Corporate Growth and Industrial Structure. Some Evidence from the Italian Manufacturing Industry*

Giulio Bottazzi[†] Elena Cefis[††] Giovanni Dosi[†]

[†] S.Anna School of Advanced Studies, Pisa, Italy

[††] Dept. of Economics, University of Bergamo, Italy

DRAFT VERSION

May 11, 2001

1 Introduction

In this work we report preliminary results of an investigation on industrial dynamics based on a decade of micro longitudinal data from four Italian industries — pharmaceuticals, primary metals, machine tools and textile — chosen as representative of quite diverse production technologies and learning modes. Here we begin addressing two sets of issues concerning (i) the shape of the size distributions and their possible inter-sectoral differences, and (ii) the characteristics of growth dynamics.

A classic reference, when dealing with the statistical properties of firm growth, is the so-called "Law of Proportionate Effect" (or "Gibrat Law") (Gibrat, 1931) entailing process of stochastic growth uncorrelated with size and basically driven by several small idiosyncratic events[1]. It represents a sort of "null hypothesis" regarding the dynamics. And it is an hypothesis that makes evolutionary economists rather uncomfortable, in that it seems at odds

*This work is a part of an ongoing research supported by the Italian Ministry of University and Research (MURST) and by the S.Anna School. It has been accomplished with the precious collaboration of the Italian Statistical Office (ISTAT) and in particular of Roberto Monducci, whose competence and dedication made the idea of a joint research enterprise possible.

[1]For discussions, following the pioneering Ijiri and Simon (1977), cf., among others, Brock and Evans (1986), Boeri (1989), Sutton (1997), Geroski (2000), Dosi et al. (1995),Marsili (2001).

both with several pieces of microeconomics evidence highlighting long-standing differences in technological and organizational competences across firms, and also with a notion of a competitive process systematically selecting within such population of heterogeneous firms[2].

While the overall evidence on "Gibrat Law" from the literature, often based on not-too-good data, is rather mixed (Dosi et al., 1995; Sutton, 1997), in Bottazzi et al. (2001) one finds, in the case of a panel including the world top pharmaceuticals firms, a rather rich structure in the statistical properties of the growth process, displaying, among others features, **(a)** "fat tails" in the distribution of growth shocks with (relatively rare) "spurs of growth" and **(b)** a significant autocorrelation of growth rates over time. An obvious issue regards the generality of such findings, which we shall indeed address in the following. Are the foregoing properties dependent upon the particular features of learning and competition of the drugs industry or, conversely, are they rather general characteristics of industrial dynamics? And, even if the latter hypothesis held true, to what extent are such characteristics influenced by industry-specific factors?

Moreover, size as such might not be the best variables upon which to condition growth events. Rather, it is much more in tune with an evolutionary idea of heterogeneity *cum* market selection[3] to search for proxies of relative degrees of firm "competitiveness" and investigate their impact on firm growth profiles. This is what we shall also do below,using labor production as proxies for production efficiencies.

In Sec. 2 we briefly describe the database and the variables under scrutiny. In Sec. 3 we discuss the evidence on size distributions, the distribution of growth shocks and their possible autocorrelation. Finally, in Sec. 4, we analyze the relationship between relative labor productivities and growth profiles.

2 The Database

This research draws upon the MICRO.1 databank developed by the Italian Statistical Office (ISTAT)[4]. MICRO.1 contains longitudinal data for a panel of several thousand Italian firms

[2]Incidentally note also that violations of Gibrat-type process of growth based on i.i.d. shocks are also implied by equilibrium models of industrial dynamics such as Jovanovic (1982) and Ericson and Pakes (1995): cf. Pakes and Ericson (1998)

[3]The evolutionary literature on industrial dynamics is rapidly expanding. Here, however, let us just mention two of the "seeding classics", namely Winter (1971) and Nelson and Winter (1982)

[4]The database has been made available to our team under the mandatory condition of censorship of any individual information.

with employment of 20 units or more over around a decade, of which for statistical consistency we utilize the period 1989-96.

In this work we are exclusively interested in the process of *internal* growth, as opposed to the growth due to mergers, acquisitions and divestments. In order to control for the latter we build "super-firms" which account throughout the period for the union of the entities which undertake such changes. So, for example, if two firms merged at some time, we consider them merged throughout the whole period. Conversely if a firm is spun off from another one, we "re-merge" them starting from the separation period.

Moreover, since the panel is open, due to entry, exit, fluctuations around the 20 employees threshold and variability in response rates we consider only the firms that are present both at the beginning and at the end of our window of observation.

Firms are classified according to their sector of principal activity[5]. For the analysis that follows, as already mentioned, we have chosen pharmaceuticals[6],primary metals[7], machine tools[8] and textiles[9] which can be reasonably taken as representative of the Pavitt's taxonomic classes identified as "science-based", "scale-intensive", "specialized supply" and "supplier dominated", respectively(Pavitt, 1984). In turn, Pavitt's categories represent an early attempt to classify different industrial sectors according to the diverse modes of generation and exploitation of novel opportunities of product and process innovation (cf. also Dosi (1988) and Marsili (2001)).

The original statistical variables we consider here are the total number of employees $L_i(t)$, sales $S_i(t)$ and value added $V_i(t)$ of "super-firm" i at time $t \in [1, \ldots, 8]$, together with labor productivity defined as $\Pi_i(t) = V_i(t)/L_i(t)$.

As discussed in Bottazzi et al. (2001), it is often convenient to analyze the normalized logarithm of those variables. For instance, regarding the number of employees we define

$$l_i(t) = \log(L_i(t)) - < \log(L_i(t)) >_i \tag{1}$$

where $< . >_i$ stands for the average over all the firms at a given time. Analogously we define "rescaled" log sales $s_i(t)$ and value added $v_i(t)$. These variables are characterized by stationary distributions[10] and allow us to treat the growth process as a stationary one. Let us denote

[5]the Italian ATECO.3 classification closely matches the ISIC one.

[6]Ateco.3: 24.4 Pharmaceuticals; 97 observations

[7]Ateco.3: 27.11 Ferrous and Non-ferrous Metals; 67 observations

[8]Ateco.3: 29.4 Machine Tools; 114 observations

[9]Ateco.3: 17.2 Textiles; 171 observations

[10]We have checked the stationarity hypothesis using Kolmogorov-Smirnov tests and we find robust evidence supporting it: the significance is always greater than .96.

