH and is of crucial importance for the present research.

One problem with FDH is that many observations may possibly be efficient because they are located in an area of the production set where there are no other observations with which they can be compared (they are, as it were, efficient by default). To circumvent this problem, we use a refinement of traditional FDH, the VP-FDH (variable-parameter FDH) proposed by Kerstens and Vanden Eeckaut (1999), which decisively reduces the problem of efficiency by default. VP-FDH is defined as the intersection of FDH technologies that impose by assumption non-decreasing and non-increasing returns to scale. First, each observation is compared not only to any other observation but also to their smaller or larger proportional replicas; then, one selects for each given observation the assumption about returns to scale that yields the highest efficiency score. While still relaxing the hypothesis of convexity (meaning that the nature of returns to scale is not restricted a priori), VP-FDH imposes more structure on the production set than traditional FDH, greatly increasing the scope for comparisons between observations, and reducing correspondingly the problem of efficiency by default.

In all frontier methods the distance of a producer from the frontier provides its measure of technical inefficiency, or, for short, its efficiency score. Typically, the (output-oriented or input-oriented) measure of Debreu-Farrell is used for this purpose. Then, output-oriented technical efficiency is equal to the maximum radial output expansion consistent with the utilisation of a given input vector (in empirical applications the inverse of this measure, bounded between zero and one, is typically utilised; we will here conform to this use), while input-oriented technical efficiency is equal to the maximum radial input reduction consistent with the production of a given output vector.

On the other hand, there are various methods in non-parametric frontier analysis to assess the nature of returns to scale on the frontier point relevant for any given producer (see the discussions in Førsund (1996), or in Kerstens and Vanden Eeckaut (1999)). In a qualitative sense, one must ascertain whether the frontier point relevant for an inefficient producer according to the variable-returns-to-scale technology must be scaled up or down to obtain the frontier point relevant for an inefficient producer according to the constant-returns-to-scale technology. In the first case, the frontier exhibits increasing returns to scale, while the contrary holds true in the opposite case. If the two frontier points coincide, the frontier exhibits constant returns to scale. There exist also some procedures that allow the derivation of quantitative measures of scale elasticities from non-parametric frontier analysis. A simple and attractive procedure, derived from Frisch's Beam variation equations, is to compute the ratio between the natural log of the output-increasing efficiency score and the natural log of the input-decreasing efficiency score (Førsund and Hjalmarsson, 1979; Førsund, 1996). This ratio is only an average measure of scale elasticity and is determined by the (generally non-measurable) magnitude of returns to scale in the two frontier points relevant for the producer taken into consideration. Hence it exists only for given data intervals (not for given points), and only for inefficient producers. It should be kept in mind that a major problem of small-sample bias arises when non-parametric frontier approaches are used (Kneip *et al.*, 1998; Gijbels *et al.*, 1999; Kittelsen, 1999). As reported in Kittelsen (1999) these approaches begin to be characterised by substantial biases for sample sizes around 100 to 150 observations. This suggested to restrict empirical analysis to industries well exceeding the 100 to 150 observations (per year) mark.

3.2 Technical Efficiency

We now proceed to give technical efficiency scores calculated separately through VP-FDH for each industry and year. The scores are subsequently used to compute the measure of scale elasticity proposed in Førsund and Hjalmarsson (1979). We present results obtained with VP-FDH, rather than with traditional FDH, because we feel that the former, by reducing the problem of efficiency by default, provides more reliable results. We rely on a production set where the output is value added, while number of blue-collar employees, number of white-collar employees and book value of fixed assets are the inputs. We could have relied on a different production set, with number of work hours and book value of fixed assets as inputs. In accordance with the empirical literature on Italian manufacturing (see, for instance, Balloni (1984); Prosperetti and Varetto (1991); Ofria (1997); Destefanis and Sena (2007)), we find however that a more satisfactory measure of the quality of labour inputs is obtained by dividing employees in blue and white collars. This makes it impossible to take into account the hours worked by each category of employees (in fact, we also utilised the alternative production set and traditional FDH, finding virtually no difference, as far as our results of interest were concerned; results are available on request). Some summary statistics of the technical efficiency scores are provided in Table 4. For the sake of brevity, we pool together all years in presenting these results.

We obtain efficiency scores with reasonably high mean and median values, and acceptably low dispersion. These results point to a satisfactory specification of the production set and lend robustness to our subsequent analysis. Indeed, we should not expect particularly low efficiency levels in our dataset, composed by firms having survived for nine years, hence belonging to the efficient and able-to-grow bunch. Being based upon an extremely reduced set of assumptions about the feasible production set, our estimates of technical efficiency may be of some interest for their own sake. For instance, the relationship between technical efficiency and firm size has often aroused interest

(TE-Input Oriented)						
Industry	Min	$1^{st}Qu$	Med	Mean	$3^{rd}Qu$	Max
DA15	0.092	0.513	0.667	0.657	0.813	1
DB17	0.121	0.536	0.667	0.666	0.800	1
DG24	0.068	0.588	0.706	0.694	0.824	1
DI26	0.104	0.591	0.72	0.706	0.836	1
DJ28	0.130	0.571	0.692	0.687	0.808	1
DK29	0.081	0.537	0.657	0.657	0.789	1
(TE-Output Oriented)						
Industry	Min	$1^{st}Qu$	Med	Mean	$3^{rd}Qu$	Max
DA15	0.013	0.406	0.537	0.550	0.689	1
DB17	0.032	0.450	0.572	0.582	0.715	1
DG24	0.038	0.522	0.651	0.642	0.778	1
DI26	0.031	0.485	0.627	0.619	0.764	1
DJ28	0.075	0.473	0.592	0.599	0.723	1
DK29	0.029	0.462	0.570	0.582	0.705	1

Table 4: **Summary statistics of technical efficiency scores. Annual scores are pooled.**

in the literature - see the recent papers by Alvarez and Crespi (2003) and Taymaz (2005) and the references therein. A proper treatment of this issue would require however a careful specification of the internal and external determinants of technical efficiency, going beyond the scope of this work.

3.3 Scale Elasticity

A problem usually besetting the empirical measurement of returns to scale is that time-series based measures conflate genuine scale effects with short-run effects mostly linked to cyclical factor utilisation. This is not likely to be the case with the present measures, fully based upon cross-sectional information for each year and industry. Indeed, it has long be recognised that cross-section based estimates tend to provide long-run responses, as most short-run effects wash out (Kuh, 1959; Baltagi and Griffin, 1984).

In order to appraise the robustness of our scale elasticity measures with respect cyclical biases, it is useful to consider Figure 2, which provides the rates of growth of Italian GDP in the years included in the analysis. Clearly, 1989 was the year closer to a cycle peak, while 1993 is the cycle trough. Toward the end of the sample the cycle picks up again.

