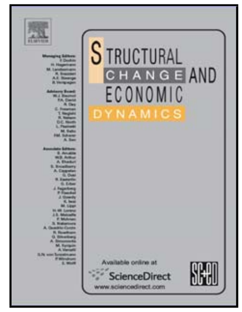

Title: Testing Gibrat's Legacy: A Bayesian Approach to Study the Growth of Firms

Authors: Elena Cefis, Matteo Ciccarelli, Luigi Orsenigo

PII: S0954-349X(07)00027-6
DOI: doi:10.1016/j.strueco.2007.04.002
Reference: STRECO 361

To appear in: *Structural Change and Economic Dynamics*

Received date: 2-2-2006
Revised date: 23-4-2007
Accepted date: 23-4-2007

Please cite this article as: Cefis, E., Ciccarelli, M., Orsenigo, L., Testing Gibrat's Legacy: A Bayesian Approach to Study the Growth of Firms, *Structural Change and Economic Dynamics* (2007), doi:10.1016/j.strueco.2007.04.002

Testing Gibrat's Legacy: A Bayesian Approach to Study the Growth of Firms*

Elena Cefis†
Università di Bergamo and T.C.Koopmans Research Institute, Utrecht University

Matteo Ciccarelli
European Central Bank

Luigi Orsenigo
Università di Brescia and CESPRI, Università Bocconi

March 2007

Abstract

Gibrat's law is a referent model of corporate growth dynamics. This paper employs Bayesian panel data methods to test Gibrat's law and its implications. Using a Pharmaceutical Industry Database (1987-1998), we find evidence against Gibrat's law on average, within or across industries. Estimated steady states differ across firms, and firm sizes and growth rates don't converge within the same industry to a common limiting distribution. There is only weak evidence of mean reversion: initial larger firms do not grow relatively slower than smaller firms. Differences in growth rates and in steady state size are persistent and firm-specific, rather than size-specific.

JEL classification nos: C11, C23, D21, L11, L25
Keywords: Gibrat's Law, Firm Growth, Pharmaceutical Industry, Heterogeneity, Bayesian Estimation

*The financial supports of the European Union (ESSY Project, 'Sectoral Systems in Europe: Innovation Competitiveness and Growth'), of EPRIS project (European Pharmaceutical Regulation and Innovation System), and the Italian Ministry of University and Research (grant n. 9913443984, and ex 60%, grant n.60CEFI06, Dept. of Economics, University of Bergamo) are gratefully acknowledged. Part of Ciccarelli's research was undertaken when he was at the University of Alicante and benefitted from funding by Ministerio de Ciencia y Tecnología and FEDER (BEC2002-03097), and Generalitat Valenciana (CTIDIB/2002/175). The views expressed in this paper are exclusively those of the authors and not those of the European Central Bank. The usual disclaimer applies.

†Corresponding Author: Elena Cefis, Università di Bergamo, Dep. of Economics, via dei Caniana 2, 24127 Bergamo, Italy. E-mail: elena.cefis@unibg.it

1 Introduction

After several years of neglect, industrial economists have recently devoted considerable attention to the processes of corporate growth. Most empirical studies in this field are based on testing the "Law of Proportionate Effects", also known as Gibrat's Law (Gibrat, 1931), which assumes that firms size follows a random walk and hence that firms growth is erratic. Consequently, there would be no convergence of firms size within or across industries, and no stable or predictable differences in growth would exist either in the short or in the long run. Rather, growth would be driven by small idiosyncratic shocks.[1]

Gibrat's Law was originally used as an explanation of the highly skewed firms size distributions. Even if the growth rate of each firm in an industry is unrelated to its current size, the variance of the firm size distribution and the level of concentration increase over time (Simon and Bonini, 1958, Ijiri and Simon 1974 and 1977). Subsequently, the Law of Proportionate Effects has become a referent model for discussing the processes of firms growth. Nevertheless, Gibrat's Law contrasts with most fundamental theories of firms growth (see Geroski, 1999). The Law of Proportionate Effects does not hold in standard models of convergence to an optimal size. More recent models of firms growth and industry dynamics assume heterogeneous firms, facing idiosyncratic sources of uncertainty and discrete events. Market selection affects firms, so that the most efficient grow while the others shrink and eventually leave the market. These models imply several violations of standard Gibrat-type processes. In particular, while (almost) all are able to generate skewed - typically log normal - distributions of firms size, they also predict that Gibrat's Law holds for large firms, while small firms - conditional on their survival - will grow faster than older, larger companies - i.e. mean reversion.[2] If firms started with equal size, they will converge to the same steady state. If firms started with different initial size, they will converge to different sizes. However, mean reversion implies that the initial differences in size should become smaller in the steady state.

Lucas (1978) suggests a model which yields a skewed distribution of firms size as a result of an optimal allocation problem, given a two inputs production function augmented with a third factor - managerial talent - unevenly distributed among the workforce. The model is compatible with Gibrat's Law only under specific conditions. . In Jovanovic's model (Jovanovic, 1982) firms are assumed to be endowed with time-invariant characteristics (i.e. efficiency parameters), whose true value is unknown to the firm. But, by producing, the firm

[1] For a recent discussion, see Sutton (1997), Bottazzi et al. (2001), and Wit (2005).

[2] Most models also predict that the variance in growth rates will fall with size and age.

learns about the true efficiency parameter. Consequently, the most efficient firms survive and grow, while the others shrink or exit from the market. Smaller, younger firms will grow faster than large ones - conditional on survival - since it takes time to learn about their efficiency, while large, older tend to remain close to their "optimal" size.

Ericson and Pakes (1995) make the same prediction. However, the "passive learning" model employed by Jovanovic is replaced by an "active learning" model, where firms are able- through investing - to improve their efficiency parameter. Also Hopenhayn's model (Hopenhayn, 1992) implies regression to the mean in firms size and convergence to different steady states sizes. Klette and Kortum (2004) develop a model which explicitly refers to Penrose's insight that firms growth is constrained by its internal resources. In Klette and Kortum model, a firm of any size adds new products by innovating, but its likelihood of success depends on knowledge capital accumulated through past innovations. Given specific assumptions on the arrival rate of new innovations, the model generates a stable skewed distribution of firm sizes, resulting from decreasing growth rates among surviving small companies while Gibrat's Law holds for large firms.

Finally, evolutionary models of firms growth based on organisational capabilities (Nelson and Winter 1982, Dosi et al., 1995; Winter et al., 2000) are intrinsically at odds with both Gibrat's Law and theories predicting convergence to the same "optimal" size". However, an evolutionary, competence-based approach is entirely consistent with the observation of the lack of any strong relationship between firms size and growth. Firms growth would be the outcome of the accumulation of new competences and market selection processes favouring more efficient and penalising less efficient companies, with size per se not attributing any intrinsic advantage or disadvantage to firms. To the extent that - and in technological regimes where - such competences evolve in a systematic and cumulative way, one would expect that firms growth processes exhibited a much stronger systematic component than predicted by Gibrat's Law. Conversely, new, small firms, if embodying superior technologies, may grow faster than old incumbents locked-in in their inferior capabilities. Thus, evolutionary models would predict convergence to different sizes, with mean reversion being highly sector-specific. Skewed distributions of firms sizes may result from aggregation of diverse firms behaviour and sectors characterised by different regimes of technological learning and market selection (Dosi et al. 1995; Dosi, 2005)

Moreover, Gibrat's Law is at odds with other observed empirical phenomena like the persistence of heterogeneity in some firms, performance measures , e.g. profits, productivity and - more controversially - innovation (see Baily and Chakrabarty, 1985; Mueller,1990;

Geroski et al., 1993; Cefis and Orsenigo, 2001; Cefis, 2003). However, the hypothesis that firms growth rates are erratic- at least for large, old firms - has often been considered a stylised fact (Geroski,1999). More generally, Gibrat's Law enters in the models and in the empirical discussion as a fundamental way of conceptualising firms growth (Klette et al., 2000; McCloughan, 1995, Geroski et al.2003) and models capable of yielding random growth (Sutton, 1997, Geroski et al., 1997; Wit, 2005).

A large body of empirical literature has explored this issue using different datasets and statistical methodologies. Typically, the the analysis starts with a simple econometric model having the following form:

$$\ln S_{it} = \beta_0 + \beta \ln S_{it-1} + u_{it} \tag{1}$$

where S_{it} is the size of firm i at time t, and u_{it} is an i.i.d. shock.

Gibrat's Law would be confirmed if the model $M_o : \beta = 1$ could not be rejected versus the alternative $M_1 : \beta < 1$. Empirical results are controversial. Some early studies (Hart and Prais, 1956; Simon and Bonini, 1958; Hymer and Pashigian, 1962) confirm the view that firms size follows a random walk ($\beta = 1$), at least for large firms (Hall, 1987; Lotti, et al., 2003). Nevertheless, a considerable body of results rejects M_o, suggesting that firms size is mean reverting (Baldwin 1995, ch.5; Baily et al., 2000) ; conditional on firm survival, average firm growth (and its variance) declines with size, holding age constant (Dunne et al., 1989; Evans, 1987a, 1987b; Hall, 1987, Caves, 1998).[3] In some versions, Gibrat's Law is considered to hold for large firms, whereas smaller companies grow faster, but with a higher variance.

This work reconsiders these issues by addressing four interrelated questions. The first two are quite conventional. First, we ask if Gibrat's law holds, by testing the random walk assumption. Second, we verify the existence and extent of mean reversion. The following two issues are simple, but less conventional. Specifically, third, we check whether firms converge to a common steady state (as implied by the mean reversion argument with firms starting with equal size) or to a firm-specific steady state size (which would occur if firms started with different initial size). Fourth, in the latter case we investigate if there is a non-negligible number of initial smaller firms able to catch-up or even to forge ahead: this result could confirm the existence of mean reversion for initially heterogeneous companies. Conversely, if initial size differences persist, the standard observation that small firms grow

[3]Differentials in growth rates have been explained with firms age (Mata, 1994; Dunne, Roberts and Samuelson, 1989)), firms size (Harhoff, Stahl and Woywode, 1998; Hart and Oulton, 1996; Hall, 1987; Evans 1987a and 1987b) or both (Farinas and Moreno, 2000).

faster than large ones would be weakened: firms growth would depend on firm-specific characteristics unrelated to size as such.

These exercises are prompted by the high degree of heterogeneity observed among firms, even in very narrowly defined industries and lines of business. Such heterogeneity might significantly influence the basic results obtained in the Gibrat's Law literature.

