Laboratory of Economics and Management
Sant'Anna School of Advanced Studies
Via Carducci, 20 - I-56127 PISA (Italy)
Tel. +39-050-883-341 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://link.sssup.it/LEM

LEM
Working Paper Series

# Innovation and Corporate Growth in the Evolution of the Drug Industry

G. Bottazzi*
G. Dosi *
Marco Lippi†
Fabio Pammolli°
and
Massimo Riccaboni°

* S. Anna School of Advanced Studies
via G. Carducci, 40 Pisa, Italy
† University of Rome, "La Sapienza", Italy
° Faculty of Economics *R.M. Goodwin*, Siena, Italy

2001/02 January 2001

ISSN (online) 2284-0400

*Final Version* *February 22, 2001*

# Innovation and Corporate Growth in the Evolution of the Drug Industry

Giulio Bottazzi [1]
Giovanni Dosi [1]
Marco Lippi [2]
Fabio Pammolli [3]
Massimo Riccaboni [1]

[1] Sant'Anna School of Advanced Studies, Pisa, Italy
[2] University of Rome, "La Sapienza", Italy
[3] Faculty of Economics *R.M. Goodwin*, Siena, Italy

*The support to the research which led to this work by the Merck Foundation (EPRIS Program), the European Union (ESSY Project, TSER, contract n. SOE1-CT 98-1116 DGXII-SOLS) and the Italian Ministry of University and Research (grant 9913443984) is gratefully acknowledged. IMS International and Glaxo Wellcome Italia kindly provided data on which our databank is partly based. Comments by Buz Brock, Paul Geroski and two anonymous referees have significantly contributed to improve upon earlier versions of this paper. The usual caveats apply.*

# 1 Introduction

This work investigates the patterns of change in the international pharmaceutical industry, in particular with respect to the industry structure and the growth processes of a large sample of incumbents, against the background of the observed patterns of innovation.

Pharmaceuticals are indeed an archetypical example of a "science-based" industry, wherein innovation — in the form of new therapeutical entities, and imitation/improvements of existing ones — is the fundamental source of competitiveness within the industry, largely shaping the dynamics of growth and decline of different firms.

As such, the industry represents also a rich domain for the analysis of the properties of microeconomic processes of growth — touching upon the dynamics of innovation arrival and imitation; the processes of inter-product and inter-firm competition; corporate diversification across markets; and the ensuing characteristics of industrial structures. These are indeed the main topics addressed in this work. In particular, we shall primarily address three (inter-related) questions, namely, first, the relationship between firm size and growth; second, more generally, the statistical properties, both at aggregate and disaggregate (market specific) levels, of corporate growth; and, third, the relations between the latter and the process of technological innovation, imitation and market competition.

Our analysis, based on longitudinal data, disaggregated down to single product markets, allows one to characterize, at least qualitatively, the systematic forces driving the evolution of the industry, the sources of heterogeneity across firms and the nature of "technological shocks" and of competition processes. Putting it another way, one is able to investigate the statistical properties which are "emergent" from specific evolutionary dynamics of heterogenous learning and market selection.

In Sec. 2 we shall briefly recall some major features of the secular evolution of the drug industry. This sets also the background of the subsequent analysis, developed in Sec. 3, addressing (i) the size distribution of the 150 top firms operating in the seven major western countries; (ii) the properties of "growth shocks" at both levels of firms as a whole and of disaggregated markets and (iii) the relationship between corporate size, diversification patterns and variances in growth rates. Finally, Section 3.5 explores the process of arrival of innovations and their impact upon corporate growth and market dynamics. Finer statistical details along a similar interpretative thread may be found in Bottazzi et al. (2000).

# 2 The evolution of the industry: an overview

The history of the international pharmaceutical industry, dating back in its origin to the 19th century, has already been extensively analyzed by several scholars[1]. Here let us just mention a few major characteristics of the processes of technological learning and market competition.

Let us begin with the latter. Competition has always centered around the discovery and introduction to the market of new products, often subject to rather quick incremental improvements, as well as to imitation and generic products competition.

Notwithstanding the historically high R&D intensity of the industry, the successful introduction of major innovations, in the form of new molecules (New Chemical Entities, thereafter NCEs) with novel therapeutical properties has always been a rather rare event. For example, Barral (1996) estimates the total number of NCEs introduced throughout the world over the period 75-94 to be 154. While major innovative breakthroughs arrive quite rarely, after the arrival they experience extremely high rates of market growth (more on this point in Grabowski and Vernon, 1992).

NCEs, however, only capture a part of innovative activities within the industry. In fact pharmaceutical innovations, broadly defined, include "inventing-around" existing molecules, new combinations among them, new ways of delivery, etc.

Remarkably, the degree to which early innovators have enjoyed an advantage in introducing later major drugs within the same family has traditionally been fairly limited (see Sutton, 1998). That, together with the coexistence of several compounds or variations thereupon targeted to the same pathology generally hinders the persistence of dominant positions in a single market (cf. also Temin, 1980). In fact, in most single markets and in the industry as a whole one observes the (*persistent*) coexistence of two basic types of firms,

[1] See, among others, Aftalion, 1959; Arora, Gambardella, 1998; Chandler, 1990; Freeman, 1982; Gambardella, 1995; Henderson,

mapping into distinct technological ensembles of competencies and competitive strategies. In a shorthand, the first group, closely corresponding to some "oligopolistic core" of the industry, undertakes what is sometimes called "pioneering R&D" (Grabowski and Vernon, 1987); generates the overwhelming majority of NCEs; when successful enjoys big, albeit not very long-lasting, first-mover advantages, and charges premium prices. The second group undertakes primarily "imitative R&D"; generates incremental innovations and more competitively priced "me-too" drugs; takes up licenses from the core and is present to different degrees in the "generic" markets, after patent expirations [2].

The qualitative historical evidence hints in fact at a long-term "ecology" of the industry relying on the competition, but also the complementarity, between two organizational populations, whose relative sizes is shaped by diverse competencies in accessing innovative opportunities (and, to some extent, also by Intellectual Property Right regimes, influencing the span and length of legal protection for temporary monopolies on innovation). Some statistical properties of such an ecology will indeed be explored below.

In a nutshell, the archetypical evolutionary story concerning each "disaggregate" market (i.e. a market aimed at one particular pathology) runs more or less as follows.

A few firms — generally from the "core" — search for NCEs with the desired properties. Some of them (a small minority) achieve the stage of clinical trials. Even among them only very few happen to immediately fulfill therapeutic efficacy and required safety standards. Many NCE prototypes, on the contrary, show various clinical shortcomings, which might sometimes be overcome by alterations to the original chemical structures, the introduction of compound combinations, or different ways of administering them. In some cases these changes are undertaken by the original discoverer of the compound itself, while frequently the project based on that particular NCE is abandoned — leaving potential room for the development of modified analogues by other firms. Moreover, even when the original innovator carries the project through to a marketable product, molecular modifications of prototypes often unable followers within any chemical/therapeutic trajectory to introduce drugs with equivalent (or even superior) pharmaceutical activities, side-effect profiles, patient tolerability, etc. [3].

Come as it may, when the first-coming NCE successfully reaches the market, it generally undergoes a very fast market diffusion (cf. the examples in Fig.s 1,2,3 below), partly through the competitive displacement of "older" drugs — whenever they exist for that particular therapeutic application — and, even more importantly, through the creation of its own fast-growing market niche. Quite soon, however, the niche i.e. the product market is invaded by competing NCEs and/or "creative analogues" which curb the growth of the early monopolist. All this might happen well before the expiration of the original patent — even if the latter event generally marks another "market shock", with generic drugs and firms expanding on the market.

Note that in the evolutionary story we have just sketched there are two basic dynamic processes at work. A first one concerns the multiplication of markets through the introduction of new families of products roughly aimed at the same therapeutic targets (either "old" targets with new methods or "new" pathologies, yet unchallenged). The second process regards competition *stricto sensu* amongst firms within each "micro" market. Clearly the growth of firm size depend on both, with the timing of entry into each "micro market" being an important factor in the combination of the two effects.

Fig. 1 to 3 illustrate the market dynamics associated with evolutionary patterns sketched above. The first case show the profile of a entirely novel market, angiotensin II antagonists, in the cardiovascular area. The new niche expands very fast allowing for the steady growth of both the first-comer and other early innovators. Here, in a sense, a fast expanding market provides "room for everyone" — and with that also expanding sizes (in that market) for all early incumbents. Fig. 2 illustrates the case of antiulcerants, with two families of products, namely H2 antagonists, the older one, and acid pump inhibitors, the new one, which over time crowds out the former. All that goes together with the fate of the two leading NCEs/products (Zantac and Losec, respectively), while new "innovative invaders" only in the late 90s begin to enter the younger niche. Finally, Fig. 3 depicts a relatively "old" market, antivirals, where the first innovative mover (Zovirax), despite a steady erosion of its market shares by late-coming innovators and analogues, maintained its dominance until patent expiration, in 1997, by which date a swarm of generic competitors entered the market [4].

