Zipf's law for firms: relevance of birth and death processes

Y. Malevergne[1,2], A. Saichev[3,4] and D. Sornette[3,5]
[1] University of Saint-Etienne – ISEAG, France
[2] EMLYON Business School – Cefra, France
[3] ETH Zurich – Department of Management, Technology and Economics, Switzerland
[4] Nizhny Novgorod State University – Department of Mathematics, Russia
[5] Swiss Finance Institute, 40, Boulevard du Pont-d'Arve, Case Postale 3, 1211 Geneva 4, Switzerland
e-mails: malevergne@em-lyon.com, saichev@hotmail.com and dsornette@ethz.ch

Abstract

Zipf's law states that, for most countries, the number of firms with size greater than S is inversely proportional to S. Most explanations start with Gibrat's rule of proportional growth but need to incorporate additional constraints and ingredients introducing deviations from it. We show that combining Gibrat's rule at all firm levels with random processes of firms' births and deaths yield Zipf's law under a "balance" condition between the average growth rate of incumbent firms and the net growth rate of investments in new entrant firms. Thus, Zipf's law can be interpreted as the signature of the long-term optimal allocation of resources that ensures the maximum sustainable growth rate of an economy. For economies with finite age, Zipf's can be recovered as a compensation between the finite-age effect and small deviations of the balance condition. We are able to account for the age-dependent hazard rate documented in the empirical literature as a result of the presence of a minimum size below which firms exit. Our results hold not only for statistical averages over ensemble of economies (i.e., in average) but also apply to a single typical economy.

JEL classification: G11, G12

Keywords: Firm growth, Gibrat's law, Zipf's law.

1 Introduction

The relevance of power law distributions of firm sizes to help understand firm and economic growth has been recognized early, for instance by Schumpeter (1934), who proposed that there might be important links between firm size distributions and firm growth. The endogenous and exogenous processes and factors that combine to shape the distribution of firm sizes can be expected to be at least partially revealed by the characteristics of the distribution of firm sizes. The distribution of firm sizes has also attracted a great deal of attention in the recent policy debate (Eurostat, 1998, for instance), because it may influence job creation and destruction (Davis et al., 1996), the response of the economy to monetary shocks (Gertler and Gilchrist, 1994) and might even be an important determinant of productivity growth at the macroeconomic level due to the role of market structure (Peretto, 1999; Pagano and Schivardi, 2003; Acs et al., 1999).

This article presents a reduced form model that provides a generic explanation for the ubiquitous stylized observation of power law distributions of firm sizes, and in particular of Zipf's law – i.e., the fact that the fraction of firms of an economy whose sizes S are larger than s is inversely proportional to s: $\Pr(S > s) \sim s^{-m}$, with m equal (or close) to 1. We consider an economy made of a large number of firms that are created according to a random birth flow, disappear when failing to remain above a viable size, go bankrupt when an operational fault strikes, and grow or shrink stochastically at each time step proportionally to their current sizes (Gibrat law).

Our contribution to the ongoing debate on the shape of the distribution of firms' sizes is to present a theory that encompasses previous approaches and to derive Zipf's law as the result of the combination of simple but realistic stochastic processes of firms' birth and death together with Gibrat's law (Gibrat, 1931). We show that Zipf's law is associated with a maximum sustainable growth of investments in the creation of new firms. An important aspect of our framework is the prediction of deviations from the pure Zipf's law (case $m = 1$) under a variety of circumstances resulting from transient imbalances between the average growth rate of incumbent firms and the growth rate of investments in new entrant firms. These deviations from the pure Zipf's law have been documented for a variety of firm's size proxies (e.g. sales, incomes, number of employees, or total assets), and reported values for m ranges from 0.8 to 1.2 (Ijri and Simon, 1977; Sutton, 1997; Axtell, 2001, among many others). Our approach provides a framework for identifying their possible (multiple) origins.

In the literature on the growth dynamics of business firms, a well established tradition describes the modification of the firm's size, over a given period of time, as the cumulative effect of a number of different shocks originated by the diverse accidents that affected the firm in that period (Kalecki, 1945; Ijri and Simon, 1977; Steindl, 1965; Sutton, 1998; Geroski, 2000, among others). This, together with Gibrat's law of proportional growth, forms the starting point for various attempts to explain Zipf's law. However, these attempts generally start with the implicit or explicit assumption that the set of firms under consideration was born at the same origin of time and live forever. This approach is equivalent to considering that the economy is made of only one single firm and that the distribution of firm sizes reaches a steady-state if and only if the distribution of the size of a single firm reaches a steady state. This latter assumption is counterfactual or, even worse, non-falsifiable.

An alternative approach to model a stationary distribution of firm sizes is to account for the fact that firms do not all appear at the same time but are born according to a more or less regular flow of newly created firms (Bonaccorsi Di Patti and Dell'Ariccia, 2004; Reynolds et al., 1994). Simon (1955) was the first to address this question (see also Ijri and Simon (1977)). He proposed to modify Gibrat's model by accounting for the entry of new firms over time as the overall industry grows. He then obtained a steady-state distribution of firm sizes with a regularly varying upper tail whose exponent m goes to one from above, in the limit of a vanishingly small probability that a new firm is created. This situation is not quite relevant to explain empirical data, insofar as the convergence toward the steady-state is then infinitely slow, as noted by Krugman (1996). More recently, Gabaix (1999) allowed for birth of new entities, with the probability to create a new entity of a given size being proportional to the current fraction of entities of that size and otherwise independent of time. In fact, this assumption does not reflect the real dynamics of firms' creation. For instance, Bartelsman et al. (2005) document that new firm entrants have a relatively small size compared with the more mature efficient size they develop as they grow. It seems unrealistic to expect a non-zero probability for the birth of a firm of very large size, say, of size comparable to the largest capitalization currently in the market[1].

The fact that firms can go bankrupt and disappear from the economy is a crucial observation that is often neglected in models. Many firms are known to undergo transient periods of decay which, when persistent, may ultimately lead to their exit from business (Bonaccorsi Di Patti and Dell'Ariccia, 2004; Knaup, 2005; Brixy and Grotz, 2007; Bartelsman et al., 2005). Simon (1960) as well as Steindl (1965) have considered this stylized fact within a generalization of Simon (1955) where the decline of a firm and ultimately its exit occurs when its size reaches zero. In Simon (1960)'s model, the rate of firms' exit exactly compensates the flow of firms' births so that the economy is stationary and the steady-state distribution of firm sizes exhibit the same upper tail behavior as in Simon (1955). In contrast, Steindl (1965) includes births and deaths but within an industry with a growing number of firms. A steady-state distribution is obtained whose tail follows a power law with an exponent that depends on the net entry rate of new firms and on the average growth rate of incumbent firms. Zipf's law is only recovered in the limit where the net entry rate of new firms goes to zero. Both models rely on the existence of a minimum size below which a firm runs out of business. This hypothesis corresponds to the existence of a minimum efficient size below which a firm cannot operate, as is well established in economic theory. However, there may be in general more than one minimum size as the exit (death) level of a firm has no reason to be equal to the size of a firm at birth. In the afore mentioned models, these two sizes are assumed to be equal, while there is *a priori* no reason for such an assumption and empirical evidence *a contrario*.

