Innovation and Corporate Dynamics: A Theoretical Framework

Massimo, Riccaboni; Jakub, Growiec and Fabio, Pammolli
University of Trento, IMT Lucca, Warsaw School of Economics, National Bank of Poland

26. March 2011

Innovation and Corporate Dynamics: A Theoretical Framework

Jakub Growiec* Fabio Pammolli† Massimo Riccaboni‡§

March 26, 2011

Abstract

We provide a detailed analysis of a generalized proportional growth model (GPGM) of innovation and corporate dynamics that encompasses the Gibrat's Law of Proportionate Effect and the Simon growth process as particular instances. The predictions of the model are derived in terms of (i) firm size distribution, (ii) the distribution of firm growth rates, and (iii-iv) the relationships between firm size and the mean and variance of firm growth rates. We test the model against data from the worldwide pharmaceutical industry and find its predictions to be in good agreement with empirical evidence on all four dimensions.

Keywords: Business firm size; firm growth distribution; Gibrat's Law; Pareto distribution; lognormal distribution, size-variance relationship.

JEL Classification: C49; L11; L25; L65.

*Institute of Econometrics, Warsaw School of Economics, Warsaw, Poland, and Economic Institute, National Bank of Poland. E-mail: jakub.growiec@sgh.waw.pl

†IMT Lucca Institute of Advanced Studies, Lucca, Italy. E-mail: f.pammolli@imtlucca.it

‡Corresponding author. DISA, University of Trento, Via Inama 5, 38100 - Trento, Italy. E-mail: massimo.riccaboni@unitn.it

§Financial support from the CERM Foundation is gratefully acknowledged. All errors belong to the authors.

1 Introduction

Empirically observed firm size distributions are formed as an outcome of the underlying firm dynamics involving entry of new firms, innovation, new product launches, growth, mergers, acquisitions, spin-outs, decline and exit. Several models have been proposed so far in the literature to account for these dynamics and thus explain the mechanisms behind the observed industry structure (Gibrat 1931, Kalecki 1945, Steindl 1965, Ijiri & Simon 1977, Jovanovic 1982, Hopenhayn 1992, Sutton 1998, Klette & Kortum 2004, Fu et al. 2005, Klepper & Thompson 2006, Bottazzi & Secchi 2006, Luttmer 2007, Rossi-Hansberg & Wright 2007). Most of them referred either to the so-called Gibrat's Law of Proportionate Effect (Gibrat 1931) or to the Simon growth process (Ijiri & Simon 1977) as useful benchmark cases (Sutton 1997).

The Gibrat's Law states that the expected value of a firm's growth rate is independent of its size, and is probably the simplest available mechanism that leads to a lognormal distribution of firm sizes. Simon and colleagues introduced, on the other hand, an "urn" scheme similar to the one originally proposed by Yule (1925): new business opportunities (balls) are assigned there to firms (urns). Incumbent firms are then assumed to capture randomly a sequence of independent "opportunities" which arise over time, each of unitary size, with a probability that is proportional to the firm's size; and there is also a constant probability that a new opportunity is assigned to a start-up firm. As opposed to the Gibrat model, the Simon growth process converges to a Pareto firm size distribution. Several other, more complex models of proportional growth have also been subsequently introduced in economics and finance (Sutton 1997, Gabaix 1999, Mitzenmacher 2004, De Wit 2005, Gabaix 2009). In most of them, the Pareto distribution is compared to an alternative represented by the lognormal distribution.

Viewed from an empirical perspective, it is however difficult to determine unambiguously whether the true firm size distributions are more consistent with the lognormal or the Pareto shape, especially in the upper tail. The debate on the shape of the firm size distribution has nevertheless intensified in the last decade (Sutton 1998, Gabaix 1999, Axtell 2001, Eeckhout 2004, Levy 2009, Eeckhout 2009) and a number of novel approaches have been proposed to discriminate among candidate

distributions and to establish the length of the Pareto upper tail (Clauset et al. 2009, Malevergne et al. 2009, Bee et al. 2011). However, since multiple generative processes may lead to the same firm size distribution and the Pareto and lognormal distributions are similar in the upper tail, no discriminatory evidence has been presented so far regarding the *dynamics* behind the observed firm size distributions.

The predictions of firm growth models should be tested on the basis of multiple stylized facts, though (Brock 1999, Klette & Kortum 2004, Klepper & Thompson 2006). In the literature, a set of empirical regularities has been repeatedly observed but – to our knowledge – rarely put together so far (Sutton 1997, Caves 1998, Coad 2009):

(1) *The size distribution of firms is highly skewed.* Gibrat showed that the size distribution of establishments is approximately lognormal for a broad range of data (Gibrat 1931, Sutton 1997). Simon and co-workers, on the other hand, argued that the observed size distributions are well approximated by a Pareto distribution, at least in the upper tail (Simon & Bonini 1958, Ijiri & Simon 1977). While the exact shape of the size distribution is still debated, the Pareto and lognormal distributions are typically retained as useful benchmarks (Hall 1987, Stanley et al. 1995, Axtell 2001, Cabral & Mata 2003, Marsili 2005, Luttmer 2007, Growiec et al. 2008).

(2) *The growth rate distribution is not Gaussian but "tent-shaped"* in the vicinity of the mean growth rate (Stanley et al. 1996, Bottazzi et al. 2001, Bottazzi & Secchi 2003, Fu et al. 2005). By looking at the entire distribution, Fu et al. (2005) have also documented the rare events of extremely large positive and negative growth shocks, thanks to which the firm growth rate distribution has power-law tails.

(3) *Smaller firms have a lower probability of survival, but those that survive tend to grow faster than larger firms.* Among larger firms, growth rates are unrelated to past growth or to firm size (Mansfield 1962, Evans 1987, Hall 1987, Dunne et al. 1989, Rossi-Hansberg & Wright 2007).

(4) *The variance of growth rates is systematically higher for smaller firms* (Hymer

& Pashigian 1962, Mansfield 1962, Evans 1987). Recently, it has been also found that the variance of firms' growth rates does decay as a power-law with size, with a power of about 1/5 (Stanley et al. 1996, Sutton 2002, Riccaboni et al. 2008, Gabaix 2011).

Viewed from the theoretical perspective, business firm growth is both the outcome of a continuous growth process at the level of products, perhaps including stochastic fluctuations à la Gibrat, as well as an outcome of capturing new business opportunities thanks to innovation, which can be modeled à la Simon. Innovation may lead both to new product launches as well as to opening new product lines, divisions, subsidiaries and plants. Firm size dynamics are also largely shaped by managerial reorganizations, mergers and acquisitions. Therefore, instead of contrasting alternative generative processes, in this article we develop a more general framework that provides an unifying explanation for the growth of business firms based on the number and size distribution of their constituent units, i.e., products, submarkets, plants or divisions (Bottazzi et al. 2001, Sutton 2002, Klette & Kortum 2004, De Fabritiis et al. 2003, Klepper & Thompson 2006, Fu et al. 2005). Specifically, we present a model of proportional growth in both the number of units and their size, from which we draw some general implications on the mechanisms which sustain business firm growth and shape the resulting firm size distributions.

The idea to decompose firms into subunits has already been the subject of some recent theorizing about industry evolution (Sutton 1998, Klette & Kortum 2004, Klepper & Thompson 2006). In particular, firms in the same industry can be differentiated according to technology they use, the products they sell, the customer segment they target and the geographic area in which they operate. Sutton (1998) and Klepper & Thompson (2006) call these different activities "submarkets". In this article, we refer to a somewhat more general notion of business "units" instead, though, interpreted as *independent submarkets*. To justify this, note that, as argued in Sutton (1998), most markets are composed of various sets of products, each of which satisfies different needs and requires distinct R&D efforts and technical know-how. Thus firms diversify their activities across submarkets even within a given market. Keeping this in mind, Sutton (1998) defines submarkets as independent groups of products on the demand side, or breaks in the chain of substitutes, but allows them to be interdependent on the R&D side. Independent

submarkets are in turn such that also the R&D activities are independent across them. In this article, we shall define business "units" as independent submarkets in the latter sense. Thus, according to our definition, a business unit is an independent subset of firm activities both on the demand side (e.g., substitution) and the supply side (e.g., scope economies in R&D).

Klette & Kortum (2004) have developed a similar model where each firm is defined as a portfolio of products. Just like in the Sutton model, by catching a new business opportunity, the innovator captures the whole market for a given product there. Klepper & Thompson (2006) also model the evolution of a given market by looking at a population of firms, each of which grows over time by taking up and losing a sequence of discrete investment opportunities. The framework discussed in the current article shares all these properties. Moreover, as in Bottazzi & Secchi (2006) we do not consider the size of new business opportunities as fixed; in our model, each business unit undergoes an independent Gibrat growth process whereas in Bottazzi & Secchi (2006) a business opportunity is any event (i.e. an innovation, market shock, managerial reorganization) that conveys a set of growth microshocks to the firm.[1]

To the best of our knowledge, all the firm growth models in the literature so far have failed in accounting for at least one of the stylized facts listed above and the variety of industry structures observed across different market settings. We hope that the current article will fill this gap.

The contribution of the current article to the literature is thus twofold. First, we demonstrate that the model discussed herein is the first one ever put forward in literature to be in good agreement with all four aforementioned empirical "stylized facts". Secondly, we are also the first to derive formally the predictions of the model on all four considered dimensions for the whole range of industry setups covered by the model. Given the large variety of obtained outcomes, these theoretical results are useful for identifying the plausible generating processes behind the dynamics of any industry, based on its observed summary characteristics such as the firm size distribution, growth rate distribution, and the relationship between innovation,

[1]When specified in continuous time, the Gibrat's growth process is a geometric Brownian motion. The Black-Scholes theory of option pricing also assumes a geometric Brownian motion of stock prices. Thus in our model firms can be seen as portfolios of investments in submarkets.

growth, and firm size.

The predictions of the model will be tested in the context of the worldwide pharmaceutical industry, which is a textbook example of an industry consisting of many independent submarkets (Sutton 1998, Matraves 1999). We shall exploit a unique dataset on yearly sales of almost one million pharmaceutical products marketed by more that 7 thousand firms in 1994-2008. Information is available both at the disaggregate level of product sales, as well as in a re-aggregated form, where each product is assigned to the firm that commercialized it. According to our model, if the market is composed of many independent submarkets such as in the case of pharmaceuticals, the firm size distribution should have a lognormal body and a Pareto upper tail. Moreover, the distribution of firm growth rates should be Laplace in the center, but with power law tails. As for the size-variance relationship, in line with the predictions of our model, the data feature a slow crossover between the two limiting cases of Gibrat and Simon growth processes.

Our empirical tests are based on the dataset from the pharmaceutical industry, where the assumption of submarket independence is particularly well justified, but they are readily generalizable to other sectors of the economy as well. More generally, in industries where each firm is a portfolio of many independent and relatively stable units, the model predicts that the Simon benchmark should be more appropriate than the Gibrat's one. Conversely, when firms consist of correlated and highly unstable units, the diversification process does not work effectively and the Gibrat benchmark should be more appropriate. Therefore, our model can be used to discriminate among plausible generating processes in different industries, in the same spirit as in Sutton (1998).

The remainder of the paper is structured as follows. Section 2 describes the theoretical framework. Section 3 derives the predictions of the model under different regimes of innovation and growth. Section 4 tests the model against the data from the pharmaceutical industry. Section 5 summarizes our main findings and concludes.

2 Theoretical framework

In this section we present the key assumptions behind the Generalized Proportional Growth model (henceforth GPGM), whose selected properties have been analyzed previously by Fu et al. (2005) and Growiec et al. (2008). The GPGM is a stochastic framework that includes the Gibrat's proportional growth model and the Simon preferential attachment growth process as particular instances and can account for the empirically observed shapes of size and growth distributions as well as the real-world size-mean growth and size-variance (scaling) relationships.

The model features proportional growth at the level of both number and size of the firm business units. Business firms are viewed as economic entities consisting of a random number of units that evolve independently of one another.[2]

Two key sets of assumptions in the model are:

- the number of units in a firm grows in proportion to its existing number of units (the Simon growth process);

- the size of each unit grows in proportion to its size, independently of other units (the Gibrat growth process).

Formally, the first set of assumptions is written as:

(1) Each firm α consists of $K_\alpha(t)$ units. At time $t = 0$ there are $N(0)$ firms of unitary size. This gives a total of $n(0) = N(0)$ units in the initial period. At each moment in time, there is a constant arrival rate μ of new units, and a constant destruction rate λ. The net arrival rate of new units $\psi \equiv \mu - \lambda$ is assumed to be positive. The number of units at time t is thus $n(t) = n(0)+\psi t$. Without loss of generality, we normalize ψ to unity.

(2) With birth probability $b \in [0, 1]$, this new unit is assigned to a new firm. With probability $1 - b$, it is assigned to an existing firm α with probability $P_\alpha = (1 - b)K_\alpha(t)/n(t)$.

The first set of assumptions should be interpreted in the following way. First of all, larger incumbent firms – that is, those having more units – can afford to

[2]See also Ijiri & Simon (1977), Sutton (1997), De Fabritiis et al. (2003).

finance larger investments in R&D. Hence, these firms should be, on average, more innovative and capture more new business opportunities resulting, for instance, in a larger flow of blueprints of new products which can be sold to the market. Even in a case where innovation is done outside of the firms present in the considered submarket, larger firms would still remain in a favorable position to grab the opportunities arriving from universities, the public sector, and other R&D institutions because of their larger budgets (Klette & Kortum 2004).