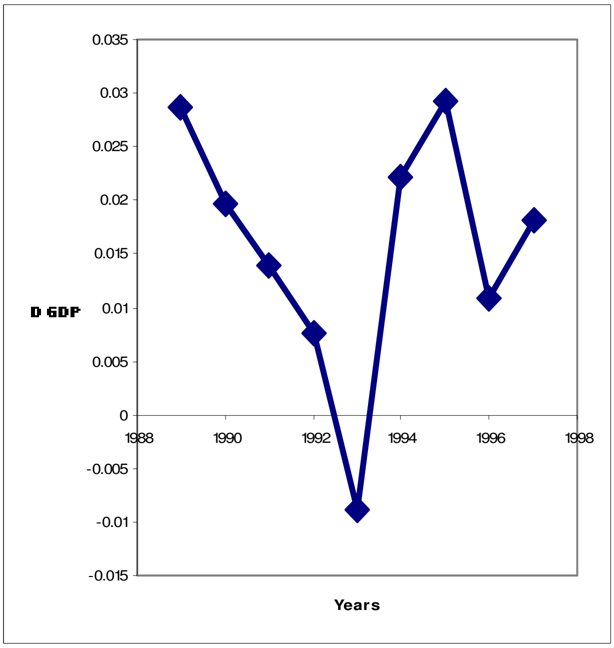

Figure 2: **Growth rates of the Italian GDP: 1989-97**

If factor-utilisation effects were not ironed out by our cross-sectional procedure, the resulting measures of scale elasticity could be cyclically biased. Hence it is important to compare their values in different periods. For the sake of brevity we carry out this comparison for 1989 (peak), 1993 (trough) and 1997 (last year, with the GDP growth rate very close to its sample mean). As can be seen from Table 5, there is no evidence of a cyclical bias in our scale elasticities. In particular, the 1993 values are not systematically higher than the other ones, as one would expect to be the case in the presence of slack inputs.

Industry	scale elasticities					
	Min	$1^{st}Qu$	Med	Mean	$3^{rd}Qu$	Max
			1989			
DA15	0.384	1	1.519	2.056	2.765	8.691
DB17	0.431	1	1.084	1.520	1.667	6.648
DG24	0.868	1	1	1.241	1.227	3.508
DI26	1.002	1.244	1.556	2.050	2.166	9.119
DJ28	0.415	1	1.202	1.604	1.960	5.957
DK29	0.121	0.782	1.212	1.501	1.899	6.522
			1993			
DA15	0.676	1	1.092	1.603	1.869	6.054
DB17	0.517	1	1.368	1.561	1.867	4.457
DG24	0.566	1	1	1.134	1.117	3.628
DI26	0.445	1	1.217	1.449	1.589	5.315
DJ28	0.610	1	1.154	1.566	1.749	6.690
DK29	0.068	0.765	1.243	1.640	2.074	6.660
			1997			
DA15	0.540	1	1.319	1.956	2.179	8.395
DB17	0.535	1	1.098	1.522	1.763	5.837
DG24	0.492	1	1	1.308	1.328	5.831
DI26	0.715	1	1.105	1.513	1.636	5.480
DJ28	0.623	1	1	1.457	1.549	5.145
DK29	0.139	0.721	1.284	1.604	2.104	6.107

Table 5: **Summary statistics of the scale elasticities (1989, 1993, 1997).**

Table 5 also makes it clear that the measurement of scale elasticities through VP-FDH produces some anomalous values. These values occur when firms are close to a very high or a very wide "step" of the VP-FDH frontier. In the following analysis the lowest and the highest 2.5% of the elasticity scores will be trimmed out of the sample.

It turns out that scale elasticity is inversely related to firm size throughout all industries, as is shown by the negative correlation coefficients reported in Table 6, where again we focus on 1989, 1993 and 1997. This inverse relationship is also clearly depicted in Figure 3 and Figure 4, leaving no doubt about its strength.

Industries	Years		
	1989	1993	1997
DA15	-0.56	-0.44	-0.44
DB17	-0.51	-0.46	-0.50
DG24	-0.46	-0.39	-0.47
DI26	-0.44	-0.55	-0.41
DJ28	-0.56	-0.56	-0.49
DK29	-0.56	-0.54	-0.48

Table 6: **Pearson correlation coefficients between firm size and scale elasticities (1989, 1993, 1997).**

In Section 2 we have remarked that a specification of the production set in terms of value added, and not of gross output, is likely to yield larger estimates of returns to scale if there are increasing returns to scale in the gross-output model (and vice versa for decreasing returns to scale), but to exhibit constant returns to scale if there are constant returns to scale in the gross-output model. This is an extremely important result for present purposes and is proved for the case of parametric estimation by Basu and Fernald (1997) and Fox *et al.* (2004), provided no correlation exists between the intermediate inputs-to-output ratio and value added. Below we provide some *prima facie* evidence against the existence of this correlation. Having no measure for intermediate inputs, we rely on (sales - value added) / sales as a proxy for intermediate inputs-to-output ratio. In Table 7 we report the correlation coefficients between value added, the inputs in the production