[2] The two basic types are not evenly distributed across countries. Firms belonging to the former come almost exclusively from the USA, Germany the UK and Switzerland, while France, Italy and Japan (not included in our sample) show up primarily in the second group.

[3] Sneader (1996) presents a detailed analysis of 244 currently employed drug prototypes, showing that out of them more than 1200 medical compounds have been derived.

[4] Note that the decrease in the total "size" of the market after 1996 is entirely due to price reductions, with quantities still

A final set of characteristics of the evolutionary dynamics of the industry that we want to recall concern the nature of product market themselves. Indeed, one observes a highly skewed distribution both of product market sizes and of intra-firm distribution of sales across products. So "a few *blockbusters* dominate the product range of all major firms" (Matraves, 1999; cf. also Sutton, 1998 and Bottazzi et al, 2000) [5].

In an extreme synthesis, the evolutionary patterns of the industry display:

- rare arrivals of major innovations (new chemical entities with novel therapeutical targets or pharmacological mechanisms) often associated with the emergence of new markets;
- a more steady activity of incremental innovation, development of therapeutic analogues, imitation, licensing;
- systematic forms of heterogeneity, even amongst incumbents, distinguishing a few rather persistent innovators from the rest of the organizational population;
- "hierarchically nested" competitive mechanisms[6] involving, at one level, innovation/imitation and market share dynamics within single product groups, and, on a longer time scale, the generation of new markets and the diversification processes across them.

Given all that, what are the resulting statistical properties of the evolution of the industry, in terms of relative sizes, rates of growth, etc.? This is what we shall explore in the following.

# 3 Patterns of corporate growth

Let us begin by noticing that the foregoing pieces of qualitative evidence suggest that the process of corporate growth is likely to display more structure than what would be predicted on the ground of uncorrelated idiosyncratic shocks, i.e. a so called Gibrat type process.

In the following, in order to disentangle the actual growth patterns, after a summary description of our database (Sec.3.1), we shall analyze the main properties of size distributions (Sec.3.2) and growth processes (Sec.3.3). Possible departures from the Gibrat hypothesis are studied by means of basic non-parametric methods. In particular we shall study: **(a)** the shape of the tails in observed growth distribution; **(b)** the time autocorrelation in growth profiles and **(c)** the relationship between size and growth. Systematic departure from the "null" (Gibrat) hypothesis are indeed identified in all the three domain of analysis.

Next, we present an interpretative framework able to account for the observed properties of growth processes, concerning in particular: **(i)** the relationship between size, diversification across different markets and growth variance (Sec.3.4), and **(ii)** the effects of products innovation, entry and imitation upon the growth of markets and of firms within them (Sec.3.5).

## 3.1 The Data Set

Our statistical analysis in the following is based on the dataset PHID (Pharmaceutical Industry Database) developed at the University of Siena. It covers top 100 companies in the seven major western markets (USA, United Kingdom, France, Germany, Spain, Italy, Canada) with 10 to 20 years of observations (depending on the variables). In this paper we aggregate the respective figures in the different national markets and consider the resulting top 150 firms [7]. Both sales figures and market shares are available for each firm from 1987 to 1997, disaggregated up to the 4-digit-level of the Anatomical Therapeutic Classification scheme (ATC) in 517 microclasses.

The database contains also detailed information about sales of 7654 drugs commercialized in the US by 57 major pharmaceutical companies. The launch date is reported for 4921 of them (64.29%). Launches are evenly distributed over the last 20 years, so we are able to track the life cycle of 1600 products over 10 years

rising.

[5] In our database the three most important products of top 5 firms account for more than 50% of the total sales. Matraves, 1999 also suggests that only the top 30 drugs worldwide cover average R&D costs.

[6] We borrow the expression from Warglien, 1995.

[7] In consequence firms which are "big" in one single national market but smaller at the aggregate level are neglected.

after their launch. Products are distinguished according to whether they are a New Chemical Entity (NCE), a patented innovation that is not an NCE, and unpatented product (including both products whose patents expired before the years under observation and products licensed from other firms).

As already mentioned, this work is focused on the processes of *internal* growth. Hence, to take into account mergers and acquisitions during the period of observation, we have constructed "super firms" which correspond to the end-of-period actual entity (so for example, if any two firms merged during the observed history, we consider them merged from the start). This procedure might bias intertemporal comparisons on actual size distributions, but it helps to highlight those changes in the distributions themselves which are due to processes of intra-market competition and inter-market diversification[8].

Finally, when studying the relationship between innovation and growth ( Sec. 3.5 below) we shall confine the analysis to the US market wherein the overwhelming majority of innovative products have been first introduced[9].

## 3.2 Size distributions

Let us begin with a descriptive analysis of size distributions. Let $S_i(t)$ be the sales of firm $i$ ($i \in [1, \ldots, 150]$) at time $t$ ($i \in [0, \ldots, 10]$). The evidence (see Bottazzi et al., 2000 for more details) shows that the ratio of the standard deviation to the mean as well as the skewness and the kurtosis are nearly constant over time. These properties imply that the "normalized size" $G_i(t) = S_i(t)/ < S_i(t) >$ (the brackets $< \ldots >$ denote the average over all firm sizes at a given year) is stationary over time[10]. This quantity is proportional to the market share when the number of firms is constant but provides two advantages: first, it can be used also to characterize distributions whenever the number of firms changes over time (while the shares distribution would yield a spurious shift of their means) and, second, it provides an easy way of comparing distributions with different number of observations[11]. In Fig. 4 we plot the distribution function for $g_i(t) = \log(G_i(t))$. The normal fit (dotted line in Fig. 4) clearly shows that the distribution possess a fatter upper tail than a Gaussian. Think of a growth process as:

$$g_i(t+1) = g_i(t) + h_i(t) \tag{1}$$

with $h_i(t)$ independent random variables. Under mild assumptions on the distribution of $h$, the size distribution should asymptotically tend to a Gaussian one[12]. We checked however that this trend is absent from data (see Bottazzi et al., 2000). As such, this property may be already considered as a piece of circumstantial evidence against a simple Gibrat model. However, let us explicitly examine the properties of the growth process both at aggregate and disaggregate levels.

## 3.3 Corporate growth dynamics

As already mentioned, a classic benchmark in the analysis of growth processes addresses the relationship between size and growth, and in particular, possible departure from the so called "Law of Proportionate Effect"[13]. A first step is to check for possible "reversion to the mean" in our data. In order to do that we estimated the model $g_i(t+1) = \beta g_i(t) + \varepsilon_i(t)$ cross-sectionally for all the years, finding values for $\beta$ statistically

[8] In particular, our analysis is meant, on purpose, to entirely wash away the effects upon industrial concentration of acquisition processes

[9] We choose to do that in order to avoid problems of international comparison between different institutional and regulatory systems.

[10] For a discussion of the accuracy of this procedure see Bottazzi et al., 2000.

[11] The analysis of $G$ is fully equivalent to the analysis of $S$ apart from a scale factor. A similar approach is adopted in Kalecki, 1945 and Hall and Prais, 1956.

[12] For the Central Limit Theorem the $n$-th normalized cumulant $\lambda_n = c_n/\sigma^n$ of the size distribution $\sum_{t=0}^{T} h_i(t)$ would behave as $\lambda_n \sim T^{1-n/2}$ where $T$ are the total time steps. In particular the third and fourth normalized cumulants, the skewness and the kurtosis, would decrease respectively as $T^{-0.5}$ and $T^{-1}$.

[13] The literature on the subject is vast and cannot be surveyed here: cf. Ijiri and Simon, 1977; and for recent critical discussion of both the evidence and the related theoretical implications, Boeri, 1989; Brock and Evans, 1986; Sutton 1997; Geroski, 2000.

equal to one, thus rejecting the "reversion" hypothesis (see Appendix 1 for more details). Consequently, one is entitled to refer to (1) and confine the analysis to the properties of $h$. Fig. 5 reports the distribution for $h$ averaged over time together with the distribution for some time steps ($t = 0, 5, 10$). As can be seen differences are small, supporting the assumption, from now onward, that the distribution of $h$ is stationary over time. However such a distribution is highly non-Gaussian (see Fig. 5) and a fit with a symmetric exponential

$$p(h = x) = \frac{\alpha}{2} e^{-\alpha |x - m|} \tag{2}$$

provides a good description of its central part. Moreover the distribution is asymmetric with fat tails corresponding to spurs of growth that are more frequent than those predictable on the ground of a Gaussian noise process (see Fig. 5).