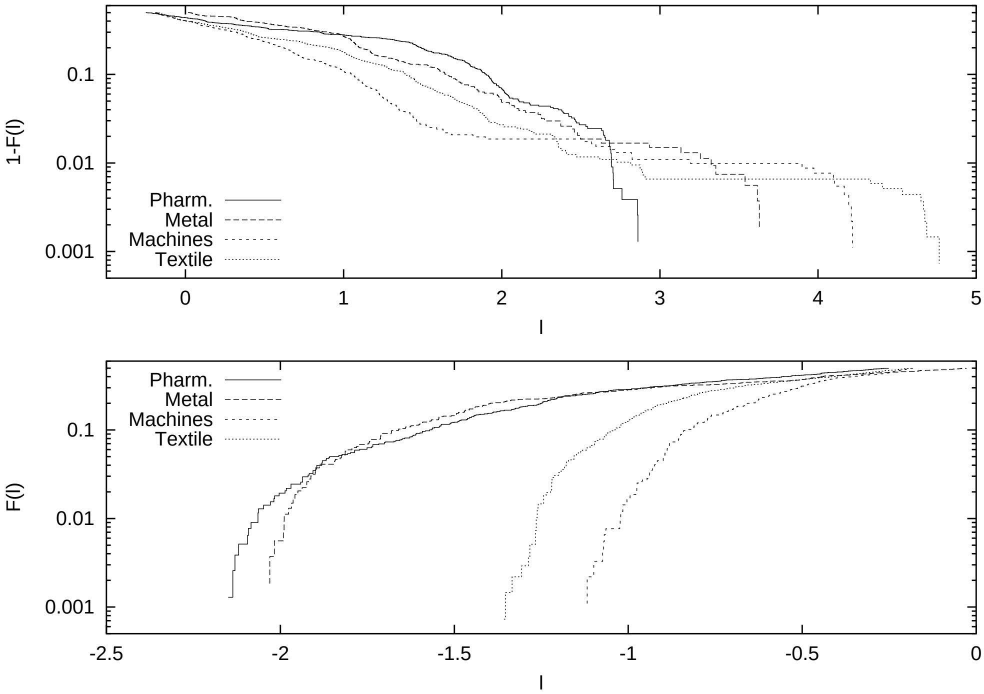

Figure 1: Upper (top) and lower (bottom) tails of the size distribution function in term of number of employees in the four sector (computed using the whole database time horizon).

the various growth shocks as

$$g_i^x(t) = x_i(t+1) - x_i(t) \tag{2}$$

where x takes the values l,s and v.

Note that through this "rescaling" procedure one washes away common trend effects due to both inflationary dynamics and real (i.e. constant price) expansion/contraction of the industry.

3 Size Distribution and Corporate Growth

Size distribution

Due to the relative low number of observations is safer to plot the distribution function rather then the probability densities. It is also handy to refer to the a "symmetric transformation" of such a distribution function. If $F(x)$ is the distribution function of the variable x, its "symmetrized" version is defined according to

$$F_s(x) = \begin{cases} F(x) & F(x) < .5 \\ 1 - F(x) & F(x) > .5 \end{cases} \tag{3}$$

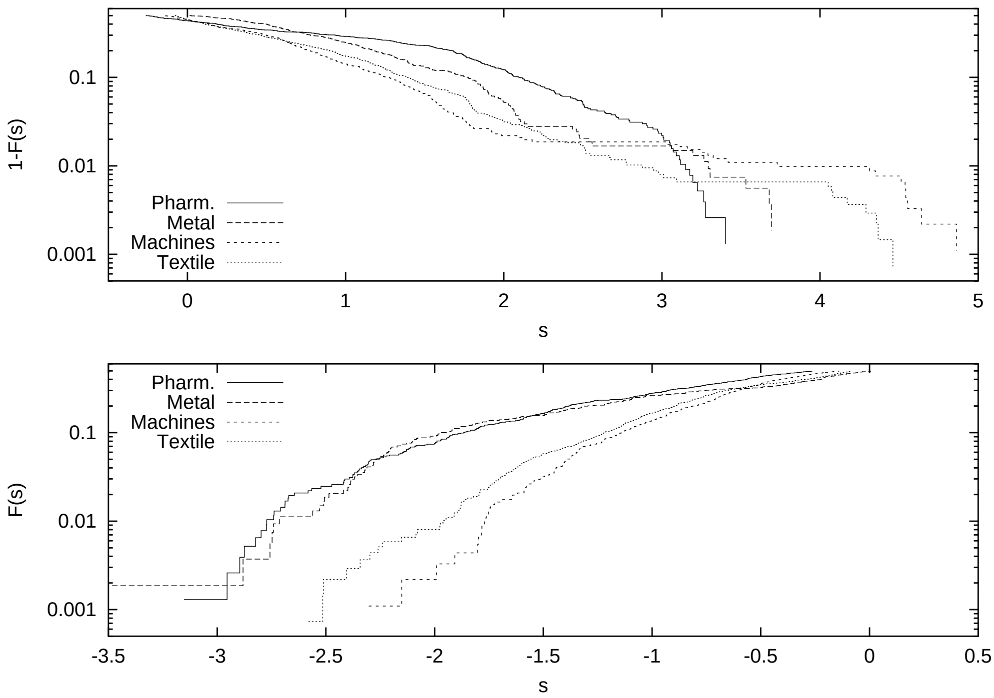

Figure 2: Upper (top) and lower (bottom) tails of the size distribution function in term of number of sales in the four sector (computed using the whole database time horizon).

under this convention, in what follows we drop by convenience the superscript d.

In Fig. 1 and Fig. 2 we show the upper and lower tails of the ("symmetrized") probability density. They display important difference both in tails and supports. The latter variable suggest different "spreads" of sizes in different industries. Concerning the former, notice that higher tails means more concentrated industries, i.e. the share of the total industry possessed by a the top fraction of firm population is higher when the probability density is asymptotically higher.

The upper tails also show some large gap which can be intuitively interpreted as a sort of "barrier" separating different segments of the industry,i.e. a core part from a fringe one. Indeed the large width of these gaps, compared to the average size of growth shocks, implies that the large majority of micro-dynamics develops separately inside the different segments, with rare events of crossing.