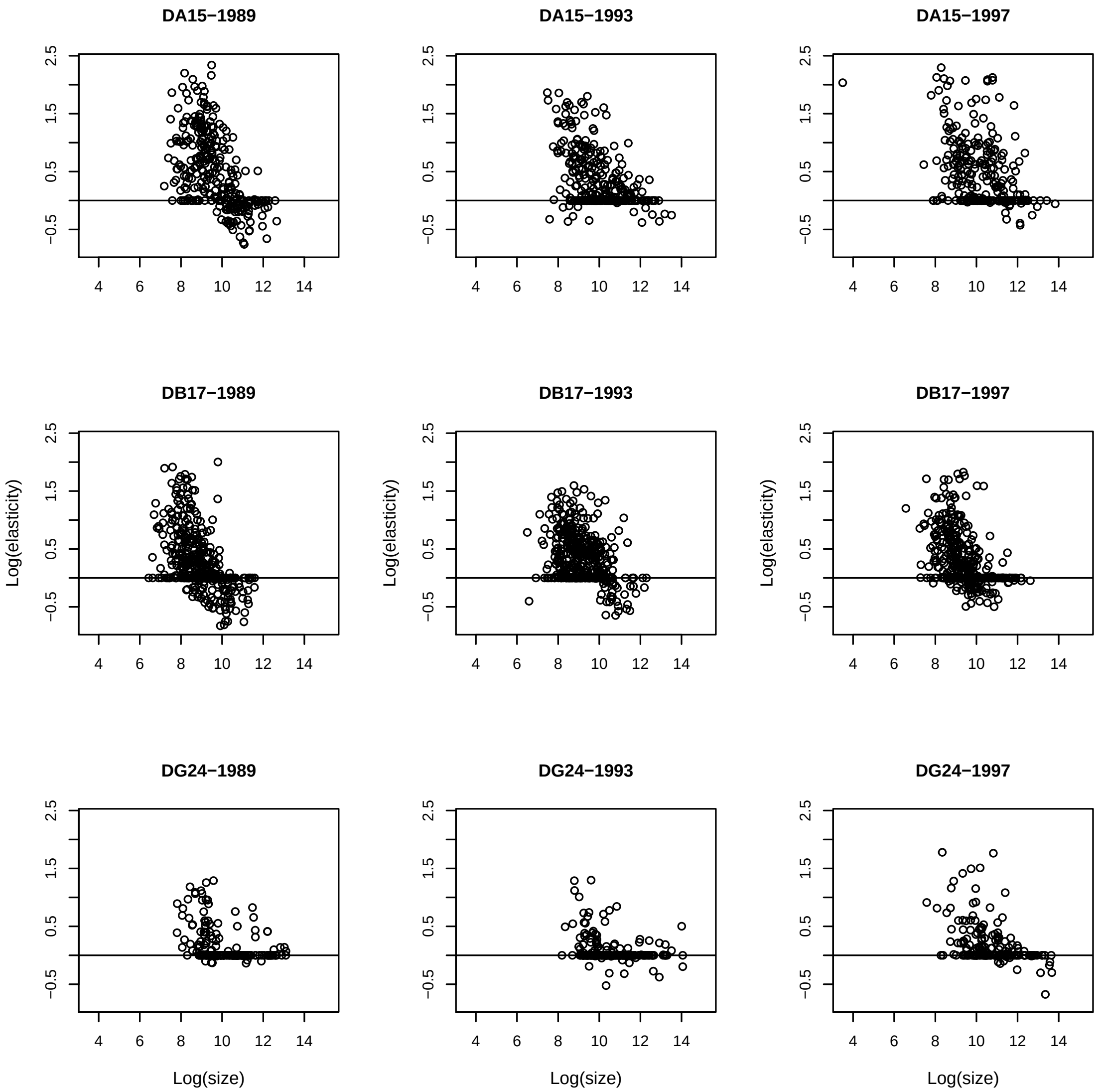

Figure 3: **The scale elasticity-size relationship. Industries DA15, DB17, DG24**

set, and this proxy.

It turns out that the correlation between value added and the intermediate inputs-to-output proxy is negligible but for two industries, DG 24 and DI 26. There too this correlation is much weaker than the correlation between value added and the inputs included in the production set. We conclude that

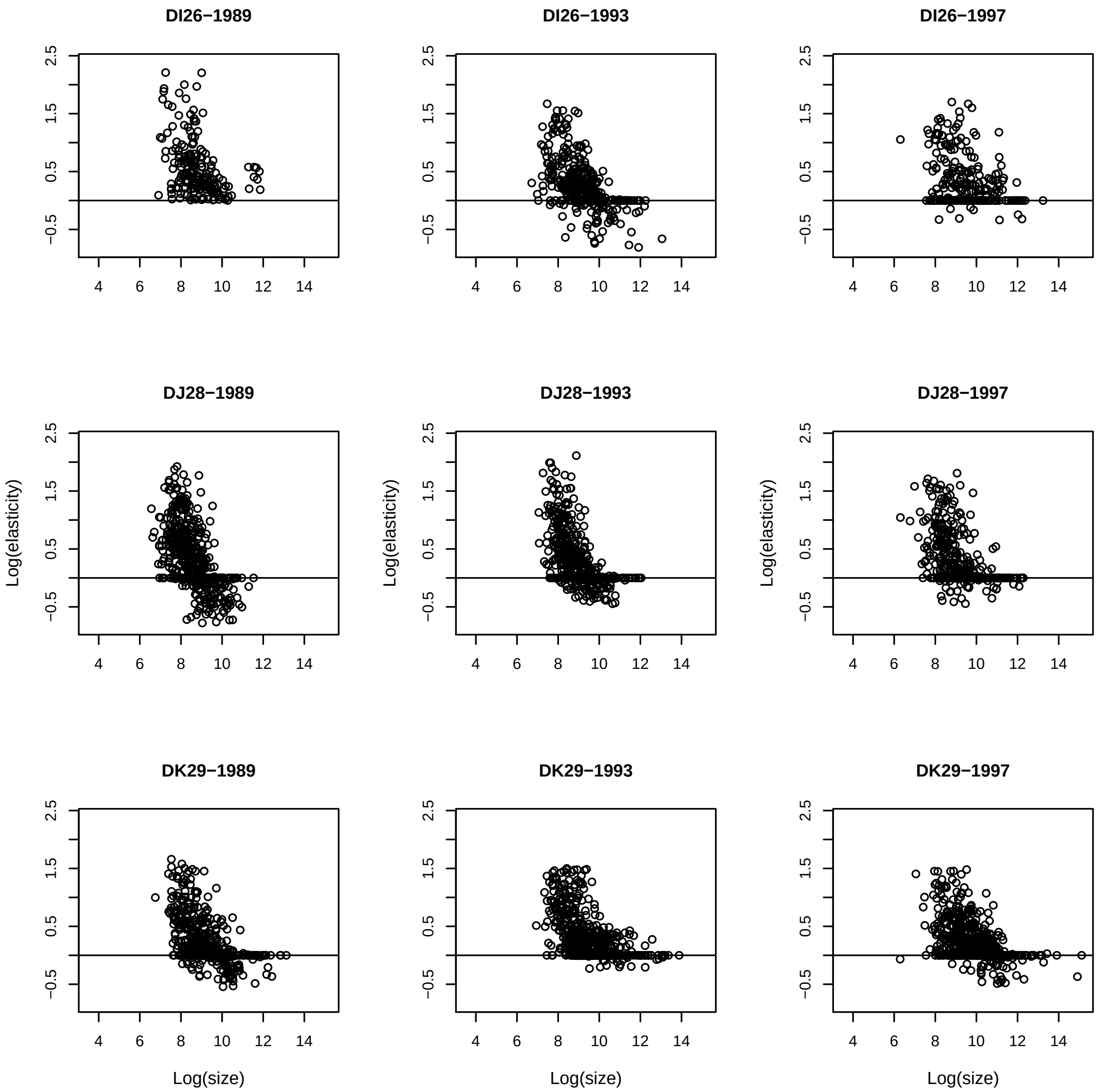

Figure 4: **The scale elasticity-size relationship. Industries DI26, DJ28, DK29**

the value-added specification of the production set is not likely to provide misleading results for our analysis. In particular, constant returns to scale are likely to be correctly characterised by our estimates.