In addition to the exit of a firm resulting from its value decreasing below a certain level, it sometimes happens that a firm encounters financial troubles while its asset value is still fairly high. One could cite the striking examples of Enron Corp. and Worldcom, whose market capitalization were supposedly high (actually the result of inflated total asset value of about $11 billion for Worldcom and probably much higher for Enron) when they went bankrupt. More recently, since mid-2007 and over much of 2008, the cascade of defaults and bankruptcies (or near bankruptcies) associated with the so-called subprime crisis by some of the largest financial and insurance companies illustrates that shocks in the network of inter-dependencies of these companies can be sufficiently strong to destabilize them. Beyond these trivial examples, there is a large empirical literature on firm entries and exits, that suggests the need for taking into account the existence of failure of large firms (Dunne et al., 1988, 1989; Bartelsman et al., 2005). To the extent that the empirical literature documents a sizable exit at all size categories, we suggest that it is timely to study a model with both firm exit at a size lower bound and due to a size-independent hazard rate. Such a model constitutes a better approximation to the empirical data than a model with only firm exit at the lower bound. Gabaix (1999) briefly considers an analogous situation (at least from a formal mathematical perspective) and suggests that it may have an important impact on the shape of the distribution of firm sizes.

To sum up, we consider an economy of firms undergoing continuous stochastic growth processes with births and deaths playing a central role at time scales as short as a few years. We argue that death processes are especially important to understand the economic foundation of Zipf's law and its robustness. We will consider two different mechanisms for the exit of a firm: (ɪ) when the firm's size

[1]We do not consider spin-off's or M&A (mergers and acquisitions).

becomes smaller than a given minimum threshold and (ıı) when an exogenous shock occurs, modeling for instance operational risks, independently of the size of the firm. The other important issue is to describe adequately the birth process of firms. As a counterpart to the continuously active death process, we will consider that firms appear according to a stochastic flow process which may depend on macro-economic variables and other factors. The assumptions underpinning this model as well as the main results derived from it are presented in section 2. Section 3 puts them in perspective in the light of recent theoretical models and empirical findings on the existence of deviations from Zipf's law. Section 4 provides complementary results which are important from an empirical point of view. All the proofs are gathered in the appendix at the end of the article.

2 Exposition of the model and main results

2.1 Model setup

We consider a reduced form model, with a first set of three assumptions, in which firms are created at random times t_i's with initial random asset values s_0^i's drawn from some given statistical distribution. More precisely:

Assumption 1. There is a flow of firm entry, with births of new firms following a Poisson process with exponentially varying intensity $\nu(t) = \nu_0 \cdot e^{d \cdot t}$, with $d \in \mathbb{R}$;

As will be clear later on, the value of the parameter ν_0 is not really relevant for the understanding of the shape of the distribution of firm sizes. In contrast, the parameter d, which characterizes the growth or the decline of the intensity of firm births, plays a key role and is directly related to the net growth rate of the population of firms

We also assume that the entry size of a new incumbent firm is random, with a typical size which is time varying in order to account for changing installment costs:

Assumption 2. At time t_i, $i \in \mathbb{N}$, the initial size of the new entrant firm i is given by $s_0^i = s_{0,i} \cdot e^{c_0 t_i}$. The random sequence $\{s_{0,i}\}_{i\in\mathbb{N}}$ is the result of independent and identically distributed random draws from a common random variable $\tilde{s}_0$ such that $\mathrm{E}\left[\tilde{s}_0^m\right] < \infty$; All the draws are independent of the entry dates of the firms.

Remark 1. Condition $\mathrm{E}\left[\tilde{s}_0^m\right] < \infty$ in Assumption 2 means that the fatness of the initial distribution of firm sizes at birth is less than the natural fatness resulting from the random growth. Such an assumption is not always satisfied, in particular in Luttmer (2007)'s model where, due to imperfect imitation, the size of entrant firms is a fraction of the size of incumbent firms.

As usual, we also assume that

Assumption 3. Gibrat's rule holds.

Assumption 3 means that, in the continuous time limit, the size $S_i(t)$ of the i^{th} firm of the economy at time $t \geq t_i$, conditional on its initial size s_0^i, is solution to the stochastic differential equation

$$dS_i(t) = S_i(t) \left(\mu \, dt + \sigma \, dW_i(t)\right) , \qquad t \geq t_i , \qquad S_i(t_i) = s_0^i . \tag{1}$$

The drift μ of the process can be interpreted as the rate of return or the ex-ante growth rate of the firm. Its volatility is σ and $W_i(t)$ is a standard Wiener process. Note that the drift μ and the volatility σ are the same for all firms.

This set of three assumptions extends Simon's model by allowing the creation of new firms at random times, instead of using fixed time steps, and, more importantly, it decouples the growth process of existing firms and the process of creation of new firms. It also departs from Gabaix (1999) and Luttmer (2007) models by considering a distribution of initial firm sizes that is unrelated to the distribution of already existing firms. As we shall see, apart from the growth rate c_0 of the typical size of a new entrant firm, the characteristics of the distribution of initial firm sizes are, to a large extent, irrelevant for the shape of the upper tail of the steady-state distribution of firm sizes, that we derive below based on the interplay between birth, death and the diffusion process (1).

We also consider two exit mechanisms, based on the following empirical facts. Referring to Bonaccorsi Di Patti and Dell'Ariccia (2004), the yearly rate of death of Italian firms is, on average, equal to 5.7% with a maximum of about 20% for some specific industry branches. Knaup (2005) examined the business survival characteristics of all establishments that started in the United States in the late 1990s when the boom of much of that decade was not yet showing signs of weakness, and finds that, if 85% of firms survive more than one year, only 45% survive more than four years. Brixy and Grotz (2007) analysed the factors that influence regional birth and survival rates of new firms for 74 West German regions over a 10-year period. They documented significant regional factors as well as variability in time: the 5-year survival rate fluctuates between 45% and 51% over the period from 1983 to 1992. Bartelsman et al. (2005) confirmed that a large number of firms enter and exit most markets every year in a group of ten OECD countries: data covering the first part of the 1990s show the firm turnover rate (entry plus exit rates) to be between 15 and 20 percents in the business sector of most countries, i.e., a fifth of firms are either recent entrants, or will close down within the year.

First of all, we assume that firms disappear when their asset values become smaller than some prespecified minimum level $s_{\min}$.