More specifically, assuming *proportional* growth in the number of units per firm means that we rule out all possible comparative advantage, or equivalently, that we impose constant returns to scale in the R&D sector.[3] Assuming positive entry ($b > 0$) implies in turn that some positive percentage of total R&D output comes from the outside of the firms present in the market. When an innovator not affiliated with any of these companies is successful, she starts up a new firm which initially consists of a single unit selling this freshly innovated product, but later, it may as well grow and sell more products.

We do not model firm exit explicitly here, but nevertheless we can still say that our parameter b captures the *net* entry rate – entry minus exit – which, in a growing economy, ought to be positive over the long run. One limitation of this simplification is that the model can account neither for firm turnover nor for the expected survival time of a firm. It is implicit that all firms live forever here (although some of them may do very badly).

The second set of assumptions in the model is:

(3) At time t, each firm α has $K_\alpha(t)$ units of size $\xi_i(t)$, $i = 1, 2, ...K_\alpha(t)$ where K_α and $\xi_i > 0$ are independent random variables. At time $t = 0$, the sizes of all units are equal, $\xi_i(0) = 1$ for all i.

(4) At time $t + 1$, the size of each unit is decreased or increased by a random factor $\eta_i(t) > 0$ so that

$$\xi_i(t + 1) = \xi_i(t)\, \eta_i(t), \tag{1}$$

[3]Sutton (1998) generalizes the Simon's model considering the case in which the probability that next opportunity is filled by any currently active firm is nondecreasing in the firm size.

where $\eta_i(t)$, the growth factor of unit i, is a random variable that is independent of all other η_i's and ξ_i's. It is assumed that $E \ln \eta_i(t) \equiv m_\eta$ and $Var(\ln \eta_i(t)) = E(\ln \eta_i(t))^2 - m_\eta^2 \equiv V_\eta$.

(5) The size of every new unit arriving at time t is drawn randomly from the distribution of unit sizes $\xi_i(t)$. Its expected size is denoted as $\bar{\xi}(t)$.

A few things must be mentioned about the second set of assumptions. First, by assuming the sizes of units to fluctuate independently of each other, we imply that each unit occupies a separate market niche.[4] Empirical evidence suggests that the variance of demand shifts at the unit level (V_η) should be substantial. Second, by requiring the fluctuations to have a purely multiplicative character, we assume that demand shifts affect all units *proportionately* and that the variance of their growth rate does not depend on their size. Third, by assuming that units cannot move between firms, we imply the existence of underlying organization capital necessary for production (Luttmer 2010), created upon starting up a firm, and whose transfer between firms is too costly to occur. Finally, the framework requires also that increases in size of existing units are independent of arrivals of new units. The average size and number of units within a firm are assumed to be independent.

The economic rationale behind this set of assumptions is the following. Firstly, in a growing economy, one should expect the average *net* growth rate of unit sales to be positive, in line with the macroeconomic "stylized facts". This we capture by assuming $m_\eta \geq 0$. Nonetheless, this does not preclude the Schumpeterian motive of creative destruction, an obsolescence effect, or the existence of product life-cycles.

Secondly, the assumption that newly arriving units are, on average, proportional in size to the already existing units, is meant to capture the disembodied component of technical change. If the overall rate of technical progress is positive, as it is if $m_\eta > 0$, then it is natural to expect that not only existing units, but also newly arriving opportunities will benefit from it. Otherwise, new units would become increasingly smaller in proportion to the established ones, and the average age of units would become the crucial factor behind firm size – an assumption

[4] Sutton (1998) calls this case the "island" model.

which is at odds with evidence, and also particularly questionable in a model which abstracts from firm and unit exit.

Despite the richness of the dynamics implied by the GPGM framework, it also has a few notable limitations. They are a consequence of the simplifying character of our above assumptions, thanks to which it remains analytically tractable. First of all, by assuming that units are attached to firms forever, we rule out the possibility of competition within independent submarkets. A case like ours could arise, e.g., if every product was fully protected by a patent (at it is, to a large but not full extent, in the pharmaceutical industry), but it cannot describe non-monopolistic submarkets where different suppliers of equivalent products compete for consumers with quantity and prices.

Secondly, the GPGM framework does not allow for market selection, i.e., it does not feature any mechanism where production of the least successful products would be discontinued, and where producers of a single (or a few) unsuccessful products would be forced to exit. In reality, however, this mechanism could provide a partial explanation to the observed exceptionally high volatility of firms with a single product.

Thirdly, this version of the model does not capture the general life-cycle patterns of products; instead, unit sizes are allowed to vary according to a scale-free multiplicative process, irrespective of the unit's age.

Fourthly, the model is not *stable*, in the sense of providing a stationary firm size and firm growth rate distribution with parameters that would be constant across time. Instead, it describes a *growing* economy; furthermore, due to the Gibrat process at the level of units, this economy is growing not only in its mean, but also in variance. In Appendix A, we put forward a modification of the GPGM, guaranteeing that it would deliver stationary firm size and growth rate distributions. The stabilization device used there builds primarily upon the results of Kalecki (1945) and a few other contributions, summarized by De Wit (2005) and Luttmer (2010). Unfortunately, it no longer features *proportional* growth at the level of units, which renders it less analytically tractable. A full analysis of such an extended model is thus clearly beyond the scope of the current article. The size–mean growth rate relationship is also affected by this change, and it is not clear which of the two setups should be preferred based on the trends observed in empirical data.

Keeping these limitations in mind, in the following section we shall derive the predictions of our model with respect to:

(1) the size distribution of firms $P(S)$; where the size of each firm is defined as $S_\alpha(t) \equiv \sum_{i=1}^{K_\alpha(t)} \xi_i(t)$;

(2) the distribution of firm growth rates $P(g)$ defined as $g_\alpha \equiv \ln\left(\frac{S_\alpha(t+1)}{S_\alpha(t)}\right)$;

(3) the size-mean growth rate relationship, summarized by the shape of $E(g|S)$ viewed as a function of S;

(4) the size-variance relationship, summarized by the parameter β in the power-law relationship of form $\sigma(g|S) \propto S^{-\beta}$.

We proceed in a systematic way. First, we list all the sub-cases of the model, each of which implies a qualitatively different mode of its behavior, and then study these cases consecutively. The cases singled out there are delineated by the assumptions on:

- the entry regime of new business opportunities: $\psi = 0$ or $\psi > 0$;
- the entry regime of firms: if $b = 0$ then all new opportunities are captured by existing firms, whereas with $b \in (0, 1)$ there is a nonzero probability that a new opportunity will assigned to a new start-up firm;
- the volatility of the unit growth rate: $V_\eta > 0$ allows the Gibrat's mechanism of proportional growth to operate at the unit level; whereas $V_\eta = 0$ switches it off, implying constant growth in unit sizes, or keeping unit sizes constant. In the latter case, the model boils down to the Simon urn model;
- the time horizon of the growth process: when it is infinite then we look at the limit distribution, otherwise, it is stopped at a finite time.

For every special case of the GPGM, we derive the predictions relevant for the stylized facts (1)–(4), presented in the introduction. We refer to already known results wherever possible.

3 Results

Before turning to the key results for our GPGM model, let us first sort out the trivial case where there is no entry of firms and units and no variance of the shocks affecting sales of business units ($b = \psi = V_\eta = 0$). In this case the initial one-point size distribution of firms $K_\alpha(0)$ is maintained across time and the growth rate distribution is degenerate; all firms grow at the same rate $m_\eta \geq 0$. An equally trivial case follows when we impose $b = 1$ and $\psi > 0$, which is equivalent to saying that each new unit goes to a new firm. If there is no variability in the unit growth rates, then again, the size distribution is a one-point distribution at all times t and each firm has the same size e^{tm_η}. The growth rate distribution is degenerate as well: all firms grow at a rate m_η. Because of the degenerate growth rate distribution, the size-variance relationship cannot be calculated in both cases.

Much more interesting are the cases where the size of existing units is allowed to vary ($V_\eta > 0$) and where both existing and new start-up firms innovate, capturing new business opportunities and thus opening up new units ($\psi > 0$). We shall first analyze the case of pure proportional growth in the size of units without innovation (the Gibrat case) and the case of growth in the number of units of a given size (the Simon case). Finally we will consider the GPGM in which both the Gibrat and the Simon growth processes are simultaneously at work.

3.1 The Gibrat case

In the case of a pure Gibrat growth process, neither new units nor new firms enter the market, but sales of each unit fluctuate idiosyncratically over time with a positive variance ($b = \psi = 0$ and $V_\eta > 0$).

Since each firm consists of exactly one unit, an application of the Central Limit Theorem to the logarithm of unit sizes yields the prediction of the firm size distribution approaching the lognormal in the limit of $t \to \infty$, regardless of the actual distribution of the growth rates $\ln \eta_i$ (Gibrat 1931, Kalecki 1945). Since due to the absence of new business opportunities, each firm has a single unit, the size distribution of firms is the exactly same as the size distribution of units: it is log-normal, $P(S) = P(\xi_i)$. The growth rate distribution, on the other hand, is the same as the distribution of $\ln \eta_i$ and there is no force capable of altering this over

time.

The size-mean growth rate and size-variance relationships found for this case is an important benchmark for further comparisons: irrespective of firm size, the mean of its growth rate is constant at $E(g|S) = Eg = m_\eta$, and variance of its growth rate is constant at $\sigma^2(g|S) = \sigma^2(g) = V_\eta$. Thus, the parameter β in the relationship of form $\sigma(g|S) \propto S^{-\beta}$ is zero. As we shall see shortly, in all other cases of the model, β will be positive.

A very similar case, with exactly the same results as for the pure Gibrat process, is obtained when arrivals of new units are allowed, but each new opportunity is assigned to a new firm ($b = 1, \psi > 0$), and there is variability at the level of unit sizes, with variance V_η.[5]

3.2 The Simon growth process without firm entry

When new business opportunities appear but they are all captured by the initially existing firms ($b = 0, \psi > 0$), and when all units grow uniformly ($V_\eta = 0$), with entering units being of the same size as the pre-existing ones – then the size of the representative unit at t is e^{tm_η} and we are in the case of an urn scheme with a fixed number of bins. Such schemes have been analyzed, among others, by Ijiri & Simon (1977), De Vany & Walls (1996), Bottazzi & Secchi (2006).[6] When $t \to \infty$, the size distribution of firms (having the same shape as the distribution of the number of units per firm since $S \propto K$ in the current case) converges to a geometric distribution (Fu et al. 2005):

$$P(S) = \frac{1}{\kappa e^{tm_\eta} - 1} \left(1 - \frac{1}{\kappa e^{tm_\eta}}\right)^S \approx \frac{1}{\kappa e^{tm_\eta}} e^{-\frac{S}{\kappa e^{tm_\eta}}}, \tag{2}$$

where $\kappa = \kappa(t) \equiv \left(\frac{N(0)+t}{N(0)}\right)$ is the average number of independent units per firm at time t. As $t \to \infty$, the average size of a firm κe^{tm_η} increases exponentially with time unless $m_\eta = 0$ when it increases only linearly with time.

[5] If one relaxes the assumption that new units are created at each time step and assumes a constant *probability* of arrival of new business opportunities over time instead, then Reed (2001) has shown that under a finite time horizon, the business size distribution will be Double Pareto and the growth distribution will be Laplace. See also Kotz et al. (2001), Bottazzi & Secchi (2006).

[6] Some authors refer to this case as the Bose-Einstein urn scheme. See also Feller (1957).

Upon the use of a continuous limit approximation – made in order to attain comparability with other cases considered in this paper – the original geometric distribution derived in (2) becomes an exponential distribution.

If $m_\eta > 0$ then the growth rate distribution of units (and hence firms, since the influx of new units can be ignored in the limit) at $t \to \infty$ is degenerate: all units grow at the same rate. If $m_\eta = 0$ then, due to a different argument, the growth rate will still tend to a degenerate distribution, this time concentrated in zero. The argument is based on the fact that all firms' sizes will tend to infinity with time if no firm entry is allowed. Thus, expected growth rate given by $(1-b)/n(t)$ will tend to zero.

The size-variance relationship is determined by the relationship between the variance of firms' growth rates and the number of their units, K. If $m_\eta = 0$ (units are of constant unitary size) then variance scales with size as $1/K$ implying $\beta = 1/2$. To see this, note that a firm with $K(t)$ units at time t will have $K(t+1) = K(t)+1$ units at $t+1$ with probability $K/n(t)$ (proportional growth) and $K(t+1) = K(t)$ units with probability $1 - K/n(t)$. Hence, the growth rate of this firm is:

$$g(t) = \begin{cases} \frac{1}{K} & \text{with probability } K/n(t), \\ 0 & \text{with probability } 1 - K/n(t). \end{cases} \tag{3}$$

It follows that the variance $\sigma^2(g|K) = \frac{1}{Kn(t)} - \frac{1}{n(t)^2} \propto \frac{1}{K}$ provided that $n(t)$ is large enough (which is for sure the case when $t \to \infty$).

If, on the other hand, $m_\eta > 0$, then the exponential growth in unit sizes will, over time, dominate the total size of the firm and the influx of new units will only have a negligible impact on its size. In such case, we would have a degenerate growth distribution and the size-variance relationship could not be calculated.

