Firm Growth Dynamics under Different Knowledge Regimes: the Case of the Pharmaceutical Industry

Pelin Demirel and Mariana Mazzucato
September 2007

NUMBER 63
Throughout the 1990s, The Open University has been developing its research capacity in economics. Economists at the OU comprise a lively and expanding group with a wide set of interests ranging from development policy to decision theory, from Marxist theories of profit to libertarian foundations of environmental policy and from econometric analysis of large data sets through institutional economics to the use of case-studies in policy formation. Nearly a 1000 students from around the world register each year to study economics courses and their needs, together with the multi-disciplinary nature of social science at the university, shape out research. Through a variety of personal and group research projects, our work makes a strong contribution to areas like business, public policy and even philosophy where sharply focused analysis can inform decision-making as well as contribute to scientific progress.

In 1999, approximately £250,000 million worth of externally funded grants (3 from the ESRC) were held by discipline members, some of whom also act as consultants to national and international bodies. Approximately half a dozen students are currently reading for doctorates with members of the discipline and we are always interested in proposals from colleagues or potential students who would like to do research with us.


The papers contain results of economic research which are the sole responsibility of the authors. Opinions expressed in these papers are hence those of the authors and do not necessarily reflect views of the University.
Firm Growth Dynamics Under Different Knowledge Regimes: 
the case of the Pharmaceutical Industry

Pelin Demirel*
Mariana Mazzucato*

Abstract:
The paper studies the dynamics of firm growth and firm size distributions in the pharmaceutical industry from 1950 to 2003 and in the biotechnology industry from the early 1980's to 2003. Growth dynamics are studied in the context of how the size composition of firms changes, how innovation patterns (patents) change, as well as locational decisions of firms (NJ vs. California). Results suggest that Gibrat's law (random growth) does not hold for the majority of the period under observation, and that the growth advantage of small pharma firms increases after the 1980's as the process of innovation becomes more ‘guided’ and scale intensive and as small firm innovation becomes more ‘persistent’. Furthermore, at the end of the 1970's a ‘bimodal’ firm size distribution emerges in the pharmaceutical industry precisely when a new division of labor between large and small firms sets in. We find that firms located in California are smaller, faster growing and more innovative than those in NJ.

Keywords: firm growth, innovation, industry dynamics, pharmaceutical

* Economics Dept, Walton Hall, The Open University, Milton Keynes, MK7 6AA 
m.mazzucato@open.ac.uk, p.demirel@open.ac.uk
**Introduction**

The paper studies firm growth dynamics in the pharmaceutical and biotechnology industries (which from now on we refer to as pharma and biotech) between 1950 and 2003. The objective is to study the properties of firm growth in a particularly innovative sector which has undergone intense changes in its knowledge base over the last 50 years. We ask whether the (time series) patterns of firm growth have changed alongside such transformations, and in particular, whether the degree to which firm growth can be described as “random” (Gibrat’s law discussed below) — as opposed to more “structured” (e.g., due to various types of increasing returns) — has changed over time. Central to this question is the different growth behaviour of small and large firms, and the co-evolution of growth dynamics with the distribution of firm size. We also briefly investigate whether the locational choices of pharma firms (i.e. being closer to a biotech cluster) affect firm growth dynamics and firm size distributions.

Results suggest that Gibrat’s law (random growth) does not hold for the majority of the period under observation, and that the growth advantage of small pharma firms increases after the 1980’s as the process of innovation becomes more ‘guided’ and scale intensive and as small firm innovation becomes more ‘persistent’. Furthermore, in the end of the 1970’s a ‘bimodal’ firm size distribution emerges in the pharma industry precisely when a new division of labor between large and small firms sets in. As regards location, we find that firms located in California are smaller, faster growing and more innovative than those in NJ.

Section 1 discusses the underlying theoretical framework used to explore firm growth dynamics. Section 2 contains a short introduction to the history of the pharma industry. Section 3 describes the dataset used in the econometric analysis. Section 4 contains a discussion of the methodology. Results and a concluding discussion are found in Section 5.

1. **Theoretical Motivation: Gibrat’s Law**

   “Gibrat’s Law” or the “Law of Proportionate Effect” is a useful starting point for exploring firm growth dynamics. This law is often used as a null hypothesis (or “straw man”) to test whether firm growth can be described by a richer structure than just a simple random walk phenomenon (Evans 1987; Geroski and Machin 1993; Sutton 1997; Dosi 2005).
A cross sectional regression model for Gibrat’s law (GL) is formulated as:

\[ y_{i,t} = \alpha + \beta y_{i,t-1} + u_{i,t}, \quad u_{i,t} = \rho u_{i,t-1} + \varepsilon_{i,t}, \quad \varepsilon_{i,t} \sim iidN(\mu, \sigma^2) \]  

where \( y_{i,t} \) (the logarithm of firm size) is regressed on a lagged \( y_{i,t-1} \) for firm \( i \) at time \( t \) and where the error term accounts for potential serial correlation in firm growth. In this paper, we prefer to estimate (1) using panel estimation techniques thanks to the long time dimension of our data. The panel regression model described in Section 4 is similar to (1) and can be interpreted in the same way.

Previous literature has tested two different versions of the law: the strong version and the weak version. The strong version of GL holds that firm growth is a result of purely stochastic shocks under constant returns to scale, i.e., that in (1) \( \varepsilon \) is an i.i.d. random variable with zero mean, and \( \beta = 1 \) for every \( i \). The strong form further states that firms grow with systematically uncorrelated random shocks that don’t persist over time (Dosi, 2005, Bottazzi, Cefis and Dosi, 2002). If \( \beta > 1 \) Gibrat’s law is violated because large firms grow faster than small firms (in the extreme case leading to monopoly). If instead \( \beta < 1 \), the law is violated because small firms grow faster than large firms, i.e., reversion to the mean. The strong form does not imply of course that there are no systematic factors behind the growth process. It simply means that random events dominate growth even though systematic factors may still be present (Hart and Oulton, 1996).