Assumption 4. There exists a minimum firm size $s_{\min}(t) = s_1 \cdot e^{c_1 \cdot t}$, that varies at the constant rate $c_1 \leq c_0$, below which firms exit.

This idea has been considered in several models of firm growth (see e.g. de Wit (2005) and references therein) and can be related to the existence of a minimum efficient size in the presence of fixed operating costs. Besides, as for the typical size of new entrant firms, we assume that the minimum size of incumbent firms grows at the constant rate $c_1 \geq 0$, so that $s_{\min}(t) := s_1 e^{c_1 \cdot t}$. But c_1 is a priori different from c_0. It is natural to require that the lower bound $\underline{s}_0$ of the distribution of $\tilde{s}_0$ be larger than s_1 and that $c_0 \geq c_1$ in order to ensure that no new firm enters the economy with an initial size smaller than the minimum firm size and then immediately disappears[2]. The condition $s_1 e^{c_1 \cdot t} < \underline{s}_0 e^{c_0 \cdot t}$ implies that the economy started at a time t_0 larger than

$$t_* = \frac{1}{c_1 - c_0} \cdot \ln\left(\frac{\underline{s}_0}{s_1}\right) < 0\ . \qquad (2)$$

We could alternatively choose $\underline{s}_0 = s_1$ so that the economy starts at time $t = 0$. Another approach, suggested for instance by Gabaix (1999), considers that firms cannot decline below a minimum size and remain in business at this size until they start growing up again. Here, we have not used this rather artificial mechanism.

Secondly, we consider that firms may disappear abruptly as the result of an unexpected large event (operational risk, fraud,...), even if their sizes are still large. Indeed, while it has been established that a first-order characterization for firm death involves lower failure rates for larger firms (Dunne et al., 1988,

[2]In fact, it seems that the typical size of entrant firms is much smaller than the minimum efficient size (Agarwal and Audretsch, 2001, and references therein). It means that two exit levels should be considered; one for old enough firms and another one for young firms. For tractability of the calculations, we do not consider this situation.

1989), Bartelsman et al. (2005) also state that, for sufficiently old firms, there seems to be no difference in the firm failure rate across size categories. Consequently

Assumption 5. There is a random exit of firms with constant hazard rate $h \geq \max\{-d, 0\}$ which is independent of the size and age of the firm.

Remark 2. As will become clear from the proof and from the exposition of the underpinnings of this result, the constraint $h \geq \max\{-d, 0\}$ is only necessary to guaranty that the distribution of firm sizes is normalized in the small size limit if there is no minimum firm size. The case $d > 0$ ensures that the population of firms grows at the long term rate d while the case $d < 0$ allows describing an industry branch that first expands, then reaches a maximum and eventually declines at the rate h. Such a situation is quite realistic, as illustrated by figure 2 in Sutton (1997) which depicts the number of firms in the U.S. tire industry. Notice, in passing, that the case $h < 0$ is also sensible. It corresponds to the situation considered by Gabaix (1999), where firms are allowed to enter with an initial size randomly drawn from the size distribution of incumbent firms.

Under assumptions 1 and 5, the average number N_t of operating firms satisfies

$$\frac{dN_t}{dt} + hN_t = \nu(t), \tag{3}$$

so that, assuming that the economy starts at $t = 0$ for simplicity, we obtain

$$N_t = \frac{\nu_0}{d+h}\left[e^{d\cdot t} - e^{-h\cdot t}\right]. \tag{4}$$

Consequently, the rate of firm birth, given by $\nu(t)/N_t$, is given by $\frac{d+h}{1-e^{-(d+h)\cdot t}} \to d+h$ for t large enough. The range of values of $d+h$ has been reported in many empirical studies. For instance, Reynolds et al. (1994) give the regional average firm birth rates (annual firm births per 100 firms) of several advanced countries in different time periods: 10.4% (France; 1981-1991), 8.6% (Germany; 1986), 9.3% (Italy; 1987-1991), 14.3% (United Kingdom; 1980-1990), 15.7% (Sweden; 1985-1990), 6.9% (United States; 1986-1988). They also document a large variability from one industrial sector to another. More interestingly, Bonaccorsi Di Patti and Dell'Ariccia (2004) as well as Dunne et al. (1988) reports both the entry and exit rate for different sectors in Italy and in the US respectively. In every cases, even if sectorial differences are reported, the average aggregated entry and exit rates are remarquably close. This suggests that d should be close to zero while h is about $4-6\%$. The net growth rate of the population of firms, given by $\frac{1}{N_t}\frac{dN_t}{dt} = \frac{\nu(t)}{N_t} - h \to d$ for t large enough, as announced after assumption 1.

2.2 Results

Equipped with this set of five assumptions, we can now define

$$m := \frac{1}{2}\left[(1-\delta+\delta_0) + \sqrt{(1-\delta+\delta_0)^2 + 4\,(\delta-\delta_0)\,\varepsilon}\right], \tag{5}$$

with

$$\delta := \frac{2\mu}{\sigma^2}\ ,\ \delta_0 := \frac{2c_0}{\sigma^2} \text{ and } \varepsilon := \frac{d+h}{\mu - c_0}\ , \tag{6}$$

and derive our main result (see appendix A.1 for the proof):

Proposition 1. *Under the assumptions 1-5, for* $t - t_* \gg \left[\left(\mu - \frac{\sigma^2}{2} - c_0\right)^2 + 2\sigma^2(d+h)\right]^{-1/2}$*, the average distribution of firm's sizes follows an asymptotic power law with tail index* m *given by (5), in the following sense: the average number of firms with size larger than* s *is proportional to* s^{-m} *as* $s \to \infty$.

Equation (5) can be rewritten as follows to make more explicit the dependence of the tail index on the basic parameters of the model μ, σ, c_0, d and h

$$m := \frac{1}{2}\left[\left(1 - 2\cdot\frac{\mu - c_0}{\sigma^2}\right) + \sqrt{\left(1 - 2\cdot\frac{\mu - c_0}{\sigma^2}\right)^2 + 8\cdot\frac{d+h}{\sigma^2}}\right]. \tag{7}$$

One can see that the tail index increases, and therefore the distribution of firm sizes becomes thinner tailed, as μ decreases and as h, c_0, and d increase. This dependence can be easily rationalized. Indeed, the smaller the expected growth rate μ, the smaller the fraction of large firms, hence the thinner the tail of the size distribution and the larger the tail index m. The larger h, the smaller the probability for a firm to become large, hence a thinner tail and a larger m. As for the impact of c_0, rescaling the firm sizes by $e^{c_0\cdot t}$, so that the mean size of entrant firms remains constant, does not change the nature of the problem. The random growth of firms is then observed in the moving frame in which the size of entrant firms remains constant on average. Therefore, the size distribution of firms is left unchanged up to the scale factor $e^{c_0\cdot t}$. Since the average growth rate of firms in the new frame becomes $\mu' = \mu - c_0$, the larger c_0, the smaller μ', hence the smaller the probability for a firm to become relatively larger than the others, the thinner the tail of the distribution of firm sizes and thus the larger m. Finally, the larger d is, the larger the fraction of young firms, which leads to a relatively larger fraction of firms with sizes of the order of the typical size of entrant firms and thus the upper tail of the size distribution becomes relatively thinner and m larger.