Pearson ρ's between value added and	Industries DA15	DB17	DG24	DI26	DJ28	DK29
book value of fixed assets	0.78	0.94	0.81	0.89	0.86	0.88
number of white-collar employees	0.89	0.92	0.92	0.88	0.87	0.91
number of blue-collar employees	0.92	0.94	0.83	0.90	0.89	0.83
(sales-value added)/sales	-0.11	0.00	-0.30	-0.23	-0.01	-0.04

Table 7: **Pearson correlation coefficients between value added, the inputs in the production set, and the proxy of the intermediate inputs-to-output ratio. Annual estimates are pooled.**

4 Returns to Scale and Firm Size Distribution

4.1 The Zipf Plot technique applied to the Pareto IV distribution

The Zipf Plot (Stanley *et al.*, 1995), is a graph of the log of the rank versus the log of the variable being analysed, in this case firm size. Let X be the random variable size, $(x_1, x_2, \ldots, x_N)$ be the vector of its realizations on a set of N firms and $F_X(x)$ its cumulative distribution function. Next, suppose that the observations are ordered from the largest to the smallest so that the index i is the rank of the i-th firm. The sample Zipf plot is the graph of $ln(i)$ against $ln(x_i)$. Further, because of the ranking,

$$\frac{i}{N} = 1 - F_X(x_i) \tag{1}$$

implying that,

$$ln(i) = ln[1 - F_X(x_i)] + ln(N) \tag{2}$$

If a theoretical probability model $\hat{F}_X(x)$ has been satisfactorily fitted on the random variable X, the estimated Zipf Plot can be superimposed to the empirical one by simply adding to the plot a graph of $ln\left[1-\hat{F}_X(x_i)\right]+ln(N)$ against $ln(x_i)$.

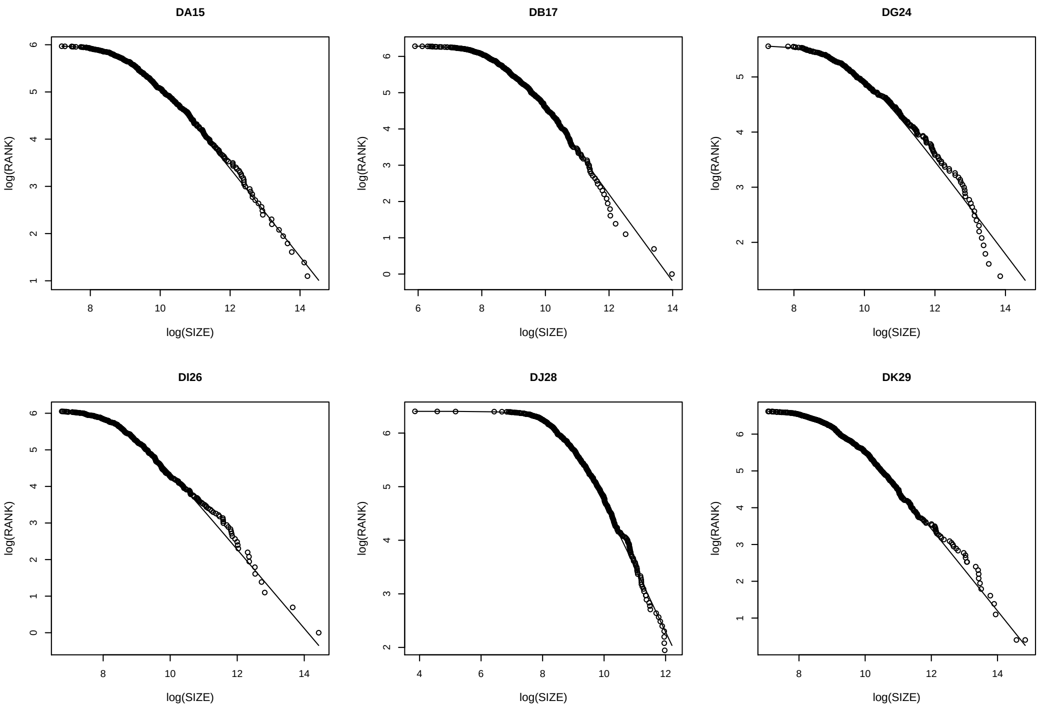

Figure 5: **Zipf Plots of the six industries. Dots represent the observed rank-size relationship, the solid line represents the relationship between the observed rank and the size estimated through the quantile function**

In Figure 5 we depict the empirical Zipf plots for the six industries together with the theoretical Zipf plot estimated under the null hypothesis of a Pareto IV distribution (details about the Maximum Likelihood estimation of the plots are provided in Crosato and Ganugi (2007)). The Pareto IV distribution is characterized by four parameters, determining respectively location

(μ), scale (σ), shape (γ and α), and is defined by

$$F_{PIV}(x) = 1 - \left[1 + \left(\frac{x-\mu}{\sigma}\right)^{\frac{1}{\gamma}}\right]^{-\alpha} \tag{3}$$

An important feature of the Pareto IV distribution is that its log-log transformation is non-linear when parameter γ differs from 1. The closer γ to zero, the stronger the curvature in the log-log chart. Still, as clearly shown by Figure 5, if in all cases the charts' left tails starkly differ from a straight line, their right tails are much closer to it. Since the linearity of the Zipf plot univocally identifies the Pareto I distribution, its presence in the right tail of the Pareto IV distribution Zipf Plot makes it clear that the same distribution satisfies the *weak Pareto law* (Mandelbrot, 1960):

$$\lim_{x\to\infty}\left[\frac{1-F_{PIV}(x)}{(x/x_{th})^{-\beta}}\right] = 1 \tag{4}$$

for $\alpha/\gamma \approx \beta$, $x_{th} \geq \mu$. This means that the Pareto IV's double log chart converges to a straight line with slope $-\alpha/\gamma$. If we set the slope of the Zipf Plot equal to $-\alpha/\gamma$:

$$\frac{dln\left[1-F_{PIV}(x)\right]}{dlnx} = -\frac{\alpha}{\gamma} \tag{5}$$

and solve for x, we can obtain the threshold x_{th} above which the Zipf Plot starts to approximate linearity. Substituting $F_{PIV}(x)$ with its expression in (3) leads to

$$-\frac{\alpha}{\gamma}\left(\frac{x-\mu}{\sigma}\right)^{\frac{1}{\gamma}-1}\left(\frac{x}{\sigma}\right) = -\frac{\alpha}{\gamma}\left[1+\left(\frac{x-\mu}{\sigma}\right)^{\frac{1}{\gamma}}\right], \tag{6}$$

therefore, provided that $\mu \neq 0$, the linearity threshold is:

$$x_{th} = \frac{\sigma^{\frac{1}{1-\gamma}} + \mu^{\frac{1}{1-\gamma}}}{\mu^{\frac{\gamma}{1-\gamma}}}. \tag{7}$$

Thus, the linearity threshold depends on all the parameters apart from α. As is easily seen by deriving expression (7) separately with respect to μ, γ and σ, x_{th} decreases for increasing values of μ and γ but increases for increasing values of σ. In particular, the threshold tends to infinity for $\mu \to 0$.

It is worthwhile to remark that the slope of the Zipf Plot, in absolute value, represents the elasticity of a reduction in the number of firms when moving to a higher size class:

$$e_s = \left| \frac{dln\left[1 - F_{PIV}\left(x\right)\right]}{dlnx} \right| = \frac{d\left[1 - F\left(x\right)\right]}{dx} \frac{x}{\left[1 - F\left(x\right)\right]} \tag{8}$$

It measures, at any size, the odds against advancing further to higher sizes in a proportionate sense (Singh and Maddala, 1976). While the Pareto I distribution has a constant e_s throughout the size range, the Pareto IV distribution's e_s increases with an increasing rate, then a decreasing rate, finally reaching the asymptote α/γ.