Upon stopping the Simon growth process without firm entry at a finite time t, we observe the following differences with respect to the results derived above. First, due to the finite-time truncation, the size distribution might have not fully converged to the exponential distribution and some relicts of the initial firm size may still be visible. Second, the growth rate distribution becomes a discrete distribution with a finite number of atoms,[7] with mean slightly above m_η. This is due to the fact that at any finite time t, the impact of influx of new units to firms cannot

[7]By an *atom* we mean an isolated point where the Cumulative Distribution Function (CDF)

be completely ignored, as it is the case in the limit $t \to \infty$. The size–mean growth rate relationship is still flat, and the size–variance relationship is still captured by the scaling parameter $\beta = 1/2$. Last three results are easily obtained by inserting $b = 0$ into the derivations done for the more general case of the Simon growth process with firm entry, discussed below.

3.3 The Simon growth process with firm entry

The more sophisticated firm growth model due to Ijiri & Simon (1977), building on the early contribution of Yule (1925), allows for net entry of new firms into the market, $b \in (0, 1)$ with $\psi > 0$ and $V_\eta = 0$. This model is thus also a special case of the GPGM. First, we shall deal with the limiting case of $t \to \infty$, in which the dynamic system at hand has been given infinite time for evolution. In this case, the distribution of the number of units per firm converges to the Pareto (power law) distribution with the exponent $1/(1-b)+1 > 2$ (Fu et al. 2005):

$$P(K) = \left(\frac{1}{K}\right)^{\frac{1}{1-b}+1} \int_0^K e^{-y} y^{\frac{1}{1-b}} dy \sim \left(\frac{1}{K}\right)^{\frac{1}{1-b}+1}, \tag{4}$$

which simplifies to $P(K) = \frac{1}{K^2}(1-(K+1)e^{-K}) \sim 1/K^2$ for $b \to 0_+$.

Since there is no randomness at the unit level, the size distribution $P(S)$ is very similar to $P(K)$. It is given by

$$P(S) = e^{\frac{tm_\eta}{1-b}} \left(\frac{1}{S}\right)^{\frac{1}{1-b}+1} \int_0^{Se^{-tm_\eta}} e^{-y} y^{\frac{1}{1-b}} dy \sim \left(\frac{1}{S}\right)^{\frac{1}{1-b}+1}, \tag{5}$$

i.e. it follows an approximate power law with an exponent $1/(1-b)+1$.

As far as the growth rate distribution, the size-mean growth and size-variance relationships are concerned, the current situation follows closely the lines of the Simon model with no firm entry. Again, if $m_\eta > 0$ then the growth rate distribution becomes a one-point distribution – the growth rate of every firm converges to $\bar{g} = m_\eta$, regardless of its size – and so in the limit, the scaling relationship cannot be calculated. Accordingly, if $m_\eta = 0$ then the scaling relationship is still summarized by $\beta = 1/2$, and the growth rate distribution still tends to concentrate in zero as $t \to \infty$. In the current case, however, this will not happen because of the average

of a distribution has a jump.

size of firms going to infinity (as there is a constant inflow of small startup firms in the current case) but rather because of the probability of getting a new unit in the firm tending to zero as the number of firms rises towards infinity.

Let us see what changes once we relax the assumption that $t \to \infty$. If the process of corporate dynamics does not evolve forever but is stopped at a finite time instead (as it must be the case in reality), we observe non-trivial truncation effects. First of all, we find that the Pareto distribution of $P(K)$ can only form in infinite time; otherwise, it is truncated with an approximately exponential cutoff (Fu et al. 2005, Yamasaki et al. 2006). More precisely, Fu et al. (2005) have shown that when t is finite, $P(K)$ is:

$$P(K) = \frac{N(0)}{N(0)+bt} P_{\text{old}}(K) + \frac{bt}{N(0)+bt} P_{\text{new}}(K), \tag{6}$$

where $P_{\text{old}}(K)$ denotes the distribution of number of units in firms present in the economy already at $t = 0$, and $P_{\text{new}}(K)$ denotes the respective distribution in firms founded later. These two distributions are given by, respectively:

$$P_{\text{old}}(K) \approx \frac{1}{\kappa} e^{-\frac{K}{\kappa}}, \tag{7}$$

$$P_{\text{new}}(K) \approx \frac{1+n(0)/t}{1-b} \left(\frac{1}{K}\right)^{\frac{1}{1-b}+1} \int_{K/(\kappa^{1-b})}^{K} e^{-y} y^{\frac{1}{1-b}} dy. \tag{8}$$

Remembering that κ is a function of t, we notice that both $P_{\text{old}}(K)$ and $P_{\text{new}}(K)$ are affected by the elapsing time.

Since all units are of equal size due to $V_\eta = 0$, and thus $S = Ke^{tm_\eta}$, the $P(S)$ distribution is proportional to $P(K)$. To compute it, it suffices to re-scale $P(K)$ given in (6). We conclude that the firm size distribution $P(S)$ is Pareto with an exponential cutoff, whereby the cutoff is obtained due to the finite-time evolution of the considered system.[8]

[8] Yamasaki et al. (2006) have estimated the parameters of this equation using the same pharmaceutical industry database as we use here. It turned out that the Pareto part (which is obtained for relatively small K) has an exponent of around 2.14 (corresponding to $b \approx 0.123$), and then the cutoff part begins around $K \approx 200$, where the Probability Density Function (PDF) $P(K)$ decays as $\exp(-0.0054K)$. Simulations carried out in that paper show that the longer is the time t of the evolution of the considered system, the later the cutoff part begins. Neither the Pareto exponent nor the cutoff slope are affected by the changes in t, though.

As regards the growth rate distribution at any finite moment in time, with $m_\eta = 0$ it is equal to

$$P(g) = \sum_{K=1}^{\infty} P(K)P(g|K), \tag{9}$$

where the conditional distribution $P(g|K)$ given by:

$$g(t) = \begin{cases} \frac{1}{K} & \text{with probability } (1-b)\frac{K}{n(t)}, \\ 0 & \text{with probability } 1-(1-b)\frac{K}{n(t)}. \end{cases} \tag{10}$$

By K we denote the size (equivalent to the number of units in the current case) of a given firm, and by $n(t)$ we denote the total number of units in the economy at time t. Hence, the distribution of firm growth rates is a discrete distribution with probability mass concentrated in atoms $\{0\} \cup \{\frac{1}{n}\}_{n=1}^{K^{max}}$, with K^{max} being the maximum number of units per firm at time t. The probability mass associated with each unit is equal to $P(K)K/n(t)$ and so the distribution converges to a degenerate distribution concentrated in zero linearly with time t. Yet due to the finite-time truncation, the impact of the arrivals of new opportunities on the growth rate distribution is not completely washed away.

If $m_\eta > 0$, then the conditional distribution $P(g|K)$ becomes:

$$g(t) = \begin{cases} m_\eta + \ln\left(1+\frac{1}{K}\right) & \text{with probability } (1-b)\frac{K}{n(t)}, \\ m_\eta & \text{with probability } 1-(1-b)\frac{K}{n(t)}. \end{cases} \tag{11}$$

In result, it is a distribution that, again, has a finite number of atoms. The probability mass associated with each unit is equal to $P(K)K/n(t)$ and thus it converges to the one-point distribution concentrated at m_η linearly with time t.

From equation (11) we also immediately infer that at any finite time t, the size-mean growth rate relationship is given by:

$$\begin{aligned} E(g|K) &= \left(1-(1-b)\frac{K}{n(t)}\right)m_\eta + (1-b)\frac{K}{n(t)}\left(m_\eta + \ln\left(1+\frac{1}{K}\right)\right) \approx \\ &\approx m_\eta + \frac{1-b}{n(t)}, \end{aligned} \tag{12}$$

and thus it is independent of K. We observe positive departures from the limit growth rate m_η if $n(t)$ is small, though. The interpretation of this finding is natural

– the flat relationship between $E(g|K)$ and K comes from the assumptions of both proportional growth at the unit level and preferential attachment of new units to larger firms, whereas positive departures from the limit growth rate if t is small come from the fact that initially most firms have only few units, and so capturing new business opportunities leads to visible jumps in their size. However, as the economy grows large, the probability that a given small firm gets such a new business opportunity converges to zero, and in consequence its *expected* growth rate goes down to m_η.

Analogously to the Simon growth process without firm entry (since $S \propto K$), the size-variance relationship is approximately $\sigma^2(g|S) \propto 1/S$ (implying $\beta = 1/2$) here. There are two differences, though: first, now the relationship cannot be calculated for any S: no firm can grow arbitrarily large in finite time; second, since the growth rate distribution is not entirely degenerate, the scaling relationship can be obtained also if $m_\eta > 0$.

3.4 GPGM with no firm entry

In the case where new units arrive at every t ($\psi > 0$) and there is also variability at the unit level ($V_\eta > 0$), the model gets substantially more complex. We shall first deal with the case of no firm entry ($b = 0$). We are then observing both the Gibrat's proportional stochastic growth process at the unit level, giving rise to a lognormal size distribution of units in the limit, and a proportional growth process at the level of firms, giving rise to an exponential distribution of the number of units per firm.

The resulting firm size distribution may be calculated using the procedure developed in Growiec et al. (2008) as

$$\mathcal{P}(S) = \sum_{K=1}^{\infty} P(K)\mathcal{P}(S|K), \tag{13}$$

where $\mathcal{P}$ denotes the complementary CDF of a random variable and P denotes its PDF. Taking the limit $t \to \infty$ and ignoring the influx of new units (which we can do because with no entry, every firm becomes arbitrarily large with time, and thus the effects of entry become negligible), in Growiec et al. (2008) we have found that

the distribution of sizes S given K – a distribution of a sum of lognormal variables[9] – can be approximated by a mixture of Slimane (2001)'s upper and lower bound:

$$\mathcal{P}(S|K) \approx 1 - \left[\Phi\left(\frac{\ln(S/K^{\gamma}) - tm_{\eta}}{\sqrt{tV_{\eta}}}\right)\right]^{K}, \tag{14}$$

with $\gamma \in [0,1]$.[10]

In the end, the PDF $P(S)$ is obtained, using $P(K)$ from (2), and denoting $\Phi\left(\frac{\ln(S/K^{\gamma})-tm_{\eta}}{\sqrt{tV_{\eta}}}\right) \equiv h(S)$, as:

$$P(S) = -\mathcal{P}'(S) = \underbrace{h'(S)}_{\text{log-normal}} \times \underbrace{\sum_{K=1}^{\infty} Kh(S)^{K-1}\lambda e^{-\lambda K}}_{\text{stretching factor}}, \tag{15}$$

where $\lambda = 1/\kappa$ is the reciprocal of the average number of units per firm.

After some manipulations we obtain that the stretching factor in the current case is uniformly *bounded* for all S:

$$\begin{aligned} &\sum_{K=1}^{\infty} Kh(S)^{K-1}\lambda e^{-\lambda K} \approx \frac{\lambda}{h(S)}\int_{1}^{\infty} Ke^{K(-\lambda+\ln h(S))}dK = \\ = \;& \lambda e^{-\lambda}\left(\frac{1}{\lambda - \ln h(S)} + \frac{1}{(\lambda - \ln h(S))^2}\right), \end{aligned} \tag{16}$$

and thus the stretching factor converges to $\frac{1+\lambda}{\lambda}e^{-\lambda}$ when $h(S) \to 1_{-}$. Hence, it is confined within the interval $(0, \frac{1+\lambda}{\lambda}e^{-\lambda})$. Boundedness of the stretching factor means that the departures from the lognormal shape of the size distribution cannot

[9]Unfortunately, the distribution of a sum of log-normally distributed random variables cannot be expressed in a closed analytical form (Slimane 2001).

[10]The parameter γ captures the distance to the upper and lower bound, within which the complementary CDF of a sum of log-normally distributed variables must be comprised (Slimane 2001). $\gamma = 0$ is the lower bound approximation, $\gamma = 1$ is the upper bound approximation, and $\gamma \in (0,1)$ captures all intermediate cases. Please note that in all the approximations, we consider γ to be a free parameter It is implicit in Slimane (2001) that for larger variances V_{η} one should expect γ to be smaller; also, for larger sizes S should γ decrease. For large firm sizes S as well as large variances V_{η}, the correct approximation might be $\gamma \approx 0$. We are however not aware of any Monte Carlo simulations aiming at assessing the true relationship between γ and the parameters of the underlying lognormal distributions and thus remain agnostic with respect to this point.

be arbitrarily large. In particular, this implies that the right tail of the distribution decays as a lognormal distribution and *not* as a power law.

As far as the growth rate distribution is concerned, an approximate result has been obtained by Fu et al. (2005):

$$P(g) = \frac{\sqrt{\kappa}}{2\sqrt{2V}} \left(1 + \frac{\kappa}{2V}(g - \bar{g})^2\right)^{-\frac{3}{2}}, \tag{17}$$

where $\bar{g} = m_\eta + V_\eta/2$ and $V = e^{tV_\eta}(e^{V_\eta} - 1)$. This result is approximate in the sense that in the course of calculating it, a Central Limit Theorem approximation has been used: $P(g|K)$ is approximated by a Gaussian distribution for all K while in theory it must be the case only for sufficiently large K. This approximation is quite precise, however, for $|g - \bar{g}| < V_\eta$, or if the assumed distribution of $\ln \eta_i$ is close to a Gaussian.

Please note that $P(g)$ is symmetric around the mean growth rate $\bar{g}$ and that it decays as a power law with an exponent of three ($P(g) \sim g^{-3}$).

As regards the size-mean growth rate relationship, Fu et al. (2005) have found that as $t \to \infty$ (and thus $n(t) \to \infty$ so the influx of new units can be ignored), the mean growth rate of firms with K units and of size S converges to $\bar{g} \equiv m_\eta + V_\eta/2$, regardless of K and S. The size-mean growth relationship is thus asymptotically flat.