As a natural consequence of proposition 1, we can assert that

Corollary 1. *Under the assumptions of proposition 1, the mean distribution of firm sizes admits a well-defined steady-state distribution which follows Zipf's law (i.e.* $m = 1$*) if, and only if,*

$$\mu - h = d + c_0 \ . \tag{8}$$

Remark 3. In an economy where the amount of capital invested in the creation of new firms is constant per unit time, namely

$$\nu(t)\cdot s_0(t) = const.\ , \tag{9}$$

we necessarily get $d + c_0 = 0$ so that the balance condition reads $\mu = h$.

To get an intuitive meaning of the condition in corollary 1, let us state the following result (see the proof in appendix A.2):

Proposition 2. *Under the assumptions of proposition 1, the long term average growth rate of the overall economy is* $\max\{\mu - h, d + c_0\}$.

The term $d + c_0$ quantifies the growth rate of investments in new entrant firms, resulting from the growth of the number of entrant firms (at the rate d) and the growth of the size of new entrant firms (at the rate c_0). The term d reflects several factors, including improving pro-business legislation and tax laws as well as increasing entrepreneurial spirit. The latter term c_0 is essentially due to time varying installment costs, which can be negative in a pro-business economy.

The other term $\mu - h$ represents the average growth rate of an incumbent firm. Indeed, considering a running firm at time t, during the next instant dt, it will either exit with probability $h\cdot dt$ (and therefore its size declines by a factor -100%) or grow at an average rate equal to $\mu\cdot dt$, with probability $(1-h\cdot dt)$. The coefficient μ can be called the conditional growth rate of firms, conditioned on not having died yet. Then, the expected growth rate over the small time increment dt of an incumbent firm is $(\mu - h)\cdot dt + O\left(dt^2\right)$. As shown by equation (72) in appendix A.2, $\mu - h$ is also the return on investment of the economy.

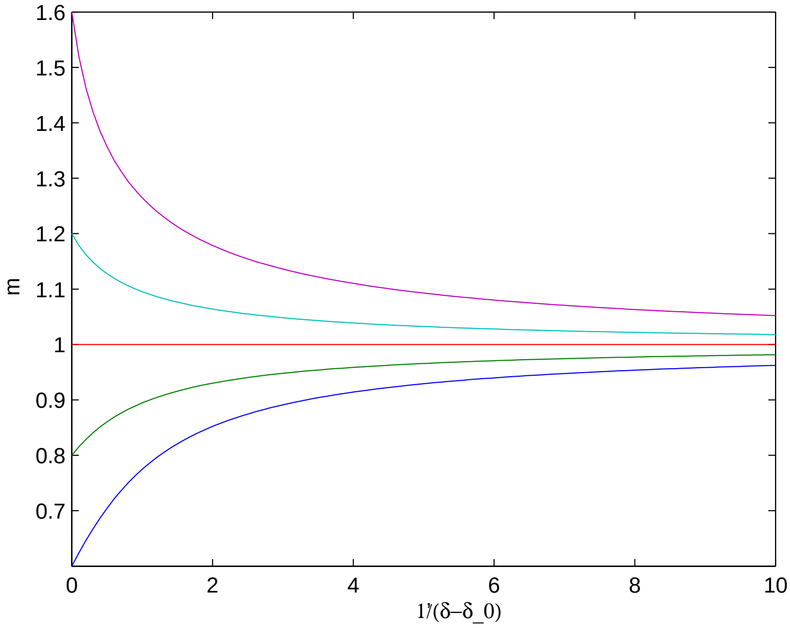

Figure 1: The figure shows the exponent m of the power law tail of the distribution of firm sizes, given by (5), as a function of $(\delta - \delta_0)^{-1}$, where $\delta = \frac{2\mu}{\sigma^2}$ and $\delta_0 = \frac{2c_0}{\sigma^2}$, for different values of the ratio $\varepsilon = \frac{d+h}{\mu - c_0}$. Bottom to top $\varepsilon = 0.6; 0.8; 1; 1.2; 1.6$.

Thus, the long term average growth of the economy is driven either by the growth of investments in new firms, whenever $d + c_0 > \mu - h$, or by the growth of incumbent firms, whenever $\mu - h > d + c_0$. The former case does not really make sense, on the long run. Indeed, it would mean that the growth of investments in new firms can be sustainably larger than the rate of return of the economy. Such a situation can only occur if we assume that the economy is fueled by an inexhaustible source of capital, which is obviously unrealistic. As a consequence, it is safe to assume $\mu - h \geq d + c_0$ on the long run. The regime $d + c_0 > \mu - h$ might however describe transient bubble regimes developing under unsustainably large capital creation (Baily et al., 2008).

Corollary 2. *In a growing economy whose growth is driven by that of incumbent firms, the tail index of the size distribution is such that $m \leq 1$.*
Along a balanced growth path, which corresponds to a maximum growth rate of the investment in new firms, the tail index of the size distribution is equal to one.

Proof. According to the natural assumption that the growth of the economy is driven by the growth of incumbent firms, i.e. $\mu - h \geq d + c_0$, we get $d + h \leq \mu - c_0$ and $\varepsilon \leq 1$ which leads to $m \leq 1$ (see illustration on figure 1); we have used assumption 5 according to which $d + h \geq 0$ hence $\mu - c_0 \geq 0$. On a balanced growth path, both investments in new firms and incumbent firms grow at the same rate $\mu - h = d + c_0$, hence the growth rate of the investment in new firms is maximum and by corollary 1 the tail index m of the size distribution equals one. □

In the present framework, the crucial parameters d, c_0, μ and h are exogenous. While this is beyond the scope of the present paper, we can however surmise that, within an endogenous theory in which the growth of investments would be naturally correlated with the growth of the firms in the economy because the success of firms generates the cash flow at the source of new investments, the balance growth condition (8) appears almost unavoidable for a sustainable development. It is quite remarkable that Zipf's law derives as the robust statistical translation of this balance growth condition.