The weak form of the law holds that given a classification of firms by size, the average and variance of growth rates do not significantly vary among classes. In other words, no relationship exists between growth and the initial size of the firm.

Empirical analyses have shown that both the strong and weak form of Gibrat’s law hold more for large firms (Hall, 1987; Evans, 1987; Geroski and Machin, 1993). Although there are systematic factors at the firm and industry level that affect the process of firm growth (e.g., innovation, advertising, demand), growth seems to be mainly affected by stochastic shocks only in the short term and in the case of new firms. Systematic factors show up more in the long-medium term (Acs and Audretsch 1990). On average, smaller
firms have a lower probability of survival, but those that survive grow proportionately faster than large firms (Marsilli, 2001).

In addition to testing the relationship between initial firm size and growth, previous research has concentrated on testing two other main assumptions of the strong version of Gibrat’s law: the lognormality of the firm size distribution (FSD) and the autocorrelation structure of firm growth. Findings reveal that empirical firm size distributions show different shapes across different industries. For example, some industries have more peaked firm size distributions. Moreover, there is evidence for bimodal FSDs in some industries which is in itself a strong violation of the log-normality assumption (Bottazzi et al. 2002, 2005; Cabral and Mata, 2003; Dosi, 2005; Fagiolo and Luzzi, 2006 and Lotti and Santarelli, 2004). Regarding the autocorrelation of growth, the findings are mixed. Some studies find no autocorrelation in growth (Geroski, 2003 and Toke and Dahl, 2004) while some find autocorrelated, and hence, persistent growth (Bottazzi et al, 2005, Abbring and Campbell, 2003).

The rest of the paper looks into these issues around firm size and growth as applied to the pharma and biotech industries, with particular attention to how such dynamics have co-evolved with changes in patenting behavior, changes in the relationship between large and small firms, and also some preliminary insights into locational decisions. Before moving on to the analysis of data, it is essential to provide a brief background of the pharma industry to familiarize the reader with what we mean by changing “knowledge regimes”.

2. A Short Insight into the History of the Pharma Industry
The origins of the pharma industry date back to the mid 19th century when it started as an extension of the chemicals industry (Malerba and Orsenigo, 2001). During the earliest days of this industry, no formal “phama” research took place, and a specialized pharma industry did not exist independent of the chemical industry until the 1930’s. The World War II period and the decade following the war witnessed increased public expenditure into this sector which turned the pharma industry into a R&D intensive industry. The pre-1980’s period is often called the “random search” phase, since a typical pharma research program would randomly screen through a massive number of molecules in search of a cure for a disease. The process of innovation was largely based on trial and
error rather than on a systematic knowledge base, hence firms could rely less on their existing knowledge for the next generation of discoveries (Gambardella 1996; Nightingale, 2000). The post-1980 period is instead referred to as the “guided search” phase because radical improvements in information technology, combinatorial chemistry, enzymology and bioinformatics (as well as increased public expenditure in pharma research and universities), radically increased the ability of scientists to understand the biological and chemical phenomena that underlie diseases as well as their ability to “design” an “ideal” molecule that would potentially cure a disease (Gambardella, 1995; Pisano, 2002, 2006; Henderson et al, 1999, Gilsing and Notebloom, 2006). Since in the guided search regime firms could build more on their existing knowledge base and research capabilities, innovation became more path-dependent and cumulative. Developments in the field of biotech are important milestones in the guided search regime since new biotech firms that entered the market in early 1980’s had a significant impact on industry practices (Henderson et al, 1999).

Furthermore, the 1980’s witnessed a new ‘innovative division of labor’ between large pharma firms and the smaller pharma and biotech firms which entered during this period (Gambardella, 1995, Lacetera and Orsenigo, 2001). Smaller firms tended to specialise in high risk, niche innovation activities and the larger firms more on marketing and distribution too costly for the smaller firms. Galambos and Sturchio (1998) argue that many small (mainly biotech based) firms found it relatively easy to enter, survive and exist during the eighties and early nineties as large and established pharma companies did not regard them as a direct threat to their market share. Indeed, the pharma industry is well known for its networks of small and large firms collaborating for research activities (Gambardella, 1995).

Our analysis in the next sections asks whether there are different firm growth dynamics at work during these different knowledge regimes.

3. Data
Two databases are used for the econometric analysis in this paper. The first dataset is an unbalanced panel dataset of 323 pharma firms that were quoted on the S&P 500 index between 1950 and 2003. The second data set is another unbalanced dataset of 563 biotech firms quoted on the S&P 500 index between 1970 and 2003 (while only very
few firms exist during the 1970’s). The databases were purchased from the S&P customized data department. They include descriptive and geographic information on pharma firms and biotech firms from 16 different countries (80% of the firms in both datasets have the USA as their country of domicile), along with industrial and financial data on these firms (all monthly data except for R&D, employment and net income which are only available on an annual basis).

Some firms have been subject to a merger or acquisition during the period under observation. This creates a potential bias as it introduces an artificial “growth” for those firms that merged with, or acquired a small firm. It also overestimates the exit rates as the acquired and merged firms are counted as exits even though their economic activity has not ceased. Given this issue, there are two options for treating the data. The first is to simply leave in the data in its raw form and ignore these events. In this case, the analysis is likely to suffer from the bias discussed above. The alternative method is Botazzi et al.’s (2001) methodology of forming “super-firms” by restructuring the dataset to adjust for mergers and acquisitions that have taken place between firms while under observation1. We have decided to utilize the first of these solutions and interpret our results with the limitations in mind. However, we have also formed an alternative dataset of super-firms so to check the robustness of our findings. The results proved robust irrespective of the dataset used.