The lower tails appear to be more homogeneous but we are also less confident about any inference on a tail "artificially bent" by a statistical threshold and burdened by proportionally more noisy response rates.

The slope of the upper tails for all sectors tend to fall for the middle-to-high size range so that the curve takes a convex shape (this is in fact analogous to what happens on Italian

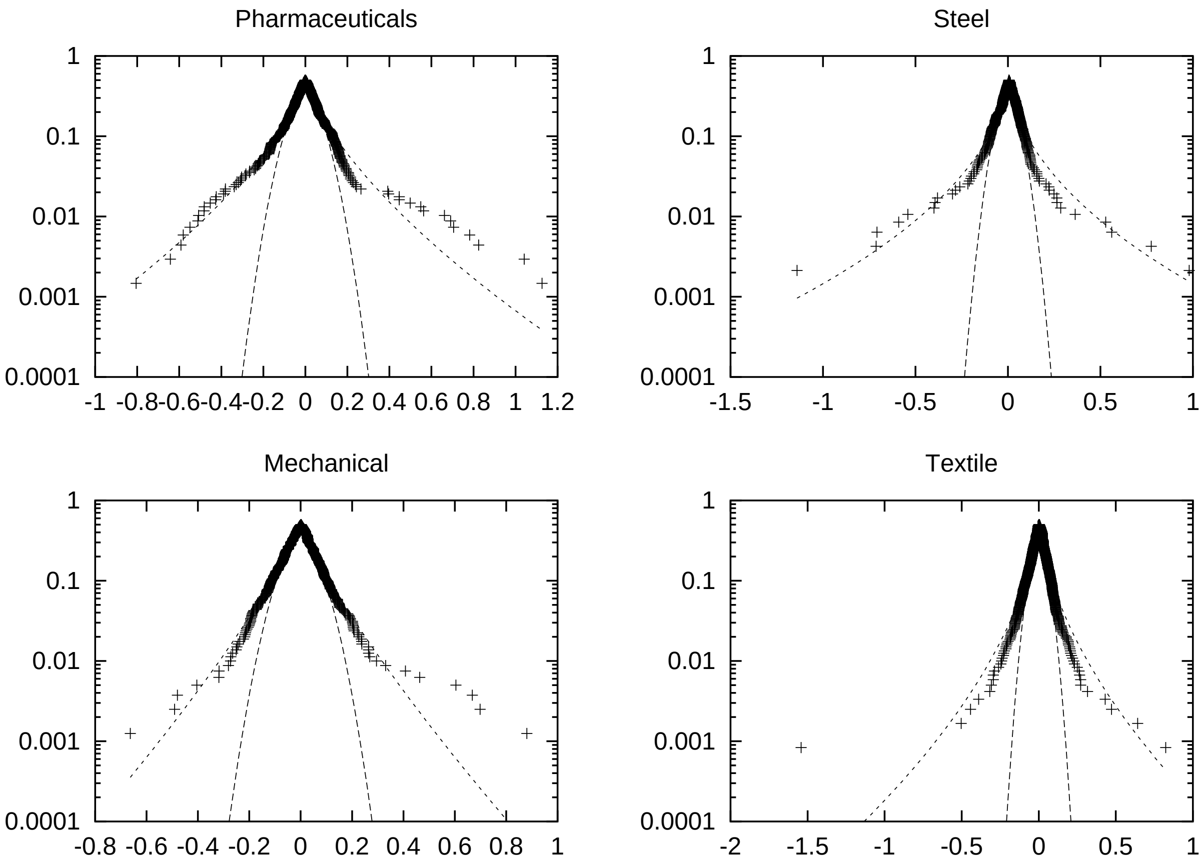

Figure 3: Probability densities for the (log) labor growth g^l in the four sectors. A normal (lower) and a generalized exponential (upper) fit are also shown. (For the parameter of the latter see Tab. 1)

data in Pareto fit to the top firms[11]) and the power-like behavior seems to be interrupted by a sudden decrease.

All this regards the general shape of all size distributions (with important "caveats" concerning the institutional specificities of the Italian case which cannot be addressed here at any details). However, note also that the rather robust inter-sectoral differences in the distributions themselves, which might come, plausibly, also from sectoral specificities, are likely to hold across different countries.

Interestingly, the representative of "science based" and "scale intensive" sectors (pharmaceuticals and primary metals respectively) are more concentrated and display smaller support, i.e. relatively lower size asymmetries. The converse hold for "specialized suppliers" and "suppliers dominated" representative (machine tools and textiles).

[11] cf. Dosi et al. (2000)