To answer these questions, we adapt a hierarchical Bayesian normal linear model (Lindley and Smith, 1972) to autoregressive panel data, i.e. data consisting of many time series generated by the same type of autoregressive model. The motivation for such statistical framework can be articulated in several points.

First, the autoregressive model is chosen for sake of consistency with previous empirical studies.

Second, the results on Gibrat's law and its implications crucially depend on the total variation,. Previous studies attempt to verify Gibrat's law using either classical modelling of time series and cross sectional data, or short-panel econometric techniques with homogeneity in parameters across units and over time. We consider these approaches as *problematic.*

First, cross section analysis ignores important information contained in unit-specific time variation in growth rates. Also, forcing the parameters to be the same across units is too restrictive. Recent studies (e.g. Pesaran and Smith, 1996) have shown that imposing the slope parameter to be homogeneous in an autoregressive panel data model distorts the estimation value of the parameter β towards the unit, irrespective of its true value, thus rendering the Gibrat's law test less powerful (see Goddard et al., 2002).

Alternatively, one can estimate equations like (1) by firm. Indeed, Geroski et al. (2003) tested Gibrat's Law on a sample of large British firms over a period of 30 years, allowing for heterogeneity in both intercept and slope of equation (1). The test is applied to each firm.Results show that firm size does not converge to either a common steady state or to a set of stable differences between firms. The main limitation of this procedure is that it can be applied only for large time spans , which can be problematic for micro data.

Second, the hierarchical model approach reduces the estimation variability and exploits coefficient similarities across firms without imposing the same population structure. Concretely, the model allows for an exchangeable scheme where the parameter vectors vary across firms, subject to a common distribution with unknown means and variance. In this sense, the model represents a satisfactory compromise between the regression model with the same coefficient for all firms and the time series regressions with different coefficients across firms.

Third, given the hierarchical model specification, the verification of Gibrat's law is based more on the autoregressive parameter of the common underlying distributions than on the identification of firm-specific coefficients. (Li, 1999).

Finally, in situations where several short time series are simultaneously modelled, the Bayesian paradigm is attractive by offerring a natural scheme for combining and weighting data from similar sources (Nandram and Petruccelli, 1997). The Bayesian estimation of this hierarchical model is also computationally straightforward due to recent advances in Bayesian statistics and Markov Chain Monte Carlo Methods (see Gelfand and Smith, 1990 and Hsiao et al., 1999).

All these considerations justify the use of the Bayesian hierarchical model whose main characteristic, complete heterogeneity, has the useful feature of exploiting all the information contained in the panel data set in a more powerful way.

Using a sample of 210 firms from a Pharmaceutical Industry Database covering the period 1987-1998, we do not find strong support for the law. Moreover, data show only weak evidence of mean reversion. Finally, differences in growth rates and in steady state size are firm-specific and, differently from the results of Geroski et al (2003), persistent. The specified model adequately fits the data and results prove to be robust to more general sets of prior information.

The paper is structured as follows. Section 1 discusses the statistical model. Section 2 describes data and comments on the estimation results. Section 3 checks the robustness of the results. Section 4 concludes

2 Model specification and Bayesian analysis

The evolution of size for all units is determined by a doubled indexed stochastic process $\{S_{it}\}$, where $i \in I$ indexes firms, $t = 0, 1, ..$ indexes time, and I is the set of the first n integers. Following Sutton (1997), if ε_{it} is a random variable denoting the proportionate rate of growth between period $t-1$ and t for firm i, then

$$S_{it} - S_{it-1} = \varepsilon_{it} S_{it-1}$$

and

$$S_{it} = (1 + \varepsilon_{it}) S_{it-1} = S_{i0} (1 + \varepsilon_{i1}) (1 + \varepsilon_{i2}) ... (1 + \varepsilon_{it})$$

Considering a short period of time, ε_{it} can be regarded as small and the approximation $\ln(1+\varepsilon_{it}) = \varepsilon_{it}$ can be justified. Hence, taking logs, we have

$$\ln S_{it} \simeq \ln S_{i0} + \sum_{t=1}^{T} \varepsilon_{it}$$

If the increments ε_{it} are independently and normally distributed, then $\ln S_{it}$ follows a random walk and the limiting distribution of S_{it} is lognormal. Therefore, the growth of the firm is unrelated to its current size and only depends on the sum of idiosyncratic shocks.

Hence, to test Gibrat's law, the vast majority of previous literature has used the following general logarithmic specification

$$\ln S_{it} = \beta_{i0} + \beta \ln S_{it-1} + u_{it} \tag{2}$$

where S_{it} is the size of firm i at time t, and u_{it} is a random variable that satisfies

$$\begin{aligned} E\left(u_{it} \mid S_{it-s}, \quad s>0\right) &= 0 \\ E\left(u_{it}u_{j\tau} \mid S_{it-s}, \quad s>0\right) &= \begin{cases} \sigma^2 & i=j,\, t=\tau \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Gibrat's law is confirmed if the hypothesis $\beta = 1$ is not rejected by the data against $\beta < 1$.

An equivalent specification used by the literature and based directly on corporate growth rates is

$$\ln \frac{S_{it}}{S_{it-1}} = \beta_{i0} + \beta_1 \ln S_{it-1} + u_{it}$$

where clearly $\beta_1 = \beta - 1$. In this case Gibrat's law is confirmed if data do not reject $\beta_1 = 0$, against $\beta_1 < 0$.[4]

In this work we follow a similar autoregressive specification, introducing three main innovations with respect to the traditional empirical approaches. First, we study the behaviour of the (log of) each unit's size relative to the average, i.e. of the variable $g_{it} = \ln\left(S_{it}/\bar{S}_t\right)$, where $\bar{S}_t$ represents the average size over all units at each time t. The use of the proportion of size g_{it} as our basic variable, instead of (the log of) plain size S_{it}, alleviates problems of serial and residual correlation, in that possible common shocks are removed by the normalisation. Moreover, the variable g_{it} can be interpreted as the firms market share. Second, we assume that even firms belonging to the same industries can differ substantially from

[4] In both cases the test is a one-tail test, the same that we use in our empirical analysis for consistency with the literature on testing for unit root. Note also that $\beta > 1$ ($\beta_1 > 0$) implies explosive growth paths, i.e., firms grow faster as they get larger. This situation is conceivable for a short period, but not indefinitely. Moreover, from a qualitative perspective its implications for market structure are similar to those of $\beta = 1$: concentration *would* increase over time, *al*though at a faster rate.

each other. This (possibly intrinsic) heterogeneity is modelled in a general way by allowing all unknown parameters to be unit-specific. Finally, the latter feature is modelled in the context of a hierarchical linear model (Lindley and Smith, 1972) estimated with Bayesian techniques. As already shown in several studies, the panel-data hierarchical-model approach uses the variability contained both in the cross-sectional and in the time-series dimensions in a powerful way, allows the implementation of the relevant tests in a natural way, and is easy to estimate, given the recent advances in Bayesian statistics and Markov Chain Monte Carlo Methods (MCMC) techniques.

For our purposes, we assume that time series realisations $\{g_{it}\}_{t=t_i}^{\mathcal{T}_i}$ for n firms $(i = 1, ..., n)$ are available, possibly of different lengths. Each series starts at time t_i and is generated by an autoregressive model of order 1 ($AR(1)$). Without lack of generality, the minimum t_i equals 1. The last observation occurs at time $\mathcal{T}_i$, for each firm. Assuming that there are no missing observation between t_i and $\mathcal{T}_i$ for each i, we let $\mathsf{T}_i = \mathcal{T}_i - t_i + 1$ denote the number of observations in the series for the ith firm. The initial conditions, g_{i0}, are observed and the subsequent estimation results are conditional on them. The following statistical model is specified for each firm i:

$$g_{it} = \alpha_i + \rho_i g_{it-1} + \eta_{it} \qquad t \geq t_i \tag{3}$$

The random variables η_{it} are assumed to be normally and identically distributed, with mean zero and variance σ_i^2, and are uncorrelated across units and over time, i.e.,

$$\eta_{it} \mid \sigma_i^2 \sim N\left(0, \sigma_i^2\right) \qquad E\left(\eta_{it}\eta_{js}\right) = 0, \quad \forall i \neq j, t \neq s \tag{4}$$

Two additional features are worth mentioning. First, our model specification allows for an intercept, α_i. Other studies (e.g. Bottazzi et. al., 2001) sometimes estimate a specification equivalent to $g_{it} = \rho g_{it-1} + u_{it}$, which, besides considering a common slope, avoids the inclusion of the specific effect α_i, as if the expected proportionate rates of growth were zero. Second, another advantage of the hierarchical model adopted here is the possibility of handling unbalanced panels in a natural way. Therefore, firms with different observations can be included in the sample, which in turn allows us to both maximise the size of cross-sectional observations and minimise the survival bias.

For the sake of simplicity, let $\theta_i = (\alpha_i, \rho_i)'$ and $x_{it} = (1, g_{it-1})'$. Equation (3) can then be written in a more compact form as

$$g_{it} = x_{it}'\theta_i + \eta_{it}, \qquad i = 1, 2, ..., n, \qquad t \geq t_i \tag{5}$$

The hierarchical structure is introduced into Equation (5) with an exchangeable assumption on the population structure

$$\theta_i \sim N(\theta_c, \Sigma_c), \qquad i = 1, 2, ..., n \tag{6}$$

where $\theta_c \left(= (\alpha_c, \rho_c)'\right)$ and Σ_c are the unknown common mean and variance-covariance matrix, respectively. The chosen prior distribution assumes that intercept and slope of the model *do not differ too much* across units, so the firm-specific parameter vector θ_i is an independent random draw from the underlying common distribution (6). The matrix Σ_c controls the variability of the firm-specific regression parameter vector θ_i. The standard linear regression model with homogenous coefficients ($\theta_i = \theta_c$) for each firm is obtained by letting Σ_c be a null matrix.