We have used firm revenues as a proxy for their size as this is the most consistently reported data in the dataset, as opposed to other potential proxies such as employment (which have many missing values).

---

1 ‘Superfirms’ are formed by merging the data of the related firms (e.g. two firms that merged in 1980) from the start date (1950). This allows one to concentrate on natural growth as opposed to other forms of growth even though the procedure of forming superfirms unavoidably introduces a bias in the size analysis that follows. The information regarding the mergers and acquisitions, name changes and subsidiary- parent affiliations were collected for each individual firm by screening six databases: Who Owns Whom, International Hoover’s Company Profiles, International Corporate Affiliations, Directory Information, ABI US Business and US Company Capsules. If there was no mention of a merger or acquisition for a particular firm in any of the six databases, it was assumed that the firm remained independent throughout the observation period and did not need to be merged under into superfirm. In our dataset, 82 firms were either subsidiaries or subject to mergers/acquisitions or name changes and were merged under 32 superfirms. Five firms were removed from the database as each was found to be exactly identical to another firm that was traded under the same name but under a different Ticker Symbol. One firm was removed from the dataset as no information was available about this firm in our resources. As a result, the final adjusted dataset includes 267 firms.
The results are reported for the unadjusted sample and the results obtained from the super-firm dataset are available from the authors on request.

3.1. Descriptive statistics

In order to test both the strong and the weak form of Gibrat's law, we divide the pharma dataset into three size classes, using ‘relative’ real revenues. The real revenues data (revenues divided by Consumer Price Index) of all firms for all the available years were sorted in ascending order and the database was divided into three equal parts so to find the cut off points for the three size classes: small, medium and large. Firms whose revenues are less than $4.99 million are classified as small firms, firms that have between $4.99 and $182.274 million dollars of revenues are classified as medium sized and firms with revenues higher than $182.274 million are classified as large firms\(^2\). We find this classification useful since it allows a firm to move between different size classes from year to year depending on its performance (i.e. a firm can be classified as small in year 1980 and it can move on to the medium size class in 1985 given that real revenues are larger than $4.99 million in 1985). The biotech dataset is treated as one class only since a very large majority of biotech firms are small and it is not meaningful to divide the industry into further size classes. Instead, we keep the biotech industry as a single category and compare it with the three pharma size classes (we believe it is particularly important not to confuse small pharma with small biotech as is done in less disaggregated studies).

Firm Numbers are reported in Figure 1, and the distribution of each firm category in the whole population through the years is reported in Figure 2. As can be clearly seen, the share of small pharma firms and biotech firms in relation to the whole population has increased over time while the share of large pharma firms has declined. In the case of small pharma firms and biotech firms, the numbers peak after 1980 (i.e., the beginning of the ‘guided search’ regime).

To see whether there is a relationship between the size classes that grew the most and the changes in their innovation behavior; Figure 3 shows the percentage of patents

\(^2\) To check for robustness we use threshold values close to these figures, and find no qualitative changes in the results.
It is clear that the share of large firms in innovative activity has declined starting from the late 1970’s while that of small and medium sized firms increased, even though at a much lower rate compared with their large counterparts. Biotech firms increased their share in patenting very significantly becoming the biggest innovators in the whole dataset. The increasing role of small and medium sized pharma firms and biotech firms in innovative activity over time might be in part due to the increased number of firms in these categories, especially after 1980. Hence, we have also looked at how the innovative behavior of firms in each size class has changed over time: the share of firms with at least one patent within each category for each year, suggests that a larger percentage of small pharma firms (15-35%) and biotech firms (40-50%) have been able to obtain at least one patent after 1980.

Finally, we look at how persistently firms patent. We tracked three-year patent spells for the four firm categories. We call a firm a “persistent innovator” if it has managed to get at least one patent granted for three consecutive years. Figure 4 reports the share of each category in relation to the total number of persistent innovators. Prior to 1980, a majority of persistent innovators came from large pharma firms. In the post 1980 period, the most significant factor is the increased persistency in innovation of biotech firms. Similarly in the same period, the share of persistent small pharma firms increased to a level at which it is almost equally likely that a persistent innovator comes from the large or small pharma firm size class. Note that although medium sized pharma firms have also increased their share of persistent innovators, they are less likely to produce persistent innovators compared with large and small size classes—suggesting a sort of U-Shape in size and persistency.

---

3 Patent data is taken from the NBER patent database (Jaffe and Trajtenberg 2001). Due to truncation problems with patent data (explained in Hall et al. 2001), we only exhibit the data to 1995. Truncation can be dealt with more rigorously using either a fixed effects approach or a structural lag model (to deal with both the changing number of patents as well as changing citations), but our use of patent data in the paper is limited. We simply cut off the data before the truncation problem becomes significant (Hall et al. 2001).
4. Methodology

Having gained preliminary insights into the changing number of firms by size class, we now look more carefully at growth dynamics by testing Gibrat’s law in its various forms discussed above. Gibrat’s law is typically tested using cross sectional Galtonian regressions (as in (1)) in which (log) firm size at time t is regressed onto initial (log) firm size (Hart and Oulton, 1996).

Studies that use the Galtonian regressions work with a large number of firms observed over a small number of years (see Botazzi et al., 2005). Our dataset has a smaller number of firms since we study only two industries (as opposed to the entire manufacturing sector). Yet our data covers a much longer time span compared with these studies and hence allows us to use panel estimation techniques.