This point is particularly relevant because according to this interpretation, the presence of a Pareto IV distribution implies that chances to improve firms' relative positions are higher for small size firms and decrease with size until they reach constancy. This property, also known as IPFR (increasing proportionate failure rate) is considered in literature as a desirable property of distributions for economic size variables, on both the empirical and the theoretical modelling side. The IPFR was first defined (Singh and Maddala, 1976) in deriving the Singh-Maddala distribution, which is known to have good performance on income distributions (McDonald, 1984), and which is a particular case of Pareto IV with $\mu = 0$. More recently, Van Den Berg (1994, 2007) stressed the importance of IPFR distributions for empirical analyses of unemployment duration and theoretical analyses of job search and job matching.

4.2 The shape of the Zipf Plot and returns to scale

The existence of a link between the shape of firm size distribution and returns to scale has been the object of some attention in the literature. Ijiri and Simon (1977) pointed out that Gibrat's Law is consistent with a regime of constant returns to scale, at least if firms are able to reach the Minimum Efficient Scale. Later, Lucas (1978) drew on this point, building a model where firms characterized by constant returns to scale grow according to Gibrat's Law. Matters are less clear-cut when considering Zipf plots characterized by non-linearities. Engwall (1972) showed that even in socialist countries the size distribution of firms was highly skewed. This was to be interpreted, according to Hjalmarsson (1974), as a support to the view that the underlying mechanism of the concentration process was basically of a technological character. Ijiri and Simon (1964, 1977) suggested that concave (downward) deviations from the straight line in the Zipf plots can still be compatible with constant returns to scale in the presence of either strong autocorrelation of growth rates, or mergers and acquisitions. However, Vining

(1976) when simulating a Gibrat-like process augmented by autocorrelation in growth rates, found a convex, rather than concave, curvature on the Zipf plot.

Hence, although the literature does not provide any clear conclusion, the presumption exists that a concave curvature in the log-log chart may stand for *... something inherent in the very nature of size that causes a progressive decline in ... [the growth rate of a firm] as it expands its activities* (Vining, 1976, p.370). Related evidence about firm growth being negatively related to size is found in a number of papers: Evans (1987); Hall (1987); Dunne *et al.* (1988); Dunne and Hughes (1994); Hart and Oulton (1996) and, for Italy, in Audretsch *et al.* (1999). This pattern is consistent with either decreasing returns to scale stepping in, or increasing returns to scale phasing out as firms grow larger.

Here we provide empirical evidence on this topic by linking explicitly the curvature of the Pareto IV distribution with the pattern of returns to scale of the production frontier, as appraised through the FDH approach. Following the above discussion we can put to test two alternative hypotheses. The first one envisages no relationship between returns to scale and firm size distribution, while according to the second hypothesis, as firms grow larger, they are systematically characterised by different returns-to-scale regimes.

We can see here a crucial advantage of FDH inasmuch as it imposes virtually no a priori structure on the pattern of returns to scale, while DEA exhibits by construction non-decreasing returns to scale for relatively smaller observations and non-increasing returns for relatively larger observations.

Turning now to the evidence, we recall that the FDH results supported the existence of an inverse relationship between elasticity of scale and size. More precisely, a careful look at Figure 3 and Figure 4 suggests that, above a given size, elasticities align themselves on the horizontal line which characterizes constant returns to scale. This tendency is observable in all industries. On the other hand, the Zipf plots depicted in Figure 5 are concave for firms up to a given size, and then become linear. Rephrasing the second hypothesis sketched above in terms of these Zipf plots, we expect constant returns to scale for firms lying in the linear part of the plot, with smaller firms experiencing increasing returns to scale.

Our research strategy is then to ascertain whether the size threshold that divides the non-linear from the linear part of the Zipf plots can also be considered a threshold in terms of returns-to-scale regimes. If so, there would be favourable evidence for the second hypothesis. In order to implement this strategy, we take the level of total assets (TA) as a measure of size, and analyse separately the distribution of elasticities for firms on each side of the linearity threshold, which has to be fixed.

The analytical threshold, x_{th}, derived in the previous section cannot be applied to all industries and years, because of finite sample effects. Small values of μ along with large values of σ push the threshold over the empirical TA domain. Still, according to Figure 5, the Zipf Plot slope becomes approximately constant for realized TA values. Therefore, in order to determine a sensible threshold, or cut-off point, we apply a sequential procedure.

The underlying idea is to estimate sequentially the Pareto I distribution tail index (β henceforth), starting from the ten largest firms[2] and adding up

[2]This choice was driven from preliminary inspection of the Zipf Plots and does not

one firm in turn. We fix the threshold when the best linear approximation for the right tail of the Pareto IV is achieved.

The sequential procedure is a three-step one. At step one, we estimate the Pareto I model moving from the right to the left of the distribution, by using the Hill estimator (HIll, 1975). Denoting by $X_1, X_2, \ldots, X_n$ the order statistics of firm TA in each industry, such that $X_1 \geq X_2 \geq \ldots \geq X_n$, the Hill estimator of the Pareto I tail index based on the k largest observations only is given by:

$$\hat{\beta}_k = \frac{1}{k}\left[\sum_{i=1}^{k} log\left(X_i\right) - log\left(X_{k+1}\right)\right]^{-1} \tag{9}$$

At step two, the sequence of $\hat{\beta}_k$ is compared with the Pareto IV's α/γ. The level of total assets corresponding to the k which returns the smallest difference between $\hat{\beta}_k$ and Pareto IV's α/γ corresponds to the linearity threshold[3], TA_{th}. This step, for the *Food products and beverage* industry, is exemplified in Figure 6, where the sequence of $\hat{\beta}_k$ along the firms is plotted against Pareto IV's α/γ.

Finally, at step three, we test (by a Chi-Squared) the null hypothesis that the fraction of data identified at step two are distributed according to a Pareto I.

The outcomes of the procedure for the first, central and final years are

affect the results.

[3]The plot of $\hat{\beta}_k$ against k is often called the Hill plot. It is used to determine the range of k which give stable $\hat{\beta}_k$ estimates. Both the β value and the threshold k are then fixed into the interval where $\hat{\beta}_k$ reaches a plateau (Weron, 2001), according to visual inspection of the Hill plot. In this paper, however, it is crucial to determine the threshold modelling all firms at once through the Pareto IV distribution, which leads to the criterium sketched in step 3.