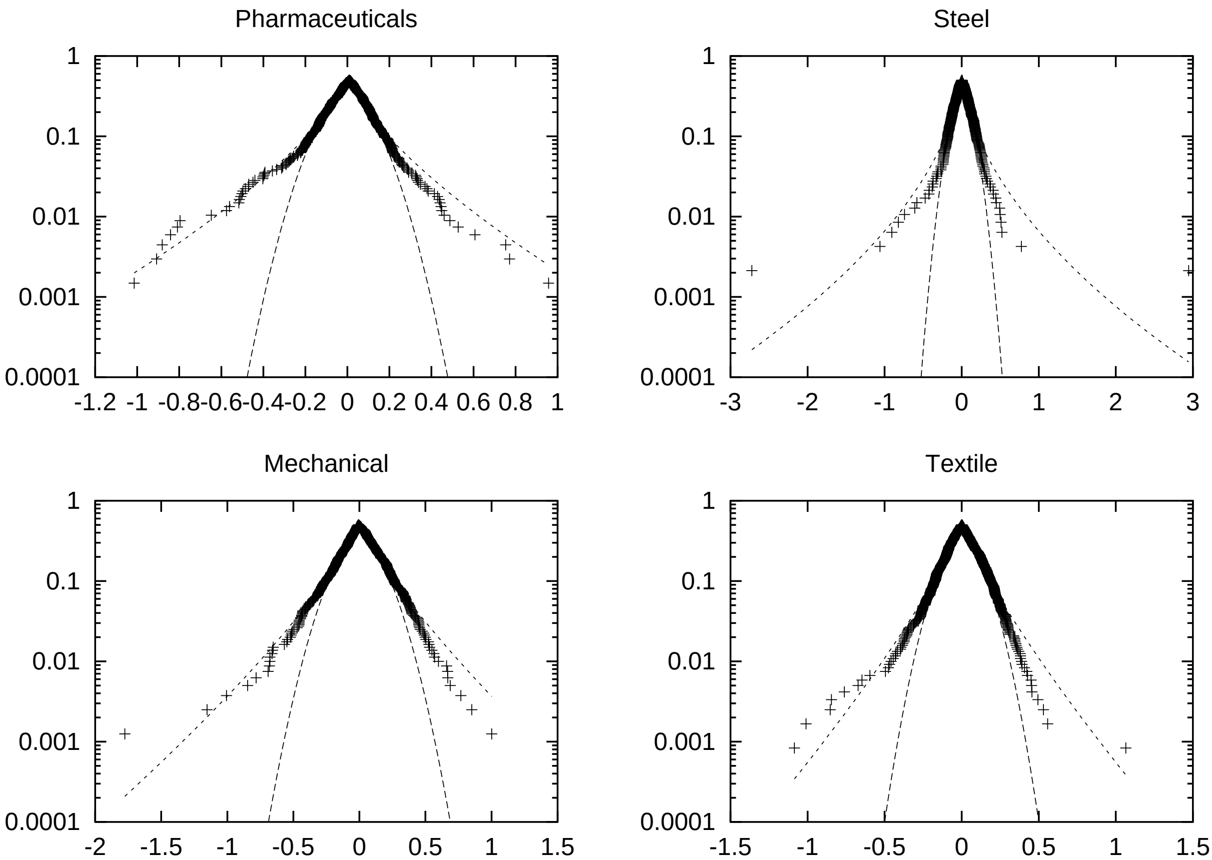

Figure 4: Probability densities for the (log) sales growth g^s in the four sectors. Again a normal (lower) and a generalized exponential (upper) fit are also shown. (For the parameter of the latter see Tab. 2)

Growth dynamics

In order to characterize the growth process, we began by checking if any relationship between size and growth were present in our data. Interestingly, both the growth means and growth variances do not display any relationship with size.[12] This suggest that the weaker form of the "Law of Proportionate Effect", prescribing the lack of any relationship between growth and size, is at work here.

However, consider as an initial benchmark for the dynamics a "stronger" Gibrat hypothesis, whereby growth shocks should be well described by a lognormal distribution[13] and compare

[12]The lack of relationship concerning growth means is a robust result which has been found many times elsewhere (cf. the evidence discussed in Sutton (1997) and Geroski (2000)). However, the presence of a negative relationship between sales and growth constitutes a quite typical feature of industrial data.

[13]This is indeed a straightforward assumption, under the Central Limit Theorem, once the idea of growth as a sequence of random shocks is accepted on every time scale.

	Pharm.	Steel	Mech.	Textile
α	8.27	9.03	10.99	10.37
β	0.58	0.39	0.77	0.49
σ^2	0.0234	0.02461	0.01161	0.01039

Table 1: Parameters from least squares fitting of the g_l distribution by a generalized exponential function.

	Pharm.	Steel	Mech.	Textile
α	6.83	6.62	5.7	7.55
β	0.62	0.46	0.73	0.76
σ^2	0.0429	0.0743	0.0658	0.0307

Table 2: Parameters from least squares fitting of the g^s distribution by a generalized exponential function.

it with the actual distribution of g^l and g^s shown in Fig. 3 and Fig. 4, respectively.

The plots clearly show how a lognormal distribution dramatically underestimates the "fatness" of the observed tails.

Let us try then to fit the data using a more fat-tailed distribution. In order to do that we use a "generalized" symmetric exponential distribution with density of the form

$$f(x) = \frac{1}{2} \frac{\beta\, \alpha^{1/\beta}}{\Gamma(1/\beta)}\; e^{-\alpha\, |x|^{\beta}} \tag{4}$$

which in turn generalizes on a similar procedure used by Stanley et al. (1997). Here $\Gamma(x)$ is the Gamma function and β represents a "tail parameter" shaping the distribution tails. The lower is β, the fatter are the tails. The "size" parameter α describes the central width of the distribution. The $2l$-th central moment of the generalized exponential reads

$$m_{2l} = \alpha^{\frac{-2l}{\beta}}\, \frac{\Gamma((2l+1)/\beta)}{\Gamma(1/\beta)} \tag{5}$$

and shows that the rescaled central moments (such as the kurtosis) do not depend on the parameter α. For $\beta = 1$ this distribution reduces to a symmetric exponential distribution.

To fit observed data we use the associated probability function:

$$F(x) = \begin{cases} \frac{1}{2}\left(1 - P(1/\beta, \alpha|x|^{\beta})\right) & x < 0 \\ \frac{1}{2}\left(1 + P(1/\beta, \alpha x^{\beta})\right) & x > 0 \end{cases} \tag{6}$$