Common criticism directed to the cross-sectional Galtonian regressions is that they do not take into account firm heterogeneity and assume that there is a common equilibrium point towards which all firms revert to if Gibrat’s law does not hold. Panel data methods have the advantage of relaxing this strict assumption of common mean reverting equilibrium for all firms by introducing fixed time and firm effects in the regressions. Moreover, the panel estimations clearly maximise the number of observations used and hence produce more reliable estimates.
As formulated by Goddard et al. (2002) the panel regression models for Gibrat’s law can conveniently be written in a similar fashion to (1), including an error term with an autoregressive structure of order one and coefficients $\alpha_i$ and $\delta_i$, namely; fixed firm and time effects:

$$y_{i,t} = (1 - \rho)\alpha_i + (\delta_i - \rho\delta_{i-1}) + \beta y_{i,t-1} + \rho(y_{i,t-1} - y_{i,t-2}) + \vartheta_{t,i},$$

(2)

where $\vartheta_{i,t} = \epsilon_{i,t} + \rho(1 - \beta)y_{i,t-2}$ (note that $\vartheta_{i,t} = \epsilon_{i,t}$ when Gibrat’s law holds and $\beta = 1$).

We introduce fixed firm effects since we believe the unobserved firm specific effects to be correlated with the right hand side variables. Following Goddard et al.’s (2002) methodology, we estimate two different panel regression models using OLS. Model 1 includes only fixed time effects (i.e. the firm specific effects are pooled $\alpha_i = \alpha$):

$$y_{i,t} = (1 - \rho)\alpha + (\delta_i - \rho\delta_{i-1}) + \beta y_{i,t-1} + \rho(y_{i,t-1} - y_{i,t-2}) + \vartheta_{i,t},$$

(Model 1)

while Model 2 includes both fixed time and firm effects:

$$y_{i,t} = (1 - \rho)\alpha_i + (\delta_i - \rho\delta_{i-1}) + \beta_i y_{i,t-1} + \rho(y_{i,t-1} - y_{i,t-2}) + \vartheta_{i,t},$$

(Model 2)

For pharma firms, we estimate Model 1 and Model 2 for each decade as well as the random search period (1950-1980) and the guided search period (1980-2003). For biotech firms, our estimation period is 1980-2003, again with results for each decade. The $\beta_1$ and $\beta_2$ coefficients are reported in Table 1 with the relevant standard errors and number of observations included in each estimation.

The limitation of using Model 1 is that it does not allow for individual firm specific effects even though the $(\beta_1 - 1)$ has a standard distribution and it is easy to test whether $\beta_1 = 1$. On the other hand, while Model 2 allows for the firm heterogeneity, $\beta_2$ has a downward bias in limited samples where $T$ is small (Greene, 2003). Our data set has a 54 year span and the downward bias is especially significant when we look at 10 year periods of the data. However as pointed out in Goddard et al. (2002), this is a common issue for most panel data methods and panel data unit root tests that deal with a similar problem to ours. Here, we demonstrate the $\beta_2$ coefficients for each decade to allow us to discuss the relative growth performance of small and large firms while we check whether Gibrat’s law holds (i.e. $\beta_2 = 1$) by using ADF type panel unit root tests which allow for individual and period specific effects and is exactly formulated like Model 2 (See Harris and Trainor, 2005 and Bentzen et al, 2005 for similar applications).
In addition to testing the strong form of Gibrat’s law, we use *kernel density estimation* techniques to explore the normality assumptions made regarding the firm size distribution and growth distribution\(^5\). Following the methodology of previous works on firm size distributions, we use a Normal Kernel function with an automatic Silverman bandwidth to estimate the FSD (Bottazzi and Secchi, 2005; Fagiolo, 2006; Higson 2002, 2004 and Lotti, 2004). The results remain robust to the choice of the kernel function.

5. Results and Discussions

Table 1 reports the results of the regressions specified in (3) estimated using OLS for the three samples: pharma, biotech and the combined dataset (pharma and biotech). \(\beta_1\) and \(\beta_2\) denote the \(\beta\) coefficients in the respective panel data models with only the period specific fixed effect and the one with both period and firm specific fixed effects.

<table>
<thead>
<tr>
<th>Observation Period (Pharma Industry)</th>
<th>(\beta_1)</th>
<th>(\beta_2)</th>
<th>Number of Cross Sections</th>
<th>Total Panel Observations (Unbalanced)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1950-1959</td>
<td>0.971438*</td>
<td>0.770695*</td>
<td>14</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>(0.01207)</td>
<td>(0.063652)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1960-1969</td>
<td>0.963329*</td>
<td>0.642905*</td>
<td>24</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>(0.009623)</td>
<td>(0.042420)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^5\) The Kernel density estimate of a series \(X\) at a point \(x\) is:

\[
f(x) = \frac{1}{Nh} \sum_{i}^{\infty} K\left(\frac{x - X_i}{h}\right)
\]

(3)

Here, \(K(.)\) is the Kernel density function and \(N\) is the number of data points in the empirical distribution. The Kernel density function \(K(.)\) determines shape of the bumps and can be chosen to be a function such as Epanechnikov, Gaussian (normal), and Uniform etc. "h" is called the “bandwidth” which is the smoothing parameter. A larger bandwidth leads to a smoother curve.
### The Biotech Industry

<table>
<thead>
<tr>
<th>Observation Period (Biotech Industry)</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>Number of Cross Sections</th>
<th>Total Panel Observations (Unbalanced)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980-1989</td>
<td>0.906267* (0.020586)</td>
<td>0.472440* (0.051818)</td>
<td>97</td>
<td>414</td>
</tr>
<tr>
<td>1990-2003</td>
<td>0.926263* (0.009979)</td>
<td>0.552511* (0.022729)</td>
<td>403</td>
<td>2415</td>
</tr>
<tr>
<td>1980-2003</td>
<td>0.923670* (0.009066)</td>
<td>0.596352* (0.019288)</td>
<td>425</td>
<td>2829</td>
</tr>
</tbody>
</table>
The Combined Dataset