We have previously argued that the stability and the shape of the size distribution can hardly be explained using the simplest Gibrat-type model in (1) with independent increments. An interesting problem concerns indeed the identification of possible sources of "dependence" in the growth process governing this industry. The first effect to analyze is the possible autocorrelation in time of firm growth. In Fig. 6 we plot the autocorrelation coefficient of the logarithmic growth $c(t, \tau)$. The plotted line is the average $\overline{c}(\tau) = \sum_t^T c(t, \tau)/T$, where $T$ is the number of years covered in our database and $\tau$ is the time lag[14]. If one considers the average correlation $\overline{c}(\tau)$, contrary to the prevailing results in the literature (for a critical discussion cf. Geroski, 2000) our data do highlight on average a significant ($p < .001$) positive autocorrelation until the second lag $\tau = 1, 2$ (while we do not dare making any claim on longer time lags)[15].

Fig. 6 refers to the growth of firms as a whole. However, in order to fully understand corporate growth, it is necessary to investigate the persistency profiles by firms within single therapeutic categories ("sub-markets"). Indeed this is a level of observation nearer to the actual competition process, where innovative shocks are likely to exert their effect. Let us look at the distribution function of the autocorrelation $\overline{c}_j(\tau)$ over the set of all the sub-markets and plot it for different $\tau$ lags. Fig. 7 presents the results for different level of disaggregation (1, 2 and 3 digits, according to the therapeutical classification; see above). As can be seen the distributions are similar for different disaggregation levels and, remarkably, sales aggregation at the firm level does *not* wash away autocorrelation in growth: on the contrary, for $\tau = 1$, *firm-level correlations* are significantly *higher* than the average correlation calculated at any disaggregation levels (this property disappears for $\tau > 1$). Moreover a positive average autocorrelation seems to survive for a long time even if of negligible magnitude[16].

Interestingly, one is able to observe also important market-specific characteristics in the competitive dynamics. In fact, as shown in Bottazzi et al., 2000, the distribution of the growth correlation of the two leading firms for each submarket is U-shaped, suggesting that different market-specific growth processes coexist within the industry, including both highly correlated and anti-correlated patterns. The inter-market heterogeneity might indeed be the effect of considering statistics over different "windows of observation" of a "technology cycle" shaping the growth process in each sub-market[17]. So, for example, one might expect positive correlation in the stages of penetration of new products and anti-correlation associated with imitative entry and patent expirations (for an illustrative example see Sec.2 above).

To further investigate the growth structure we have computed the one-year transition matrices of the stochastic variable $h$ both at the aggregate and at the 3-digit-submarket levels. Having controlled the robustness of our results for different numbers of bins[18] (ranging from 20 to more than 200) without revealing remarkable variations, let us present the analysis at the highest level of resolution supported by data. More

[14] The data points relative to the same lag $\tau$ but different initial time $t$ are dispersed, partly due to the relatively small number of observations and partly for a seemingly "true" difference in the growth from time to time (indeed there are points separated by more than 3 standard deviations). In any case, the hypothesis $c(t, 1) = 0$ is rejected with a significance greater than .001 in 7 over 9 time steps $t$.

[15] Notice, however, that one ought to be cautious about the procedure of taking the average correlation as an estimate of a "true" (stationary) correlation, due to the high dispersion of $c(t, \tau)$ for different initial times $t$.

[16] We searched also for a possible dependence of the autocorrelation $\overline{c}_j(\tau)$ on the size $M_j$ of the sub-markets themselves, but no evidence of any such dependence has been found.

[17] Our statistics cannot discriminate between purely random micro processes and more systematic competition processes observed at random times.

[18] The "bins" are the different intervals of values (quantiles) by which the data are partitioned.

precisely, the actual rate of variation over time of the aggregate standardize growth ($h_i$) and disaggregate ones ($h_{ij}$) have been uniformly divided into 50 quantiles and every firm has been assigned to them in each year. Since the transition matrices of $h$ do not change substantially over time (as well as the distribution of $h$ itself, see Fig. 5), we analyze jointly the one-year transition probability matrices. Two transition matrices have been computed, at the aggregate ($\Pi_a$) and at the submarket level ($\Pi_d$)

$$[h_i(t)] = \Pi_a[h_i(t-1)] \quad (3)$$

$$[h_{ij}(t)] = \Pi_d[h_{ij}(t-1)] \quad (4)$$

where $[h_{i(j)}(t)]$ are column vectors of binned growth at time $t$ and each row of $\Pi_a$ and $\Pi_d$ represent the conditional probability vector of moving through the grid. Fig. 8 and Fig. 9 show the three-dimensional plot of transition matrices $\Pi_a$ and $\Pi_d$. To interpret the graphs, take any point on the $[h_{i(j)}(t-1)]$ axis and look in the direction parallel to the other axis in order to trace out the probability density describing the transition to different parts of the growth distributions (more details on discrete stochastic kernel analysis are in Quah, 1996). If the graphs pile up on the positive sloped diagonal this may be interpreted as evidence of high persistence and "inertia". Actually, Fig. 8, in line with our previous findings, shows a significant autocorrelation in growth rates at the aggregate firm level: more specifically, an oblique shape clearly emerges which may be fitted by a line with slope 0.38 passing through the mean values of $h_i(t)$ and $h_{ij}(t-1)$. Over time, the growth distribution converge toward its mean value where most of the observations are concentrated. However note that autocorrelation appears to be highly dependent on relatively rare events of sustained growth represented by the spikes on the top right quadrant of the transition matrix. Remarkably, qualitatively inspection shows that the majority of these growth events corresponds to big innovation breakthroughs.

The companion Fig. 9, computed at the submarket level enables us to disentangle growth autocorrelation in terms of single markets dynamics. Two different regimes are clearly distinguishable. The products that at time $t-1$ experienced a growth above the average are sharply divided into two groups: some of them keep their growth pace unchanged, while most of them subsequently drop on the mean. Conversely, the mean growth level represents an absorptive state for slow (below the mean) growing products too, but the phase transition is considerably smoother. New products bust into the market and grow swiftly for a short period of time, then a cluster of analogue drugs enter (and possibly licensing begins, too). As a result, innovative drugs growth slows down. After a while, all incumbents tend to growth approximately at the same rate, even if with highly asymmetric shares, in favor of the early movers.

Given the foregoing features of growth processes, at both corporate and disaggregate market levels, let us suggest some elements of an interpretation drawing upon possible forms of heterogeneity across firms, in particular with reference to diversification profiles and innovative patterns.

### 3.4 Size, Diversification across sub-markets and Gibrat violations

The tangled and "classic" question concerning the possible dependence of the distributions of the growth rates on the initial firm size is, of course, not exhausted by the forgoing (negative) findings on correlation measures between growth and size. While the existence of correlation in a Gibrat-type test between growth and size would be sufficient to reveal dependence, the converse does not hold: dependence might just show up at higher moments of the conditional distributions. In Fig. 10 we plot the moments of the growth distribution for different size classes built considering, at each time step, all the firms with size in a given range.

In line with a few contributions in the literature[19], as Fig. 10 shows, no dependence of mean growth appears, and neither does dependence in autocorrelation. However a clear pattern emerges concerning the variance of growth rates, decreasing with increasing size. Fitting the relation between growth variance and size with an exponential law

$$\sigma(h) \sim e^{\beta g} \quad (5)$$

we obtain a value $\beta \sim .2 \pm 0.02$ which is striking similar to the one found in other analyses on different datasets (cf. Stanley et al., 1996, Lee et al., 1998).

The rather unique possibility offered by the PHID database is to break down sales until the 4-th digit of the ATC code allowing, as mentioned, the identification of sub-markets that are "specific" enough to be

[19] cf. Boeri, 1989; Evans, 1987; Geroski, 2000 and Geroski et al. 1998; Hall, 1987; Hymer, Pashigian, 1962; Mansfield, 1962; Sutton, 1997 among others.

considered the *loci* of competition among firms, and also a more accurate evaluation of the relationship between diversification, variance of growth and size.

Were one to assume that firms are collections of independent elementary lines of business, roughly of the same size, whose number is proportional to the overall size of the firm, then the Law of Large Numbers would predict a relation between variance of aggregate growth and firm size of the form $var(h) \sim g^{0.5}$. However, both the existing literature and our data (see Fig. 10), show a *lower* exponent.

In the literature, this departure from the predictions of the Law of Large Numbers is typically imputed to some interrelation between the "components" that make up each firm[20] (cf. Boeri, 1989; Stanley et al., 1996).

Conversely, as we shall do, one could relax the assumption of (unobserved) elementary components and measure the actual relationship between aggregate firm size and the number and size of its lines of business. In fact, our analysis shows that the Law of Large Numbers does explain the observed relationship between size and variance of growth if one consider as elementary lines of business the different submarkets in which each firm operates. This result comes from two observations: first, that the correlation across sub-market is negligible and second, that the number of active sub-markets of a given firm increase, on average, with its size following a non linear scaling law.