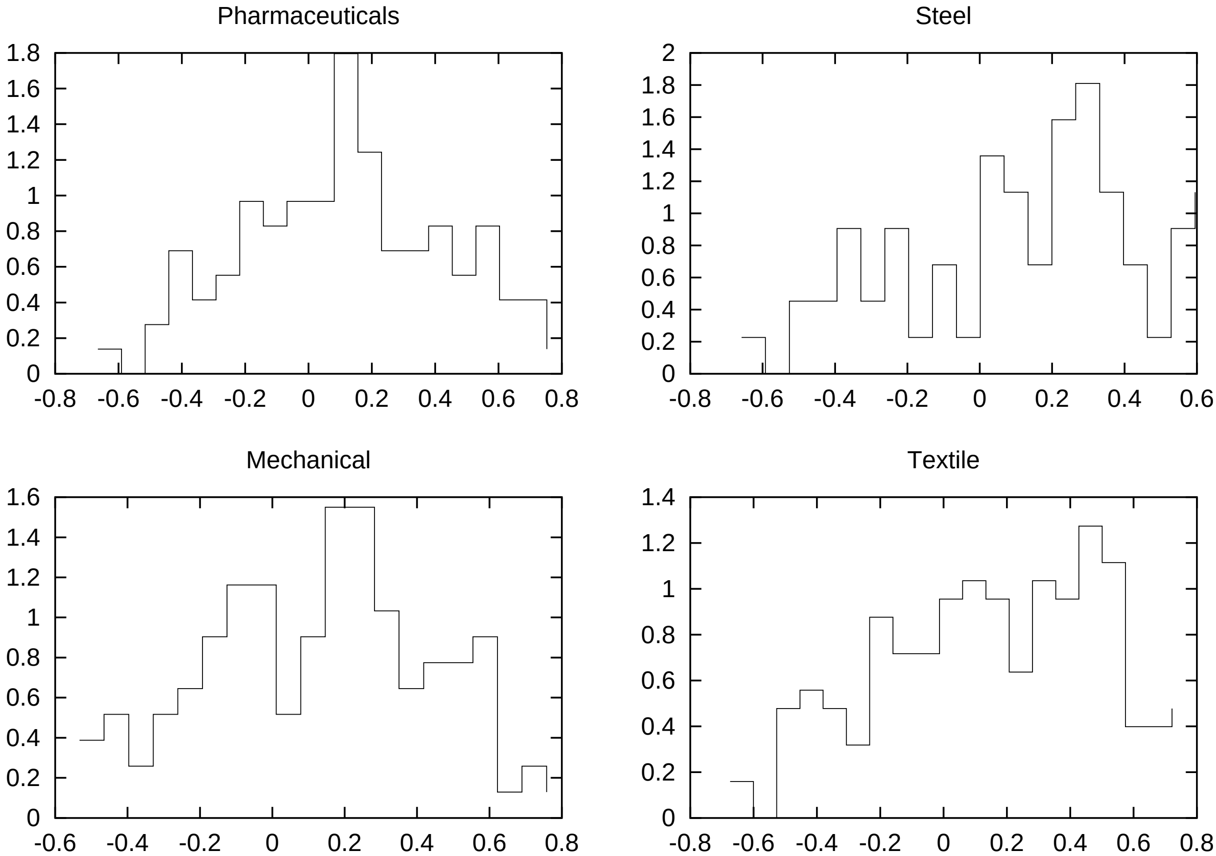

Figure 5: Observed frequency for the autocorrelation coefficient of the g^{l}'s growth. See Tab. 3 for the mean values and the results of a Kolmogorov-Smirnov comparison with bootstrapped values.

where $P(a, x)$ stands for the incomplete Gamma function:

$$P(a,x) = \frac{1}{\Gamma(a)} \int_0^x e^{-t}\, t^{a-1}\, dt \tag{7}$$

In general the distribution (6) provides a good description of the observed frequencies on a wide range of values. Interestingly, the major drawback comes from a remarkable asymmetry of the growth distribution between positive and negative parts, both for the sales and the employees variables, in some of our sectors.

In Tab. 1 and Tab. 2 we report the result of a OLS fitting procedure of (6) on the observed frequencies. As one sees, they display rather different values of β across sectors, and between employment vs. sales growth within sectors. Moreover, they show impressively different "scales" for the growth shocks (as captured by parameter α).

Begin by noticing that growth appears to be in general more "lumpy" in terms of employees rather than sales, as one can easily see comparing the β values of Tab. 1 with those of Tab. 2.

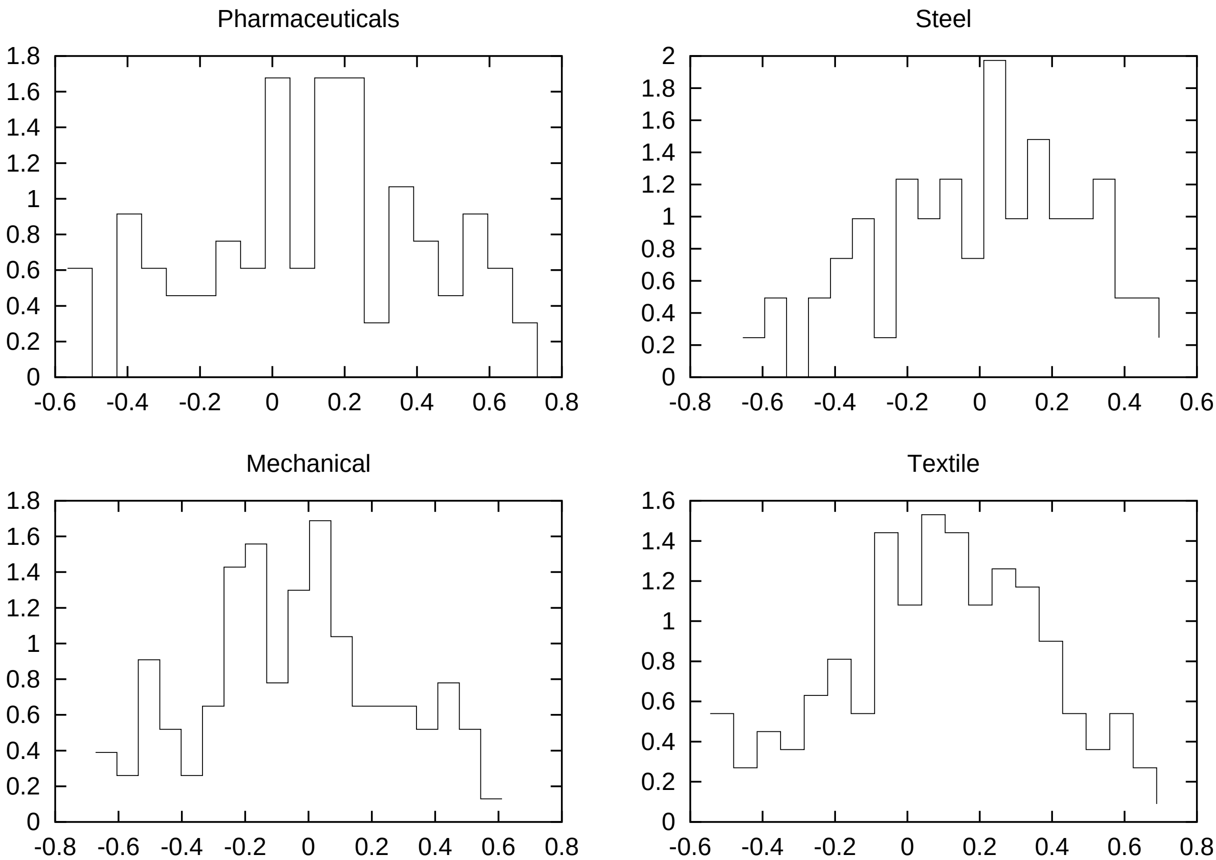

Figure 6: Observed frequency for the autocorrelation coefficient of g^s's growth. See Tab. 4 for the mean values and the results of a Kolmogorov-Smirnov comparison with bootstrapped values.