<table>
<thead>
<tr>
<th>Observation Period (Pharma and Biotech)</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>Number of Cross Sections</th>
<th>Total Panel Observations (Unbalanced)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980-1989</td>
<td>0.958349* (0.006745)</td>
<td>0.503227* (0.031475)</td>
<td>201</td>
<td>1040</td>
</tr>
<tr>
<td>1990-2003</td>
<td>0.960097* (0.005353)</td>
<td>0.553687* (0.016902)</td>
<td>642</td>
<td>3981</td>
</tr>
<tr>
<td>1980-2003</td>
<td>0.959451* (0.004422)</td>
<td>0.606987* (0.013530)</td>
<td>692</td>
<td>5021</td>
</tr>
</tbody>
</table>

*Significant at 1%

For each period in Table 1, $\beta_1$ is significantly smaller than 1, indicating that Gibrat’s law is violated because small firms grow faster than large firms. $\beta_2$ coefficients are smaller than $\beta_1$ due to the downward bias associated with this estimator. In each case, Hausman tests (not reported here) show that the firm specific fixed effects are indeed valid, rendering $\beta_2$ a more reliable estimator for our purposes. Due to the non-standard distribution of the estimator, we use the ADF-Fisher panel unit root test discussed in Section 3 to test for the validity of Gibrat’s Law. In Table 2, results of the panel unit root tests for firm growth ($y_{i,t} - y_{i,t-1}$) are reported for the different datasets and each decade under observation.

Table 2: ADF Fisher Unit Root Tests

<table>
<thead>
<tr>
<th>Period (Pharma Industry)</th>
<th>ADF- Fisher Chi Square</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1950-1959</td>
<td>20.513</td>
<td>0.8451</td>
</tr>
<tr>
<td>1960-1969</td>
<td>40.489</td>
<td>0.4487</td>
</tr>
<tr>
<td>1970-1979</td>
<td>88.136</td>
<td>0.0105</td>
</tr>
<tr>
<td>1980-1989</td>
<td>169.291</td>
<td>0.0004</td>
</tr>
<tr>
<td>1990-2003</td>
<td>440.957</td>
<td>0.0000</td>
</tr>
<tr>
<td>1950-1980</td>
<td>192.772</td>
<td>0.0000</td>
</tr>
<tr>
<td>1981-2003</td>
<td>583.279</td>
<td>0.0000</td>
</tr>
<tr>
<td>Period (Biotech Industry)</td>
<td>ADF- Fisher Chi Square</td>
<td>P value</td>
</tr>
<tr>
<td>1980-1989</td>
<td>164.871</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
In the case of pharma, Table 2 shows that the first two decades of our dataset (1950-1959 and 1960-1969) are characterised by a unit root (i.e., Gibrat’s law holds) while the rest of the periods for the pharma industry as well as the biotech industry samples are characterised by a stationary growth process where Gibrat’s law is violated because small firms grow faster. Interestingly, no qualitative differences in these results were found for the random search period (1950-1980) and the guided search period (1980-2003): Gibrat’s law does not hold for either of these periods.

**Figure 5**

![Beta Coefficient Plotted Over Time](image)

Figure 5 shows β coefficients plotted over time in the two panel regression models in Table 1 for the pharma industry sample. The coefficients tend to follow similar movements over years and as clearly seen in the case of $\beta_2$, the coefficient declines over time. A smaller β suggests that especially after 1970’s growth dynamics have favoured smaller firms over large firms *even more than before*. This is an interesting finding in light of the changes in the ‘innovative division of labor’ referred to above (with smaller firms investing in more risky niche innovation projects). The fact that small pharma and biotech firms have increased in number during the guided search regime
and they have also become more active and persistent innovators (see Figure 4), suggests that the new division of labor in the guided search regime may have produced a stronger correlation between the growth of small firms and their innovation (which we examine in more detail in our future work).

The fact that we find $\beta$ higher before the 1980’s, with a peak in the 1950’s, finds support in the literature on Gibrat’s Law which has found clear evidence for the law ($\beta>1$) only in the 1950’s—due to the greater presence in that period of large firms in the general economy (Hart 1962; Prais 1974; Singh and Whittington 1975).

**Firm Size Distribution**

The parametric methodology we have used so far has been criticized by Bottazzi et al. (2005) for its rigid assumptions regarding the normality of growth rates. The fact that we have found a $\beta$ lower than 1, casts further doubt on the normality assumptions. As discussed in Section 1, several studies have identified divergence from normality both in the growth and size distribution of firms. Hence we now explore the shape of the firm size distribution and growth distributions using a non-parametric estimation technique: the Kernel density estimation. Figure 6 shows the Kernel estimation of firm size (log) distributions for the pharma, biotech and the combined samples for four representative years (1985, 1990, 1995 and 1999).

**Figure 6: Kernel Density Estimates for Firm Size Distributions**

![Kernel Density Estimates for Firm Size Distributions](image-url)
The most striking result is that the Firm Size Distribution (FSD) for the pharma industry has a bi-modal character unlike the normal distribution. Bottazzi and Secchi (2005) have noticed this peculiarity of the FSD for the pharma industry and their multi-modality tests formally confirm the bi-modal shape of the FSD. They conclude that MODE 1 (on the right tail) is due to a stable “core” of the industry that has persistently held a dominant position in the industry. Studying the evolution of the FSD over time, it is apparent that the emergence of Mode 2 coincides with the early 1970’s and the shape becomes stable, after 1980.

Note: The Normalised Firm Size is plotted in each estimation (Normalised Firm size= Firm Size at time t-Average firm size in the sample at time t)
While the FSD for the biotech industry is more unimodal, the combined dataset of pharma and biotech firms shows a bimodal character giving further support to the possibility of a “core” and “fringe” structure in which small firms (both pharma and biotech) coexist with large pharma firms.