The size-variance relationship will be dealt with in two steps. In the first step, we shall derive the probability distribution of partitioning total firm size S into $K = 1, 2, 3, ...$ units. This means that for each given firm size S, we will calculate the posterior probability distribution of $P(K|S)$ using the Bayes' law. In the second step, we will use the law of total variance to infer from $P(K|S)$ and $\sigma^2(g|K)$ the overall size-variance relationship $\sigma^2(g|S)$.

Regarding the first step, the posterior distribution of partitions $P(K|S)$ is obtained, with the use of the results described above as well as the Slimane (2001)

bounds approximation, as:

$$P(K|S) = \frac{P(S|K)P(K)}{P(S)} = \tag{18}$$

$$= \frac{\lambda e^{-\lambda K} K \left[\Phi\left(\frac{\ln(S/K^{\gamma})-tm_{\eta}}{\sqrt{tV_{\eta}}}\right)^{K-1}\right] \exp\left(-\frac{(\ln(S/K^{\gamma})-tm_{\eta})^2}{2tV_{\eta}}\right)}{\sum_{K=1}^{\infty} \lambda e^{-\lambda K} K \left[\Phi\left(\frac{\ln(S/K^{\gamma})-tm_{\eta}}{\sqrt{tV_{\eta}}}\right)^{K-1}\right] \exp\left(-\frac{(\ln(S/K^{\gamma})-tm_{\eta})^2}{2tV_{\eta}}\right)}.$$

In the second step, to calculate the size-variance relationship in the total distribution of firm sizes, we shall use (i) the partition derived just above, and (ii) the relationship between the variance of the growth rate given $\sigma^2(g|K)$ and K itself. As far as (ii) is concerned, our starting point are the results obtained in Fu et al. (2005), signifying that this relationship should be well approximated by

$$\sigma^2(g|K) = \frac{e^{tV_{\eta}}(e^{V_{\eta}} - 1)}{K} \equiv \frac{V}{K}. \tag{19}$$

Equation (19) implies a $1/K$ scaling relationship between variance and the number of units, exactly the same as in the Simon case where unit sizes are deterministic. This is however by no means a robust result and thus at least three remarks must be made here. First, the relationship summarized in (19) must hold for large K but need not hold for small K such as $K = 1$ or $K = 2$. Simulations of the model show that for small K, the scaling is in fact better approximated by $\sigma^2(g|K) \approx 1/K^{2\tilde{\beta}}$ with $\tilde{\beta}$ depending on V_{η} ($\tilde{\beta} \to 1/2$ when $V_{\eta} \to 0$ and $\tilde{\beta} \to 0$ when $V_{\eta} \to \infty$). Second, empirical observations tend to suggest that already at the level of $\sigma^2(g|K)$, the scaling relationship is markedly flatter than $1/K$ (Fu et al. 2005, Riccaboni et al. 2008). Third, the reason to use this approximation in the analysis is that, as we shall see shortly, even with the counterfactual, too steep $1/K$ scaling (see also the discussion in Section 3.6), it is still likely that the overall size-variance relationship $\sigma^2(g|S)$ predicted by the model will be flatter than $1/S$, in line with the empirical evidence.

Knowing the posterior distribution of partitions $P(K|S)$, and we can obtain the size–variance relationship by using the law of total variance,

$$\sigma^2(g|S) = E\sigma^2(g|K,S) + \sigma^2(E(g|K,S)), \tag{20}$$

and using the fact that $E(g|K,S) = m_\eta + V_\eta/2$ independently of K and S so that the second term in the sum is zero. The results are the following:

$$\begin{aligned}
\sigma^2(g|S) &= E\sigma^2(g|K,S) = \\
&= \sum_{K=1}^{\infty} P(K|S)\sigma^2(g|K,S) = \qquad (21)\\
&= \frac{\sum_{K=1}^{\infty} V\lambda e^{-\lambda K}\left[\Phi\left(\frac{\ln(S/K^\gamma)-tm_\eta}{\sqrt{tV_\eta}}\right)^{K-1}\right]\exp\left(-\frac{(\ln(S/K^\gamma)-tm_\eta)^2}{2tV_\eta}\right)}{\sum_{K=1}^{\infty}\lambda e^{-\lambda K}K\left[\Phi\left(\frac{\ln(S/K^\gamma)-tm_\eta}{\sqrt{tV_\eta}}\right)^{K-1}\right]\exp\left(-\frac{(\ln(S/K^\gamma)-tm_\eta)^2}{2tV_\eta}\right)}.
\end{aligned}$$

The infinite series in equation (21) do not have closed-form sums, but our subsequent numerical work reveals that the underlying size-variance relationship follows an approximate power law ($\sigma^2(g|S) \propto S^{-\beta}$) and that β can, in principle, take any value in the range $(0, 1/2)$, depending on the values of V_η and λ. The zero limit is converged to when the variance $V_\eta \to \infty$ and the $\beta = 1/2$ limit works well for very small variances V_η, in accordance with the $\beta = 1/2$ result we have obtained for the case $V_\eta = 0$ (Riccaboni et al. 2008).[11] The dependence of the size-variance relationship on λ is much less pronounced than on V_η.

Upon stopping the GPGM without firm entry at a certain finite time t, we observe the following differences with respect to the results presented above. First, as shown in numerical simulations, the firm size distribution will still carry some relicts of the initial distribution assumed for $t = 0$. Second, the growth rate distribution will also be somewhere in between the distribution (17) and the assumed distribution of $\ln \eta_i$ (converging to the former one with time t). Third, regarding the size–mean growth relationship, due to the fact that the impact of influx of new units on the growth rate distribution cannot be completely ignored if time is finite, we shall observe $E(g|K,S)$ not as being flat, but falling with K, S, and $n(t)$, and converging only gradually to the limit value $\bar{g} = m_\eta + V_\eta/2$. Fourth, regarding the size–variance relationship, the scaling relationship will still be described with $\beta \in (0, 1/2)$. The last two results are obtained as a special case of the full GPGM with firm entry described below, by taking $b = 0$.

[11] When $V_\eta \to \infty$, $\gamma = 0$ must be used, and conversely, when $V_\eta \to 0$, $\gamma = 1$ must be used.

3.5 Full GPGM with firm entry

Allowing the entry of both new units and firms ($b > 0$, $\psi > 0$), we finally arrive at the full GPGM which is based on a mixture of a proportional growth process with net entry at the level of firms, and a Gibrat growth process with net entry at the level of their constituent units ($V_\eta > 0$). As we shall see in the next section, the predictions obtained for this case align with stylized facts (1)–(4) very well and are also in good agreement with empirical evidence present in our dataset on the worldwide pharmaceutical industry.

In terms of the firm size distribution, a slight generalization of the results presented in Growiec et al. (2008) gives the following result for $t \to \infty$:

$$P(S) = -\mathcal{P}'(S) = \underbrace{h'(S)}_{\text{log-normal}} \times \underbrace{\sum_{K=1}^{\infty} h(S)^{K-1} \left(\frac{1}{K}\right)^{\frac{1}{1-b}} \int_0^K e^{-y} y^{\frac{1}{1-b}} dy}_{\text{stretching factor}}. \tag{22}$$

The stretching factor is again uniformly bounded for each given $b > 0$ (as it was in the case of $b = 0$). However, as $b \to 0_+$, these upper bounds diverge to infinity now, signifying that in the case of very low but positive entry, the stretching factor can in fact be arbitrarily large, giving rise to an approximate power law decay of $P(S)$ for very large S.

Indeed, when $h(S) \to 1_-$, then the stretching factor is approximately equal to $\sum_{K=1}^{\infty} \left(\frac{1}{K}\right)^{\frac{1}{1-b}} \approx \int_1^{\infty} \left(\frac{1}{x}\right)^{\frac{1}{1-b}} dx = \frac{1-b}{b}$ which tends to infinity as $b \to 0_+$.

As far as the growth rate distribution is concerned, we shall refer to Fu et al. (2005) as well as Buldyrev et al. (2007) for the approximate result in the case $|g - \bar{g}| < \sqrt{V_\eta}$:

$$P(g) = \frac{1}{1-b} \frac{1}{\sqrt{2\pi V}} \int_0^{\infty} e^{-y} y^{\frac{1}{1-b}} \left(\int_y^{\infty} e^{-\frac{(g-\bar{g})^2 K}{2V}} K^{-\frac{1}{2}-\frac{1}{1-b}} dK \right) dy, \tag{23}$$

which simplifies when $b \to 0_+$ to

$$P(g) = \frac{2V}{\sqrt{(g-\bar{g})^2 + 2V}(|g-\bar{g}| + \sqrt{(g-\bar{g})^2 + 2V})^2}. \tag{24}$$

The Fu growth rate distribution (24) combines a Laplace cusp at $g \approx \bar{g}$ and power-law wings, decaying as g^{-3} when $|g - \bar{g}| \to \infty$. In the intermediate range of g, there is a crossover from a Laplace distribution to a power law.

Because with stochastic fluctuations at the level of units and unit entry, the mean firm growth rate converges to $\bar{g} = m_\eta + V_\eta/2$, and thus is positive in the long run even if $m_\eta = 0$, the limiting size–mean growth rate relationship is necessarily flat for all $m_\eta \geq 0$. Furthermore, as the total number of units in the economy $n(t) \to \infty$ with time, this convergence occurs uniformly for all K and S.

As far as the size-variance relationship is concerned, the situation is quite similar to the one observed with $b = 0$ and unit entry. The only difference is that $P(K)$ is now Pareto instead of being exponential. In sum, the size-variance relationship is:

$$\begin{aligned} \sigma^2(g|S) &= E\sigma^2(g|K,S) \approx \qquad (25) \\ &\approx \frac{\sum_{K=1}^{\infty} V \left(\frac{1}{K}\right)^{\frac{1}{1-b}+1} \left[\Phi\left(\frac{\ln(S/K^\gamma)-tm_\eta}{\sqrt{tV_\eta}}\right)^{K-1}\right] \exp\left(-\frac{(\ln(S/K^\gamma)-tm_\eta)^2}{2tV_\eta}\right)}{\sum_{K=1}^{\infty} \left(\frac{1}{K}\right)^{\frac{1}{1-b}} \left[\Phi\left(\frac{\ln(S/K^\gamma)-tm_\eta}{\sqrt{tV_\eta}}\right)^{K-1}\right] \exp\left(-\frac{(\ln(S/K^\gamma)-tm_\eta)^2}{2tV_\eta}\right)}. \end{aligned}$$

The approximation comes from using Slimane (2001) bounds for the sum of lognormally distributed variables as well as replacing the true $P(K)$ with a pure power law distribution. The latter approximation is valid for large S, which is the range we are particularly interested in.

Again, the infinite series defined above do not offer closed-form sums, but we can figure out numerically what is the relationship between $\sigma^2(g|S)$ as defined in (25), and S as such (see Figure 1).

Again, we see that the β slope in $\sigma^2(g|S) \propto S^{-\beta}$ falls with V_η. We also confirm that $\beta \to 1/2$ when $V_\eta \to 0$ and $\beta \to 0$ when $V_\eta \to \infty$. The only difference we find between the scaling relationships predicted within this case and the case without firm entry is that here we see a substantial non-linearity in the plot (i.e. a substantial departure from the power law) in the range of small sizes S, especially when the variance V_η is large. Similar results have been obtained numerically by Riccaboni et al. (2008).

Let us now elaborate the most complex of our cases: with both unit and firm entry ($b > 0, \psi > 0$), stochastic fluctuations at the level of units ($V_\eta > 0$), and only a finite time of evolution.

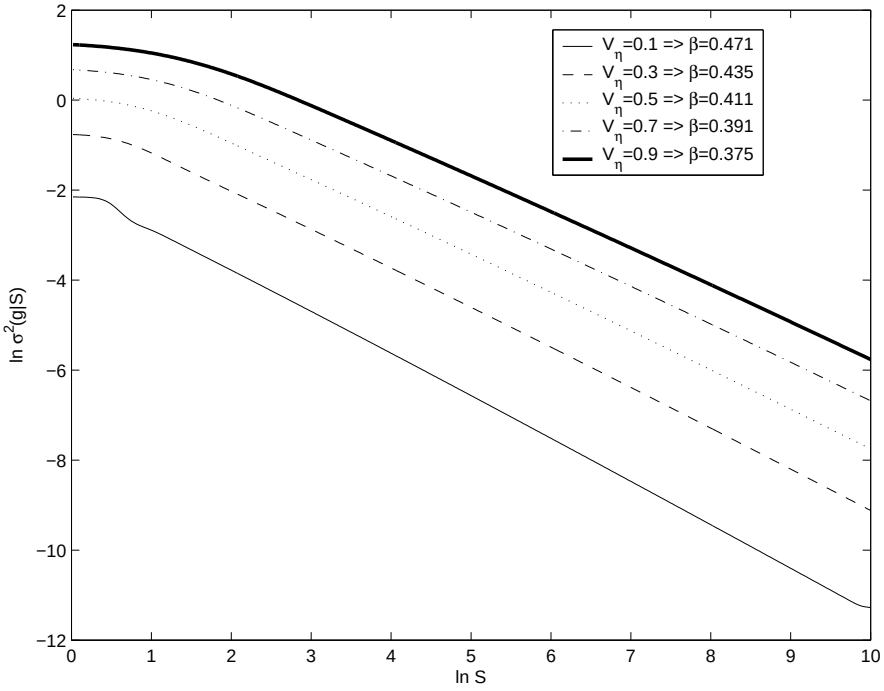

Figure 1: The size-variance relationship as a function of $V_\eta > 0$ in GPGM with firm entry. Assumed parameter values: $m_\eta = 0.001, b = 0.1$ as well as $t = 1$ (in fact, this does not matter alone, only relative to V_η) and $\gamma = 0.95$ (γ must be close to 1 since tV_η is relatively small in the example).