A full implementation of the Bayesian approach is achieved here using the Gibbs sampler (e.g. Gelfand et al, 1990 for illustration of general models, and Nandram and Petruccelli, 1997 for an application to autoregressive time series panel data), a recursive Monte Carlo method which requires only knowledge of the full conditional posterior distribution of the parameters. The analysis requires the specification of a prior for θ_c, Σ_c and σ_i^2. Assuming independence, as is customary in the literature, we take

$$\begin{aligned} & p\left(\theta_c, \Sigma_c^{-1}, \{\sigma_i^2\}_{i=1}^n\right) \\ \propto \quad & p(\theta_c) \cdot p\left(\Sigma_c^{-1}\right) \cdot \prod_{i=1}^n p\left(\sigma_i^2\right) \end{aligned} \tag{7}$$

to have a Normal-Wishart-Inverse Gamma structure:

$$\theta_c \sim N(\mu, C) \tag{8}$$

$$\Sigma_c^{-1} \sim W\left(s_o, S_o^{-1}\right) \tag{9}$$

$$\sigma_i^2 \sim IG\left(\frac{v}{2}, \frac{\delta}{2}\right) \tag{10}$$

The notation $\Sigma_c^{-1} \sim W\left(s_o, S_o^{-1}\right)$ means that the matrix Σ_c^{-1} is distributed as a Wishart with scale S_o^{-1} and degrees of freedom s_o, while $\sigma_i^2 \sim IG\left(\frac{v}{2}, \frac{\delta}{2}\right)$ denotes an inverse gamma distribution with shape $\nu/2$ and scale $\delta/2$. The hyperparameters μ, C, s_o, S_o, v and δ have to be specified by the researcher. Concretely, μ is the prior mean of the common mean vector θ_c; C controls the dispersion of our prior belief around θ_c: the larger the C matrix, the weaker the prior information on θ_c; s_o, the degrees-of-freedom parameter of the Wishart, controls the dispersion of Σ_c^{-1}, and S_o^{-1} the corresponding location: the bigger is s_o relative to the size of the cross-section, n, and the smaller is S_o, the smaller is the prior mean of Σ_c^{-1} making the prior on θ_i more informative and shrinking θ_i more towards the common mean θ_c; finally v and δ control the shape and the scale of the prior distribution for σ_i^2: a less informative prior is obtained by letting v and δ become smaller. In the following section, results are reported under four different prior specifications, by varying these hyperparameters in a reasonable range.

Under this prior assumption, the conditional posterior distributions of the parameters $\left(\left\{\theta_i, \sigma_i^2\right\}_{i=1}^n, \theta_c, \Sigma_c\right)$ are straightforward to derive.[5] Bayesian point estimates and other quantities of interest are obtained by taking appropriate averages over the *useful* Gibbs sample draws, i.e. those draws for which convergence to the marginal posterior distributions has been achieved. Gibrat's law can be tested by comparing the model $M_o : \rho_c = 1$ against $M_1 : \rho_c < 1$. A finding that ρ_c is not statistically different from 1 would confirm the law as holding over time and across firms. A third model specification, $M_2 : \rho_i = 1$ $(i = 1, ..., n)$ can also be examined. A finding that ρ_i $(i = 1, ..., n)$ is not statistically different from 1 would then be considered as the Gibrat's law holding uniformly over individual firms. The comparisons are based on the Bayes Factor (BF), that is the ratio of marginal data densities under alternative models. These quantities are computed using the Gibbs output as in Chib (1995).

Further implications of the law can then separately be examined.

First, we can compare the speed of adjustment $(1 - \rho_i)$ of each unit to its own steady state, with the respective initial conditions, g_{i0}, which is a question related to the mean reversion argument and the decrease in the variance of the firm size over time. Using the Gibbs draws for the individual ρ_i, we compute $\Pr\{\rho_i < 1 \mid Y\}$, i.e. the posterior probability that the autoregressive coefficient for the ith firm is lower than the unit. The mean reversion is checked by comparing this quantity with the initial conditions. A negative relationship will provide evidence in favour of the mean reversion.

Also, we verify whether steady states are all equal across firms by comparing $H_o : \mathsf{SS}_i = \mathsf{SS}_j$ against $H_1 : \mathsf{SS}_i \neq \mathsf{SS}_j, \forall i \neq j$, where SS_i is the steady state of firm i. Finally, if steady states are not common, the model specification is easily used to verify whether the long-run differences across firms are transitory or permanent, i.e. whether there is persistence in size differences. The latter can be done by comparing the posterior distribution of the steady states to the initial conditions. In particular, we compute the following probabilities: $p_1 = (1/n_1) \sum_i \mathcal{P}_i$, and $p_2 = (1/n_2) \sum_i \mathcal{Q}_i$, where $\mathcal{P}_i = \Pr\{\mathsf{SS}_i < 0 \mid g_{i0} < 0, Y\}$ is the

[5] The derivation of the posterior distributions can be found in the Working Paper version available at http://www.uu.nl/content/05-02.pdf. There the reader can also find a sensitivity analysis to investigate how much our results change when we use other reasonable probability models.

probability that the posterior steady state size is lower than the average, given that the initial size is lower than the average, and $\mathcal{Q}_i = \Pr\{\mathsf{SS}_i > 0 \mid g_{i0} > 0, Y\}$ is the probability that the posterior steady state size is greater than the average, given that the initial size is greater than the average. n_1 and n_2 are the number of firms which started below and above the average respectively. The higher p_1 and p_2, the more *attractive* are the initial conditions and the more persistent are the initial differences in size. The complementary probabilities, $1 - p_1$ and $1 - p_2$ will then provide the *transition* probabilities of going from low to high and from high to low size respectively. A visual inspection of the size persistence argument is also easily obtained by comparing both the average posterior steady state for each firm ($\overline{\mathsf{SS}}_i = \left(\bar{L} - L + 1\right)^{-1} \sum_{l=\bar{L}}^{L} \mathsf{SS}_i^{(l)}$) and the unconditional probability $q = (1/n) \sum_i \mathcal{R}_i$, where $\mathcal{R}_i = \Pr\{\mathsf{SS}_i < 0 \mid Y\}$, with the initial conditions.

A measure of model fit is also reported, in order to check that the model provides an adequate fit to the data. We use the Bayes p-value described in Gelman et al. (1995, Ch. 6 and 12) and compare simulated values from the posterior predictive distribution of replicated data to the observed data. Major failures of the model typically correspond to extreme tail-area probabilities (less than 0.01 or more than 0.99).[6]

3 The Data

The pharmaceutical industry constitutes a particularly interesting testbed for Gibrat's Law. In fact, pharmaceuticals might be considered an ideal case where the process of firms growth should behave in accordance with the Law of Proportionate Effects, as a consequence of the peculiar role and nature of innovation in this industry. It is well known that pharmaceuticals is a highly innovation-intensive industry. Moreover, the innovative process in this sector has often been described and conceptualised as a pure "lottery model", whereby previous innovations (even in a particular submarket) do not influence in any way current and future innovation in the same or in other submarkets (Sutton, 2002).

Data come from the PHID (Pharmaceutical Industry Database) dataset, developed under the EPRIS Program at C.E.R.M. (Center for the Economic Analysis of Competitiveness, Markets, and Regulation, Rome). The database provides longitudinal data for a sample of 210 firms in the seven largest western markets (France, Germany, Italy, Spain, UK, Canada, and USA) during the period 1987-1998. Values are in thousands of Pound Sterlings at constant 1998 exchange rate. The values reported in the database regard only

[6] A more detailed description of the model fit can be found in the Working Paper version available at http://www.uu.nl/content/05-02.pdf.

the pharmaceutical component of the companies total sales. The companies included in the dataset result from selecting the largest 100 companies (in terms of sales) in each national market. Since most of these companies are active in two or more national markets, cleaning for the double-counting was necessary. We then obtain a total of 210 companies for the seven largest western markets. The dataset has been constructed by aggregating the values of the sales of these firms in the different national markets: therefore, sales for each firm stand for the sum of their sales in each of the national markets in which they are active. It is important to emphasise that the panel is unbalanced since processes of entry and exit are explicitly considered.

A few comments of the data are in order here.

We use sales as proxy for firm size not only because of data availability, but also because sales are usually considered the best available proxy of firm size in pharmaceuticals and in some recent studies on firms growth (Hart and Oulton, 1996; Geroski et al., 1997, Higson et al., 2002). Alternative measures are more difficult to obtain on a longitudinal basis and suffer of a number of drawbacks. In particular, the use of employment - even if available - would unduly increase the lumpiness of the growth process, especially as it concerns divestiture, opening (or closure) of plants, R&D labs, etc., particularly as they occur at the international level.[7]

The use of a broad geographical definition for the relevant market and the consideration of the international firm - i.e. the sum of the sales in each national market - as the unit of analysis is justified in our view by the global nature of competition in the pharmaceutical industry. Many firms operate in different countries at the same time, when it concerns R&D, production and marketing. More importantly, successful drugs are sold worldwide and firms growth depends crucially on the ability to be present in different countries at the same time. Hence, the world market - as approximated by the seven large countries for which data were available - seems to us an appropriate level of aggregation for capturing the locus of competition.

In terms of products, the market is defined here at the level of one single class at the 4 digit level of the Standard Industrial Classification, i.e. pharmaceutical products. In this respect, the definition of the market is quite narrow. It could be argued that the pharmaceutical industry is actually constituted by a collection of several (independent) submarkets or therapeutic categories, definable at extremely fine levels of disaggregation. Yet, firms

[7]Employment, assets, sales, market value, and value added are some of the most common measures of company size. For a discussion on their advantages and their limits, see Hart and Oulton (1995).

growth in this industry is fundamentally dependent on the process of diversification on a variety of submarkets (Sutton, 2002; Bottazzi et al. 2001; Henderson et al., 1999). Thus, focusing the analysis on a single (or few) submarket(s) would imply missing an essential driver of firms growth.

In this paper we exclusively focus on the process of *internal* growth of firms. In order to control for mergers and acquisitions during the period of observation, we construct "virtual-firms". These are firms actually existing at the end of the period, for which we construct backward the series of their data in the case they merged or made an acquisition. Hence, if two firms merged during the period, we consider them merged from the start, summing up their sales from the beginning. This procedure might introduce a bias in the intertemporal comparison of firms size distributions along time, but it has the advantage of emphasising the changes in the distributions that derive strictly from intra-market competition (Bottazzi et al. 2001, p.1168). Furthermore, by constructing "virtual firms" we avoid mixing the phenomenon of pure (greenfield) entry with entry due to mergers. In fact, there is evidence that greenfield entrants are smaller than average firms (Baldwin et al., 1995; Acs, 1996), whereas entrants resulted from a merger are usually larger than average, like in the case of the entry of Novartis in 1996 following the exit of Ciba-Geigy and Sandoz in 1995. Our database confirms this finding: the greenfield entrants have smaller size than the average and generally are the minimum values of the size distributions.

The methodology used to construct the dataset supports the use of the international market as the relevant locus of competition.Though the smallest firms might seem under-represented, however, our dataset is not constituted only or even mainly by large companies. Descriptive statistics reported in the Working Paper version show that the sample includes several small and medium sized companies.[8] In fact, the firms size distribution is skewed towards the smallest firms of the sample, since the skewness is always significantly positive and the median is always much smaller than the mean.