It is impossible to asses, on the grounds of our data, the extent to which such a "lumpiness" is due to institutional features of the Italian labor market. Certainly, however, there is a strong technological component, where metal and textiles stand out as the sector with fatter tails (expecially on the positive side).

More generally, concerning the asymmetries between positive and negative shocks, note that *positive* tails tend to be relatively fatter in terms of employments while *negative* tails are fatter in terms of sales (suggesting the possibility of rather large "competitive disasters").

Autocorrelation

Computing the autocorrelations coefficient for all the firms in a given sector for a given variable, one obtains a rather wide distribution. The average of this distribution, under the assumption that different firm histories represent different realizations of the same random process, should

	Pharm.	Steel	Mech.	Textile
mean	0.0789	0.093	0.095	0.123
σ	0.320	0.327	0.306	0.351
significance (P val.)	0.0085	0.00034	9.511010^{-05}	4.6710^{-08}

Table 3: Mean and standard deviation of the observed autocorrelations coefficients in labor growth for the four sectors. The significance of the Kolmogorov-Smirnov comparison test between the observed distributions and the distributions obtained with randomly resampled (bootstrapped) growth shocks is also shown.

	Pharm.	Steel	Mech.	Textile
mean	0.085	-0.016	-0.066	0.124
σ	0.327	0.284	0.305	0.351
significance (P val.)	0.0042	0.815	0.029	4.6710^{-8}

Table 4: Mean and standard deviation of the observed autocorrelations coefficients in sales growth for the four sectors. (cf. Tab. 3)

be the best estimate of the process autocorrelation. In fact, it turns out to be relatively close to zero (the mean and standard deviation of the autocorrelation frequencies distribution are reported in Tab. 3 and Tab. 4, for the employees and sales variables respectively.) Moreover, even if their difference from zero are, in some sectors for some variables, marginally significant, their small value (around .01) should not produce any perceivable effect on the short time horizon characterizing our database.

However here the assumption of "identity" amongst firms appears rather questionable. One can indeed compare the observed frequencies distributions, shown in Fig. 5 and Fig. 6 for the labor and sales variables respectively, with the ones obtained from a dataset made of "artificial" firms, obtained randomly extracting "growth histories" from the set of all the observed growth shocks. If one perform such a "bootstrap" sampling and compute again on these "artificial" datasets the autocorrelation distribution, a different shape is obtained. One can quantify this difference by performing a Kolmogorov-Smirnov test comparison between the "artificial" and the observed distributions, and look at the obtained significance p-value (i.e. the probability the observed difference between the distribution is simply a matter of chance). As can be seen in Tab. 3 and Tab. 4, the p-value is in many cases so low to lead to a clear rejection of the "identity" hypothesis between the growth processes of different firms.

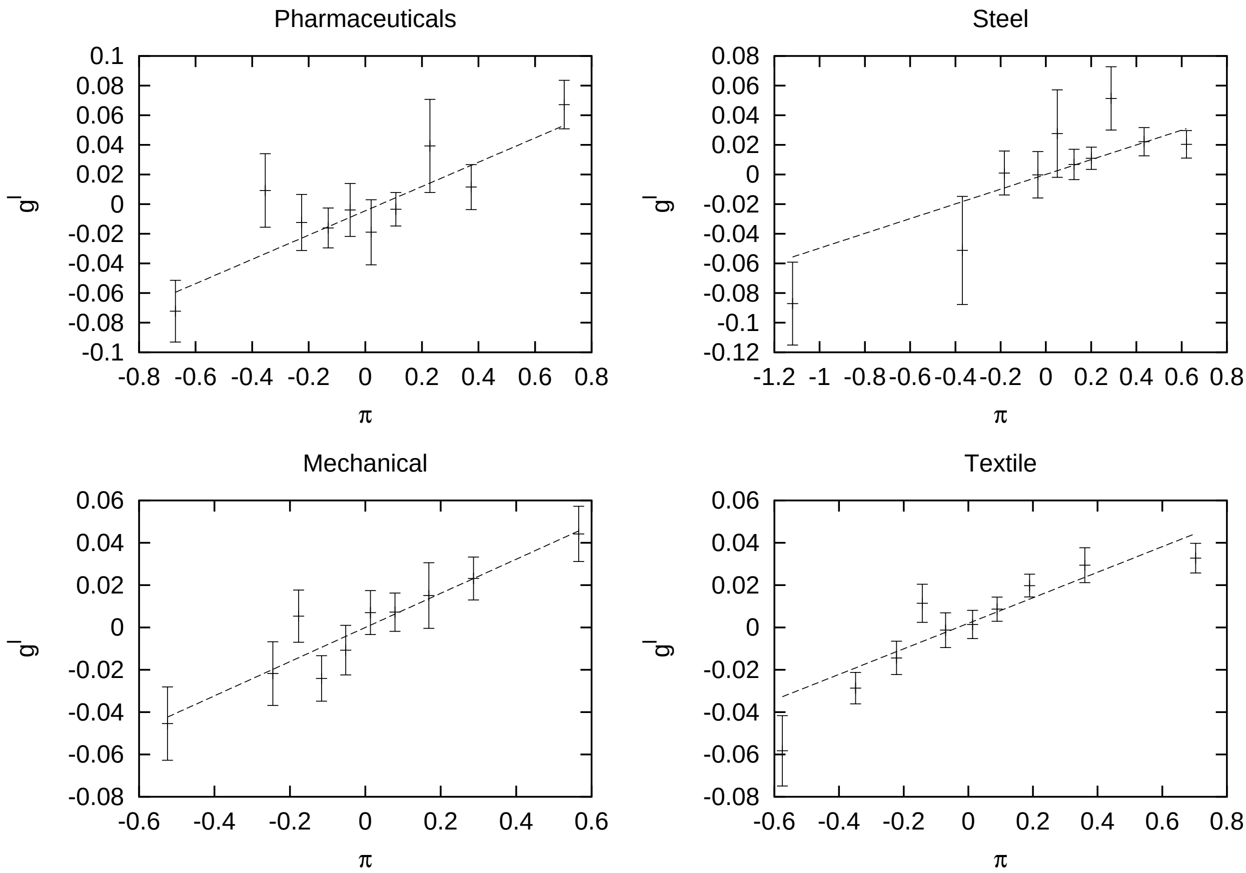

Figure 7: Regression of the employees growth g^l against (relative) productivity π. The data are distributed according to the latter in 10 equipopulated bins. The regression parameters are reported in Tab. 5.