As far as $P(K)$ is concerned, in such case we observe the Pareto distribution with an exponential cutoff, summarized in (6). The overall size distribution $P(S)$ is quite similar to the one obtained in the case of $V_\eta > 0$ and $t \to \infty$ and is given by

$$P(S) = -\mathcal{P}'(S) = \underbrace{h'(S)}_{\text{log-normal}} \times \underbrace{\sum_{K=1}^{\infty} P(K) K h(S)^{K-1}}_{\text{stretching factor}}, \tag{26}$$

with $P(K)$ as in (6). Hence, the stretching factor is in the current case a convex combination of the stretching factors obtained for the two aforementioned cases of (i) GPGM without firm entry, and (ii) GPGM with firm entry, as $t \to \infty$. The parameters of this convex combination are given by $\frac{N(0)}{N(0)+bt}$ and $\frac{bt}{N(0)+bt}$, respectively. The stretching factor in (26) is thus bounded, but the greater is t and the smaller is b, the larger is its magnitude. The stretching factor may become arbitrarily

large only when $b \to 0_+$ and $t \to \infty$. This means that the size distribution $P(S)$ is essentially log-normal in the current case, with a possible approximately power law departure for large S, but nevertheless eventually decaying as a log-normal distribution, which is an effect of the finite-time truncation.

The growth rate distribution, $P(g) = \sum_{K=1}^{\infty} P(K)P(g|K)$, is too complex in this case to be computed analytically. To circumvent this problem, we have carried out a series of numerical exercises. These exercises confirm that as time t passes, the growth rate distribution evolves gradually from $P(\ln \eta_i)$ at $t = 0$ (when all firms have single units like in a pure Gibrat process) to a distribution exhibiting power-law wings (when the role of "old" firms is still important and when all firms are still relatively small in terms of K), and finally, when $t \to \infty$, to the distribution summarized in (23), exhibiting a Laplace cusp and power-law wings decaying as g^{-3}.

Needless to say, the assumptions made in relation to the distribution of $\ln \eta_i$ are crucial for the shape of the growth rate distribution $P(g)$ at any finite time t.[12]

As regards the dependence between firm size and its mean growth rate, finite-time truncation enables us finally to observe significant departures from the generic flat relationship inherent in the discussed *proportional* growth model. These departures are especially pronounced if K, S and t (and thus $n(t)$) are small, because then the increases in firm size due to catching new business opportunities are most clearly visible. In the light of our assumptions, the expected growth rate conditional on S and K is computed as:

$$\begin{aligned} E(g|S,K) &= \left(1-(1-b)\frac{K}{n(t)}\right) E\left(\ln\left(\frac{\sum_{i=1}^{K}\xi_i\eta_i}{S}\right)\middle|S\right) + \\ &+ (1-b)\frac{K}{n(t)} E\left(\ln\left(\frac{\sum_{i=1}^{K}\xi_i\eta_i + \bar{\xi}(t)}{S}\right)\middle|S\right) \approx \\ &\approx \left(1-(1-b)\frac{K}{n(t)}\right)\left(m_\eta + \frac{V_\eta}{2}\right) + (1-b)\frac{K}{n(t)}\ln\left(e^{m_\eta+\frac{V_\eta}{2}} + \frac{\bar{\xi}(t)}{S}\right). \end{aligned} \tag{27}$$

[12] As we shall show in the following section, growth rates of individual units are distributed according to an approximate exponential power distribution with the shape parameter $\varphi \in (0,1)$, and thus they exhibit even sharper spikes at $\bar{g}$ and fatter tails than the Laplace distribution (Buldyrev et al. 2007). This implies that the convergence to the $P(g)$ distribution as in (23) should come from the direction of distributions having sharper spikes at $\bar{g}$ than the Laplace.

The formula above is always larger than $\bar{g} = m_\eta + V_\eta/2$, albeit it converges to this limiting value with $K/n(t) \to 0$ as well as with $S \to \infty$. A numerical quantification of the magnitude of departures from this limit is provided in Figure 2. The interpretation of this result is the following. First, at a given moment in time, firm growth is faster than average growth of its constituent units only when it is assigned a new unit of size that is not negligible in comparison to the firm's size. This is likely only when the probability that this firm is assigned a new unit is not negligible, i.e. when $n(t)$ is low. Second, as opposed to the Simon model with $V_\eta = 0$, the current case detaches firm size from the number of units it has. For a given $n(t)$, high growth rates are thus particularly likely for firms that have many small units, so that K is large compared to S. In the end, the probability of being assigned a new unit is proportional to K, and growth rates are computed with S in the denominator. Under the assumption $S \propto K$, the size–mean growth rate relationship becomes flat again.[13]

The size-variance relationship in the current case is also a mixture of the results we have obtained in the cases: (i) GPGM without firm entry, and (ii) GPGM with firm entry, as $t \to \infty$. An additional complication comes from the fact that with finite time of evolution of the system and thus a non-negligible impact of unit entry on the firms' mean growth rate, the second component in the law of total variance, $\sigma^2(E(g|S,K))$, is generally positive, especially if K, S and $n(t)$ are small. It converges to zero fast with these three variables, though. Hence, at least for large S, the power law scaling relationship between $\sigma^2(g|S)$ and S holds closely also in the current case. Numerical simulations confirm that the exponent $\beta \in (0, 1/2)$ depends both on the time of system evolution t and on the variance V_η.

[13] A selection mechanism might also be at work in real-world data, which is not accounted for in the GPGM which abstracts from firm exit. This mechanism implies that the observed size–mean growth relationship is downward sloping partly because it is computed *conditional on survival*: if a small firm observes a negative shock, its size may fall below a certain survival threshold and thus it may be driven out of the market, whereas a large firm will likely survive. Hence, even if the size–mean growth rate relationship is actually flat, we may perceive it as downward sloping in any dataset that includes surviving firms only.

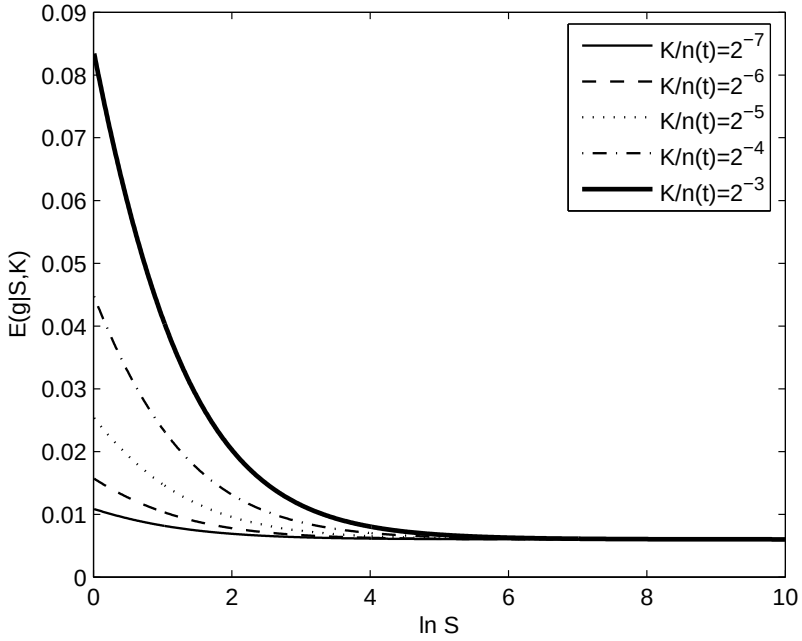

Figure 2: The size–mean growth rate relationship as a function of the number of units per firm K and firm size S in the GPGM with firm entry and a finite time of evolution. Assumed parameter values: $m_\eta = 0.001, V_\eta = 0.01, b = 0.1$ as well as $t = 1$ (t matters only in relation to m_η and V_η). These values imply $\bar{\xi} = e^{t(m_\eta + V_\eta/2)} = 1.006$.

3.6 Qualifications of the results

The results derived above have a few shortcomings. First, there is a shortage of analytical results for the case when the evolution of the system is stopped at finite time ($t < \infty$). Second, in the case with $b \in (0, 1)$, $\psi > 0$, and $V_\eta > 0$, we rely on imprecise approximations for the $P(g|K)$ distribution which cannot be expressed in a closed form for small K. These approximations are especially hurting when firms with a small number of units K constitute a large percentage of the total firm population. Third, since a closed form for the PDF of a random variable which is a sum of K lognormally distributed variables does not exist either, we rely on approximations, such as the one due to Slimane (2001), also when dealing with the distribution $P(S|K)$ of firm sizes. Fourth, the γ parameter, capturing the

distance from the Slimane's lower- and upper-bound approximations is assumed constant in our calculations (depending only on V_η), but in reality it would rather decrease with S. At the same time, with finite t, there are finite-size cutoffs (in the size distributions) and crossovers from the assumed $\ln \eta_i$ distribution to the limit distribution (24) with a Laplace body and power-law tails $\sim g^{-3}$.

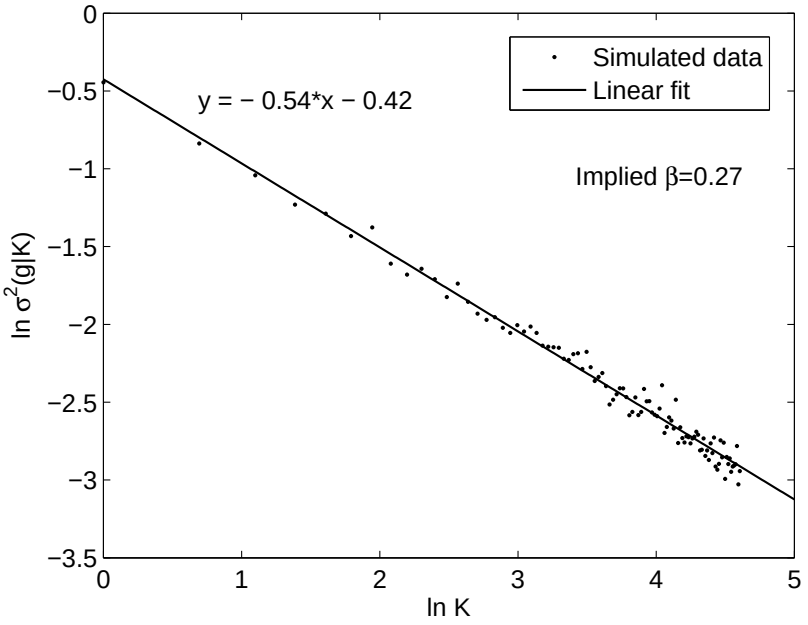

Figure 3: The scaling relationship between K and $\sigma^2(g|K)$. The $1/K$ scaling fails when V_η is relatively small or when small firms make up a large share of the distribution. In this simulation, the scaling parameter α turns out to be $\alpha = 0.27$ and not 1/2. units' growth rates $\ln \eta_i$ are assumed to be Laplace-distributed with $m_\eta = 0.001$ and $V_\eta = 0.36$.

Some of the problems indicated above have been resolved using simulative methods. One of such problems is the lack of analytical results for the growth rate distribution $P(g)$ when the dynamic process is stopped at finite time. The outcome has been already presented above.

Another example is the problem with the relationship between the number of units in firms and the variance of their growth rates $\sigma^2(g|K)$. A Central Limit Theorem approximation used by Fu et al. (2005) implies a $1/K$ scaling relation,

apparent in (19). The equation (19) is valid only as $K \to \infty$, however. Our numerical simulations (see Figure 3) show that for small K, the scaling is in fact better approximated by $\sigma^2(g|K) \approx 1/K^{2\tilde{\beta}}$ with $\tilde{\beta}$ depending on V_η ($\tilde{\beta} \to 1/2$ when $V_\eta \to 0$ and $\tilde{\beta} \to 0$ when $V_\eta \to \infty$). As we have also confirmed numerically, a flatter scaling relationship $\sigma^2(g|K)$ implies, other things constant, a flatter size-variance relationship $\sigma^2(g|S)$, i.e. a smaller $\tilde{\beta}$ implies a smaller β.

From Figure 3, we see in particular that when $V_\eta = 0.36$, the $\tilde{\beta}$ parameter implied by the model is around 0.27, not 0.5 as suggested by (19). Within the same parametrization, we find that the implied β exponent in the size-variance relationship $\sigma^2(g|S)$ falls accordingly from 0.426 (using $\tilde{\beta} = 0.5$) to 0.23 (using $\tilde{\beta} = 0.27$).

4 Empirical findings

Let us now test the predictions of our model in the context of the worldwide pharmaceutical industry. To this end, we shall first note that the pharmaceutical industry is characterized by a positive net inflow of both new units ($\psi > 0$) and firms ($b > 0$). Secondly, a unit is naturally defined here as a molecular entity. New molecular entities are products developed by innovator companies, which after undergoing clinical trials translate into drugs that cure specific diseases. The number of new molecular entities approved by the US Food and Drug Administration and similar agencies in other countries is widely used as a measure of innovation in pharmaceuticals (Pammolli et al. 2011). Since molecular units have different therapeutic properties, they cannot be substituted, and thus they can be credibly analyzed as independent submarkets (Sutton 1998). The whole pharmaceutical industry can be viewed as an aggregation of many independent units. Moreover, the sales of each unit are extremely volatile over the product lifecycle ($V_\eta > 0$), especially after patent expiry (Magazzini et al. 2004). These structural features of the pharmaceutical industry imply that the full GPGM model should apply in this case.