As expected, the variable we construct, $g_{it} = \ln\left(S_{it}/\bar{S}_t\right)$ washes away the increasing trend of the total sales over time, and removes the possible shocks common to all the firms in the industry.[9] In fact, the ratio of the standard deviation to the mean as well as to the skewness and the kurtosis are nearly constant over time.Finally, notice that the minimum number of observations in the series for each firm is $\mathsf{T}_i = 2$ (for just one firm).

[8] The descriptive statistics can be found in the Appendix of the Working Paper version available at http://www.uu.nl/content/05-02.pdf.

[9] See the variable Ln_dev87...98 in the descriptive statistics of the Working Paper version.

4 Estimation results

In this section we present the empirical results. They are shown in Figures 1-9 and Tables 1-3.

The estimation results are reported under four prior specifications. Table 1 describes the chosen hyperparameters.

Table 1. Prior information

prior	v	δ	s_o	S_o	μ	C
A	10	1	$k+2$	$20 \cdot I$	$\breve{\mu}$	$\breve{C}$
B	$\mathsf{T}_i + 10$	$\hat{\delta}_o$	$k+50$	$10 \cdot I$	$\tilde{\mu}$	$\tilde{C}$
C	5	1	$k+2$	$20 \cdot I$	$\tilde{\mu}$	$100 \cdot I$
D	5	1	$k+50$	$10 \cdot I$	$\tilde{\mu}$	$\tilde{C}$

The notation of the table is as follows:

1. $\breve{\mu} = 0.5\left(\bar{\mu} + \mu_{pool}\right)$, $\breve{C} = 0.5\left(\bar{C} + C_{pool}\right)$
2. $\bar{\mu} = (1/N)\sum_i \hat{\theta}_i$, with $\hat{\theta}_i = \left(X_i'X_i\right)^{-1} X_i'G_i$, $\mu_{pool} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'Y$, where $\mathbf{X} = diag\left(X_1, ..., X_n\right)$;
3. $\bar{C} = (1/N)\sum_i \left(\hat{\theta}_i - (1/N)\sum_i \hat{\theta}_i\right)\left(\hat{\theta}_i - (1/N)\sum_i \hat{\theta}_i\right)'$, and $C_{pool} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}$;
4. $\tilde{\mu} = (0 \;\; 0.96)'$ and $\tilde{C} = diag\left(10, 0.006\right)$;
5. $\hat{\delta}_o = (1/N)\sum_i \left(G_i - X_i'\breve{\mu}\right)'\left(G_i - X_i'\breve{\mu}\right)$.

The hyperparameters have been chosen following simple criteria and mixing sample information, and previous empirical studies both on Gibrat's law and on hierarchical models. Prior A can be regarded as moderately non informative at all levels of the hierarchical structure. Prior B is more informative than prior A in all respects. First, it has a bigger s_o relative to the size of the cross-section, and a smaller S_o: both features imply a smaller prior mean for Σ_c^{-1} making the prior on θ_i more informative and shrinking θ_i more towards the common mean θ_c. Second, the mean and the variance of μ_2, i.e. the hyperprior mean of ρ_c, have been taken from the previous empirical estimates of ρ_c in the literature as reported in Goddard et al. (2002, pp.417, table 1). We fit an empirical distribution on this estimates and then compute its mean and variance. Figure 1 reports the histogram of these estimates. Finally, the shape ν_o and the scale δ_o of the prior on the variance have been chosen to match on average the sample variance.

Prior C is less informative than priors A and B at all levels of the hierarchy. Prior D is a mixture of priors B and C.

The Gibbs sampling is run in 4 cycles of 5000 iterations. Results are based on the last 5000 iterations, therefore discarding the first 15000 draws. Convergence has been checked following the method proposed by Brooks and Gelman (1998). It has been achieved already after 10000 iterations, using different seeds of the random number generator and different initial values of the unknown parameters.

Table 2 reports the mean and the 90% central part of the posterior distribution of ρ_c, the model testing as explained above (M_o vs. M_1, M_o vs. M_2, M_1 vs. M_2, and H_o vs. H_1), the marginal likelihood, i.e., the posterior density of the data, $(\ln(\hat{m}(y)))$, computed as in Chib (1995) and the Bayes p-value (B-p) to check the model fit of the data under the four priors.

As a general comment, notice that Prior A fits the data well –according to the particular test quantity chosen– and produces the highest posterior density as well. Therefore, most of the results discussed below, especially those concerning mean reversion and persistence of size differences, are based on the output generated under Prior A. Results based on the other priors are qualitatively identical and therefore omitted to avoid replications.

Table 2. Estimation and testing

prior	ρ_c (5% – 95%)	$\ln(BF_{01})$	$\ln(BF_{02})$	$\ln(BF_{12})$	$\ln(BF_{ss})$	$\ln(\hat{m}(y))$	B-p
A	0.92 (0.85 – 0.99)	−6.19	420.94	427.13	−823.66	2201.03	0.59
B	0.94 (0.88 – 1.01)	−6.32	135.30	142.21	−835.45	2163.02	0.03
C	0.93 (0.85 – 1.02)	−4.72	453.53	458.25	−818.20	1977.23	0.97
D	0.96 (0.91 – 1.01)	−4.94	187.87	192.81	−725.42	2110.99	0.99

Notes: BF_{01} compares $M_o : \rho_c = 1$ against $M_1 : \rho_c < 1$; BF_{02} compares $M_o : \rho_c = 1$ against $M_1 : \rho_i = 1, i = 1, .., n$; BF_{12} compares $M_1 : \rho_c < 1$ against $M_1 : \rho_i = 1, i = 1, .., n$; BF_{ss} compares $H_o : SS_i = SS_j$ against $H_1 : SS_i \neq SS_j, \forall i \neq j$. $\ln(\hat{m}(y))$ is the log marginal posterior density. *B-p* is the Bayes p-value (Gelman et al.,1995).

4.1 Does Gibrat's law hold?

Table 1 shows that under all priors the numerical estimate of ρ_c is different from 1. This can also be seen in the plots of the posterior densities of ρ_c (Figure 2). It is worth noting that in almost all cases the posterior distribution contains 1. However, the values of the Bayes factor BF_{01}, which compare $M_o : \rho_c = 1$ against $M_1 : \rho_c < 1$, are always lower than 1. To interpret these numbers, one can compute the highest prior probability to assign to model M_1 in order to obtain posterior odds in favour of M_o. If π is the prior probability of model M_1, the Posterior Odds ratio is defined as the product of the prior odds ratio and the Bayes Factor:

$$PO = \frac{1-\pi}{\pi} BF.$$

Therefore, the highest prior probability to assign to model M_1 in order to obtain posterior odds (just) in favour of M_o is $\pi^* = 1/\left(1 + \exp\left(1 - \ln\left(BF_{01}\right)\right)\right)$. Hence, for instance, under prior C one should assign at most $\pi^* = 0.00327$ to M_1 for the data to change the conclusion and make the posterior inference favourable to M_o. Such a small probability would imply an implausible prior odds ratio of 304.8 in favour of M_o. We consider these numbers as clear evidence against model M_o as compared to M_1. The former is however strongly favoured when compared to model M_2, meaning that the posterior density of the sample data is much higher when we impose an average random walk across firms than when we impose the same assumption to all individual firms. Finally, M_2 is *a fortiori* not favoured when compared to model M_1.

Overall, these findings do not confirm that Gibrat's law holds on average, over time and across firms. The histogram of $\bar{\rho}_i$, the posterior mean of ρ_i $(i = 1, ..., n)$ averaged across firms (notation above), provides a first visual inspection of the finding that several firms are far from following Gibrat's Law (Figure 3). An interesting issue is to check which firms do follow the law. A further straightforward analysis shows clearly that large firms have a posterior distribution of ρ_i centred on unity. Figure 4 (chart a) plots the posterior distribution of ρ_i for top 10% firms, i.e. firms whose initial size is in the top decile of the initial distribution of sizes. The picture confirms the intuition that large firms do follow Gibrat's law, in line with previous finding of the literature (e.g. Hall 1987, Lotti et al.2003). The marginal likelihood $(\ln \hat{m}(y))$ of the model under prior A and the restriction that $\rho_i = 1$ for the top 10% firms is equal to 2146.58. The same figure is only 1849.38 when we restrict all the other firms to have $\rho_i = 1$. The positive Bayes factor resulting from the difference between the former and the latter provides odds in favour of the hypothesis that the size of large firms follows a random walk, even though the model without any restrictions (and under the same prior) has a higher posterior density (2201.03), as shown in Table 1. This implies that the sample data prefer a model where no random walk restriction is imposed on any firms.

On the contrary, if we look at the firms whose initial size belongs to the first decile of the distribution (bottom 10%, Figure 4b) we see that the posterior distributions are not centred on the unity, confirming that initially small firms do not follow Gibrat's law and have a higher speed of convergence to the steady state than initially large firms, which may have already reached their own steady states. These claims are also summarised in Figure 5 (charts "a" and "b"), which show the scatter plots of the average posterior speed of convergence versus the initial sizes. In chart "b" the top 10% and the bottom 10% have been excluded from the sample. The figures confirm that for very small firms the average convergence speed is far away from zero, while for very large firms is around zero (chart "a"). The evidence shows no clear pattern when the remaining firms are considered (chart "b").

In sum, these results seem to confirm previous findings that Gibrat's Law holds only for very large firms but not for the others. The Law appears to hold only for 15% of the firms in our sample. Moreover, the failure of Gibrat's Law and the observation of higher convergence rates for small firms do not necessarily imply mean reversion.

4.2 Do data show mean reversion?

From the same plots an initial assessment of the mean reversion argument can be drawn. Do smaller firms have a higher speed of convergence than larger firms? As argued above, this seems to be true only when the first and the last decile of the initial size distribution are compared. If we take the "extreme-size" firms out of the sample,there is no evidence of a negative relation between speed of convergence and initial size.