**Concentration and Instability**

To further explore the possibility of the core-fringe structure in the pharma industry, we plot the 4 firm concentration ratio (CR4) in Figure 7 we also consider how the variance of firm shares changes over time in Figure 8. The variance of firm shares is calculated based on the Herfindahl Hirschman index (HHI) (Martin 1998):

\[
\sigma^2 = \sum_{i=1}^{n} \left( \frac{1}{n} - s_i \right)^2 = H - \frac{1}{n}
\]

where H is the standard Herfindahl index.

Hence the statistical variance of firm shares, \( \sigma^2 = \frac{H - 1/n}{n} \).

The CR4 shows that the share of the top 4 firms has declined over time, consistent with the rise of a fringe (MODE 2) that consists of small firms which, according to our results, have a growth advantage over large firms.

Figure 7
In fact, the result that small firms have a growth advantage over larger firms (especially after the 1980’s) provides some insight into why MODE 2 emerged.

Further investigation of the firms in MODE 1 and MODE 2 reveal more interesting facts.
First of all, the identity of the firms in MODE 1 persists over time i.e. around 80% of the firms that are in MODE 1 in mid 1970’s are still in MODE 1 in late 1990’s. This is true even though certain periods are characterized by a shakeup of market shares. Most of the instability in market shares occurs during particularly innovative periods. Figures 9 and 10 show that the market share instability index (tracking absolute changes in market shares over time) in both pharma and biotech were in fact highest during the period in which citation weighted patents (a proxy for the importance and/or radicalness of innovations) were highest—providing support to the idea of Tushman and Anderson (1986) that ‘competence destroying innovations’ which disrupt market structure. While CR4 ratio shows a pretty stable market concentration especially after 1970’s, the market instability index for the top 10 firms in the pharma industry (Figure 11) indicates an increased shake up of the market shares for the large firms. Hence, even though the core exists and is stable, there is still strong competition between the core players.

6 The market share instability index introduced in Hymer and Pashigian (1962) is calculated as:

\[ I = \sum_{t=1}^{n} \left| s_{it} - s_{i,t-1} \right| \]

where \( s_{it} \) = the market share of firm \( i \) at time \( t \).
Location
We find signs of a locational separation between the firms of the two modes. We have considered the top two most populated locations in our pharma industry database: New Jersey (11% of the pharma dataset) and California (19%). The New Jersey (NJ) cluster consists of large and old firms while the California (CA) cluster consists of much smaller and younger firms, founded close to the well known biotech cluster in Sorrento Valley, San Diego. The investigation of firm growth rates in these two regions in relation to average industry growth (as shown in Figures 12 and 13) reveal that for most of the period CA firms grow much faster compared with NJ firms, confirming our finding of the significant growth advantage of small firms.

Figure 12: Firm Growth Rates: CA

![Firm Growth: CA vs. Industry Average](image)

Figure 13: Firm Growth Rates: NJ

![Firm Growth: NJ vs Industry Average](image)
Moreover, as Figure 14 shows, as well as showing higher growth performance, there are more firms that patent in the CA cluster compared with the NJ cluster, indicating that CA firms are more innovative on average.

**Figure 14**

![Bar chart showing patenting rates in NJ, CA, and Other regions.]

The bimodality structure we observe for the whole of the pharma industry does not apply to either of the regions. The FSD in both the NJ and CA clusters is unimodal suggesting the absence of a fringe-core structure *within* either of the clusters. However, a closer look into the firms in MODE 1 and 2 reveals that MODE 1 has a significant number of NJ firms (33%-42% of the MODE 1 firms are NJ firms) while between 90% and 100% of the CA firms are MODE 2 firms. Hence, the bimodality structure might have potential roots in the locational choices of firms. Small pharma firms in MODE 2 located closer to the innovative biotech cluster in CA grow faster and innovate more while the established, slower growth NJ firms reside in MODE 1.

The interesting question is how the core firms in MODE 1 persist, notwithstanding the faster growth of small firms. Do they have better firm specific capabilities (innovative and other), or is it a matter of slowness in the “selection mechanism” as Lotti and Santarelli (2004) argue? Or perhaps the changes in the innovative division of labour (Gambardella and Arora 1994) have allowed the core large firms to find a new role in the
industry (e.g., as marketers/distributors of innovations which have their source in smaller firms), which allows them to persist even without being particularly innovative. How do the locational choices of pharma industry affect the FSD? Does being located close to biotech clusters give a special advantage to small pharma firms which seem to behave similar to biotech? These questions remain to be answered.

**Conclusions:**

The initial expectation of this paper was that the period of random screening might manifest itself in the form of random growth. Similarly, more path dependent innovation in the guided search regime should lead to more systematic and persistent growth. Surprisingly, the econometric analysis reveals that there is no such difference between the growth dynamics of the two knowledge regimes. Indeed, we reject the null hypothesis of Gibrat’s law for both periods with some evidence of random growth during the first two decades of the random search regime. Some interesting subtleties emerge:

First we find support for a negative size-growth relationship in both the pharma and biotech industries. Smaller firms in the pharma industry grow faster than larger firms starting from 1970’s. The specific characteristics of how innovation changed in the pharma industry from the 1970’s onwards provides us with insights into this change.

Second, we find that the size distribution of firms in the pharma industry cannot be described by a normal distribution. It is a bimodal distribution, indicating the simultaneous existence of a core set of firms and a fringe (Mode 1 and Mode 2). The biotech industry, on the other hand, shows a more similar character to the normal distribution. The combined sample of pharma and biotech firms also shows a bimodal FSD. The dynamics between the core and the fringe might have to do with changes in the innovative division of labour, or also due to changes in mergers and acquisitions.

Our further investigation of the two modes in the pharma industry reveals that the bimodality might be related to the locational choices of pharma firms. Those that prefer the California region, which is close to the biotech cluster in Sorrento Valley, reside in Mode 2, grow and innovate more compared with the rest of the pharma firms. Those that are founded in the New Jersey region form a persistent core in Mode 1, and are
relatively slower in their growth and innovative performance compared to the rest of the industry.