Remark 4. Our theory suggests two simple explanations for the empirical evidence that the exponent m is close to 1. Either the investment in new firms is close to its maximum sustainable level so that the balance condition is approximately satisfied, or the volatility σ of incumbent firms sizes is large. Indeed, according to equation (7), the tail index m goes to one as σ goes to infinity irrespective of the values of the parameters d, μ, c_0 and h. In fact, the larger the volatility, the larger the tolerance to the departure from the balance condition. Indeed, expanding relation (7) for σ large, we get

$$m = 1 - 2 \cdot \frac{\mu - h - c_0 - d}{\sigma^2} + 4 \cdot \frac{(d+h)\,(\mu - h - c_0 - d)}{\sigma^4} + O\left(\frac{1}{\sigma^6}\right), \tag{10}$$

and for small departures from the balance condition

$$m = 1 - \frac{2}{1 + 2\frac{d+h}{\sigma^2}} \cdot \frac{\mu - c_0 - h - d}{\sigma^2} + \frac{8\frac{d+h}{\sigma^2}}{\left(1 + 2\frac{d+h}{\sigma^2}\right)^3} \cdot \left(\frac{\mu - c_0 - h - d}{\sigma^2}\right)^2. \tag{11}$$

2.3 Calibration to empirical data

According to Dunne et al. (1988, table 2), the relative size of entrant firms to incumbent firms seems to have slightly declined during the period 1963-1982 in the US. According to our model, the ratio of the average size of entrant firms to the average size of incumbent firms is, for large enough time t,

$$\frac{s_0 \cdot e^{c_0 \cdot t}}{\Omega_t / N_t} \sim \begin{cases} \frac{\mu - h - d - c_0}{d+h} \cdot e^{-(\mu - h - d - c_0)\cdot t}, & \text{provided that } \mu - h > d + c_0, \\ \frac{1}{d+h} \cdot \frac{1}{t}, & \text{provided that } \mu - h = d + c_0, \\ \frac{d + h - \mu + c_0}{d+h}, & \text{provided that } \mu - h < d + c_0, \end{cases} \tag{12}$$

where Ω_t is the average size of all incumbent firms (see Appendix A.2) and N_t is the average number of incumbent firms, at time t. The fact that Dunne et al. (1988) observe a slight decay in the relative size of entrant firms to incumbent firms suggests that the condition of sustainable growth $\mu - h > d + c_0$ holds. Under this hypothesis, the calibration of the first alternative in (12) by OLS gives, on an annual basis,

$$\mu - h - d - c_0 = 1.8\% \; (1.2\%) \qquad \text{and} \qquad \frac{\mu - h - d - c_0}{d + h} = 28\% \; (3\%). \tag{13}$$

The figures within parenthesis provide the standard deviations of the estimates. As a consequence, the alternative hypothesis $\mu - h - d - c_0 \leq 0$ cannot be rejected at any usual significance level and we cannot affirm that Dunne's data corresponds to the regime $\mu - h - d - c_0 > 0$.

Under the second hypothesis $\mu - h = d - c_0$, the second alternative in (12) leads to test the null hypothesis that the slope of the OLS regression of the logarithm of the size of entrant firms relative to the size of incumbent firms against the logarithm of time is equal to -1. Instead, we estimate a slope equal to -0.086 (0.068), which is therefore not significantly different from zero. Thus, the second alternative is rejected.

According to the third alternative in (12), the size of entrant firms relative to the size of incumbent firms is constant over the period under consideration. To formally test this hypothesis, we perform the OLS regression of the size of entrant firms versus to the size of incumbent firms against time. We find that the hypothesis of a time dependent ratio of the size of entrant firms relative to the size of incumbent firms is rejected at any usual significance level. We thus have to conclude that the third alternative actually holds and we get

$$\frac{d+h-\mu+c_0}{d+h} = 25\%. \tag{14}$$

With the figures $d = 0$ and $h = 5\%$ obtained from Dunne et al. (1988), we obtain $d + h - \mu + c_0 = 1.25\%$ and $\mu - c_0 = 3.75\%$. Thus, the balance condition is not strictly satisfied but the observed departure from the balance condition remains weak.

To sum up, reasonable estimates of the key parameters are $h = 4-6\%$, $d = \pm 0.5\%$, $\mu - c_0 = h \pm 2\%$. As for σ, Buldyrev et al. (1997) report the standard deviations of the growth rates in terms of sales, assets, cost of goods sold and plant property and equipment for US publicly-traded companies. Buldyrev et al. (1997) find that σ ranges typically between 30% to 50%. Based upon this set of figures, relation (7) leads to a tail index m ranging between 0.7 and 1.3, in agreement with the range of values usually reported in the literature.

Proposition 1 states that the asymptotic power law of the distribution of firm sizes can be observed if the age of the economy is large compared with $\left[\left(\mu - \frac{\sigma^2}{2} - c_0\right)^2 + 2\sigma^2(d+h)\right]^{-1/2}$. With the set of parameters above, this corresponds to economies whose age is large compared to 5 to 12 years.

3 Discussion

3.1 Comparison with Gabaix's model

Corollary 1 seems reminiscent of the condition given by Gabaix (1999), which relies on the argument that, because they are all born at the same instant, firms grow – on average – at the same rate as the overall economy. Consequently, when discounted by the global growth rate of the economy, the average expected growth rate of the firms must be zero. Applied to our framework, and focusing on the distribution of *discounted* firm sizes, this argument would lead to $\mu = h$, with $d = c_0 = c_1 = 0$ in order to match Gabaix's assumptions. Gabaix (1999)'s condition would thus seem to be equivalent to our balance condition for Zipf's law describing the density of firms' sizes to hold.

Actually, this reasoning is incorrect. Consider the case where $\mu > h$, such that the global economy grows at the average growth rate $r_G = \mu - h$. Gabaix (1999) proposed to measure the growth of a firm in the frame of the global economy. In this frame, the new conditional average growth rate of the firm is $\mu' = \mu - r_G = h$, which indeed would suggest that the balance condition is *automatically* obeyed when μ is replaced by μ'. But, one should notice that μ' is a transformed growth rate, and not the true rate. The average growth rate $r_G = \mu - h$ of the global economy is micro-founded on the contributions of all growing firms. It would be incorrect to insert μ' in the statements of Proposition 1, as μ' is the effective growth rate resulting from the change of frame, while our exact derivation requires the parameters μ and h for Proposition 1 to hold. As such, nothing in our model automatically sets the growth rate μ of firms to their death rate h, contrarily to what happens in Gabaix (1999)'s model. The main difference that invalidates the application of Gabaix (1999)'s argument is the stochastic flow of firm's births and deaths.