More formally, let $S_{i,j}(t)$ be the size of firm $i$ in sub-market $j$ at time $t$, and $S_i(t) = \sum_j S_{i,j}(t)$ its total size. The aggregate growth of each firm can be written as

$$\frac{S_i(t+1)}{S_i(t)} = \sum_j \frac{S_{i,j}(t+1)}{S_i(t)} \,. \tag{6}$$

We have computed the correlation of $S_{i,j}(t+1)/S_i(t)$ for all firms in all submarkets, obtaining a distribution sharply centered around zero[21]. For any practical purpose, the growth on different sub-markets can then be considered uncorrelated and, therefore, the variance of the aggregate growth is the sum of the variances of growth in each sub-market.

Defining $R_{i,j}(t) = S_{i,j}(t+1)/S_{i,j}(t)$ and $\Delta_{i,j}(t) = N_i(t)\, S_{i,j}(t)/S_i(t)$ where $N_i(t)$ is the number of sub-markets in which firm $i$ operates at time $t$ (active sub-markets), the variance of the growth of the "normalized" size $G$ becomes

$$\mathrm{var}_{i,t}[H_i(t)] = \mathrm{var}_{i,t}\left[\frac{G_i(t+1)}{G_i(t)}\right] = \sum_j \mathrm{var}_{i,t}\left[\frac{M(t)}{M(t+1)} R_{i,j}(t) \frac{\Delta_{i,j}(t)}{N_i(t)}\right] \tag{7}$$

where $M(t)$ is the average (aggregate) size of firm at time $t$ and the ratio $M(t)/M(t+1)$ is a normalization factor (proportional to the rate of growth of the total industry). Here $\mathrm{var}_{i,t}$ denotes the variance of the distribution obtained using the complete panel (all firms at all time steps).

In (7) the contribution of each submarket factorizes in three terms, namely, first, $R_{i,j}(t)$, the actual growth of the firm $i$ in sub-market $j$; second, the inverse number of active markets $1/N_i(t)$ and, third, $\Delta_{i,j}(t)/N_i(t)$, a weighting coefficient describing the "diversification asymmetry" of firm $i$[22]. It happens that the mean and variance of the distribution of $R_{i,j}(t)$ and $\Delta_{i,j}(t)$ obtained using different size bins do not show any clear dependence on the average size of the firms in each bin (see Bottazzi et al., 2000). Therefore, the number of active submarkets $N_i(t)$ must be the sole responsible for the observed dependence of variance over the aggregate size. Fitting on a log-log scale the average number of active submarkets for each bin against the average size of the bin one obtains a slope $\alpha = .39 \pm .02$ and an intercept $q = 6 \pm .12$ (see Fig. 11). The Law of Large Numbers would predict a relation between the exponent in Fig. 5 and the slope in Fig. 11 of the form $\beta = -\alpha/2$ which is in perfect agreement with our evidence[23].

Summarizing, our evidence shows (i) that the number of submarkets in which a firm operates increases non linearly with firm size and (ii) that such a number fully accounts for the observed relationship between growth variance and size.

[20] Note that imposing a simple correlation in components growth is not enough to explain the low value of $\gamma$.

[21] With a standard deviation of $.388\ 10^{-4}$ and an average deviation of $.24\ 10^{-5}$

[22] This term capture the asymmetry in the contribution of each submarket to the overall sales of the firm. If the firm $i$ at time $t$ is symmetrically diversified over its active submarkets, the distribution of $\Delta_{i,j}(t)$ in $j$ is centered around 1, otherwise it is more broadly distributed.

[23] Notice that there is a weak relationship between the variance of $\Delta_{i,j}(t)$ and the aggregate size. A linear fit provides a slope of .09 that is however negligible compared to the effect of the number of active submarkets.

the relation provided by the Law of Large Numbers is valid as long as one considers the actual number of sub-markets a firm operates in. In order to demonstrate this statement, it has been necessary to rule out two possible sources of functional dependence between aggregate growth variance and size, namely first, the possibility that the mean and the variance of firm growth in individual sub-markets depends (on average) on its total size and, second, the possibility that the diversification pattern of a firm varies (on average) with its size[24].

In turn, however, a puzzling implication of these findings (already pointed out by Boeri, 1989) is that a large firm is more "risky" than a collection of smaller ones: a notional investor would face a lower risk by diversifying its portfolio in $N$ (independent) firms of size $S$ rather than betting on one single firm of size $N\,S$.

From an interpretative perspective all this militates against the hypothesis that diversification is driven by risk minimizing considerations. Rather, the evidence may be plausibly interpreted in terms of *competence-driven* diversification processes, in presence of knowledge spillovers across products and lines of search. In fact, as formally discussed in Bottazzi, 2000, observed diversification patterns can be essentially described by a stochastic branching process; its economic interpretation may be plausibly grounded in the incremental development of knowledge bases, driving the exploration of an expanding range of products/markets.

## 3.5 Innovation and growth

A major tenet of evolutionary theories of industrial dynamics is indeed the general conjecture that the processes of technological innovation and imitation are major drivers of industrial dynamics and also of the competitive fate of individual firms[25]. How does our evidence bear on this proposition?

Let us beginning by considering the process of introduction of innovative drugs, both New Chemical Entities and patented products[26], in the US market. Consider first the distribution of NCE launches over the population of firms throughout the 11 years of observation (1987-97). In fact, the number of NCEs that a firm introduces over a given period may be understood as one proxy for its "degree of innovativeness", and as such contribute to reveal possible underlying forms of heterogeneity across firms in their ability to innovate.

As a benchmark, let us model what would happen with technologically homogeneous firms. In these circumstances, as a first approximation, one may consider the arrival of different NCEs as independent events. This means that given a set of $N$ NCEs introduced by population of $F$ firms, the probability to find a firm which introduced exactly $k$ NCEs is given by the binomial distribution:

$$p_{\mathrm{M.B.}}(k) = \binom{N}{k} \left(\frac{1}{F}\right)^{k} \left(1 - \frac{1}{F}\right)^{N-k} \quad . \tag{8}$$

As shown in Fig. 12, this model (known as Maxwell-Boltzmann statistics) provides a poor description of the observed frequencies.

Indeed the assumption of random independent assignments is at odds with the qualitative evidence on whole families of research projects conducted by each firm over several years, often entailing knowledge spill-overs across them. Hence, one may conjecture some correlation amongst arrivals of NCE due to learning effects across individual research projects. In order to empirically check this hypothesis one ought to check whether the random assignments of innovations to individual firms concern indeed "packets" of NCEs rather then single products. Under the assumption of equiprobability of the packet sizes the appropriate statistics, known in physics as the Bose-Einstein statistics (Reichl, 1980) consider the probability to find a firm who introduced exactly $k$ NCE:

$$p_{\mathrm{B.E.}}(k) = \frac{\binom{F+N-k-2}{N-k}}{\binom{F+N-1}{N-1}}. \tag{9}$$

As can be seen from Fig. 12 the latter distribution provides an excellent description for low to average NCE numbers and only fails for the very large assignments. It is also interesting to notice that the propensity to introduce NCE is not monotonic in the size of the innovative firms themselves (so that for example the upper

[24] Both these possibilities are actually discussed in the literature as possible sources of violation to the Gibrat law (see for instance Hart, Prais, 1956).

[25] For theoretical arguments, cf., among others, Nelson and Winter, 1982 and Dosi et al., 1995 and for qualitative historical discussions Freeman, 1982 and Pavitt, 1999.

[26] Note that the former are a small subset of the latter.

tail in the distribution in Fig. 12 does not feature the biggest firms in the industry). This evidence hints at some underlying forms of heterogeneity in search competences and/or search orientation (e.g. biased in favor or against the quest for relatively major innovations). Together, the widespread occurrence of "clustering" in the arrivals of NCEs does suggest the importance of firm-specific learning effects across different research projects.

Given the foregoing evidence, a crucial question concerns if (and how much) the introduction of NCE during a firm history affects its growth performance. One can start by partitioning the set of firms depending on the number of NCEs they have introduced during the time window of observation. In Bottazzi et al., 2000 we show that no relationship appears between the number of NCEs and firm performances: indeed "more innovative" firms do not seem to gain, on average, market shares with respect to "less innovative" ones.

As another, broader, proxy for the innovative capability, let us consider the *patent intensity* of each given firm, defined as the share of patented products present in its products portfolio[27]. Again, we observe that, first, the distribution of patent intensity on the population of firms is very heterogeneous; second, bigger firms tend to show lower than average patent intensity in their product portfolios, while some of the smaller ones have a value near to unity; and, third, no systematic relationship appears between the structure of product portfolios and growth performance. Taken together, all the foregoing pieces of evidence suggest that firms embody rather idiosyncratic bundles of products, characterized by varying degrees of innovativeness, without however systematic effects of the "technological ID" of the firm itself upon its global growth performances (more on this in Bottazzi et al., 2000).