The argument against strong evidence of mean reversion stands also when considering the entire posterior distribution of ρ_i and not just its posterior mean. Figure 6 shows the relation between the initial condition and the posterior probability of ρ_i being lower than unity, i.e., $\Pr(\rho_i < 1 \mid Y)$. In chart "a" all firms are considered; in chart "b" we again exclude the top and the bottom 10%. As from the previous scatter plots, it can be argued that it is indeed true that for very small firms $\Pr(\rho_i < 1 \mid Y)$ is quite high while for very large firms the same probability is low. However, when considering only the central 80 percent firms of the initial size distribution, this relation is very weak. We therefore cannot express a posterior confidence in favour of a mean-reversion claim. Finally, to the extent that mean reversion actually operates,the process is still very slow indeed, given the observed values of the speed of convergence (Figures 5a and 5b)

4.3 Are the steady states equal?

But do firms sizes converge to the same steady state, as it would be implied by a strict interpretation of the mean reversion argument? Table 2 (column 6: $ln(BF_{ss})$) shows that under all priors the Bayes factor overwhelmingly favours the model where steady states are not restricted to be equal. The evidence is reinforced by Figure 7, where the posterior distributions of the steady states of 10% randomly chosen firms are plotted. Both the Bayes factor and the chart support the evidence that the firms in the sample have very different steady states, confirming once more not only that Gibrat's Law does not hold on average, but also that firms do not converge to the same size. These results suggest that firm-specific characteristics are very important in determining firms growth.

4.4 Do initial differences in size persist?

Gibrat's Law would imply that initial size differences would not tend to persist. The mean reversion argument would suggest that persistence of size differentials should be quite low, as small firms grow faster than large ones, even if the steady state sizes are different. Here, we check the persistence of size differences. Figure 8 plots the posterior mean of the steady states versus the initial conditions. The positive relation favours the conclusion that differences across firms are persistent,depending strongly on their initial size.

In order to consider both the uncertainty on the steady state resulting from the model and the posterior mean, we further investigate the relation between the entire posterior distributions of steady states and the initial condition. First, dividing the initial sample into firms below and above the average, we plot the scatter points of the unconditional probability $q = (1/n)\sum_i \mathcal{R}_i$, where $\mathcal{R}_i = \Pr\{\mathsf{SS}_i < 0 \mid Y\}$, against the initial conditions. Figure 9 reinforces our first preliminary conclusion on the persistence, namely that most of the smallest firms have a high posterior probability of remaining below the average, while for almost all largest firms the same probability is negligible. Second, we compute the posterior probabilities $p_1 = (1/n_1)\sum_i \mathcal{P}_i$, and $p_2 = (1/n_2)\sum_i \mathcal{Q}_i$, where $\mathcal{P}_i = \Pr\{\mathsf{SS}_i < 0 \mid g_{i0} < 0, Y\}$ is the probability that the posterior size steady state is lower than the average, given that the initial size is lower than the average, and $\mathcal{Q}_i = \Pr\{\mathsf{SS}_i > 0 \mid g_{i0} > 0, Y\}$ is the probability that the posterior steady state size is greater than the average, given that the initial size is greater than the average. Results give $p_1 = 0.77$ and $p_2 = 0.80$, meaning that the probability of remaining in the same initial position is almost 80 per cent for both states (below and above the average) or, that there is only a 20 percent probability for a firm which starts below (above) the average to reach a steady state above (below) the average.

The same analysis can be refined by considering more quantiles of the initial size distribution and not just the mean. Table 3 reports the average probabilities that a firm in a certain quartile of the initial conditions distribution will end up in a quartile of the steady states distribution. For instance, cell (1,1) reports the average probability of a firm initially in the first quartile to have its steady state again in the first quartile; cell (2,1) reports the average probability of a firm initially in the first quartile to have its steady state in the second quartile; and so on. Therefore, the average probability of remaining in the same initial position is reported on the main diagonal, whereas the off diagonal elements represent the *transition* probabilities of moving from one state to another.

Probabilities in Table 3 reinforce previous results: initial conditions in firms size matter for the firms position in the steady state distribution. The probabilities of remaining in the same quartile of the distribution are always larger than those of moving into other quartiles, especially in the upper part of the distribution. It is worth noting that the state showing the highest persistence is the last quartile: firms initially larger than the average are very likely to end up also in steady states much larger than the average.

In sum, these results indicate not only that Gibrat's Law is violated in our sample, but also that mean reversion is very weak: smaller firms tend to remain small and larger firms tend to remain large. Moreover, the speed of convergence to such steady state is very slow.

Table 3. Persistence of differences

	Initial	$\leq 25\%$	$25\% - 50\%$	$50\% - 75\%$	$> 75\%$
Steady state					
$\leq 10\%$		0.5219	0.3184	0.1251	0.0372
$25\% - 50\%$		0.2652	0.4582	0.2577	0.0240
$50\% - 75\%$		0.1296	0.1506	0.5263	0.1840
$> 90\%$		0.0833	0.0728	0.0909	0.7548

5 Summary and concluding remarks

The results of this paper can be summarised as follows:

(i) The main assertion of Gibrat's law that growth rates are erratic is not true on average, across firms and over time. The estimated average speed of adjustment is far from being zero on average when both cross-sectional and time-series information is used. However, the finding that the growth of initially very large firms follows a random walk is confirmed here.

(ii) Data show only a very weak evidence of mean reversion. Even if on average $\rho_c < 1$, this does not necessarily mean that initially larger firms grow slower than smaller firms. Our analysis shows that the relative speed of convergence of smaller firms is not necessarily higher than the one of larger firms, except in the extreme tails of the distribution.

(iii) More important, firm sizes do not converge to a common limiting distribution but to firm-specific steady state size: estimated steady states differ across units . This fact does not imply *per se* that firm size drifts unpredictably over time, as argued by some authors (see Geroski, 1999, and Geroski at al. 2003). It is true that a unit root in the process of firm size implies divergence, but the reverse causality does not necessarily hold, as shown in this paper.

(iv) Initial conditions are important determinants of the estimated distribution of steady states. Initial size differences do not seem to disappear over time and when they do, the rate is very slow. Thus, a firm with an initial size below the average will narrow the gap somewhat with respect to larger firms, but it does not seem to increase its relative size in the cross sectional distribution. In other words, differences in firms size persist.

(vi) The model used does not show failings in fitting to data. Moreover, results are unchanged using alternative models.

In sum, our results contradict two basic implications of both Gibrat's Law and the "generalised" mean reversion argument: almost no correlation is observed between initial size and speed of adjustment, while a strong correlation is found between initial size and the steady state. Thus, there seem to be systematic differences in growth rates among firms that are not size-specific and may depend on other firm-specific features that are not observable in our data. Given that these results are sufficiently robust to different prior specifications, they create space for investigating further the determinants of firms growth. Most likely, size is not the only variable growth should be conditioned on. Other sources of heterogeneity (age being a primary - but certainly not the only - candidate) may more plausibly be responsible for differential growth rates of firms over time. In particular it would be interesting to explore some common features across clearly divergent/convergent firms, as well as the role of other variables in explaining the cross sectional dispersion in estimated steady states. Finally, the mechanisms through which market selection operates in promoting the growth and the decline of firms should also be explicitly modelled and tested.

At a more general level, the results of this paper strengthens once more the argument that extreme attention has to be given to treating heterogeneity appropriately in econometric models.

References

[1] Acs, Z.J. (ed.) (1996), *Small Firms and Economic Growth,* Elgar Reference Collection, International Library of Critical Writings in Economics, vol. 61; Cheltenham, U.K.

[2] Baily, M.N. and A.K. Chakrabarty (1985), 'Innovation and Productivity an US Industry', *Brookings Papers on Economic Activity*, 2, pp.609-632.

[3] Baily, M.N., C. Hulten and D. Campbell (2000), 'Productivity Dynamics in Manufacturing Plants', in D.B. Audretsch and S. Klepper,(eds.) *Innovation, Evolution of Industry and Economic Growth.* Volume 2. Elgar Reference Collection. International Library of Critical Writings in Economics, vol. 118. Cheltenham, U.K. and Northampton, Mass., Elgar.

[4] Baldwin, J.R, P.K. Gorecki, R. Caves, T. Dunne and J.Haltiwanger (1995), *The Dynamics of Industrial Competition: A North American Perspective*, Cambridge, Cambridge University Press.

[5] Bottazzi, G., G. Dosi, M. Lippi, F. Pammolli and M. Riccaboni, (2001), 'Innovation and Corporate Growth in the Evolution of the Drug Industry', *International Journal of Industrial Organization*, 19, pp.1161-1187.

[6] Brooks, S. P. and A. Gelman (1998), 'General methods for monitoring convergence of iterative simulations', *Journal of Computational and Graphical Statistics,* 7, 434-455.

[7] Caves, R. E. (1998), 'Industrial Organization and New Findings on the Turnover and Mobility of Firms', *Journal of Economic Literature*, 36, December, pp. 1947-1982.

[8] Cefis, E. (2003), 'Is There Persistence in Innovative Activities?', *International Journal of Industrial Organization*, 21(4), pp.489-515.

[9] Cefis, E. and L. Orsenigo (2001), 'The Persistence of Innovative Activities. A Cross-Countries and Cross-Sectors Comparative Analysis', *Research Policy*, 30(7), pp.1139-1158.

[10] Chib, S. (1995), 'Marginal Likelihood from the Gibbs output', *Journal of the American Statistical Association*, 90, 1313–1321.

[11] Dosi, G. (2005), 'Statistical Regularities in the evolution of Industries', SSSUP Working Paper 2005/17, Sant'Anna School of Adnaced Studies, Pisa, Italy.

[12] Dosi, G., O. Marsili, L. Orsenigo and R. Salvatore (1995), 'Technological Regimes, Selection and Market Structure', *Small Business Economics*, 7, pp. 1-26.

[13] Dunne, T., M.J. Roberts and L. Samuelson (1989), 'The Growth and Failure of US Manufacturing Plants', *Quarterly Journal of Economics,* 104(4), pp.671-698.

[14] Ericson, R. and A. Pakes (1995), 'Markov-Perfect Industry Dynamics: a Framework for Empirical Work', *Review of Economic Studies*, 62(1), pp.53-82.

[15] Evans, D.S. (1987a), 'The Relationship between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries', *Journal of Industrial Economics*, 35(4), pp. 567-81.

[16] Evans, D.S. (1987b), 'Tests of Alternative Theories of Firm Growth', *Journal of Political Economy*, 95(4), pp. 657-74.

[17] Farinas, J.C. and L. Moreno (2000), 'Firms' Growth, Size and Age: A Nonparametric Approach', *Review of Industrial Organization*, 17(3), pp. 249-65.

[18] Gelfand, A.E. and A.F.M. Smith (1990), 'Sampling–based approaches to calculating marginal densities', *Journal of the American Statistical Association,* 85(410):398–409.

[19] Gelman, A., J.B. Carlin, H.S. Stern and D.B. Rubin (1995), *Bayesian Data Analysis*, Chapman & Hall, London.

[20] Geroski, P.A. (1999), 'The growth of Firms in Theory and in Practice', in N. Foss and V. Mahnke (eds.), *Competence, Governance, and Entrepreneurship*, Oxford, Oxford University Press.