It is important to understand that Gabaix (1999)'s derivation of Zipf's law relies crucially on a model view of the economy in which *all firms are born at the same instant*. Our approach is thus essentially different since it considers the flow of firm births, as well as their deaths, which is more in agreement with empirical evidence. Note also that the available empirical evidence on Zipf's law is based on analyzing *cross-sectional* distributions of firm sizes, i.e., at specific times. As a consequence, the change to the global economic growth frame, argued by Gabaix (1999), just amounts to multiplying the value of each firm by the same constant of normalization, equal to the size of the economy at the time when the cross-section is measured. Obviously, this normalization does not change the exponent of the power law distribution of sizes, if it exists. Furthermore, elaborating on Krugman (1996)'s argument about the non-convergence of the distribution of firm sizes toward Zipf's law in Simon (1955)'s model, Blank and Solomon (2000) have shown that Gabaix (1999)'s argument suffers from a more technical problem. Based on the demonstration that the two limits, the number of firms $N \to \infty$ and $s_{\min}/\mathrm{E}[s] \to 0$[3] (or equivalently the limit of large times $t \to \infty$) are non-commutative, Blank and Solomon (2000) showed that Zipf's exponent $m = 1$ as obtained by Gabaix (1999)'s argument requires (i) taking the long time limit $s_{\min}/\mathrm{E}[s] \to 0$ over which the economy made of a large but finite number N firms grows without bounds, while simultaneously obeying the condition (ii) $N \gg \exp[\mathrm{E}[s]/s_{\min}]$. The problem is that conditions (i) and (ii) are mutually exclusive. Blank and Solomon (2000) showed that this inconsistency can be resolved by allowing the number of firms to grow proportionally to the total wealth of the economy.

Our approach allows to overcome Blank and Solomon (2000)'s objection to Gabaix's model. Indeed, in Gabaix's model, the balance condition ensures that the size of any firm goes to zero almost surely. Thus, since all the firms are born at the same time, it is necessary to introduce a repelling lower bound in order to let some of them keep a non-vanishing size. Otherwise, the size distribution degenerates. In contrast, our approach assumes an unremitting flow of firms' birth. So, even if our balance condition implies that firm sizes almost surely go to zero (or to the lower admissible firm size) as in Gabaix's model, our economy always exhibits a significant number of firms whose size is arbitrarily large so that the convergence toward the limit distribution always occurs, whether or not Blank and Solomon's condition holds. We discuss this issue in more details in the next section where we investigate the effect of the finite age of the economy on the shape of the distribution of firm sizes. In addition, since Proposition 1 also takes into account the case where the rate of firms' births increases exponentially as $\nu(t) = \nu_0 e^{d \cdot t}$, when $d > 0$, our analysis provides a generalization of Blank and Solomon (2000)'s results.

3.2 Comparison with Luttmer's model

Based upon structural models, an important modeling strategy has been developed, starting from Lucas (1978) and evolving to the more recent Luttmer (2007, 2008) or Rossi-Hansberg and Wright (2007a,b) models. The distribution of firm sizes then appears as one of the properties of a general equilibrium model, which depends on different industry parameters. In these models, Zipf's law is obtained as a limit case, needing a rather sharp fine tuning of the control parameters. Rossi-Hansberg and Wright's model is a "one firm" model as in Gabaix (1999) and is therefore subjected to the same restrictions. We do not discuss further this model in light of the results of our reduced form model. In contrast, the assumptions underpinning Luttmer's model match the assumptions under which proposition 1 holds, which motivates a closer comparison.

Luttmer (2007) considers an economy of firms with different ages. For a firm of age a, its size S_a follows a geometric Brownian motion

$$d \ln S_a = \mu \cdot da + \sigma \cdot dW_a \tag{15}$$

where the drift μ and volatility σ are derived from a micro-economic model and are related to the price elasticity of the demand for commodity $\frac{\beta}{1-\beta}$, to the rate at which the productivity of entering firms grows over time θ_E, to the trend of log productivity for incumbent firms θ_I, and to the volatility of the

[3]The term $\mathrm{E}[s]$ refers to the statistical average over the finite but large population of N firms.

productivity σ_Z

$$\mu = \frac{\beta}{1-\beta}\left(\theta_I - \theta_E\right), \qquad \sigma = \frac{\beta}{1-\beta}\sigma_Z. \tag{16}$$

Due to the presence of fixed costs, incumbent firms exit when their size reaches a constant minimum size b and in this case only. In our notations, this implies $c_1 = 0$ and $h = 0$. In addition, Luttmer assumes that the overall number of incumbent firms grows at a rate $\eta > \mu + \sigma^2/2$ so that the size of a typical incumbent firm is not expected to grow faster than the population growth rate. Within our framework, the number of firms grows, on the long run, at the rate d, so that we have the correspondence $d = \eta$. Finally, Luttmer considers that firms enter either with a fixed size or with a size taken from the same distribution as the incumbent firms; consequently, in our notations, we have $c_0 = 0$. Then, by application of proposition 1, we conclude, as in Luttmer (2007, section III.B), that the size distribution of firms follows a power law with a tail index given by

$$m = -\frac{\mu}{\sigma^2} + \sqrt{\left(\frac{\mu}{\sigma^2}\right)^2 + 2\frac{\eta}{\sigma^2}}\,. \tag{17}$$

Notice that Luttmer only considers the long term distribution of firm sizes, while our result allows considering the transient regime which eventually leads to the power law. In particular, accounting for the transient regime avoids resorting to the assumption $\eta > \mu + \sigma^2/2$. Indeed, in Luttmer's model, this assumption ensures that the tail index m remains larger than one so that, detrended by the overall growth of the number of firms given by $e^{+\eta t}$, the average firm size is finite. This is a natural requirement if the economy is assumed to be finite. This latter assumption is more questionable in economies of infinite duration. In contrast, when finite time effects are considered as in our framework, the average firm size is always finite for finite times, since the density of firm sizes decays faster than any power law beyond the intermediate asymptotic described by the power law, whether $\eta > \mu + \sigma^2/2$ or not. In other words, the power law is truncated by a finite time effect, as derived in appendix A.1. As time increases, the truncation recedes progressively to infinity, thus enlarging the domain of validity of the power law. The exact power law distribution is attained therefore for asymptotically large times (we come back to this point latter on in section 4.2). Therefore, the constraint $\eta > \mu + \sigma^2/2$ is not necessary anymore in our framework, since there is no reason for the average firm size to remain finite at infinite times when the size of the overall economy becomes itself infinite.

The endogeneization of the growth rate of the productivity of entrant firms performed by Luttmer in the second part of his article does not match our assumptions, so that we cannot proceed further with the comparison of his results with ours. Indeed, in Luttmer's case, the upper tail of the distribution of entrant firms behaves as the tail of the distribution of incumbent firms, so that assumption 2 is not satisfied.

4 Miscellaneous results

4.1 Distribution of firms' age and declining hazard rate

Brudel et al. (1992), Caves (1998, and references therein) or Dunne et al. (1988, 1989), among others, have reported declining hazard rates with age. Under assumption 5, the hazard rate is constant, which seems to be counterfactual. However, we now show that the presence of the lower barrier below which firms exit allows to account for age-dependent hazard rate.