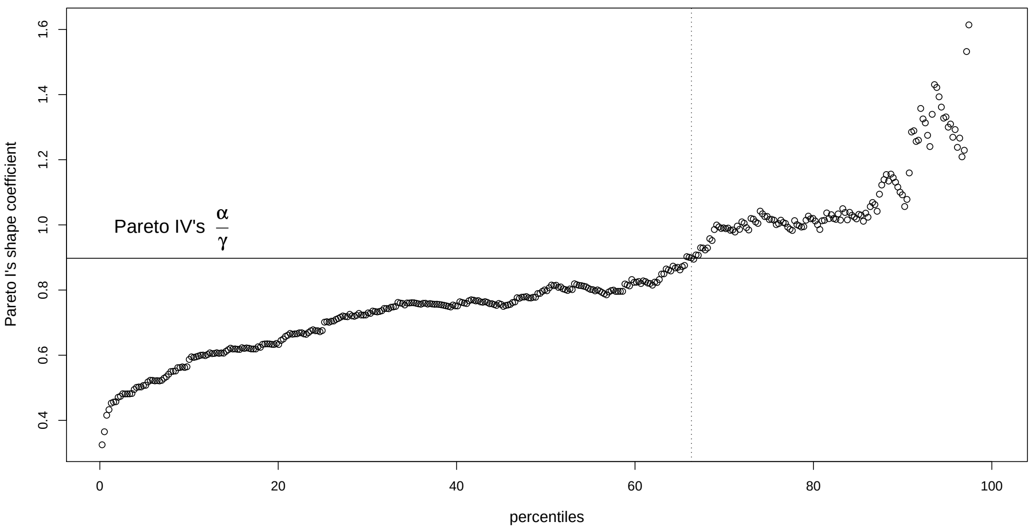

Figure 6: **Fixing of the linearity threshold (vertical line) over the sequence of estimated $\hat{\beta}_k$ (dots). Food products and beverage (DA15).**

summarized in Table 8. As can be seen, the threshold shifts both between industries and, within industries, over time. Still, thresholds in the second column indicates that the Pareto I distribution covers medium or large firms in all cases apart from industries DA15 and DG24 in 1997[4].

Comparison between the fourth and the fifth column clearly highlights the proximity between the Pareto I shape parameter and the ratio of the Pareto IV's α and γ. Finally, the p-values reported in the last column validate a satisfactory fit of the Pareto I distribution for firms with TA exceeding the cutoff point. In most cases, these p-values widely exceed the 5% level of significance[5].

After the cutoff point has been determined for each year, firms can be

[4] About the definition of medium and large firms see Table 3. Obviously the TA thresholds of Table 8 were converted in Euro millions to fit that definition.

[5] Estimates for all years are available on request

Year	TA_{th}	perc-linear	P1-$\hat{\beta}$	$P4\,(\alpha/\gamma)$	p.value
			DA15		
1989	21,581	33.4	0.899	0.897	0.692
1993	42,254	26.1	0.940	0.937	0.669
1997	119,375	13.6	1.071	1.070	0.771
			DB17		
1989	13,589	27.6	1.157	1.154	0.660
1993	15,512	30.2	1.203	1.197	0.966
1997	25,685	23.2	1.161	1.162	0.014
			DG24		
1989	32,028	34.5	0.856	0.860	0.010
1993	60,814	31.3	0.897	0.896	0.477
1997	159,192	15.6	1.118	1.109	0.078
			DI26		
1989	11,040	31.5	1.001	1.001	0.824
1993	14,342	35.4	1.004	1.020	0.481
1997	24,207	28.4	0.963	0.964	0.715
			DJ28		
1989	6,845	35.3	1.209	1.209	0.306
1993	7,895	40.3	1.108	1.107	0.103
1997	13,808	33.5	1.174	1.173	0.469
			DK29		
1989	13,389	35.2	0.951	0.948	0.099
1993	26,443	25.8	1.077	1.077	0.587
1997	39,629	23.7	1.076	1.077	0.608

Table 8: **Estimates of the Pareto I model on the linear tail of the Pareto IV model. The linearity threshold in the Pareto IV Zipf Plot is expressed both in terms of the total assets level (TA-th, million Lire) and of the percentage of firms on the right of the threshold (perc-linear). The fourth and the fifth column relate repectively to the Pareto I's β and to the ratio of Pareto IV's α and γ. The sixth column illustrates the p-values from the Chi-Squared test of the null that the data on the right of the threshold are distributed according to the Pareto I distribution**

straightforwardly classified as belonging either to the concave or the linear part of the plot. To analyse the type of returns to scale prevailing for firms on the left and on the right of the boundary, we divide the elasticities into three classes, precisely below 0.95, from 0.95 to 1.05 and above 1.05, representing decreasing, constant and increasing returns to scale, respectively (the three cases are labelled DRS, CRS and IRS). The main results for this taxonomy are reported in Table 9, where all years have been pooled together.

Industries	Zipf Plot concave part				Zipf Plot linear part			
	DRS	CRS	IRS	tot	DRS	CRS	IRS	tot
DA15	1.5%	30.0%	68.5%	100%	18.0%	43.4%	38.6%	100%
DB17	7.3%	23.1%	69.6%	100%	34.1%	45.5%	20.4%	100%
DG24	2.5%	45.7%	51.8%	100%	9.5%	75.3%	15.1%	100%
DI26	5.9%	29.4%	64.7%	100%	25.5%	47.7%	26.9%	100%
DJ28	5.8%	19.0%	75.2%	100%	29.5%	47.1%	23.4%	100%
DK29	2.9%	27.5%	69.6%	100%	18.6%	63.7%	17.7%	100%

Table 9: **Percentage of firms in concave and linear part according to returns-to-scale regime, all years**

As can be seen from Table 9, firms in the concave part clearly exhibit increasing returns to scale in all industries. On the contrary, firms staying in the linear part are mostly concentrated in the $[0.95, 1.05]$ interval, hence sharing a tendency to constant returns to scale.

Accordingly, our evidence favours the second hypothesis outlined above. The Pareto IV distribution is consistent with different returns-to-scale regimes: as firms grow larger, they tend to be characterised in turn by increasing and then constant returns to scale. Most firms in the concave part experience an IRS regime, while the linear tail mainly includes firms facing CRS in all the six industries.

An important question is whether the closed structure of our survey influences the association of the concave and convex parts with different returns-to-scale regimes. There is evidence that, in a closed sample, small survivors show higher growth than their population counterparts. As a consequence the dominance of the IRS regime in the concave part may be an artefact of sample selection. To shed evidence on this, we consider again our returns-to-scale taxonomy for the final year of the sample in Table 10. We surmise that as the smaller survivors - especially those endowed with higher growth potential - grow over the years, at least some of them shift from the IRS to the CRS regime. Hence, the evidence from the final years should be the least affected by sample selection.