Note that all this does not imply homogeneous market dynamics of the three groups of products (NCEs, patented drugs and not patented ones) in the 4-digit submarkets in which they compete. Let us consider $S_{i,j,k}$, the sales of the product $k$ of firm $i$ in submarket $j$ and define the "normalized" sales with respect to all other products in the same 4-digit submarket as

$$G_{i,j,k} = \frac{S_{i,j,k}}{< S_{i,j,k} >_{i,k}} \quad . \tag{10}$$

Fig. 13 reports the mean and variance of $G_{i,j,k}$ for three categories of products (i.e. NCEs, patented products and all others) as a function of the time elapsed since their introduction (so that on the x-axis one reads the "market age" of each product). Here one observes a pronounced "market cycle" of NCEs which tend to "hit the market big" and decline relatively soon thereafter, with a burst in the variance (intuitively, a burst in the competition with the ensuing high turbulence in market shares) early in their life cycle. The much more numerous family of patented drugs follows, on average, a "market cycle" loosely similar to NCEs but much less pronounced changes in both their means and variances. Finally, not patented drugs appears to be highly stable and occupy from the start their long term market position.

In all that, innovation does indeed drive the evolution of each submarkets, but the competitive regime is not such as to guarantee a sustained competitive advantage and a systematic above-average growth to the individual innovators either in the affected submarkets or for the firms as a whole. Rather, one may think of some analogue to population-level mixed-strategy equilibria (here, populations of innovative vs. imitative and "old" products) which persistently coexists within submarkets and also within single firms. Innovations continue to upset this population but imitations, analogues developements, etc. are fast enough to curb any long term advantage to specific products.

# 4 Conclusions and conjectural implications for the theory

In this work we have explored the statistical properties of the dynamics of an industry — pharmaceuticals — whose long-term evolution is fueled by innovation, imitation and creation of new markets, trying to identify the possible links between the fundamental features of such evolutionary patterns and the quantitative evidence on corporate growth.

Here — as well as in several studies of this *genre* — a benchmark of departure has been the so-called Gibrat Law. However, such a "law", as Brock (1999) emphasizes, "is useful as a rough approximation to the unconditional distribution of rates of growth of firm sizes, which is especially pertinent to illustrating the degree of accuracy of the "Law of Proportionate Growth"; [however] it has poor power to discriminate across

[27] Recall that patents, often based on "creative analogues", new ways of combining existing NCEs, etc. are much more frequent than NCEs

different plausible stochastic processes that might fit the stochastic dynamics of firm growth" (p. 432). Our data, breaking down firms dynamics over highly disaggregated product markets, as well as complementary pieces of evidence on innovation and competitive patterns, allows to discriminate the finer structure of the growth processes and their links with size distributions, on one hand, and innovative activities, on the other.

The evidence show, together, a) fat tails in the distribution of growth shocks, present at all levels of aggregation, with (relatively rare) big "spurs of growth"; b) a significant autocorrelation in growth rates, again at all aggregation levels; c) a fall of the variance of growth rate with size (in line with previous findings) which, at closer inspection, is entirely dependent on diversification patterns, in turn plausibly shaped by the "competence scope" of each firm; d) different "lifecycles" of diverse types of products (defined in term of their degrees of innovativeness) displaying equally diverse growth profiles; e) persistent form of heterogeneity across firms in terms of innovative output which however do not appear to effect their comparative growth performances.

Our results, on the negative side, allow us to rule out some interpretations of the growth processes as sum of independent events. For example, the fat-tailed growth distribution and its departure from a Gaussian one even in its central part is at odds with a "pure Gibrat process". In this respect note that if Gibrat dynamics were a strict description of the process this should apply to all time scales (e.g. on monthly or weekly bases, etc.). But then, pushing the reasoning to the extreme, "years" should display a much more Gaussian profile — for the Central Limit Theorem — irrespectively of the original distribution of events.

On the positive side, our evidence, first, is well in tune with the conjecture (Geroski, 2000) that the time scale of arrival of "big" growth impulses associated with the arrival of major innovations — i.e., in the case of pharmaceuticals, New Chemical Entities — is different from the scale on which corporate growth is measured (i.e. accounting years). Such shocks are quite rare, are persistently generated by a relatively small number of innovators (indeed a subset of the population of top incumbents considered in this work), but any one innovator is unlikely to hit the same market twice. Moreover, NCEs a few times *create new markets.*

Hence, the overall, industry-wide, growth dynamics is likely to be the mixing of two different underlying evolutionary processes. The first, driven by major, rather rare, innovations often entails the generation of new market niches (new therapeutic targets, etc.). The second ("faster") process is associated with imitation, development of analogue drugs, incremental therapeutic improvements, etc., and drives the competition process within already existing markets.

Our analysis reveals also persistent forms of heterogeneity across firms. First, the autocorrelation in firm growth, increasing with the scale of observation, does indeed hint at some significant firm-specific structure in the growth process, possibly related, we conjecture, with firm-specific organizational competences in the search and introduction of products in different markets. Second, firms systematically differ in their innovative propensity (either when measured in term of NCEs and of patented products).

However, the diversity in the technological profile of each firm does not appears to influence long term growth performance. Rather, our evidence appears to support some sort of "ecology" of heterogeneous firms (and of products at different stages of their lifecycle within single firms) holding some sort of long-term evolutionary complementarity.

Further corroborations of this interpretation will involve also conditioning the observed dynamics upon finer proxies for the technological characteristics of firms and upon the "stages" of market development. We would like to consider this work just as an initial exploration of links between some basic features of industrial evolution — so far empirically analyzed in largely qualitative manners — on the one hand, and "emergent" statistical properties of industrial structures and growth dynamics, on the other.

# Appendix 1 - A linear test of the Gibrat Law

A common procedure to analyze growth processes, to which we shall adhere, is to check for departures from the so-called Gibrat Law, that is for growth patterns deviating from proportionality of mean growth to size. The "law" may be stated in different but statistically equivalent forms[28].

In the version proposed by Kalecki (1945) and adopted, among others, by Hall and Prais (1956) and Chesher (1979), the Gibrat linear test refers to the dynamics of the deviations of natural logarithms of firm sizes from their means ($g_{i(j)}(t) = s_{i(j)}(t) - < s_{i(j)}(t) >$) and it is meant to provide an estimate of the

[28]cf. among others Mansfield; 1962; Ijiri and Simon, 1977; Geroski, 2000; Sutton, 1997.

divergence/convergence of the size distribution toward its mean. In that vein, one tests the model

$$g_i(t) = \beta g_i(t-1) + \varepsilon_i(t) \tag{11}$$

and one typically concludes that the Gibrat Law is satisfied if the OLS estimator of $\beta(t)$ is close to unity.

We apply this analysis on our panel, testing the model cross-sectionally for each time step. Let $\beta(t)$ be the OLS estimation of the coefficient in (11) at time $t$:

$$\beta(t) = \frac{\sigma_g(t)}{\sigma_g(t-1)}\rho_g(t) \tag{12}$$

where $\sigma_g^2(t) =< g_i(t)^2 >$ is the variance of $g$ at time $t$ and $\rho_g(t) =< g_i(t)g_i(t-1) >$ is the autocorrelation.

From (12) it immediately follows that the variance of size distribution decreases at time $t$ when $\beta(t) < \rho(t) \leq 1$. Tab. 1 reports the statistics resulting from the tests carried out over the period 1987-1997. As shown there, $\beta(t)$, and $\xi(t) = \sigma_g^2(t)/\sigma_g^2(t-1)$ are always very close to unity, meaning that size does not exert a significant influence on expected growth, with the variance of the size distribution remaining constant over time. In fact, given these results, one is entitled, as we did, to refer constantly to (1) and analyze the distribution of $h$ defined there (since this analysis is fully equivalent to the analysis of errors $\varepsilon$ of (11)).

Let $h_{i,j}(t) = g_{i,j}(t) - g_{i,j}(t-1)$ be the (logarithmic) growth of the firm $i$ in the sub-market $j$. Remember that $g$ is defined as the logarithm of the "normalized" size $G$. Here the $G_{i,j}(t)$ is normalized using the size of the $j$-th sub-market $M_j(t)$. We then test the analogous to (11) at the disaggregate level[29]. Fig. 14 reports the values of $\beta_j(t)$ and $\xi_j(t)$ for each sub-market $j$. In line with the aggregate results, the median of $\beta_j(t)$ stays quite close to unity over time, but the distribution of $\beta_j(t)$ display a remarkable degree of heterogeneity among sub-markets, ranging from 0.82 to 1.14. On the contrary, the median of $\xi_j(t)$ distribution is constantly above unity meaning that the variance of firm sizes in individual sub-markets increases (on average) over time.

# Bibliography


Aftalion J., 1959, History of the International Chemical Industry, University of Pennsylvania Press, Philadelphia.

Arora A., Gambardella A., 1998, "The Chemical Industry: Evolution of Industry Structure", in Arora A., Landau R., Rosenberg N., eds. Chemicals and Long Run Economic Growth, Wiley, New York.