Let us denote by θ the age of a firm at time t, i.e., the firm was born at time $t - \theta$. Expression (51) in appendix A.1 allows us to derive the probability that, at time t, a firm older that θ is still alive, which corresponds to the distribution of firm ages. Indeed denoting by $\tilde{\Theta}_t$ the random age of the considered firm at time t,

$$\Pr\left[\tilde{\Theta}_t > \theta\right] = \int_{s_{\min}(t)}^{\infty} \frac{1}{s}\varphi\left[\ln\left(\frac{s}{s_{\min}(t)}\right); t, \theta\right] ds, \tag{18}$$

where $\frac{1}{s}\varphi\left[\ln\left(\frac{s}{s_{\min}(t)}\right); t, \theta\right]$ is the size density of firms of age θ at time t. Some algebraic manipulations give

$$\begin{aligned}
\Pr\left[\tilde{\Theta}_t > \theta\right] &= \frac{1}{2}\left[\operatorname{erfc}\left(-\frac{\ln\rho(t) + (\delta - 1 - \delta_0)\tau}{2\sqrt{\tau}}\right)\right. && (19)\\
&- \left.\rho(t)^{1-\delta+\delta_0} \cdot \operatorname{erfc}\left(\frac{\ln\rho(t) - (\delta - 1 - \delta_0)\tau}{2\sqrt{\tau}}\right)\right], && (20)\\
& && (21)
\end{aligned}$$

with $\tau := \frac{\sigma^2}{2}\theta$.

Accounting for the independence of the random exit of a firm with hazard rate h (assumption 5) from the size process of the firm, the "total" hazard rate reads

$$\begin{aligned}
\mathcal{H}(t,\theta) &= = h - \frac{d\ln\Pr\left[\tilde{\Theta}_t > \theta\right]}{d\theta}, && (22)\\
&= h + \frac{\ln\left(\frac{s_0(t)}{s_{\min}(t)}\right)\cdot\left(\frac{s_0(t)}{s_{\min}(t)}\right)^{-\frac{1-\delta+\delta_0}{2}}\cdot\exp\left[-\frac{\ln^2\left(\frac{s_0(t)}{s_{\min}(t)}\right)+(1-\delta+\delta_0)^2\tau^2}{4\tau}\right]}{\operatorname{erfc}\left(-\frac{\ln\frac{s_0(t)}{s_{\min}(t)}+(\delta-1-\delta_0)\tau}{2\sqrt{\tau}}\right)-\left(\frac{s_0(t)}{s_{\min}(t)}\right)^{1-\delta+\delta_0}\cdot\operatorname{erfc}\left(\frac{\ln\frac{s_0(t)}{s_{\min}(t)}-(\delta-1-\delta_0)\tau}{2\sqrt{\tau}}\right)} && (23)
\end{aligned}$$

assuming, for simplicity, that the random variable $\tilde{s}_0$ reduces to a degenerate random variable s_0. Expression (23) shows that the failure rate actually depends on firm's age. It also depends explicitly on the current time t through the ratio $\frac{s_0(t)}{s_{\min}(t)}$.

Let us focus on the case $c_0 = c_1$ ($\delta_0 = \delta_1$), which corresponds to the same growth rate for $s_0(t)$ and $s_1(t)$. This allows considering arbitrarily old firms since, according to (2), the starting point of the economy can be $t_* = -\infty$. We then obtain the limit result

$$\mathcal{H}(t,\theta) \stackrel{\theta\to\infty}{\longrightarrow} \begin{cases} h, & \delta - \delta_0 - 1 \geq 0, \\ h + \dfrac{(\delta - \delta_0 - 1)^2\,\sigma^2}{8}, & \delta - \delta_0 - 1 < 0\,. \end{cases} \tag{24}$$

Differentiating the age-dependent hazard rate given by (23) with respect to θ and using the asymptotic expansion of the error function (Abramowitz and Stegun, 1965), we get

$$\partial_\theta\mathcal{H}(t,\theta) = \begin{cases} -\dfrac{(\delta-\delta_0-1)^2\cdot\sigma^2}{8}\cdot\mathcal{H}(t,\theta)\cdot\left[1+O\left(\dfrac{1}{\theta}\right)\right], & \delta-\delta_0-1\geq 0, \\ -\dfrac{12}{(\delta-\delta_0-1)^2\cdot\sigma^2\cdot\theta^2}\mathcal{H}(t,\theta)\cdot\left[1+O\left(\dfrac{1}{\theta}\right)\right], & \delta-\delta_0-1<0, \end{cases} \tag{25}$$

which shows that the total failure rate decreases with age, at least for large enough age, in agreement with the literature.

4.2 Deviations from Zipf's law due to the finite age of the economy

Considering, for simplicity, that $\tilde{s}_0$ is a degenerate random variable such that $\Pr[\tilde{s}_0 = s_0] = 1$, we can determine the deviations from the asymptotic power law tail of the mean density of firm sizes (given explicitly by (65) in appendix A.1) due to the finite age of the economy. For this, it is convenient to study the s-dependence of the mean number of firms whose sizes exceeds a given level s:

$$N(s,t) = \int_s^\infty g(s',t)ds' \; . \tag{26}$$

Zipf's law corresponds to $N(s,t) \sim s^{-1}$ for large s.

All calculations done, defining $s_0(t) := s_0 e^{c_0 \cdot t}$ as being the initial size of an entrant firm at time t when $\Pr[\tilde{s}_0 = s_0] = 1$, we obtain the number of firms whose normalized size $\frac{s}{s_0(t)}$ is larger than κ at the standardized age $\tau := \frac{\sigma^2}{2}\theta$,

$$\mathcal{N}(\kappa,\tau) = B_-\,\kappa^{-\varrho_-} + B_+\,\kappa^{+\varrho_+} - C \;, \tag{27}$$

where

$$\begin{aligned}
B_- &:= \frac{1}{2\alpha(\eta)\varrho_-}\left[\operatorname{erfc}\left(\frac{\ln\kappa - \tau\alpha(\eta)}{2\sqrt{\tau}}\right) - \left(\frac{s_0(t)}{s_{\min}(t)}\right)^{-\alpha(\eta)} \operatorname{erfc}\left(\frac{\ln(\kappa\rho^2) - \tau\alpha(\eta)}{2\sqrt{\tau}}\right)\right] , \\
B_+ &:= \frac{1}{2\alpha(\eta)\varrho_+}\left[\operatorname{erfc}\left(\frac{\ln\kappa + \tau\alpha(\eta)}{2\sqrt{\tau}}\right) - \left(\frac{s_0(t)}{s_{\min}(t)}\right)^{\alpha(\eta)} \operatorname{erfc}\left(\frac{\ln(\kappa\rho^2) + \tau\alpha(\eta)}{2\sqrt{\tau}}\right)\right] , \\
C &:= \frac{1}{2\eta}e^{-\eta\tau}\left[\operatorname{erfc}\left(\frac{\ln\kappa - \tau\alpha}{2\sqrt{\tau}}\right) - \left(\frac{s_0(t)}{s_{\min}(t)}\right)^{-\alpha} \operatorname{erfc}\left(\frac{\ln(\kappa\rho^2) - \tau\alpha}{2\sqrt{\tau}}\right)\right] ,
\end{aligned} \tag{28}$$

and $\varrho_\pm := \frac{1}{2}\left[\alpha \pm \alpha\,(\eta)\right]$, $\alpha := \delta - 1 - \delta_0$, $\alpha(\eta) := \sqrt{\alpha^2 + 4\eta}$, $\eta := \frac{\sigma^2}{2}\,(d+h)$.