[21] Geroski, P., S.J. Machin and J. Van Reenen (1993), 'The Profitability of Innovating Firms', *RAND Journal of Economics*, 24(2), pp. 198-211.

[22] Geroski, P., S.J. Machin and C.F. Walters (1997), 'Corporate growth and profitability', *The Journal of Industrial Economics*, 45, 2, June, 171-189.

[23] Geroski, P.A., S. Lazarova, G. Urga and C.F. Walters (2003), 'Are Differences in Firm Size Transitory or Permanent?', *Journal of Applied Econometrics,* 18(1), pp.47-59.

[24] Gibrat, R. (1931), *Les inequalites economiques.* Librairie du Recueil Sirey, Paris.

[25] Goddard, J., J. Wilson and P. Blandon (2002), "Panel tests of Gibrat's law for Japanese manufacturing", *International Journal of Industrial Organization,* 20, 415-433.

[26] Hall, B. (1987), 'The Relationship between Firm Size and Firm Growth in the U.S. Manufacturing Sector', *Journal of Industrial Economics*, 35(4), pp. 583-606.

[27] Harhoff, D., K. Stahl and M. Woywode (1998), 'Legal Form, Growth and Exit of West German Firms-Empirical Results for Manufacturing, Construction, Trade and Service Industries', *Journal of Industrial Economics*, 46(4), pp. 453-488.

[28] Hart, P.E. and N. Oulton (1996), 'Growth and Size of Firms', *The Economic Journal*, 106(Sept.), pp1242-1252.

[29] Hart, P.E. and N. Oulton (1995), 'Growth and Size of Firms', National Institute of Economic and Social Research Discussion Paper No.77.

[30] Hart, P.E. and S.J. Prais (1956), 'The Analysis of Business Concentration: A Statistical Approach', *Journal of the Royal Statistical Society*, 119, series A, pp.150-191.

[31] Henderson, R., L. Orsenigo, G.P. Pisano (1999), "The Pharmaceutical Industry and the Revolution in Molecular Biology: Interactions among Scientific, Institutional, and Organizational Change", in Mowery, D.C., Nelson, R.R. (eds.) *Sources of industrial leadership: Studies of seven industries.* Cambridge; New York and Melbourne: Cambridge University Press, pp. 267-311.

[32] Higson, C., S. Holly and P. Kattuman (2002), 'The Cross-Sectional Dynamics of the US Business Cycle: 1950-1999', *Journal of Economic Dynamics and Control*, 26, pp.1539-1555.

[33] Hopenhayn H. (1992), 'Entry, Exit and Firm Dynamics in Long Run Equilibrium', *Econometrica*, 60(5), pp.1127-1150.

[34] Hsiao, C., M.H. Pesaran and A.K. Tahmiscioglu (1999), 'Bayes estimation of short run coefficients in dynamic panel data models', in Hsiao et al. (eds.) *Analysis of panels and limited dependent variable models: in honor of G.S. Maddala*, Cambridge, Cambridge University Press.

[35] Hymer, S. and P. Pashigian (1962), 'Firm Size and the Rate of Growth', *Journal of Political Economy,* 70(4), pp.556-569.

[36] Ijiri, Y. and H. Simon (1974), 'Interpretations of Departures from the Pareto Curve Firm-Size Distribution', *Journal of Political Economy*, 82(2), part 1, pp.315-331.

[37] Ijiri, Y. and H. Simon (1977), *Skew Distribution and the Sizes of Business Firms*, Amsterdam, North-Holland Publishing.

[38] Jovanovic, B. (1982), 'Selection and the Evolution of Industry', *Econometrica*, 50(3), pp.649-70.

[39] Kadiyala, K.R. and S. Karlsson (1997), 'Numerical methods for estimation and inference in Bayesian VAR-models', *Journal of Applied Econometrics*, 12(2), pp. 99-132.

[40] Klette, T.J., M. Jarle and Z. Griliches (2000), 'Do Subsidies to Commercial R&D Reduce Market Failures? Microeconometric Evaluation Studies', *Research Policy*, 29(4-5), pp. 471-95.

[41] Klette, T.J. and S. Kortum (2004) 'Innovating Firms and Aggregate Innovation', *Journal of Political Economy*, 112(5), pp. 986-1018.

[42] Li, K. (1999), 'Testing Symmetry and Proportionality in PPP: A Panel-Data Approach', *Journal of Business and Economic Statistics*, 17, pp. 409-418.

[43] Lindley, D.V. and A.F.M. Smith (1972), 'Bayes estimates for the linear model' (with discussion), *Journal of the Royal Statistical Society*, B, 34, pp. 1–41.

[44] Lotti, F., E. Santarelli and M. Vivarelli (2003), 'Does Gibrat's Law Hold among Young, Small Firms?', *Journal of Evolutionary Economics*, 13(3), pp. 213-35.

[45] Lucas, R.E. Jr. (1978), 'On the Size Distribution of Business Firms', *Bell Journal of Economics*, 9, pp. 508- 523

[46] Mata, J. (1994), 'Firm Growth during Infancy', *Small Business Economics*, 6(1), pp. 27-39.

[47] McCloughan, P.(1995), 'Simulation of Concentration Development from Modified Gibrat Growth-Entry-Exit Processes', *Journal of Industrial Economics*, 43(4), pp.405-433.

[48] Mueller, D.C. (ed.) (1990), *The Dynamics of Company Profits: an International Comparison,* Cambridge, MA, Cambridge University Press.

[49] Nelson, R. and S. Winter (1982), *An Evolutionary Theory of Economic Change,* Cambridge, MA, The Bellknap Press of Harvard University Press.

[50] Pesaran, M.H. and R. Smith (1996), 'Estimating Long Run Relationships from Dynamic Heterogeneous Panels', *Journal of Econometrics*, 68, pp.79-113.

[51] Simon, H. and C. Bonini (1958), 'The Size Distribution of Business Firms', *American Economic Review*, 58(4), pp.607-617.

[52] Sutton, J. (1997), 'Gibrat's Legacy', *Journal of Economic Literature*, 35, pp.40-59.

[53] Sutton, J. (2002), 'The Variance of Firm Growth Rates: the 'Scaling' Puzzle', *Physica A*, 312, pp.577-590.

[54] Winter, S. G., Y.M. Kaniovski and G. Dosi (2000), 'Modeling Industrial Dynamics with Innovative Entrants', *Structural Change and Economic Dynamics*, 11(3), pp. 255-93.

[55] Wit de, G. (2005), 'Firm Size Distributions. An Overview of Steady-State Distributions resulting from firm dynamics models', *International Journal of Industrial Organization*, 23, pp.423-450.

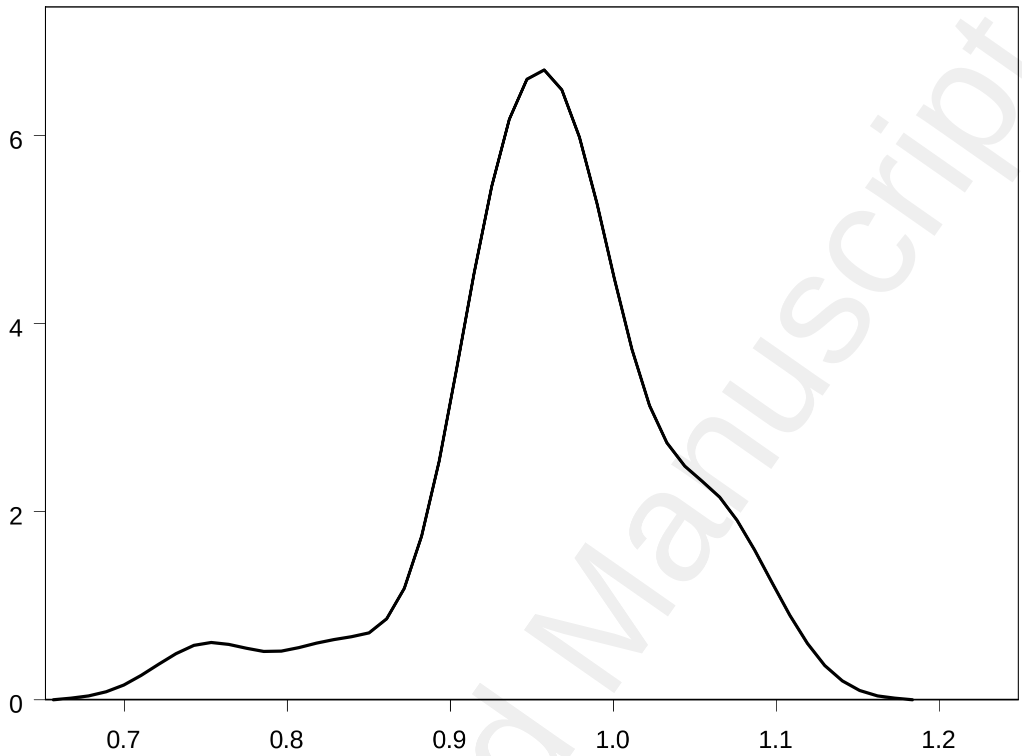

Figure 1. Histogram of previous estimates of ρ_c. *Source: Goddard et al. (2002)*