The pharmaceutical industry database (PHID) at IMT Lucca, upon which we base our analysis, is a unique dataset which records sales figures of the 916 036 drugs commercialized by 7 184 pharmaceutical firms in 21 countries from 1994 to

Variable	Model
K	-.0035
	(.0029)
τ_α	.1212***
	(.0120)
Time Dummies	yes***
Firm Dummies	yes***
N	8,092
R^2	0.017

Robust standard errors in parenthesis

*** statistically significant at 1% level; ** at 5%; * at 10%

Table 1: Fixed effects panel regression of the relationship between the average growth of firm units, the number of units and firm age, marginal effects.

2008, covering the whole size distribution of units and firms and monitoring the flows of entry and exit at both levels. Firms capture new business opportunities by launching new units on the market and the size of each firm is defined as the sum of the sales of their units: $S_\alpha(t) = \sum_{i=1}^{K_\alpha(t)} \xi_i(t) = \langle \xi_\alpha(t) \rangle K_\alpha(t)$ where $\langle \xi_\alpha(t) \rangle$ is the average size of units in firm α at time t.

Before we proceed to empirical tests of implications of our model, let us first check if assumption (3) holds. This amounts to verifying if, controlling for firm age, the average unit size $\langle \xi_\alpha(t) \rangle$ is independent from the number of units $K_\alpha(t)$. To measure firm age, we use the age in years τ_α of the oldest molecule a firm has still on the market at time t. Table 1 shows that the average unit size increases with company age, but indeed does not depend on K, thus verifying our assumption.

Since the two growth processes are thus arguably independent and the cross-correlation of growth across units is weak (Sutton 2002, Riccaboni et al. 2008), we can now test the predictions of the model with respect to: (1) the size distribution of companies, $P(S)$; (2) the distribution of firm growth rates, $P(g)$; (3) the relationship between firm size and its mean growth rate and (4) the size-variance relationship, summarized by the parameter β in the power-law relationship of form $\sigma(g|S) \propto S^{-\beta}$.

Since the same dataset has been previously analyzed many times with our other coauthors (Bottazzi et al. 2001, De Fabritiis et al. 2003, Fu et al. 2005, Yamasaki et al. 2006, Buldyrev et al. 2007, Growiec et al. 2008, Riccaboni et al. 2008, Bee et al. 2011), we rely on previous results whenever possible.

4.1 Size distribution

Our model predicts that if the size distribution of units is approximately lognormal, and the distribution of units among firms $P(K)$ is a power-law with an exponential cut-off, then the firm size distribution should be a lognormal multiplied by a stretching factor can be arbitrarily large, giving rise to an approximate power law decay of $P(S)$ for very large S. By using the same data as we use here, Growiec et al. (2008) have found that the unit distribution is indeed approximately lognormal, whereas Yamasaki et al. (2006) have revealed that the $P(K)$ is a power-law with an exponential cut-off. Thus the firm size distribution should depict a power-law upper tail (Growiec et al. 2008).

Several tests to detect a Pareto tail have been recently developed (Bee et al. 2011). The list of most widely applied ones includes the uniformly most powerful unbiased (UMPU) test of the Pareto against the lognormal (Malevergne et al. 2009), the Hill test of Clauset et al. (2009) (CSN), and the maximum entropy (ME) test due to Bee et al. (2011). The results of these tests are summarized in Table 2. It turns out that the Pareto tail of the distribution of firm sizes spans the top 1400 ranks (19.49%) according to the ME test, 1 200 to 1 300 for the UMPU test (16.70% to 18.10%) and 900 for CSN (12.53%). These results are completely different than the ones obtained at the unit level, where with ME and CSN, the threshold is found between ranks 8 and 9 thousands (0.87% and 0.98%), and with the UMPU test, the rank is approximately equal to 300 (0.03%). Hence, disaggregated data show that the Pareto tail is most likely confined to the last percentile of the distribution, whereas at the firm level, it is markedly more pronounced.

Thus we can conclude that, as shown in Growiec et al. (2008) and Bee et al. (2011), the power-law upper tail includes at least top 12.53% of pharmaceutical firms. On the contrary, the unit size distribution does not have a Pareto tail

Firms	UMPU	CSN	ME
ranks	1200-1300	900	1400
min S (USD,thousand)	4121-3529	7330	2939
%	16.70-18.10	12.53	19.42
slope	n.a.	.532	.601
Units	UMPU	CSN	ME
ranks	300	8000-9000	8000-9000
min S (USD, thousand)	17686	6792-5919	6792-5919
%	0.03	0.87-0.98	0.87-0.98
slope	n.a.	1.038	1.021

Table 2: UMPU, ME and CSN tests of the tail behavior of the size distribution at the firm and unit levels.

(only up to 0.98% of the largest units could be Pareto distributed). In sum, the predictions of our model regarding the size distribution of business firms are strongly supported by empirical evidence.

4.2 Growth distribution

In the empirical application, the firm growth rate $P(g)$ is defined as the yearly growth rate of firm sales. To capture this, we use two different measures of firm growth. The first is given by $g_\alpha = \ln\left(S_\alpha(t+1)/S_\alpha(t)\right)$. The second is $g^*_\alpha = \left(S_\alpha(t+1) - S_\alpha(t)\right)/S_\alpha(t)$. The two measures are equivalent for growth rates close to zero, which can be reasonably assumed by taking a short time period and a fixed number of units K.

In this subsection we refer to the first measure of growth. Under this definition, we compute the growth rate distributions for both firms and units and run a set of maximum likelihood estimates (MLE).[14] As candidate distributions we consider: the Gaussian, Laplace, Exponential Power, and the "Fu" distribution characterized in equation (24). The Exponential Power Distribution has the form:

$$P(g) = \frac{\kappa}{2\mu\Gamma(\frac{1}{\kappa})} \exp -(|g-\mu|/\sigma)^\kappa, \tag{28}$$

[14]See also Buldyrev et al. (2007).

where the parameter $\sigma > 0$ is the scale parameter, whereas $\kappa > 0$ is the shape parameter[15]. By varying the exponent κ, it is possible to describe the Gaussian as well as platikurtic and leptokurtic distributions. For $\kappa = 2$ the distribution is Gaussian. For $\kappa = 1$ we obtain a Laplace distribution with a standard deviation σ. The Fu distribution summarized in (24) has only one parameter, V.

Table 3 reports both the Kolmogorov-Smirnov (KS) and the Anderson-Darling (AD) statistics for the four considered distributions. The result is that despite the Fu distribution has only one free parameter, it outperforms the Gaussian and the Laplace fit and it is slightly better than the three-parameter Exponential Power Distribution in the body, while the Anderson-Darling test shows that the Exponential Power Distribution provides a better fit of the tails.

Distribution	μ	σ	κ	V	KS	AD
Gaussian	-.056	1.212	-	-	19.221	n.a.
Conf. int.	*-.056*	*1.200*	-	-		
	-.058	*1.224*	-	-		
Laplace	.059	.569	-	-	3.777	190.080
Conf. int.	*.059*	*.561*	-	-		
	.060	*.577*	-	-		
Exponential Power	.047	.148	.525	-	2.638	.054
Conf. int.	*.046*	*.135*	*.513*	-		
	.047	*.161*	*.538*	-		
Fu distribution	-	-	-	.555	1.845	.099
Conf. int.	-	-	-	*.533*		
	-	-	-	*.577*		

Table 3: Maximum Likelihood Estimates (MLE) of the 1-year firm growth distribution. "Fu distribution" denotes the distribution summarized in (24) which has a Laplace cusp and power-law wings $\sim g^{-3}$.

To better investigate the tail behavior of the growth distribution we use the Hill estimator (Clauset et al. 2009). Table 4 shows that the growth distribution indeed depicts power-law wings (about 6.7% of the total growth events are power-law

[15]We applied a non-parametric methodology to identify the robust estimator of the location of the distributions and shifted the data prior to computing estimators presented in Table 3.

distributed) $P(g) \sim g^{-3}$, as predicted by our model.

Tail	slope	xmin	KS	Exc. %
Positive	3.0225	2.1632	.0644	3.3934
Negative	3.0903	0.9302	.0494	3.2975

Table 4: Hill estimator of the tail behavior of the firm growth distribution $P(g) \sim g^{-3}$.

In sum, the Fu distribution provides a better fit to our data than any alternative candidate distribution.[16] Therefore we can conclude that, as predicted by our model, the growth distribution has a Laplace body and power law wings. We have thus shown so far that our model simultaneously accounts for the shape of the firm size distribution and the distribution of firm growth rates.

4.3 The relationship between firm size and mean growth rate

In this subsection, we shall measure firm growth as g^*_α, for two reasons. First, we prefer to use g^*_α to facilitate comparisons with previous findings in the literature (Mansfield 1962, Rossi-Hansberg & Wright 2007). Second, by using g^*_α we can also apply the Mansfield's correction for firm exit and put $g^*_e = -1$ for firms e that leave the market.

Figure 4 illustrates the relationship between firm size and mean growth rate. Firms are grouped into ten bins of equal number there. When considering the whole set of pharmaceutical companies ($K > 0$), we find a negative relationship between growth and size, in line with the stylized fact (3) but against the predictions of our theory. For companies with more than three units ($K > 3$), growth rates are independent of size, however, just as the GPGM implies. Since most of the small companies have less that three units, on average small companies grow more than large ones. To better investigate the effect of the number of units on firm growth rates, in Table 5 we split the firms in four groups ($K = 1, 2, 3$ and $K > 3$) and count how many of them capture new units in a given year ($\Delta K > 0$) or leave the market.

[16]The Levy distribution is also ruled out since the tails decay with a power > 2.

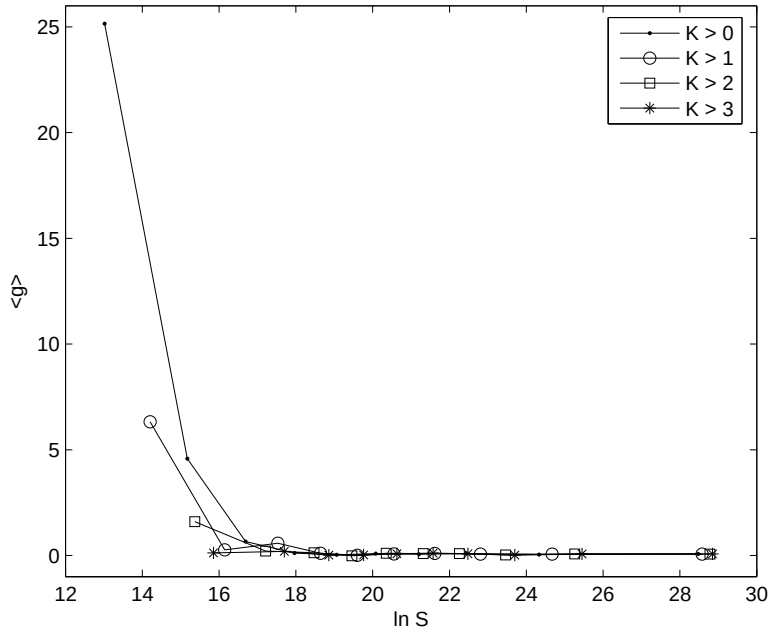

Figure 4: The relationship between the logarithm of firm size (S) and its mean growth rate (g).

The average growth rate of a firm with a single unit is thus almost 50 times bigger than the average growth rate of a company with more than three units. Furthermore, among companies with one unit those that capture new business opportunities grow 100 times more than others. In the pharmaceutical industry this can happen for instance when a biotech start-up company with one molecule in the market for a restricted population of patients launches a new blockbuster drug.[17] Rare spurs of very fast growth are thus due to a discontinuous process of innovation-led growth through the capturing of new business opportunities.[18].

Companies with one unit have also a far higher exit probability (13.17% versus 0.20% for companies with $K > 3$). To control for the selection bias, we compute

[17]For instance, MedImmune's FluMist vaccine was first approved by the Food and Drug Administration for a restricted population of patients in 2003. Then in 2007 a new version of the product (CAIV-T) has been authorized for a far bigger market

[18]When the median growth rate is considered instead of the mean, the relationship is flat for all K.

Number of firms, by K and ΔK				
	$\Delta K \leq 0$	$\Delta K > 0$	Exit	Total
$K = 1$	2388	171	388	2947
$K = 2$	766	109	38	913
$K = 3$	402	115	9	526
$K > 3$	1516	1534	6	3056
Percentage distribution, by K and ΔK				
	$\Delta K \leq 0$	$\Delta K > 0$	Exit	Total
$K = 1$	81.03%	5.80%	13.17%	100%
$K = 2$	83.90%	11.94%	4.16%	100%
$K = 3$	76.43%	21.86%	1.71%	100%
$K > 3$	49.61%	50.20%	0.20%	100%
Average Growth Rate, by K and ΔK				
	$\Delta K \leq 0$	$\Delta K > 0$	Total	Total, Exit= -1
$K = 1$	5.15	483.83	37.28	32.24
$K = 2$	0.56	9.90	1.71	1.60
$K = 3$	0.13	3.44	0.85	0.82
$K > 3$	0.07	1.33	0.70	0.69

Table 5: Average growth rates of companies by number of units (K). For the case $\Delta K <= 0$, the growth rates after the Mansfield's correction are: 4.28 ($K = 1$), 0.48 ($K = 2$), 0.11 ($K = 3$), 0.06 ($K > 3$).

the Mansfield's correction. This correction only partially attenuates the selection effect, though. Table 6 reports the complementary log-log (C-Log-Log) estimates of the hazard probability to exit for companies with a different number of units, average unit size, and of a different age. This analysis confirms that firms with more units have a lower probability to exit. The average unit size has also a positive effect on survival probability, whereas firms' age is far less significant. Though preliminary, this result suggests that the age effect on firm survival could be mediated by the innovation process and the capturing of new business opportunities, as in the Klette & Kortum (2004) model.