Dosi (2002) notes that rejecting Gibrat’s law is indeed good news for evolutionary economists because it provides evidence for some sort of a *structure* in the industry which is a natural outcome of the persistent heterogeneity of firms and the competitive market selection among them. Yet it is difficult to speak about ‘structure’ without better understanding the underlying changes in production and technology and how these have impacted the dynamics between small and large firms.

In fact, our work has contributed precisely to this goal of better understanding the structural dynamics underlying changing growth dynamics. We find that small pharma firms grow faster than large firms, *more so* in the guided search regime, which the case study literature links with changes in innovation networks between large and small firms (Gambardella 1995). Not only does this help us better understand the degree to which Gibrat’s law does not hold (i.e. it holds less so in the guided search phase), but also helps us better understand the possible source of the bimodal firm size distribution. We observe that the second mode in the pharma industry starts during the early 1970’s and becomes very prominent from the 1980’s. Our analysis suggests that the significant advantage of smaller firms (in growth) may have led to the emergence of a mass of smaller firms that provide an alternative to the well established firms located at the industry core. Moreover, locational dynamics could be important factors in understanding the structure behind the bimodality.

These findings probe further research questions that should aim to unveil the relationship between innovative performance and growth performance as well as the relationship between small, large pharma firms and the biotech firms. Future research will concentrate on understanding whether there are any differences in the growth dynamics of firms in relation to the quantity and quality (particular characteristics) of their innovation.
References:


Bothner, M. S. (2005), Relative size and firm growth in the global computer industry. Industrial and Corporate Change 14, 617-638.


Mazzucato, M. (2003), Risk, Variety and Volatility: Innovation, Growth and Stock


**Titles available in the series:**

Number 1  Valuing the environmental impacts of open cast coalmining: the case of the Trent Valley in North Staffordshire  
*Andrew B Trigg and W Richard Dubourg, June 1993*

Number 2  Scarcity and stability in a very simple general equilibrium model  
*Vivienne Brown, February 1994*

Number 3  A conflict model, with rational expectations, of the disinflation of the early 1980s  
*Graham Dawson, February 1994*

Number 4  Foreign Investment, Globalisation and International Economic Governance  
*Grahame Thompson, May 1994*

Number 5  Testing the Small Country Hypothesis for Developing Countries  
*Jonathan Perraton, December 1994*

Number 6  The Discovery of ‘Unpaid Work’: the social consequences of the expansion of ‘work’  
*Susan Himmelweit, June 1995*

Number 7  Exit, Voice and Values in Economic Institutions  
*Graham Dawson, June 1995*

Number 8  Residential Summer Schools Attendance and Students’ Assessed Performances on Open University Foundation Courses  
*Alan Gillie and Alan Woodley, June 1995*

Number 9  Putting Words into People’s Mouths? Economic Culture and its Implications for Local Government  
*Maureen Mackintosh, December 1995*

Number 10  What is a Fair Wage? A Critique of the Concept of the Value of Labour-Power  
*Susan Himmelweit, December 1995*

Number 11  The Origin of the Poverty Line  
*Alan Gillie, December 1995*

Number 12  The Determinants of Product and Process Innovations  
*Roberto Simonetti, Daniele Archibugi, Rinaldo Evangelista, February 1996*

*Roberto Simonetti, February 1996*

Number 14  Utilities vs. Rights to Publicly Provided Goods: Arguments and Evidence from Health-Care Rationing  
*Paul Anand and Allan Wailoo, January 2000*

Number 15  Proceeding to the Paddling Pool: The Selection and Shaping of Call Centre Labour  
*George Callaghan and Paul Thompson, January 2000*

Number 16  Doing ‘Qualitative Research’ in Economics: Two Examples and Some Reflections  
*Elizabeth Hill and Gabrielle Meagher, November 1999*

Number 17  Veblen, Bourdieu and Conspicuous Consumption  
*Andrew B Trigg, January 2000*
Number 18  The Effect of Idiosyncratic Events on the Feedback between Firm Size and Innovation  
  Mariana Mazzucato, January 2000

Number 19  Non-market relationships in health care  
  Maureen Mackintosh and Lucy Gilson, January 2000

Number 20  Selling pollution and safeguarding lives: international justice, emissions trading and the Kyoto Protocol  
  Graham Dawson, October 2000

Number 21  Entrepreneurship by Alliance  
  Judith Mehta and Barbara Krug, September 2000

Number 22  A disorderly household - voicing the noise  
  Judith Mehta, October 2000

Number 23  Sustainable redistribution with health care markets? Rethinking regulatory intervention in the Tanzanian context  
  Maureen Mackintosh and Paula Tibandebage, November 2000

Number 24  Surplus Value and the Keynesian Multiplier  
  Andrew B Trigg, October 2000

Number 25  Edwards Revised: Technical Control and Call Centres  
  George Callaghan and Paul Thompson, November 2000

Number 26  Social Norms, Occupational Groups and Income Tax Evasion: A Survey In The UK Construction Industry  
  Maria Sigala, November 2000

Number 27  Procedural Fairness in Economic and Social Choice: Evidence from a Survey of Voters  
  Paul Anand, December 2000

Number 28  Alternative rationalities, or why do economists become parents?  
  Susan Himmelweit, December 2000

Number 29  Agglomeration and Growth: A Study of the Cambridge Hi-Tech Cluster  
  Suma Athreye, December 2000

Number 30  Sources of Increasing Returns and Regional Innovation in the UK  
  Suma Athreye and David Keeble, January 2001

Number 31  The Evolution of the UK software market: scale of demand and the role of competencies  
  Suma Athreye, September 2000

Number 32  Evolution of Markets in the Software Industry  
  Suma Athreye, January 2001