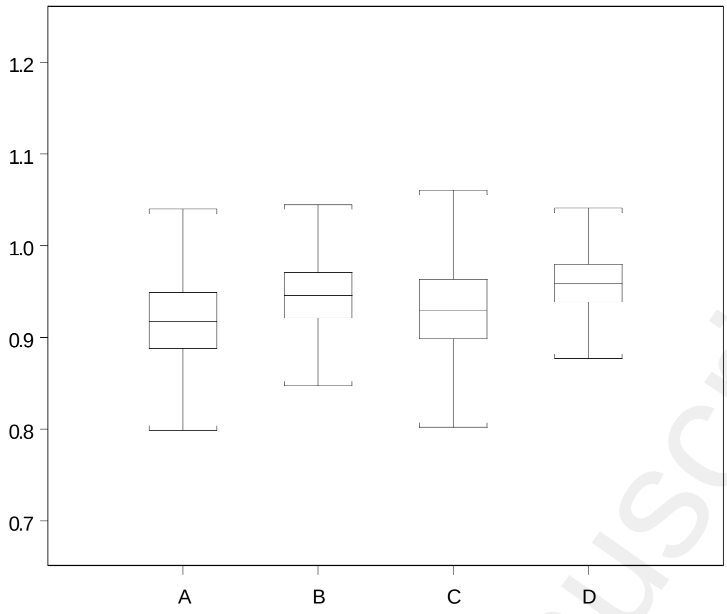

Figure 2. Posterior distribution of ρ_c under the four priors

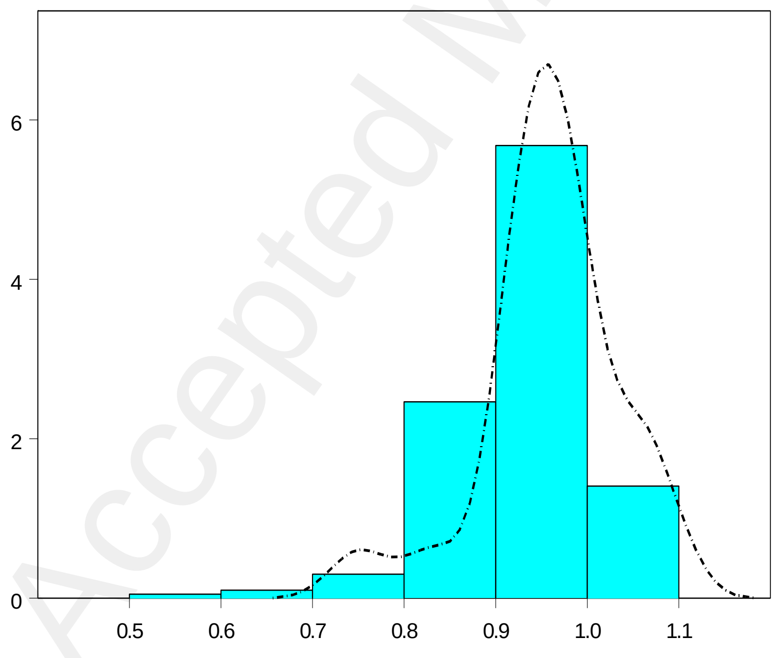

Figure 3. Histogram of posterior mean of ρ_i

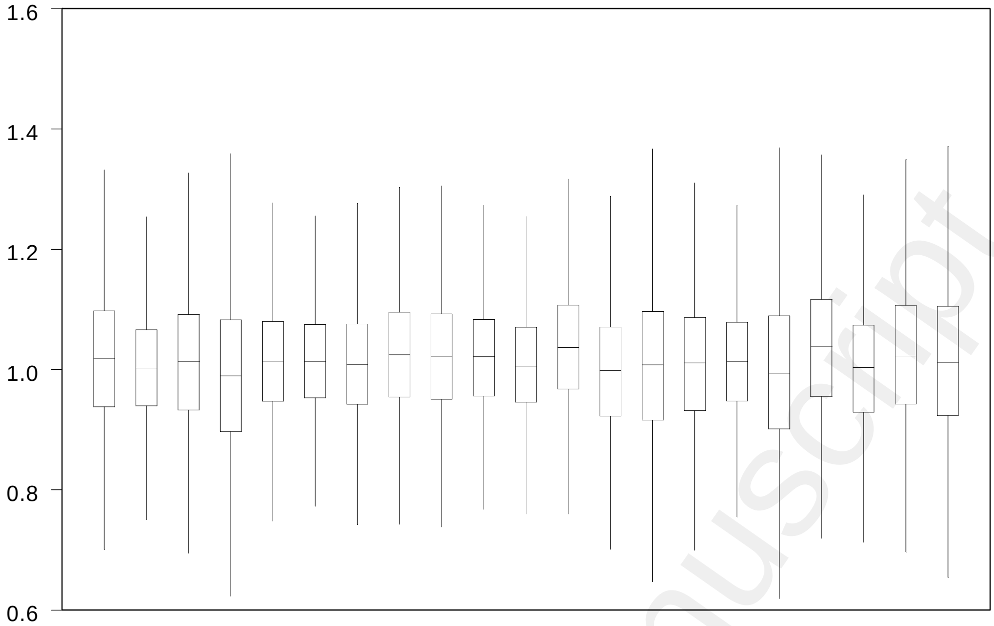

Figure 4a. Posterior distribution of ρ_i*. Top 10% firms*

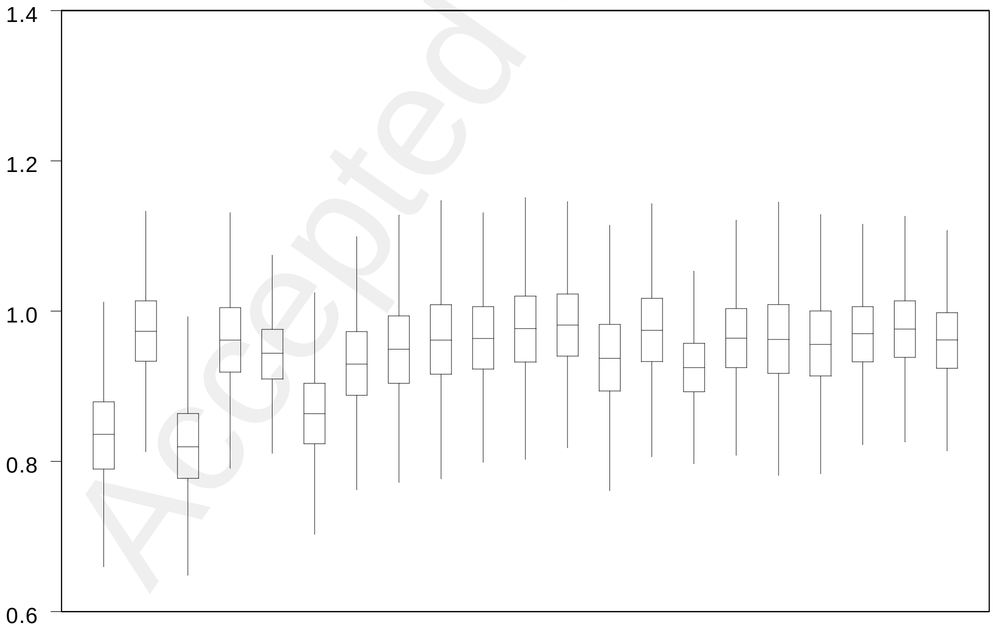

Figure 4b. Posterior distribution of ρ_i*. Bottom 10% firms*

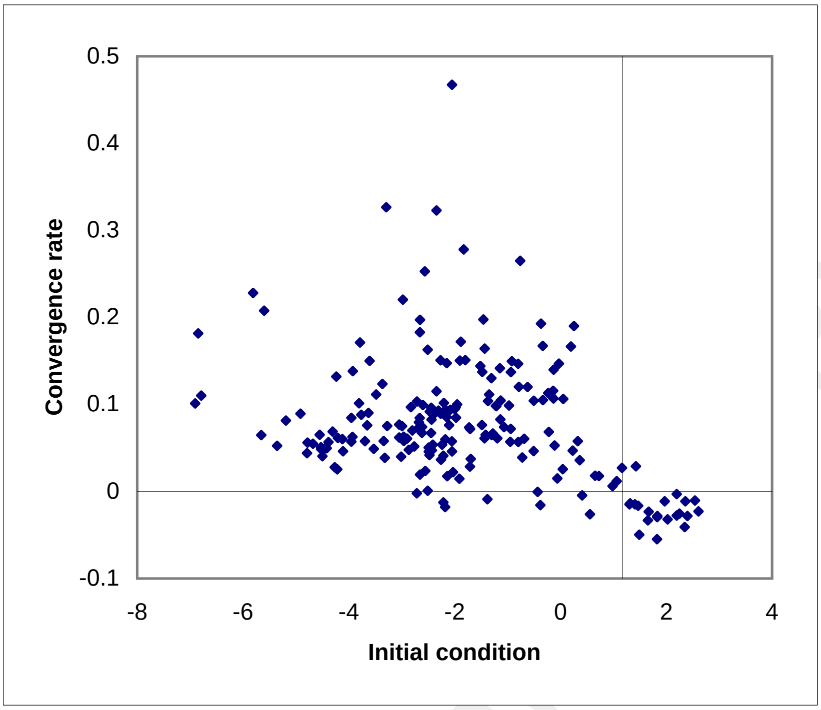

Figure 5a. Speed of convergence: all firms

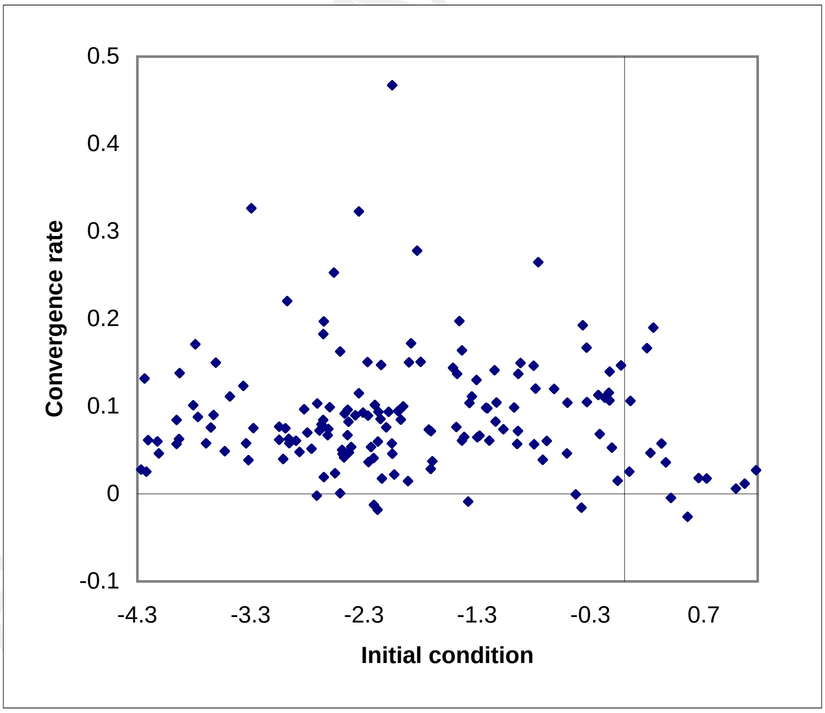

Figure 5b. Speed of convergence: excluding 10% tails

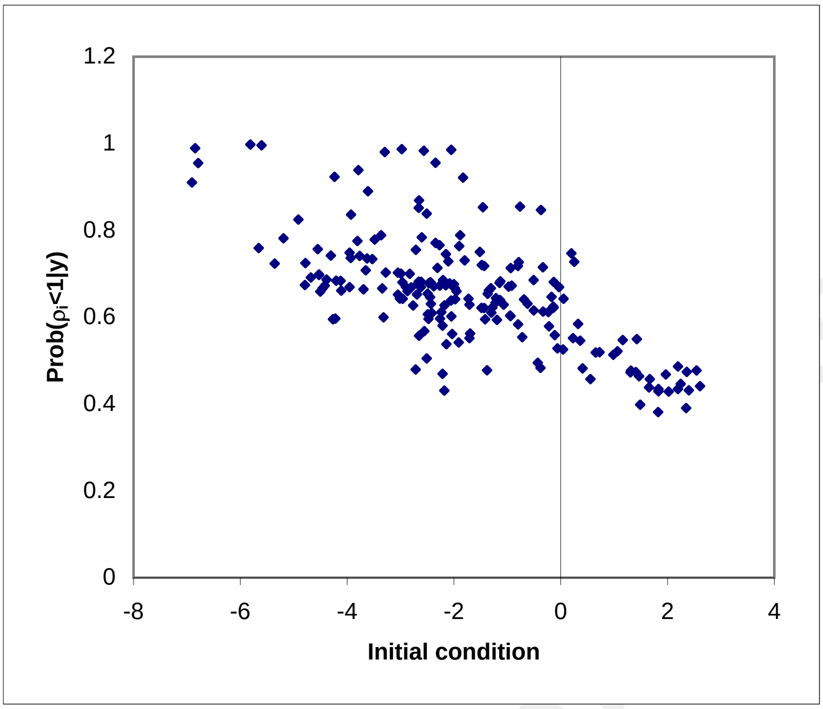

Figure 6a. Mean reversion. All firms.