All in all, we find that the downward sloping relationship between firm growth

Variable	C-Log-Log	C-Log-Log (RE)	C-Log-Log (RE, pop. averaged)
K	.4196***	.3565***	.3569***
	(.0379)	*(.0380)*	*(.0384)*
$< \xi >_\alpha$	.9835***	.9826***	.9827***
	(.0018)	*(.0019)*	*(.0019)*
τ_α	1.044	1.178**	
	(.0509)	*(.0779)*	
Time Dummies	no	no	yes
Firm Dummies	no	yes***	yes***
N	8201	8201	8201
Log lik.	-1192.7	-1180.2	-1174.5

Robust standard errors in parenthesis

*** statistically significant at 1% level; ** at 5%; * at 10%

Table 6: Survival probability, complementary log-log regressions

and size among small firms is driven primarily by innovation and selection. Among firms that sell more than 3 products, however, the size–mean growth rate relationship is essentially flat, in line with the predictions of the model.

4.4 The variance of firm growth rates

As for the size variance relationship, our model predicts that it crucially depends on the partition of firm sales into units. If firms have $P(K)$ units and $V_\eta = 0$, the Law of Large Numbers applies precisely and $\sigma(K) \sim K^{-\beta}$, where $\beta = 1/2$. On the contrary, if each firm consists of a single unit only and $V_\eta > 0$, then the scaling of the size–variance relationship disappears and $\beta = 0$. When both mechanisms are at work, the speed of the crossover depends on the skewness of $P(K)$. At one extreme, if all companies have the same number of units, $\beta = 0$ and there is no crossover. On the contrary, if $P(K)$ is power-law distributed, for a wide range of empirically plausible V_η, β is far from $1/2$ and statistically different from zero.

In the pharmaceutical industry, we find the size-variance scaling coefficient β to be $\approx 1/5$ (see Figure 5). More generally, it has been found in the related literature

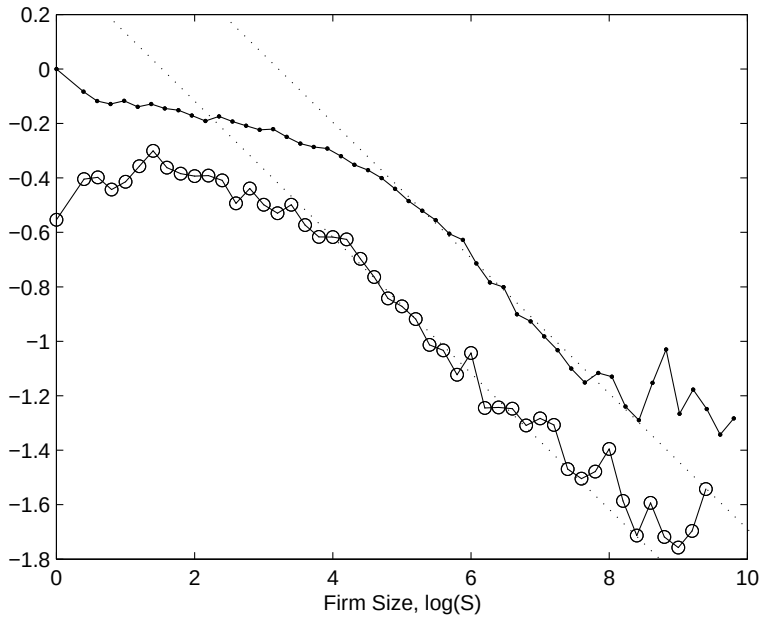

Figure 5: The standard error of firm growth rates (σ) (circles), and the share of the largest units ($1/K_e$) (dots) versus the size of the firm (S). The flattering of the upper tail is due to some large companies with unusually large units. A reference line with slope 1/5 is also reported.

that the relationship between the size and the variance of firm growth rates follows an approximate power-law behavior $\sigma(S) \sim S^{-\beta(S)}$ where S is the firm size and the exponent $\beta(S) \approx 1/5$ is weakly dependent on S (Stanley et al. 1996, Bottazzi et al. 2001, Sutton 2002, Riccaboni et al. 2008). Riccaboni et al. (2008) have shown how a model of proportional growth which treats firms as classes composed of various number of units of variable size, can explain this size–variance dependence. In general, their model predicts that $\beta(S)$ must exhibit a crossover from $\beta(0) = 0$ to $\beta(\infty) = 1/2$. As shown in Figure 5, the reason why the variance does not scale as predicted by the Law of Large Numbers has to do with the skewed size distribution of units in the firm's portfolio. In fact, the size-variance relationship scales as the share of the firm's largest unit. These findings are in good agreement with the implications of the GPGM framework.

5 Summary and concluding remarks

In this article, we have provided a few important findings regarding the theoretical predictions of the Generalized Proportional Growth model (GPGM) and its empirical relevance. In the first respect, the results obtained analytically are summarized briefly in Table 7. When a firm consists of exactly one unit, the GPGM boils down to the standard Gibrat growth process leading to a lognormal size distribution of business firms. Conversely, when the size of units is fixed or grows deterministically ($V_\eta = 0$), we get the Simon model of firm dynamics that leads to a Pareto firm size distribution. As we gradually allow for more complexity in the considered system, resultant size distributions, growth rate distributions, and size-variance relationships get more pronounced as well. Most importantly, however, the case with both unit and firm entry ($b \in (0, 1), \psi > 0$), a positive variance of multiplicative unit-specific shocks V_η, and a finite time truncation, gives a good fit to the observed data from the pharmaceutical industry *on all four considered dimensions*: the firm size distribution, the firm growth rate distribution, as well as the size–mean growth rate and size–variance relationships. Since multiple candidate generative processes can explain a single stylized fact, a good explanatory mechanism should match a larger set of empirical facts. In this light, the GPGM discussed here has proved to be successful.

Case	Key assumptions	Size distribution	Growth distrib.	Size-mean growth	Size-variance
Deterministic growth	$b = 0$, $\psi = 0$, $V_\eta = 0$	one-point	one-point	$\bar{g} \equiv m_\eta$	n.a.
Det., new firms only	$b = 1$, $\psi > 0$, $V_\eta = 0$	one-point	one-point	$\bar{g} \equiv m_\eta$	n.a.
Pure Gibrat	$b = 0$, $\psi = 0$, $V_\eta > 0$	lognormal	always $\ln \eta_i$	flat	$\beta = 0$
Gibrat, new firms only	$b = 1$, $\psi > 0$, $V_\eta > 0$	lognormal	always $\ln \eta_i$	flat	$\beta = 0$
Simon, no firm entry	$b = 0$, $\psi > 0$, $V_\eta = 0$, infinite time	exponential	one-point	$\bar{g} \equiv m_\eta$	$\beta = 1/2$ or n.a.
Simon, no f.e., finite t	$b = 0$, $\psi > 0$, $V_\eta = 0$, finite time	$\approx$ exponential	atoms $\to$ one-point	flat, but $E(g) > \bar{g}$	$\beta = 1/2$
Simon	$b \in (0, 1)$, $\psi > 0$, $V_\eta = 0$, infinite time	Pareto	one-point	$\bar{g} \equiv m_\eta$	$\beta = 1/2$ or n.a.
Simon, finite t	$b \in (0, 1)$, $\psi > 0$, $V_\eta = 0$, finite time	Pareto, exp cutoff	atoms $\to$ one-point	flat, but $E(g) > \bar{g}$	$\beta = 1/2$
GPGM, no firm entry	$b = 0$, $\psi > 0$, $V_\eta > 0$, infinite time	$\approx$ lognormal	$\sim g^{-3}$	flat	$\beta \in (0, 1/2)$
GPGM, no f.e., finite t	$b = 0$, $\psi > 0$, $V_\eta > 0$, finite time	$\approx$ lognormal	$\ln \eta_i \to \sim g^{-3}$	$E(g\|S, K) \searrow \bar{g}$ with S & K	$\beta \in (0, 1/2)$
GPGM	$b \in (0, 1)$, $\psi > 0$, $V_\eta > 0$, infinite time	lognormal, Pareto tail	Fu dist.	flat	$\beta \in (0, 1/2)$
GPGM, finite t	$b \in (0, 1)$, $\psi > 0$, $V_\eta > 0$, finite time	stretched lognormal	$\ln \eta_i \to$ Fu dist.	$E(g\|S, K) \searrow \bar{g}$ with S & K	$\beta \in (0, 1/2)$

Table 7: Summary of the results. "Fu dist." denotes the distribution summarized in (24) which has a Laplace cusp and power-law wings $\sim g^{-3}$. The "$\to$" sign signifies the direction of convergence as t increases. $\searrow$ signifies convergence from above with a potentially substantial impact of transitional dynamics.

According to our findings, the Simon model of growth in the number of business units per firm with positive entry, combined with the Gibrat-type model of proportional growth in units sizes, turns out to be a reliable and powerful generative mechanism able to explain a variety of findings related to corporate dynamics, observed in microeconomic data.

More research is still needed to test the GPGM in different industries and against other stylized facts concerning the relationship between firm size, growth, and age. It might also be the case that due to some potential misalignments of the model's predictions with further characteristics of the data, the model should be generalized or modified. Equally importantly, further work is also required to provide sound economic microfoundations behind the stochastic assumptions of this and related articles, in particular ones that could account for competition within submarkets and lifecycles.

However, we believe that future research will be able to discriminate among the growth regimes at work in different industries and countries over time only by combining simple and general theoretical frameworks, akin to the one described in this paper, with a rigorous empirical strategy.

References

Axtell, R. (2001), 'Zipf Distribution of US Firm Sizes', *Science* **293**(5536), 1818.

Bee, M., Riccaboni, M. & Schiavo, S. (2011), Pareto versus Lognormal: A Maximum Entropy Test, Working paper, Department of Economics working paper 02-11, University of Trento.

Bottazzi, G., Dosi, G., Lippi, M., Pammolli, F. & Riccaboni, M. (2001), 'Innovation and Corporate Growth in the Evolution of the Drug Industry', *International Journal of Industrial Organization* **19**(7), 1161–1187.

Bottazzi, G. & Secchi, A. (2003), 'Why are Distributions of Firm Growth Rates Tent-Shaped?', *Economics Letters* **80**(3), 415–420.

Bottazzi, G. & Secchi, A. (2006), 'Explaining the Distribution of Firm Growth Rates', *The RAND Journal of Economics* **37**(2), 235–256.

Brock, W. (1999), 'Scaling in Economics: A Reader's Guide', *Industrial and Corporate Change* **8**(3), 409.

Buldyrev, S. V., Growiec, J., Pammolli, F., Riccaboni, M. & Stanley, H. E. (2007), 'The Growth of Business Firms: Facts and Theory', *Journal of the European Economic Association* **5**, 574–584.

Cabral, L. & Mata, J. (2003), 'On The Evolution of the Firm Size Distribution: Facts and Theory', *American Economic Review* **93**(4), 1075–1090.

Caves, R. (1998), 'Industrial Organization and New Findings on the Turnover and Mobility of Firms', *Journal of Economic Literature* **36**(4), 1947–1982.

Clauset, A., Shalizi, C. & Newman, M. (2009), 'Power-Law Distributions in Empirical Data', *SIAM Review* **51**(4), 661–703.

Coad, A. (2009), *The Growth of Firms: A Survey of Theories and Empirical Evidence*, Edward Elgar Publishing.

De Fabritiis, G., Pammolli, F. & Riccaboni, M. (2003), 'On Size and Growth of Business Firms', *Physica A* **324**(1-2), 38–44.

De Vany, A. & Walls, W. (1996), 'Bose-Einstein Dynamics and Adaptive Contracting in the Motion Picture Industry', *The Economic Journal* **106**(439), 1493–1514.

De Wit, G. (2005), 'Firm Size Distributions:: An Overview of Steady-State Distributions Resulting from Firm Dynamics Models', *International Journal of Industrial Organization* **23**(5-6), 423–450.

Dunne, T., Roberts, M. & Samuelson, L. (1989), 'The Growth and Failure of US Manufacturing Plants', *The Quarterly Journal of Economics* **104**(4), 671.

Eeckhout, J. (2004), 'Gibrat's Law for (All) Cities', *American Economic Review* **94**(5), 1429–51.

Eeckhout, J. (2009), 'Gibrat's Law for (All) Cities: Reply', *American Economic Review* **99**(4), 1676–83.

Evans, D. (1987), 'The Relationship between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries', *The Journal of Industrial Economics* **35**(4), 567–581.

Feller, W. (1957), *An Introduction to Probability Theory and its Applications*, John Wiley.

Fu, D., Pammolli, F., Buldyrev, S. V., Riccaboni, M., Matia, K., Yamasaki, K. & Stanley, H. E. (2005), 'The Growth of Business Firms: Theoretical Framework and Empirical Evidence', *Proceedings of the National Academy of Sciences* **102**(52), 18801–18806.