Barral E., 1996, 20 Years of Pharmaceutical Research Results Throughout the World, Rhone Poulenc Foundation, Antony Cedex, France.

Boeri T., 1989, "Does Firm Size Matter?", Giornale degli Economisti e Annali di Economia, 48, pp. 477 - 495.

Bottazzi G., Dosi G., Lippi M., Pammolli L., Riccaboni M., 2000, "Processes of corporate growth in the evolution of an innovation-driven industry. The case of pharmaceuticals.", LEM Working Paper, Sant'Anna School of Advanced Studies, Pisa.

Bottazzi G., 2000, "Diversification of Firms and the Law of Proportionate Effect", LEM Working Paper, Sant'Anna School of Advanced Studies, Pisa.

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

Brock W. A., 1999, "Scaling in Economics: A Reader's Guide", Industrial and Corporate Change, 8, 3, pp. 409-446.

Chandler A., 1990, Scale and Scope, Harvard University Press, Cambridge, Ma.


[29] We considered the 180 sub-markets including at least 20 firms out of the 302 covered by PHID. They account for the 85.6% of the pharmaceutical market. Furthermore, we checked the invariance of our results considering sub-markets with different numbers of competitors.

Chesher A., 1979, "Testing the Law of Proportionate Effect", Journal of Industrial Economics, 27, 4, pp. 403-411.

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., Malerba F., Marsili O., Orsenigo L., 1997, "Industrial Structures and Dynamics: Evidence, Interpretations and Puzzles", Industrial and Corporate Change, 6, 1, pp. 3 - 24.

Evans D., 1987, "Tests of Alternative Theories of Firm Growth', Journal of Industrial Economics, 95, pp. 657-674.

Freeman, C., 1982, The Economics of Industrial Innovation, Francis Printer, London.

Gambardella A., 1995, Science and Innovation. The US Pharmaceutical Industry During the 1980s, Cambridge University Press, Cambridge.

Geroski P. A., Walters C., Urga G., 1998, "Are Differences in Firm Size Transitory of Permanent?", mimeo, London Business School, London.

Geroski P. A., 2000, "The Growth of Firms in Theory and Practice", in N. Foss and V. Mahnke Eds, Competence, Governance and Entrepreneurship, Oxford University Press, Oxford and N.Y.

Grabowski H., Vernon J., 1987, "Pioneers, Imitators, and Generics. A Simulation Model of Schumpeterian Competition", Quarterly Journal of Economics, 102, pp. 491-525.

Grabowski H., Vernon J., 1992, "Brand Loyalty, Entry, and Price Competition in Pharmaceuticals after the 1984 Drug Act", Journal of Law and Economics, 35, pp. 311-350.

Hart P.E., Prais S.J., 1956, "The Analysis of Business Concentration: A Statistical Approach", Journal of the Royal Statistical Society, 119, Series A, pp. 150-191.

Hall B.H., 1987, "The Relationship Between Firm Size and Firm Growth in the US Manufacturing Sector", Journal of Industrial Economics, 35, 4, pp. 583-606.

Henderson R.M., Orsenigo L., Pisano G., 1999, "The Pharmaceutical Industry and the Revolution in Molecular Biology: Interactions among Scientific, Institutional, and Organizational Change", in Mowery D., Nelson R. R., eds., Sources of Industrial Leadership, Cambridge University Press, Cambridge, pp. 267-311.

Hymer S., Pashigian P., 1962, "Firm Size and Rate of Growth", Journal of Political Economy, 72, 6, pp. 556-569.

Ijiri Y., Simon H.A., 1977, Skew Distributions and the Sizes of Business Firms, North-Holland, New York.

Kalecki M., 1945, "On the Gibrat Distribution", Econometrica, 13, 2, 161-170.

Lee Y., Nunes Amaral L.A., Canning D., Meyer M., Stanley H.E., 1998, "Universal features in the growth dynamics of complex organizations", Cond.Mat. 9804100

Mansfield D.E., 1962, "Entry, Gibrat's Law, Innovation, and Growth of the Firms", American Economic Review, 52, 4, pp. 1024-1051.

Matraves C., 1999, "Market Structure, R&D and Advertising in the Pharmaceutical Industry", The Journal of Industrial Economics, 47, 2, pp. 169-194.

Nelson R. R., Winter S. G., 1982, An Evolutionary Theory of Economic Change, The Belknap Press of Harvard University Press, Cambridge, Ma.

Pavitt K., 1999, "Technology, Management and Systems of Innovation", Edward Elgar, Cheltenham.

Quah D. T., 1996, "Convergence Empirics across Economies with (some) Capital Mobility", Journal of Economic Growth, 1, pp. 95-124.

Reichl L.E., 1980, A Modern Course in Statistical Physics, Edward Arnold, London.

Sneader W., 1996, Drug Prototypes and their Exploitation, Wiley, New York.

Stanley M.H.R., Nunes Amaral L.A., Buldyrev S.V., Havlin S., Leschhorn, H., Maass P., Salinger M.A., Stanley H.E., 1996, "Scaling behavior in economics: empirical results and modelling of company growth",

Nature, 319, p. 804.

Sutton J., 1997, "Gibrat's Legacy", Journal of Economic Literature, 35, pp. 40-59.

Sutton J., 1998, Technology and Market Structure, Theory and History, MIT Press, Cambridge, Ma.

Temin P., 1980, Taking your Medicine: Drug Regulation in the United States, Harvard University Press, Cambridge, Ma.

Warglien M., 1995, "Hierarchical Selection and Organizational Adaptation", Industrial and Corporate Change, 4, pp. 161-186.

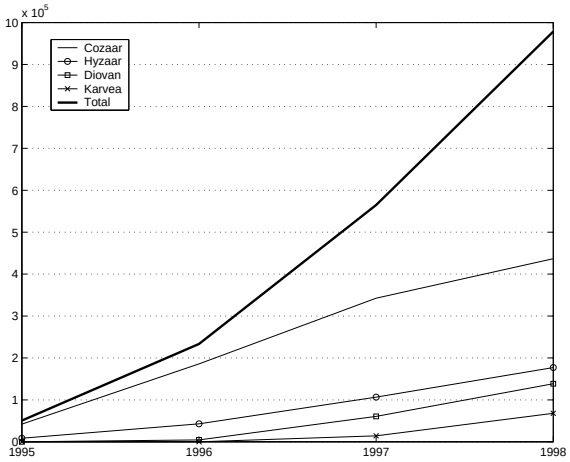


Figure 1: Angiotensin-II Antagonists: Size of the market and firms/products sales, 1987-1998 (US.$ millions)

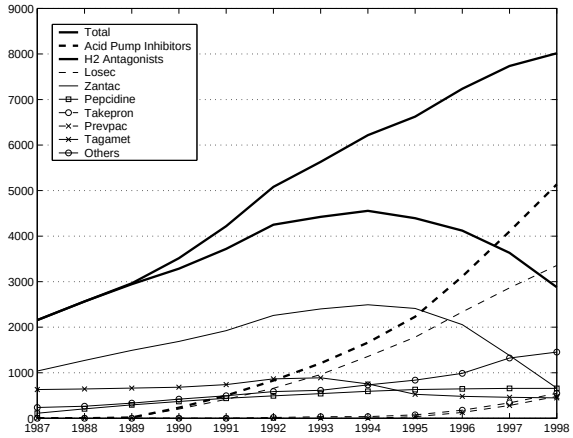


Figure 2: Antiulcerants: Acid pump inhibitors and H2 antagonists: Size of the markets and firms/products sales, 1987-1998 (US.$ millions)

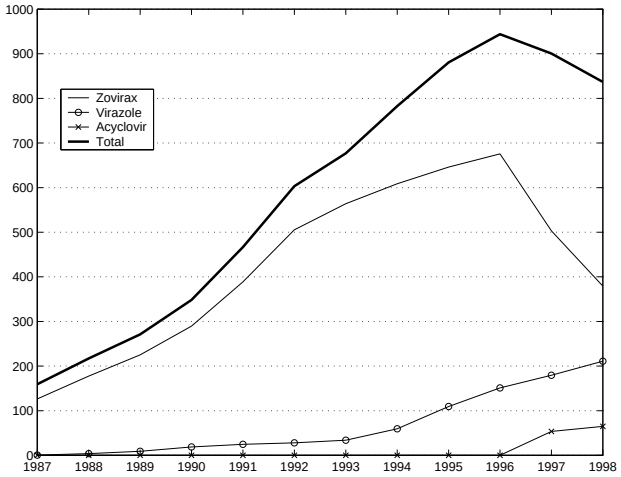


Figure 3: Antivirals, excl. vaccines: Size of the market and firms/products sales, 1987-1998 (US.$ millions)

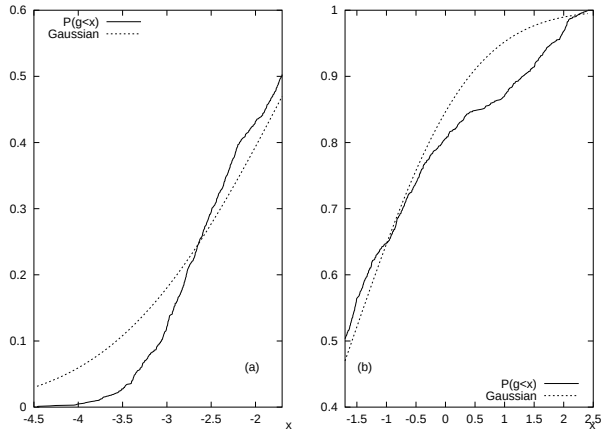


Figure 4: Distribution function of firm sizes (lower half in **(a)**, upper half in **(b)**). A fit with the normal distribution is also shown (dotted line).