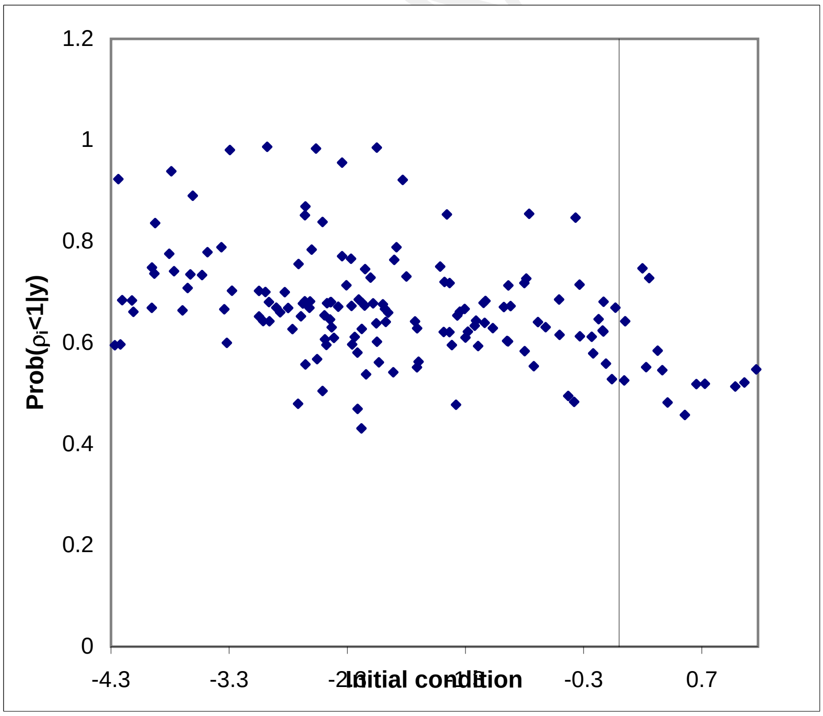

Figure 6b. Mean reversion without 10% tails

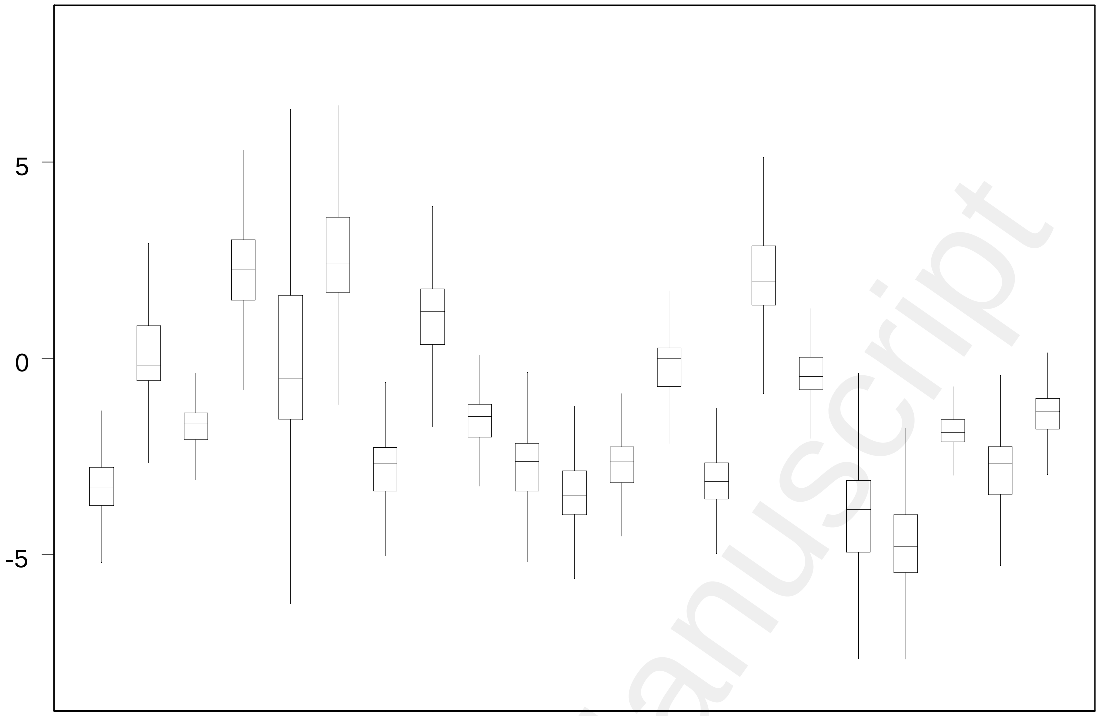

Figure 7. Posterior distributions of steady state. Randomly selected firms.

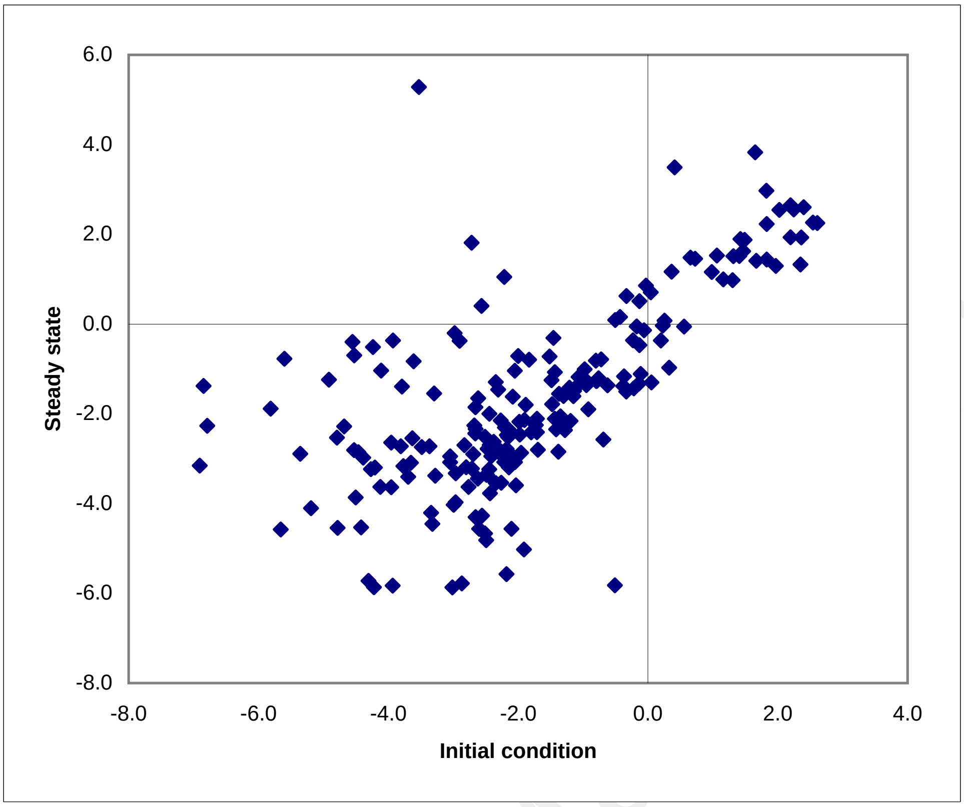

Figure 8. Average steady states and initial conditions

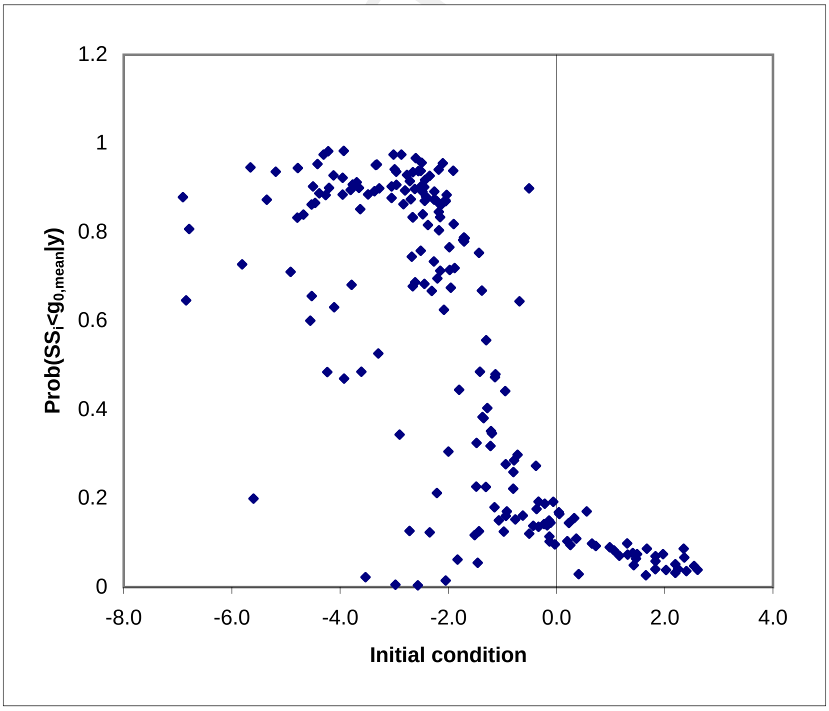

Figure 9. Persistence of differences