4 Value Added

Recall the definition of Π and of the rescaled (log) variable π. Thus π represents a "relative" productivity. One then may plot employees and sales growth for different bins over the π distribution[14]. Results for the employees case are shown in Fig. 7. Here, a clear positive relationship appears, supported also by exercise of linear regression of the employment growth vs. relative productivity, whose estimated coefficients are reported in Tab. 5. The positive relationship appears quite robust and, noteworthy, rather homogenous for the different sectors.

Such a dynamic is well in tune with a "replicator-type" process of market selection whereby, in probability, firms with above-average productivity tend to expand and that below the average tend to shrink. However, in our data, we observe that this relationship disappears when considering firm growth as measured in terms of sales or values added rather then

[14] "Bin" stand for a quantile in the distribution of the population in the variable at hand

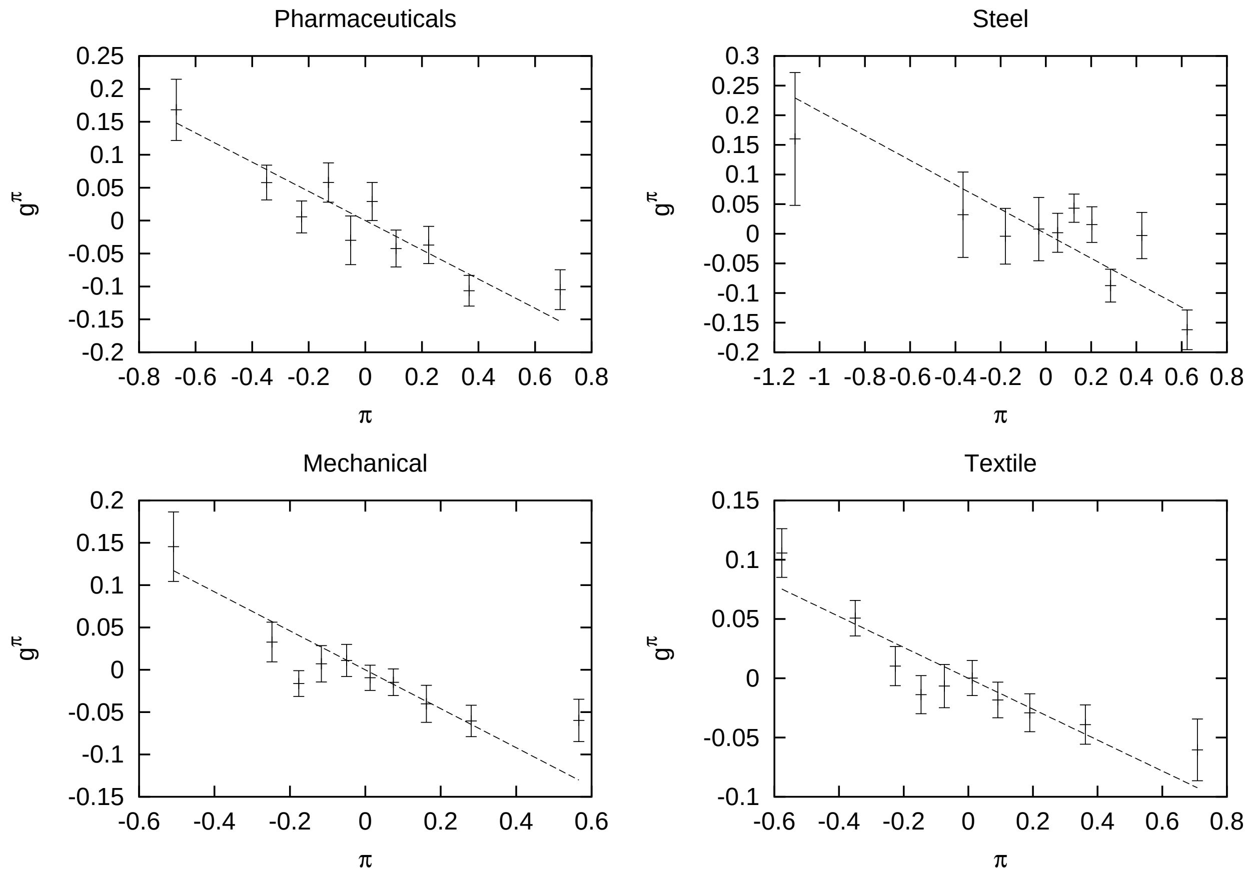

Figure 8: Regression of the productivity growth g^π on the average productivity π. The data are distributed according to the latter in 10 equipopulated bins. The regression parameters are reported in Tab. 6.

employees[15].

Next, let us consider the dynamics of relative productivity itself which in fact are in clear violation of any "Law of Proportionate Effect" but rather display powerful convergence to the mean tendencies.

In Fig. 8 we show the average productivity growth for different productivity bins. A clear inverse relationship emerges, where more productive firms are on average doomed to see their productivity decreased relatively to the industry average the next year. This is consistent with a general process of *learning* and *imitation* amongst firms which leads to *capabilities diffusion* over the industry: systematic "catch up" abilities of the technological followers persistently wash away any position of (temporary) leadership in production efficiency.

[15]This is indeed a puzzle that we intend to explore in future works

	Pharm.	Steel	Mech.	Textile
r	.081	.049	.08	.06
σ_r	.015	.012	.012	.01
r/σ_r %	18	25	15	16

Table 5: The slope r and the asymptotic standard error as obtained with an OLS linear regression of the employees growth vs. labor productivity.

	Pharm.	Steel	Mech.	Textile
r	-.013	-0.23	-0.2	-.22
σ_r	.015	0.023	0.031	.025
r/σ_r %	12	10	15	11

Table 6: The slope r and the asymptotic standard error as obtained with an OLS linear regression of the productivity growth vs. actual productivity.