Gabaix, X. (1999), 'Zipf's Law: An Explanation', *Quarterly Journal of Economics* **114**(3), 739–767.

Gabaix, X. (2009), 'Power Laws in Economics and Finance', *Annual Review of Economics* **1**, 255–93.

Gabaix, X. (2011), 'The Granular Origins of Aggregate Fluctuations', *Econometrica* . forthcoming.

Gibrat, R. (1931), *Les Inégalités Économiques; Applications: aux Inégalités des Richesses, à la Concentration des Entreprises, aux Populations des Villes, aux Statistiques des Familles, etc., d'une Loi Nouvelle, la Loi de l'Effect Proportionnel*, Recueil Sirey, Paris.

Growiec, J., Pammolli, F., Riccaboni, M. & Stanley, H. E. (2008), 'On the Size Distribution of Business Firms', *Economics Letters* **98**, 207–212.

Hall, B. (1987), 'The Relationship Between Firm Size and Firm Growth in the US Manufacturing Sector', *The Journal of Industrial Economics* **35**(4), 583–606.

Hopenhayn, H. (1992), 'Entry, Exit, and Firm Dynamics in Long Run Equilibrium', *Econometrica* **60**(5), 1127–50.

Hymer, S. & Pashigian, P. (1962), 'Firm Size and Rate of Growth', *The Journal of Political Economy* **70**(6), 556–569.

Ijiri, Y. & Simon, H. (1977), *Skew Distributions and the Sizes of Business Firms*, North-Holland, Amsterdam.

Jovanovic, B. (1982), 'Selection and the Evolution of Industry', *Econometrica* **50**(3), 649–670.

Kalecki, M. (1945), 'On the Gibrat Distribution', *Econometrica* **13**(2), 161–170.

Klepper, S. & Thompson, P. (2006), 'Submarkets and the Evolution of Market Structure', *The RAND Journal of Economics* **37**(4), 861–886.

Klette, T. & Kortum, S. (2004), 'Innovating Firms and Aggregate Innovation', *Journal of Political Economy* **112**(5), 986–1018.

Kotz, S., Kozubowski, T. & Podgorski, K. (2001), *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*.

Levy, M. (2009), 'Gibrat's Law for (All) Cities: Comment', *American Economic Review* **99**(4), 1672–75.

Luttmer, E. (2007), 'Selection, Growth, and the Size Distribution of Firms', *The Quarterly Journal of Economics* **122**(3), 1103–1144.

Luttmer, E. (2010), 'Models of Firm Heterogeneity and Growth', *Annual Review of Economics* **2**, 547–576.

Magazzini, L., Pammolli, F. & Riccaboni, M. (2004), 'Dynamic Competition in Pharmaceuticals', *The European Journal of Health Economics* **5**(2), 175–182.

Malevergne, Y., Pisarenko, V. & Sornette, D. (2009), Gibrat's Law for Cities: Uniformly Most Powerful Unbiased Test of the Pareto Against the Lognormal, Working paper, Swiss Finance Institute Research Paper Series.

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

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

Matraves, C. (1999), 'Market Structure, R&D and Advertising in the Pharmaceutical Industry', *Journal of Industrial Economics* **47**(2), 169–94.

Mitzenmacher, M. (2004), 'A Brief History of Generative Models for Power Law and Lognormal Distributions', *Internet mathematics* **1**(2), 226–251.

Pammolli, F., Riccaboni, M. & Magazzini, L. (2011), 'The Productivity Crisis in Pharmaceutical R&D', *Nature Reviews Drug Discovery* . forthcoming.

Reed, W. (2001), 'The Pareto, Zipf and Other Power Laws', *Economics Letters* **74**(1), 15–19.

Riccaboni, M., Pammolli, F., Buldyrev, S., Ponta, L. & Stanley, H. (2008), 'The Size Variance Relationship of Business Firm Growth Rates', *Proceedings of the National Academy of Sciences* **105**(50), 19595.

Rossi-Hansberg, E. & Wright, M. (2007), 'Establishment Size Dynamics in the Aggregate Economy', *The American Economic Review* **97**(5), 1639–1666.

Simon, H. A. & Bonini, C. P. (1958), 'The Size Distribution of Business Firms', *The American Economic Review* **48**(4), 607–617.

Slimane, S. (2001), 'Bounds on the Distribution of a Sum of Independent Lognormal Random Variables', *IEEE Transactions on Communications* **49**(6), 975–978.

Stanley, M., Buldyrev, S., Havlin, S., Mantegna, R., Salinger, M. & Stanley, H. (1995), 'Zipf Plots and the Size Distribution of Firms', *Economics Letters* **49**(4), 453–457.

Stanley, M. H. R., Amaral, L. A. N., Buldyrev, S. V., Havlin, S., Leschhorn, H., Maass, P., Salinger, M. A. & Stanley, H. E. (1996), 'Scaling Behaviour in the Growth of Companies', *Nature* **379**, 804–806.

Steindl, J. (1965), *Random Processes and the Growth of Firms: A Study of the Pareto Law*, Griffin, London.

Sutton, J. (1997), 'Gibrat's Legacy', *Journal of Economic Literature* **35**(1), 40–59.

Sutton, J. (1998), *Technology and Market Structure: Theory and History*, The MIT Press Cambridge, Mass.

Sutton, J. (2002), 'The Variance of Firm Growth Rates: The 'Scaling' Puzzle', *Physica A* **312**(3-4), 577–590.

Yamasaki, K., Matia, K., Buldyrev, S. V., Fu, D., Pammolli, F., Riccaboni, M. & Stanley, H. E. (2006), 'Preferential Attachment and Growth Dynamics in Complex Systems', *Physical Review E* **74**(3), 35103.

Yule, G. (1925), 'A Mathematical Theory of Evolution, Based on the Conclusions of Dr. JC Willis, F.R.S.', *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character* **213**, 21–87.

A A generalized GPGM allowing for stable firm size and firm growth rate distributions

As we have indicated in the main text, it is possible to design a mechanism that would guarantee that the firm size and firm growth rate distributions obtained from the model would converge to distribution with a fixed mean and variance. One of such possibilities consists in replacing Assumption (4) of the GPGM with the following assumption, following the advice of Kalecki (1945):

(4') At time $t+1$, the size of each unit is raised to the power $1-\alpha$ and then decreased or increased by a random factor $\eta_i(t) > 0$ so that

$$\xi_i(t+1) = (\xi_i(t))^{1-\alpha}\,\eta_i(t), \qquad \alpha \in (0,1), \tag{29}$$

where $\eta_i(t)$ is a random variable that is independent of all other η_i's and ξ_i's. It is assumed that $E\ln\eta_i(t) \equiv m_\eta$ and $Var(\ln\eta_i(t)) = E(\ln\eta_i(t))^2 - m_\eta^2 \equiv V_\eta$.

The above equality can be alternatively interpreted as an assumption that the size of each unit is multiplicatively affected by a random variable $\tilde{\eta}_i(t) = (\xi_i(t))^{-\alpha}\eta_i(t)$. In such case, the mean growth rate at the level of units will no be longer independent of their size; it will systematically decline with size instead.

The mechanism by Kalecki (1945) can be justified in terms of random exit (Luttmer 2010). Indeed, if one augments the original assumption that the net entry rate of units ψ with a positive probability of unit decline and exit α, due to reasons unrelated directly to the firm's sales (e.g., technological obsolescence), then this can be reflected in a positive α in eq. (29). We would then see "creative destruction" effects, absent in our original specification: the expected growth rate of units would decline with their age, and the faster is the aging of existing units (higher α), the higher would be the entry rate of new ones, captured by $\psi_K = \mu - \lambda + \alpha$.

As far as the innovation process is concerned, thanks to which firms capture new business opportunities and new start-up firms enter the market, it does not generate systematic increases in variance. To obtain stationary size and growth rate distributions, it is thus enough to de-trend the variables (De Wit 2005). This can be achieved, e.g., by considering the distribution of number of units K relative to the average number of units per firm, or relative to the total number of units in the economy.

A.1 Implications for the growth process at the level of units

When describing the implications of the current change in assumptions, let us begin with the growth process at the level of units, or equivalently – at the level of firms, provided that the economy is in the pure Gibrat regime ($b = 0$, no unit entry nor firm entry). We will then pass to the discussion of the pure Simon case where no variance at unit level is allowed. These two limiting cases will provide us with two bounds, within which the final results will be confined.

As opposed to the proportional growth case discussed in the main text, the current growth process guarantees convergence of unit sizes over time to a stationary lognormal distribution with mean $m = \lim_{t\to\infty} E \ln \xi_i(t) = m_\eta/\alpha$ and variance $V = \lim_{t\to\infty} Var(\ln \xi_i(t)) = \frac{V_\eta}{\alpha(2-\alpha)}$.

The distribution of unit growth rates, that is $\ln\left(\frac{\xi_i(t+1)}{\xi_i(t)}\right) = \ln\left(\frac{\eta_i(t)}{(\xi_i(t))^\alpha}\right)$, converges to a distribution that is a convolution of the lognormal distribution and the assumed distribution $\ln \eta_i$. Its mean converges to zero (so that the distribution is stable), and its variance converges to $\frac{2}{2-\alpha}V_\eta$.

Thanks to the Kalecki (1945) mechanism, the size–mean growth rate distribution is now *negative* for all unit sizes ξ_i, according to the functional form: $E(\ln \tilde{\eta}_i|\xi_i) = -\alpha \ln \xi_i + m_\eta$. There is thus a fixed slope implicit in this relationship, equal to $-\alpha$.

The size–variance relationship is however still flat, just like in the standard Gibrat case, because $Var(\ln \tilde{\eta}_i|\xi_i) = Var(\ln \eta_i)$ when ξ_i is given.

A.2 Implications for the unit entry process

As far as the unit entry process is concerned, captured by the arrivals of new units according to the Simon model, it is enough to define it in re-scaled units. To see this, consider switching off all variation at unit level by assuming $V_\eta = 0$. If $m_\eta = 0$, the growth rate distribution converged then to a one-point distribution concentrated at zero already in the model discussed in the main text. If $m_\eta > 0$, on the other hand, then it converged to a one-point distribution concentrated at the growth rate of units, m_η. Yet now, thanks to the Kalecki (1945) mechanism, the average units size is growing at a constant rate m_η but converging to m_η/α over

time according to $\ln \xi_i(t) = \bar{\ln}\xi(t) = \ln \bar{\xi}(0)(1-\alpha)^t + m_\eta/\alpha$. Hence, the growth rate distribution converges then to a one-point distribution concentrated at zero also if $m_\eta > 0$.

The firm size distributions summarized in (2) and (5) can be therefore also used in the "stabilized" model, once one replaces K with $\frac{K}{\kappa(t)\bar{\xi}(t)}$. Furthermore, in the Simon case with firm entry, the distribution of K itself converges to a stationary Pareto distribution as $t \to \infty$, and $\bar{\xi}(t) \to m_\eta/\alpha$, so in that case no normalization is necessary.

The firm growth rate distribution is naturally again a two-point distribution. However, this time it takes the form:

$$g(t) = \begin{cases} m_\eta - \ln \bar{\xi}(t) + \ln\left(1 + \frac{1}{K}\right) & \text{with probability } (1-b)\frac{K}{n(t)}, \\ m_\eta - \ln \bar{\xi}(t) & \text{with probability } 1 - (1-b)\frac{K}{n(t)}. \end{cases} \tag{30}$$

Hence, using the result $\bar{\ln}\xi(t) = \ln \bar{\xi}(0)(1-\alpha)^t + m_\eta/\alpha$, we obtain that the expected growth rate $E(g|K)$ converges to zero linearly with $t \to \infty$, irrespective of K and m_η:

$$E(g|K) \approx \frac{1-b}{n(t)} - \ln \bar{\xi}(0) \cdot \alpha(1-\alpha)^t, \tag{31}$$

and so the growth rate distribution converges to a one-point distribution concentrated at zero. By the same token, the size–mean growth rate relationship is flat in the Simon case.

As far as the size–variance relationship is concerned, we get:

$$Var(g|K) \approx \frac{1-b}{n(t)} \cdot \frac{1}{K} + \frac{2(1-b)}{n(t)} \ln \bar{\xi}(0) \cdot \alpha(1-\alpha)^t + \left(\ln \bar{\xi}(0) \cdot \alpha(1-\alpha)^t\right)^2 \tag{32}$$

and hence $Var(g|K) \propto 1/K$, so the scaling relationship of the Simon model, captured by $\beta = 1/2$, holds also when units evolve according to the Kalecki (1945) process.

A.3 The range of attainable results

The results presented above indicate the range of results one could expect in the general case:

- The firm size distribution should be a stationary distribution being a mixture of lognormal distributions, where the mixing distribution is either exponential or Pareto. No differences with respect to GPGM should be expected here.

- The firm growth rate distribution should be a distribution with zero mean and a fixed variance. Its shape will likely be similar to the one obtained for GPGM, but we do not have proof for that.

- The size–mean growth rate relationship should be downward sloping with a slope coefficient between $-\alpha$ (characterizing the size–mean growth rate relationship at the level of units) and zero (pertaining to the size–mean growth rate relationship if $V_\eta = 0$). This is the key discrepancy between the current specification of the model and the GPGM.

- The size–variance relationship should be characterized by the slope coefficient β between zero (characterizing the size–variance relationship at the level of units) and 1/2 (pertaining to the size–variance relationship if $V_\eta = 0$). Qualitatively, no differences with respect to GPGM should be expected.