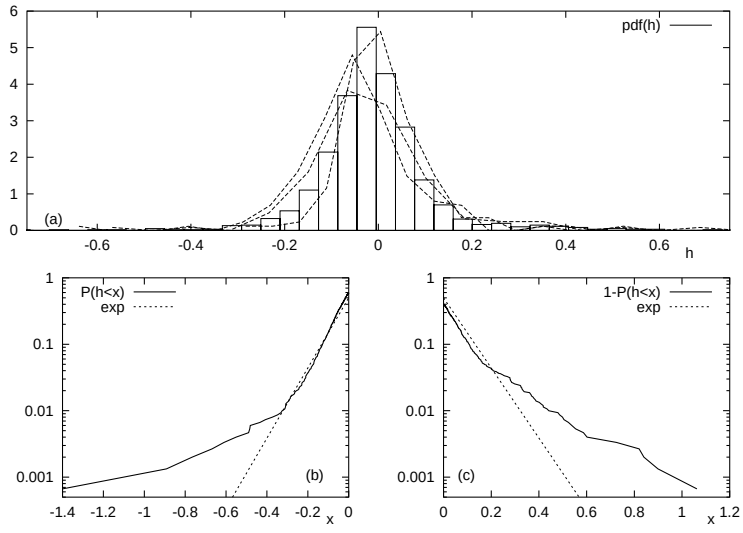


Figure 5: **(a)** Probability density for the growth obtained with bins (quantiles) of 100 values. The bars show the "average" distribution and the dotted lines the distributions at different time steps; **(b)** Distribution function for $h$ (lower half). **(c)** Distribution function for $h$ (upper half). The fit is performed with the exponential distribution

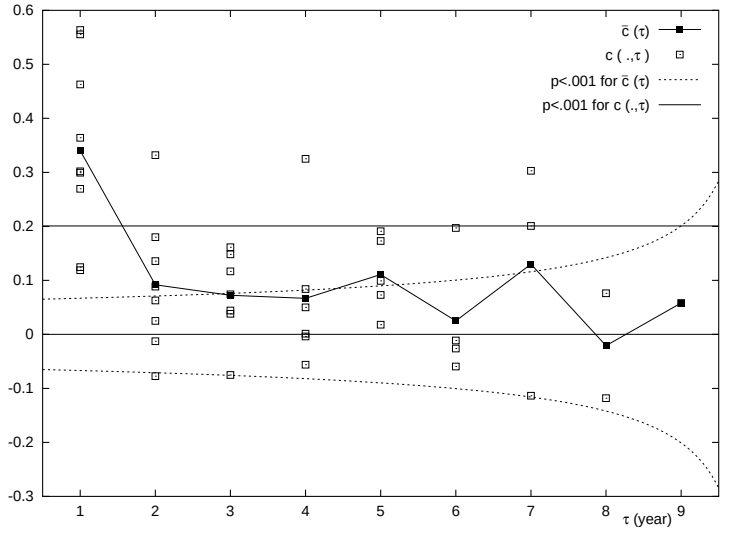


Figure 6: Time autocorrelation of the firm growth. The points are the values $c(t, \tau)$ for different time $t$ plotted against $\tau$. The line is the mean value $\bar{C}(\tau)$. The significance line for $p = .001$ (about 2.46 standard deviations) is plotted for both the single points and the average.

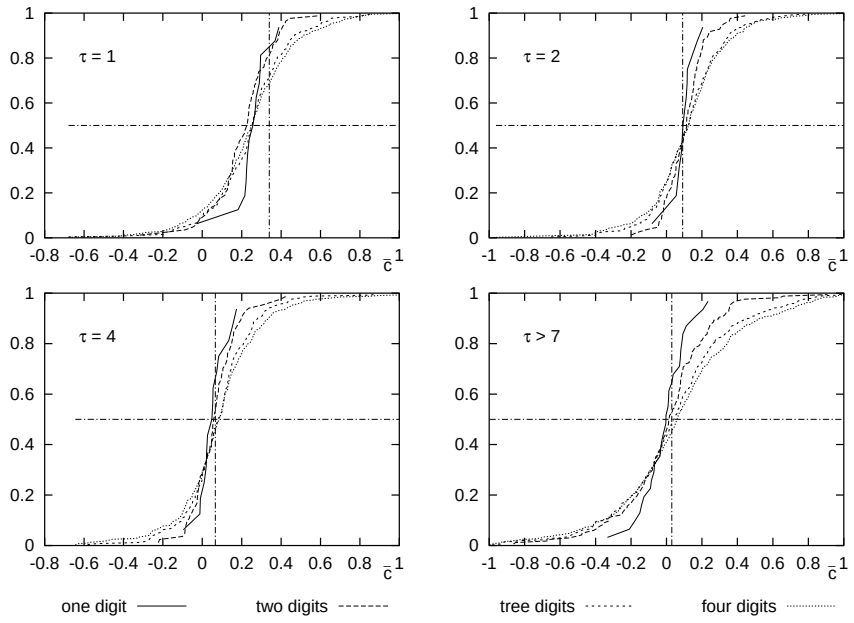


Figure 7: Distribution function of the autocorrelation coefficients $\overline{c}_j(\tau)$. Different aggregation levels and time lags $\tau$ are plotted. The vertical lines correspond to "aggregate" values at the firm-level. The horizontal line is centered at 0.5.