5 Conclusions

As already mentioned, this work is just a preliminary study within a wider search of the statistical regularities of industrial dynamics. As such it suggest both relatively robust insights into the nature of the underlying evolutionary process and also some challenging puzzles. In these conclusions let us just mention some of the latter.

First, our data, in tune with a lot of the evidence reviewed in Geroski (2000), lack any strong autocorrelation in the growth process.

This is particularly puzzling since also in our data one finds abundant evidence of systematic heterogeneity across firms. First, as discussed above, we find at least circumstantial evidence of differences across firms in the generating processes of growth shocks. Second, and even more important, our data display striking differences across firms in production efficiencies.

Why shouldn't these asymmetries in efficiency be reflected in more systematic selection processes autocorrelated over time? Part of the answer might rest in the differences in the time scales at which productivity shocks arrive vis-à-vis the time scale at which market adjustments take place. After all, we have in the real world asynchronous processes of adjustments in production technologies, prices and market shares which might be badly reflected by an "artificial" sampling over one-year time periods (This is also akin the hypothesis put forward by Geroski (2000)).

However, we are not convinced that this is by any means the whole story. A lot of evidence from the literature suggests that profits tends to be asymmetrically distributed and that such asymmetries are persistent over time. In future works we intend to check whether these properties apply also to our data and whether they are systematically correlated with asymmetric in efficiency. If that were the cases, one would have to draw also far-reaching implications regarding the patterns of competition. In the last resort, one would be forced to conclude that asymmetric efficiencies do not translate so much in systematic "replication-type" dynamics in the relative sizes of output but primarily in differential abilities to generate profits (and possibly affect relative sizes in the longer term only through the impact of profitability upon investment).

Another puzzle regards the evidence stemming form our data of any lack of relationship between growth variance and size - contrary to a lot of previous evidence from the literature, and contrary also to our findings on the international pharmaceutical industry (Bottazzi et al., 2001; Bottazzi, 2000). In particular, in the latter work we propose an explanation of the negative variance-size relationship grounded into diversification patterns (for a similar interpretation on the American manufacturing industry cf. Stanley et al. (1997)).

In brief, Bottazzi (2000) shows that the number of markets in which a firm diversifies bears a (less than proportional) relation with size and than the underlying dynamics is a (plausibly, competence-driven) branching process. In turn diversification across (uncorrelated) markets fully explains the observed coefficients of the negative relation between growth variance and size. The lack of such a relationship in our Italian data might be interpreted on the ground of different, possibly complementary, phenomena.

First, it might well be that diversification plays a relatively weaker role in Italian firms[16]. Second, it could be the even when diversification occurs, it affects lines of business whose demand profiles tend to be highly correlated. Come as it may, the determinants of the variance in growth profiles is yet another challenging issue ahead.

[16]Note that this could well be the case if the diversification of business groups occurred through the formation of formally separate legal entities (cf. the discussions, unfortunately in Italian, in Balconi (1996) and Barca (1997)).

References

Balconi, M. (1996), "Sistemi informativi e crescita delle imprese. Un' interpretazione del recente sviluppo di gruppi medio-grandi in Italia e una rivisitazione teorica di crescita e dimensione", Economia e Politica Industriale, 92, pp. 35-78.

Barca, F. (1997)(ed), "Storia del Capitalismo Italiano dal dopoguerra ad oggi", Roma, Donzelli.

Boeri, T., (1989) "Does Firm Size Matter?", Giornale degli Economisti e Annali di Economia, 48.

Bottazzi G., (2000) "Firm Diversification and the Law of Proportionate Effect", LEM Working Paper, S. Anna School of Advanced Studies. `http://lem.sssup.it`

Bottazzi G., Dosi G.,Lippi M.,Pammolli F., Riccaboni M., (2001) "Innovation and Corporate Growth in the Evolution of the Drug Industry" International Journal of Industrial Organization, forthcoming publication.

Brock W. A., Evans D., (1986) The Economics of Small Businesses: Their Role and Regulation in the U.S. Economy, New York, Holmes and Meier.

Dosi, G. (1988) "Sources, Procedures and Microeconomic Effect of Innovation", Journal of Economic Literature, XXVI, 3, pp. 1120-1171

Dosi G., Marsili O., Orsenigo L., Salvatore R., (1995) "Learning, Market Selection and the Evolution of Industrial Structures" Small Business Economics, 7, pp. 411-

Dosi G., Riccaboni M., Varaldo R., (2000) "LA Dinamica dei Rapporti tra Grande e Piccole Imprese: Alcune Riflessioni tra Teoria e Analisi Storica" LEM Working Paper, S. Anna School of Advanced Studies. `http://lem.sssup.it`

Ericson, R. and Pakes, A., (1995) "Markov-perfect industry dynamics: a framework for empirical work" Rev. Econ. Stud. 62 pp. 53-82

Geroski P. A., (2000), "The Growth of Firms in Theory and Practice" in Foss N. and Malinke V. eds., New directions in economic strategy research, Oxford University Press

Gibrat R., (1931) "Les inégalités économiques", Paris, Librairie du Recueil Sirey.

Ijiri Y., Simon H.A., (1977) Skew distributions and the sizes of business firms, Amsterdam, North Holland.

Jovanovic, B., (1982) " Selection and Evolution in Industry", Econometrica, 50 pp.649-670.

Marsili, O. (2001) The Anatomy and Evolution of Industries, Cheltenham (UK), Edward Elgar Pub.

Nelson, R. and Winter S., (1982) An Evolutionary Theory of Economic Change, Cambridge, Mass. The Belknap Press of Harvard University Press

Stanley M. H. R., Nunes Amaral L. A., Buldyrev S. V., Havlin S., Leschhorn, H., Maass P., Salinger M. A., Stanley H. E., (1997) "Scaling behavior in economics: empirical results and modelling of company growth", Nature, 319.

Sutton J., (1997) "Gibrat's Legacy" Journal of Economic Literature, XXXV, pp.40-59

Pakes, A., and Ericson, R. (1998) "Empirical Implications of Alternative Models of Firm Dynamics" J. of Econ. Th., 79, pp. 1-45

Pavitt, K. (1984) Sectoral Patterns of Technical Changes: Towards a Taxonomy and a Theory, Research Policy, 13, pp. 343-75

Winter, S. (1971) "Satisficing, Selection, and the Innovating Remnant", Quart. J. of Eco. pp. 237-261