Industries	Zipf Plot concave part				Zipf Plot linear part			
	DRS	CRS	IRS	tot	DRS	CRS	IRS	tot
DA15	1.8%	32.9%	65.3%	100%	13.1%	55.3%	31.6%	100%
DB17	12.1%	22.8%	65.1%	100%	17.9%	73.8%	8.3%	100%
DG24	2.1%	46.5%	51.4%	100%	20.7%	65.5%	13.8%	100%
DI26	2.5%	39.8%	57.7%	100%	3.8%	55.7%	40.5%	100%
DL28	4.6%	35.2%	60.2%	100%	6.6%	77.0%	16.4%	100%
DK29	2.8%	31.6%	65.6%	100%	19.7%	56.7%	23.6%	100%

Table 10: **Percentage of firms in concave and linear part according to returns-to-scale regime, 1997 only**

The evidence from Table 10 shows that sample selection is not likely to be the driving force behind our results. In line with our expectations, the share of IRS firms in the final year is generally lower than in Table 9 (only firms in the linear part of DI26 and DK29 except to this), but different parts of the Zipf plot are still distinctly associated with different returns-to-scale regimes. IRS firms are always the majority of firms in the concave part of the plot, while the dominance of CRS firms in the linear part is even stronger. We thus feel entitled to claim that the survivor bias does not materially affect the plot-scale elasticity nexus, which is the main focus of our analysis.

5 Concluding Remarks

In the literature on firm-size distribution, the relationship between technology and firm size has been treated mostly from a theoretical point of view. In this paper we provide empirical evidence on this topic analysing six industries from Italian manufacturing.

In previous research Crosato and Ganugi (2007) have pointed out that the size distribution of firms in Italian manufacturing industries cannot be satisfactorily modelled by means of the Lognormal and Pareto I distributions. On the contrary, a good fit is achieved through the Pareto IV distribution, a more general Paretian model. As can be easily shown through the Zipf plot, the Pareto IV model possesses a Pareto I right-hand tail, linear in double log-scale. In this paper we link explicitly the curvature of the Pareto IV distribution with the returns to scale of the production frontier, as appraised through the non-parametric Free Disposal Hull (FDH) approach, in the version refined by Kerstens and Vanden Eeckaut (1999). Utilisation of FDH is of crucial importance, since this approach imposes very little a priori structure on the pattern of returns to scale.

We find an inverse relationship between scale elasticity and firm size. Building on this result, we investigate the connection between elasticities and the firm size distribution shape. Once determined the linearity threshold, we find that firms occupying the concave part of the Pareto IV generally exhibit increasing returns to scale in all industries. On the contrary, firms in the linear Pareto I-type part clearly show constant returns to scale. The presence of two different technological regimes is consistent with the existence of different regimes of growth along the Pareto IV distribution. Furthermore, since both constant returns to scale and linearity of the Zipf plot are compatible with Gibrat's Law, their joint presence in the right tail of the size distribution supports the validity of the Law for larger firms of Italian manufacturing (Lotti *et al.*, 2004).

A possible problem with our evidence is that we rely on a closed panel, exposed in principle to survivor bias. A very rich literature - among others Evans, 1987; Hall, 1987; Dunne et al., 1989; Dunne and Hughes, 1994; Hart and Oulton, 1996; Audretsch et al., 1999 - suggests however that the concave part of the Zipf plot is not simply an artefact of sample selection bias. Similarly, according to our evidence, sample selection is not the driving force behind the association of the concave and linear parts of the plot with increasing and constant returns to scale.

We believe that our results are of some interest and could shed novel light, among other things, on the relationships between Gibrat's Law and the Minimum Efficient Scale, the output level above which production is characterised by constant long-run average costs (Bain, 1956; Lyons, 1980). In future work we intend to extend the study of the convergence of the Pareto IV to the Pareto I distribution to other datasets, analysing why the linearity threshold changes across industries and time. Interestingly, some recent contributions have found that concavity of the Zipf plot is affected by the innovation process. According to Marsili (2005), concavity is linked to the cumulative nature of innovation and to the dominance of product upon process innovation. Other hypotheses to be considered in this ambit are the influence of the age of the firm on its growth regime and of company groups or conglomerates on the pattern of returns to scale.

References

Alvarez, R. and Crespi, G. (2003). 'Determinants of Technical Efficiency in Small Firms', *Small Business Economics*, 20, 3, 233–244.

Arnold, B. (1983). *Pareto Distributions*, International Co-operative Publishing House, Fairland, ME.

Audretsch, D. and Santarelli, E. and Vivarelli, M. (1999). 'Start-up Size and Industrial Dynamics: some Evidence from Italian Manufacturing', *International Journal of Industrial Organization*, 17, 965–983.

Bain, J.S. (1956). *Barriers to New Competition: Their Character and Consequences in Manufacturing Industries*, Harvard University Press.

Balloni, V. (1984). 'Misure e spiegazioni dell'efficienza del sistema industriale: considerazioni di carattere metodologico', in: *Produttività e competitività nell'industria italiana, AA.VV.*, il Mulino, Bologna.

Baltagi, B.H. and Griffin, J.M. (1984). 'Short and Long Run Effects in Pooled Models', *International Economic Review*, 25, 3, 631–645.

Basu, S. and Fernald, J.G. (1997). 'Returns to Scale in US Production: Estimates and Implications', *Journal of Political Economy*, 249–283.

Bottazzi,G. and Grazzi, M. and Secchi, A. (2005). 'Characterizing the Production Process: a Disaggregated Analysis of Italian Manufacturing Firms', *Rivista di Politica Economica*, I-II, 243–270.

Cefis, E. and Dosi, G. and Bottazzi, G. and and Secchi, A. (2003). 'Invariances and Diversities in the Evolution of Manufacturing Industries', Technical Report 03-17, Utrecht School of Economics, discussion Paper.

Clementi, F. and Gallegati, M. (2005). 'Pareto's Law of Income Distribution: Evidence for Germany, the United Kingdom, and the United States', http://www.citebase.org/abstract?id=oai:arXiv.org:physics/0504217.

Crosato, L. and Ganugi, P. (2007). 'Statistical Regularity of Firm Size Distribution: the Pareto IV and Truncated Yule for Italian SCI Manufacturing', *Statistical Methods and Applications*, 16, 1, 83–115.

Dagum, C. (1977). 'A New Model of Personal Income Distribution: Specification and Estimation', *Economie Appliquée*, 30, 413–437.

Deprins, D. and Simar, L. and Tulkens, H. (1984). 'Measuring labor-efficiency in post offices', in: *The performance of public enterprises: concepts and measurement*, Marchand, M. and Pestieau, P. and Tulkens, H., eds., North Holland, Amsterdam, 243–267.

Destefanis, S. and Sena, V. (2007). 'Patterns of Corporate Governance and Technical Efficiency in Italian Manufacturing', *Managerial and Decision Economics*, 28, 1, 27.

Dunne, P. and Hughes, A. (1994). 'Age, Size, Growth and Survival: UK Companies in the 1980s', *The Journal of Industrial Economics*, 42, 2, 115–140.

Dunne, T. and Roberts, M. and Samuelson,L. (1988). 'Patterns of Firm Entry and Exit in US Manufacturing Industries', *Rand Journal of Economics*, 19, 495–515.

Engwall, L. (1972). 'Inequality of Firm Sizes in Different Economic Systems', *Zeitschrift für Nationalökonomie*, 32, 4, 449–460.

Evans, D.S. (1987). 'Tests of Alternative Theories of Firm Growth', *The Journal of Political Economy*, 95, 4, 657–674.