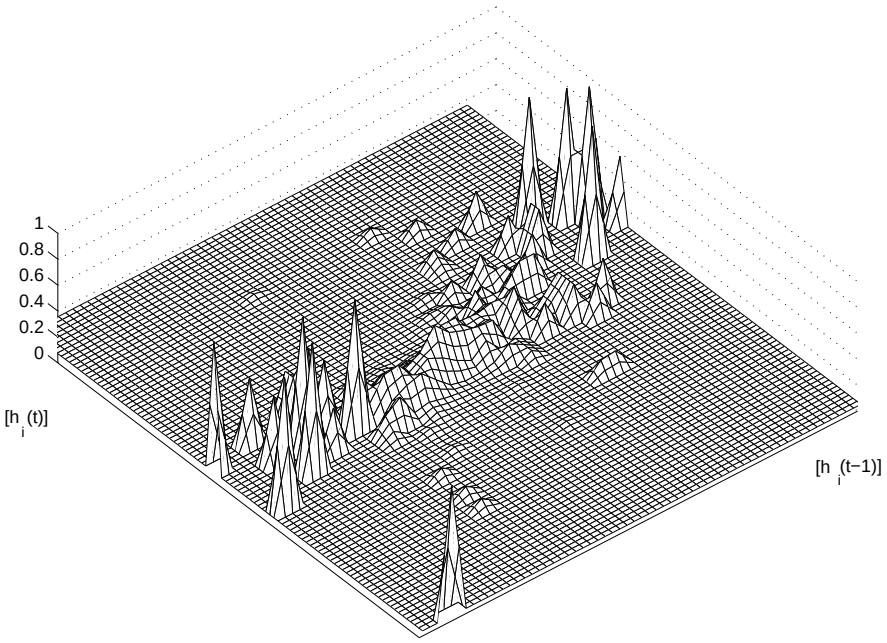


Figure 8: Growth transition matrix at the aggregate level

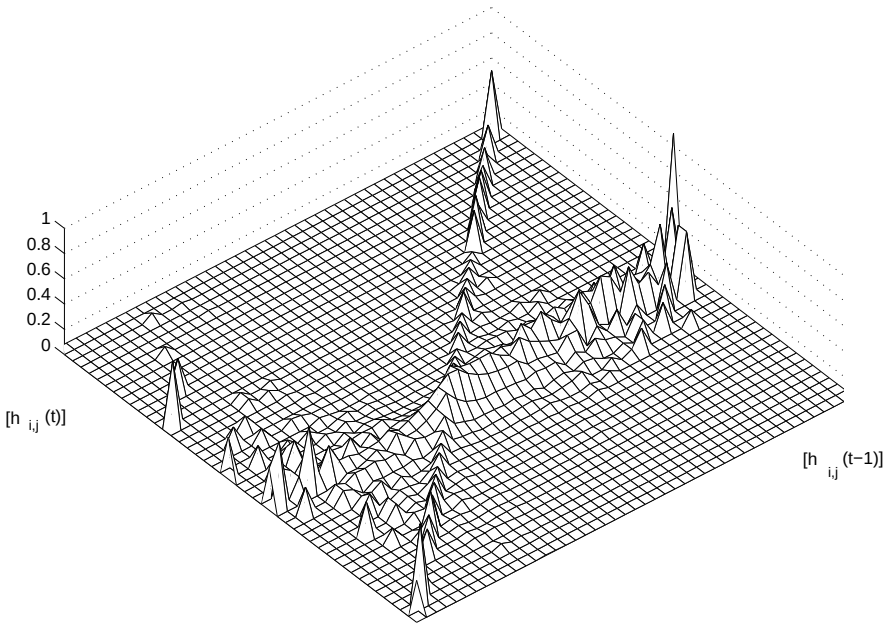


Figure 9: Growth transition matrix at the submarket level

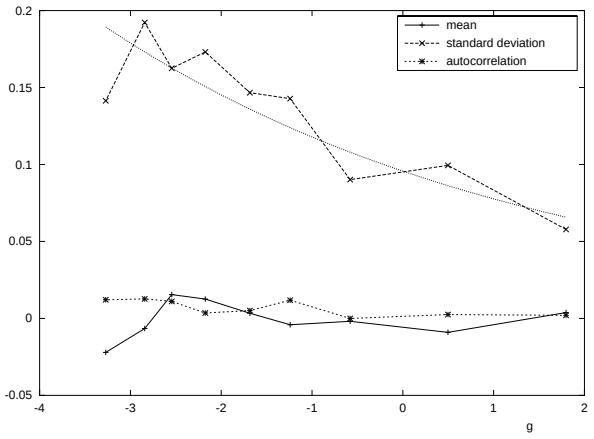


Figure 10: Mean, standard deviation and autocorrelation of growth $h$ computed for different size bins plotted againts the average size in the bin. The exponential fit to the standard deviation (see (5)) gives a value $\beta = -0.20 +/- 0.03$

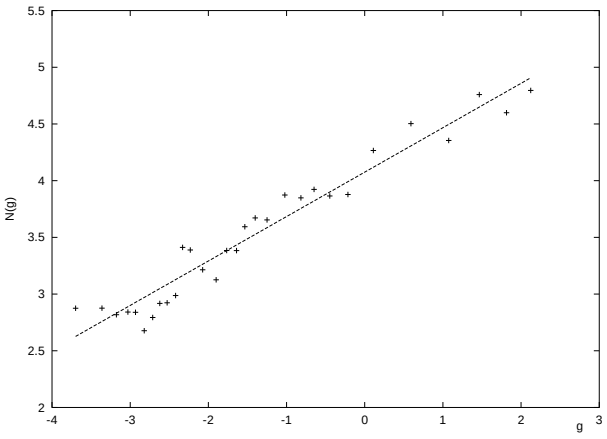


Figure 11: Number of submarkets a firm operates in vs. firm's size (log-log scale)

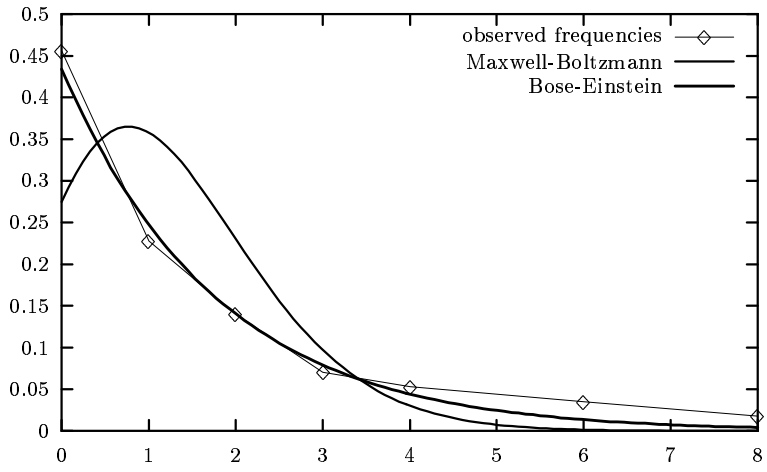


Figure 12: The frequencies of total NCEs introduced over the firms population (x-axis: firms introducing 0, 1, ..., 8 NCEs; y-axis: frequencies thereof).

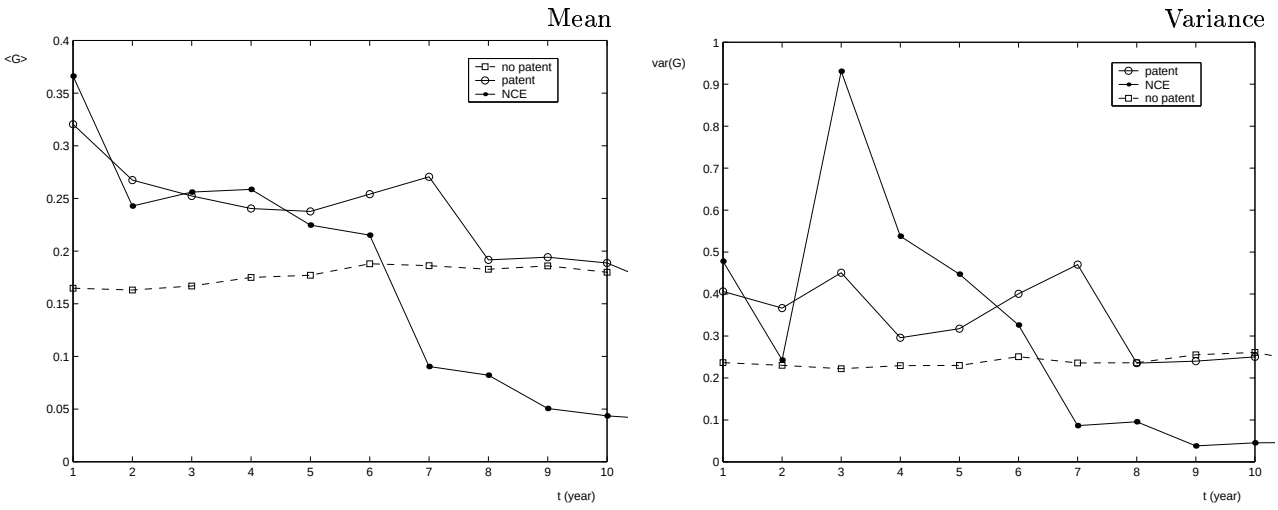


Figure 13: Mean and vaiance of sales disribution for NCEs, patented and no patented products. At any given time we consider only submarkets with at least 10 products. The product sales are rescaled by the average sales of the submarket where the product is launched.

| t | $\beta$ | $\xi$ |
|---|---|---|
| 1988 | 0.982 | 0.998 |
| 1989 | 0.991 | 0.989 |
| 1990 | 1.006 | 1.006 |
| 1991 | 1.008 | 0.958 |
| 1992 | 1.017 | 1.004 |
| 1993 | 0.997 | 0.982 |
| 1994 | 1.004 | 0.938 |
| 1995 | 1.001 | 0.953 |
| 1996 | 1.000 | 0.959 |
| 1997 | 1.005 | 0.978 |

Table 1: Gibrat test results - aggregate level, one-year time lag. $R^2$ are always greater than 0.98.

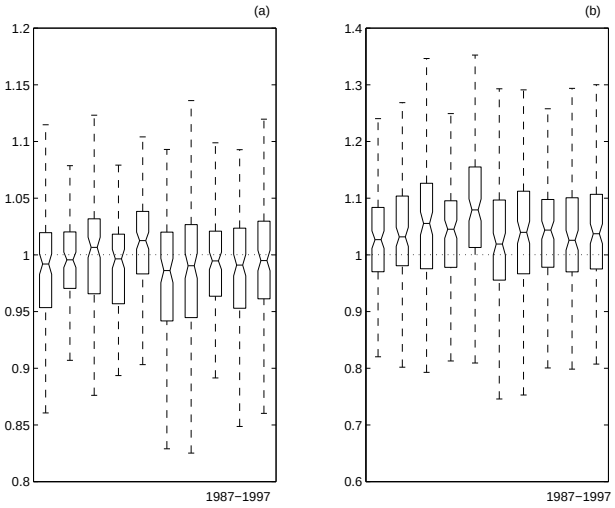


Figure 14: Gibrat test at the 3-digits submarket level, one year time lag. Boxplot **(a)** reports the distributions of $\beta_j(t)$ while **(b)** depicts the distributions of $\xi_j(t)$. (cf. (11))