Number 33  Specialised Markets and the Behaviour of Firms: Evidence from the UK’s Regional Economies  
  Suma Athreye and David Keeble, January 2001

Number 34  Markets and Feminisms  
  Graham Dawson, January 2001

Number 35  Externalities and the UK Regional Divide in Innovative Behaviour  
  Suma Athreye and David Keeble, January 2001

Number 36  Inequality and redistribution: analytical and empirical issues for developmental social policy  
  Maureen Mackintosh, March 2001
<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
<th>Authors</th>
<th>Authors' Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>37</td>
<td>Modelling the Dynamics of Industry Populations</td>
<td>Mariana Mazzucato and P A Geroski</td>
<td>January 2001</td>
</tr>
<tr>
<td>39</td>
<td>The Determinants of Stock Price Volatility: An Industry Study</td>
<td>Mariana Mazzucato and Willi Semmler</td>
<td>February 2001</td>
</tr>
<tr>
<td>40</td>
<td>Surplus Value and the Kalecki Principle in Marx’s Reproduction Schema</td>
<td>Andrew B Trigg</td>
<td>March 2001</td>
</tr>
<tr>
<td>41</td>
<td>Risk, Variety and Volatility in the Early Auto and PC Industry</td>
<td>Mariana Mazzucato</td>
<td>March 2003</td>
</tr>
<tr>
<td>42</td>
<td>Making visible the hidden economy: the case for gender impact analysis of economic policy</td>
<td>Susan Himmelweit</td>
<td>August 2001</td>
</tr>
<tr>
<td>43</td>
<td>Learning and the Sources of Corporate Growth</td>
<td>Mariana Mazzucato and P A Geroski</td>
<td>June 2001</td>
</tr>
<tr>
<td>44</td>
<td>Social Choice, Health and Fairness</td>
<td>Paul Anand</td>
<td>September 2002</td>
</tr>
<tr>
<td>45</td>
<td>The Integration of Claims to Health-Care: a Programming Approach</td>
<td>Paul Anand</td>
<td>November 2002</td>
</tr>
<tr>
<td>46</td>
<td>Pasinetti, Keynes and the principle of Effective Demand</td>
<td>Andrew B Trigg and Frederic S Lee</td>
<td>June 2003</td>
</tr>
<tr>
<td>47</td>
<td>Capabilities and Wellbeing: Evidence Based on the Sen-Nussbaum Approach to Welfare</td>
<td>Paul Anand, Graham Hunter and Ron Smith</td>
<td>January 2004</td>
</tr>
<tr>
<td>48</td>
<td>Entry, Competence-Destroying Innovations, volatility and growth:</td>
<td>Mariana Mazzucato</td>
<td>June 2004</td>
</tr>
<tr>
<td></td>
<td>Lessons from different industries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Taking risks with ethical principles: a critical examination of the ethics of ‘ethical investment’</td>
<td>Graham Dawson</td>
<td>November 2004</td>
</tr>
<tr>
<td>50</td>
<td>Innovation and Idiosyncratic Risk: an Industry &amp; Firm Level Analysis</td>
<td>Mariana Mazzucato and Massimiliano Tancioni</td>
<td>November 2005</td>
</tr>
<tr>
<td>51</td>
<td>Industrial Concentration in a Liberalising Economy: a Study of Indian Manufacturing</td>
<td>Suma Athreye and Sandeep Kapur</td>
<td>October 2004</td>
</tr>
<tr>
<td>52</td>
<td>Creating Competition? Globalisation and the emergence of new technology producers</td>
<td>Suma Athreye and John Cantwell</td>
<td>October 2005</td>
</tr>
<tr>
<td>54</td>
<td>Does International Trade Transfer Technology to Emerging Countries? A Patent Citation Analysis</td>
<td>Elif Bascavusoglu</td>
<td>August 2006</td>
</tr>
<tr>
<td>Number</td>
<td>Title</td>
<td>Authors</td>
<td>Date</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>55</td>
<td>Stock Price Volatility and Patent Citation Dynamics: the case of the</td>
<td>Mariana Mazzucato and Massimiliano Tancioni</td>
<td>December 2006</td>
</tr>
<tr>
<td></td>
<td>pharmaceutical industry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>Violent Crime, Gender Inequalities and Well-Being: Models based on a</td>
<td>Paul Anand and Cristina Santos</td>
<td>January 2007</td>
</tr>
<tr>
<td></td>
<td>Survey of Individual Capabilities and Crime Rates for England and</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wales</td>
<td></td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>Innovation and Firm Growth in High-Tech Sectors: A Quantile</td>
<td>Alex Coad (CES-Matisse) and Rekha Rao (LEM)</td>
<td>January 2007</td>
</tr>
<tr>
<td></td>
<td>Regression Approach</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Profile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>Prices of Production are Proportional to Real Costs</td>
<td>Ian Wright</td>
<td>January 2007</td>
</tr>
<tr>
<td>60</td>
<td>Temporary Work in Tuscany: a Multinomial Nested Logit Analysis</td>
<td>Lorenzo Corsini (Pisa University) and Marco Guerrazzi (Pisa University)</td>
<td>May 2007</td>
</tr>
<tr>
<td>61</td>
<td>Wage Bargaining in an Optimal Control Framework: A Dynamic</td>
<td>Marco Guerrazzi (Pisa University)</td>
<td>June 2007</td>
</tr>
<tr>
<td></td>
<td>Version of the Right-to-Manage Model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>Innovation and Knowledge Spillovers in Developing Countries</td>
<td>Elif Bascavusoglu</td>
<td>July 2007</td>
</tr>
<tr>
<td>63</td>
<td>Stock Price Volatility and Patent Citation Dynamics: the case of the</td>
<td>Mariana Mazzucato and Massimiliano Tancioni</td>
<td>September 2007</td>
</tr>
<tr>
<td></td>
<td>pharmaceutical industry</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>