Farrell, M.J. (1957). 'The Measurement of Productive Efficiency', *Journal of the Royal Statistical Society, Series A*, 70, 253–281.

Førsund, F. (1996). 'On the Calculation of Scale Elasticity in DEA Models', *Journal of Productivity Analysis*, 7, 283–302.

Førsund, F. and Hjalmarsson, L. (1979). 'Generalised Farrell Measures of Efficiency: An Application to Milk Processing in Swedish Dairy Plants', *Economic Journal*, 89, 294–315.

Fox, K.J. and Hill, R.J. and Diewert, W.E. (2004). 'Identifying Outliers in Multi-Output Models', *Journal of Productivity Analysis*, 22, 1, 73–94.

Fried, H. and Lovell, C.A.K. and Schmidt, S.S. (Eds.) (1993). *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford University Press, London.

Gibrat, R. (1931). *Les Inégalités Economiques*, Librairie du Recueil Survey, Paris.

Gijbels, I. and Mammen, E. and Park, B.U. and Simar, L. (1999). 'On estimation of monotone and concave frontier functions', *Journal of the American Statistical Association*, 94, 220–228.

Hall, B.H. (1987). 'The relationship between firm size and growth in the U.S. Manufacturing sector', *Journal of Industrial Economics*, 35, 4, 583–606.

Hart, P.E. and Oulton, N. (1996). 'Growth and Size of Firms', *The Economic Journal*, 106, 438, 1242–1252.

Hart, P.E. and Oulton, N. (1997). 'Zipf and the Size Distribution of Firms', *Applied Economics Letters*, 4, 4, 205–206.

HIll, B.M. (1975). 'A Simple General Approach to Inference About the Tail of a Distribution', *The Annals of Statistics*, 3, 1163–1174.

Hjalmarsson, L. (1974). 'The size distribution of establishments and firms derived from an optimal process of capacity expansion', *European Economic Review*, 5, 123–140.

Ijiri, Y. and Simon, H.A. (1964). 'Interpretations of departures from the Pareto curve firm-size distributions', *Journal of Political Economy*, 82, 315–331.

Ijiri, Y. and Simon, H.A. (1977). *Skew Distributions and the Sizes of Business Firms*, North-Holland Publishing Company, Amsterdam-New York-Oxford.

Kalecki, M. (1945). 'On the Gibrat distribution', *Econometrica*, 13, 161–170.

Kerstens, K. and Vanden Eeckaut, P. (1999). 'Estimating returns to scale using non-parametric deterministic technologies: a new method based on goodness-of-fit', *European Journal of Operational Research*, 113, 206–214.

Kittelsen, S.A.C. (1999). 'Monte Carlo simulations of DEA efficiency measures and hypothesis tests', Memorandum 9, University of Oslo, Department of Economics, Norway.

Kneip, A. and Park, B.U. and Simar, L. (1998). 'A note on the convergence of nonparametric DEA estimators for production efficiency scores', *Econometric Theory*, 14, 783–793.

Kuh, E. (1959). 'The Validity of Cross-Sectionally Estimated Behavior Equations in Time Series Applications', *Econometrica*, 27, 2, 197–214.

Lotti, F. and Santarelli, E. and Vivarelli, M. (2004). 'Gibrat's Law and Market Selection', Discussion Papers on Entrepreneurship, Growth and Public Policy 28, Max Planck Institute of Economics.

Lucas, R.E. (1978). 'On the Size Distribution of Business Firms', *Bell Journal of Economics*, 9, 508–523.

Lyons, B. (1980). 'A New Measure of Minimum Efficient Plant Size in UK Manufacturing Industry', *Economica*, 47, 185, 19–34.

Mandelbrot, B. (1960). 'The pareto-levy law and the distribution of Income', *International Economic Review*, 1, 79–106.

Mansfield, E. (1962). 'Entry, Gibrat's Law, Innovation, and the Growth of Firms', *American Economic Review*, 52, 5, 1023–1051.

Marsili, O. (2005). ‘Technology and the Size Distribution of Firms: Evidence from Dutch Manufacturing’, *Review of Industrial Organization*, 27, 4, 303–328.

McDonald, J.B. (1984). ‘Some Generalized Functions for the Size Distribution of Income’, *Econometrica*, 52, 3, 647–664.

Ofria, F. (1997). *Divari regionali di produttività nell’industria manifatturiera italiana: cause ambientali e aziendali*, FrancoAngeli.

Okuyama, K. and Takayasu, M. and Takayasu, H. (1999). ‘Zipf’s Law in Income Distribution of Companies’, *Physica A*, 269, 125–131.

Prosperetti, L. and Varetto, F. (1991). *I differenziali di produttività nord-sud nel settore manifatturiero: Un’analisi microeconomica*, Il Mulino.

Simon, H.A. (1955). ‘On a Class of Skew Distribution Functions’, *Biometrika*, 52, 227–230.

Simon, H.A. and Bonini, C.P. (1958). ‘The Size Distribution of Business Firms’, *American Economic Review*, 48, 607–617.

Singh, S.K. and Maddala, G.S. (1976). ‘A Function for Size Distribution of Incomes’, *Econometrica*, 44, 963–970.

Stanley, M.H.R. and Buldyrev, S.V. and Havlin, S. and Mantegna, R.N. and Salinger, M.A. and Stanley, H.E. (1995). ‘Zipf Plots and the Size Distribution of Firms’, *Economics Letters*, 49, 453–457.

Steindl, J. (1965). *Random Processes and the Growth of Firms*, Griffin, London.

Stoppa, G. (1990). 'A New Model for Income Size Distributions', in: *Income and Wealth Distribution, Inequality and Poverty: Proceedings of the 2nd International Conference on Income Distribution by Size: Generation, Distribution, Measurement and Applications*, Dagum C. and Zenga M., eds., Springer, New York, Berlin, London and Tokyo, 33–41.

Taymaz, E. (2005). 'Are Small Firms Really Less Productive?', *Small Business Economics*, 25, 5, 429–445.

Tulkens, H. (1993). 'On FDH efficiency analysis: some methodological issues and applications to retail banking, courts and urban transit', *Journal of Productivity Analysis*, 4, 183–210.

Van Den Berg, G.J. (1994). 'The Effects of Changes of the Job Offer Arrival Rate on the Duration of Unemployment', *Journal of Labor Economics*, 12, 3, 478–498.

Van Den Berg, G.J. (2007). 'On the uniqueness of optimal prices set by monopolistic sellers', *Journal of Econometrics*, 141, 482–491.

Vining, D.R. (1976). 'Autocorrelated Growth Rates and the Pareto Law: a Further Analysis', *Journal of Political Economy*, 84, 369–380.

Voit, J. (2001). 'The Growth Dynamics of German Business Firms', *Advances in Complex Systems*, 4, 1, 149–162.

Weron, R. (2001). 'Levy-stable distributions revisited: tail index > 2 does not exclude the Levy-stable regime', *International Journal of Modern Physics C*, 209–223.