Figure 2 shows the mean cumulative number $\mathcal{N}(\kappa,\tau)$ of firms as a function of the normalized firm size κ, for $\mu = c_0$ and $h = -d > 0$ satisfying to the balance condition of corollary 1, for $s_0 = 100 \cdot s_{\min}$, $c_0 = c_1 = 0$ and reduced times $\tau = 5, 10, 50$. As expected, the older the economy, the closer is the mean cumulative number $\mathcal{N}(\kappa,\tau)$ to Zipf's law $\mathcal{N}(\kappa,\infty) \sim \kappa^{-1}$. The downward curvatures of the graphs for all finite τ's show that the apparent tail index can be empirically found larger than 1 even if all conditions for the asymptotic validity of Zipf's law hold.

This effect could provide an explanation for some dissenting views in the literature about Zipf's law. The two recent influential studies by Cabral and Mata (2003) and Eeckhout (2004)[4] have suggested that the distribution of firms size could be well-approached by the log-normal distribution, which exhibits a downward curvature in a double-logarithmic scale often used to qualify a power law. Our model shows that a slight downward curvature can easily be explained by the partial convergence of the distribution of firm sizes toward the asymptotic Zipf's law due to the finite age of the economy.

It is interesting to note that two opposing effects can combine to make the apparent exponent m close to 1 even when the balance condition does not hold exactly. Consider the situation where $\varepsilon = \frac{d+h}{\mu - c_0} < 1$. For $\varepsilon < 1$, figure 1 shows that $m < 1$ for all values δ. But, figure 2 shows that the distribution of firm sizes for a finite economy is approximately a power law but with an exponent larger than for the asymptotic regime of an infinitely old economy. It is possible that these two deviations may cancel out to a large degree, providing a nice apparent empirical Zipf's law.

[4] See the comment by Levy (2009) which suggests that the extreme tail of the size distribution is indeed a power law, the reply by Eeckhout (2009) and the confirmation by Malevergne et al. (2009).

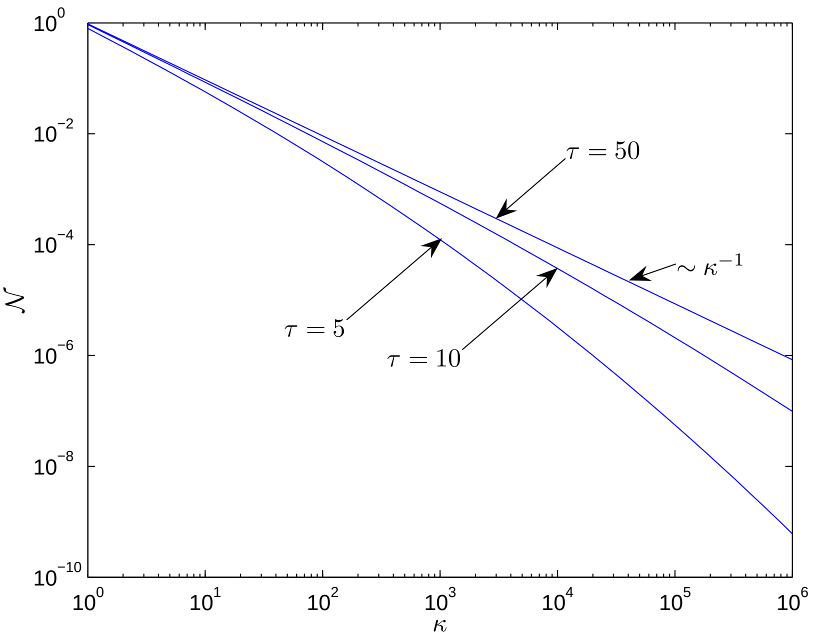

Figure 2: The figure quantifies the deviations from Zipf's law resulting from the finite age of the economy, by showing the mean number $\mathcal{N}(\kappa, \tau)$ of firms of normalized size $s/s_0(t)$ larger than κ as a function of κ, for parameters $\delta = \delta_0$, $d + h = 0$ (satisfying the balance condition), and for $s_0 = 100 \cdot s_{\min}$, $c_0 = c_1 = 0$ and reduced times $\tau = 5; 10; 50$ as defined in (47). The exact asymptotic Zipf's law $\sim \kappa^{-1}$ is also shown for comparison.

4.3 Representativeness of the mean-distribution of firm size

All our results have been established for the average number $N(s,t)$ of firms whose size is larger than s, where the average is performed over an ensemble of equivalent statistical realizations of the economy. Since empirical data are usually sampled from a single economy, it is important to ascertain if the average Zipf's law accurately describes the distribution of single typical economies. The answer to this question is provided by the following proposition whose proof is given in appendix A.3.

Proposition 3. *Under assumptions 1 and 2, the random number $\tilde{N}(s,t)$ of firms whose size is larger than s in a given economy follows a Poisson law with parameter $N(s,t)$ (defined in (39) with (44) and (66)):*

$$\Pr\left[\tilde{N}(s,t)=n\right]=\frac{N(s,t)^n}{n!}e^{-N(s,t)}. \tag{29}$$

As a consequence of proposition 3, we state

Corollary 3. *Under the assumptions of proposition 3, the variance of the average relative distance $\frac{\tilde{N}(s,t)}{N(s,t)}-1$ between the number of firms in one realization and its statistical average is given by*

$$\mathrm{E}\left[\left(\frac{\tilde{N}(s,t)}{N(s,t)}-1\right)^2\right]=\frac{1}{N(s,t)}. \tag{30}$$

Proof. The left hand side of the equation above is nothing but the variance of $\tilde{N}(s,t)$ divided by $N(s,t)^2$. Since, $\tilde{N}(s,t)$ follows a Poisson law, $\mathrm{Var}\,\tilde{N}(s,t)=N(s,t)$, hence the result. □

To give a quantitative illustration, let us consider firms whose sizes evolve according to the pure Geometric Brownian Motion, i.e., $\Pr[\tilde{s}_0=s_0]=1$, $c_0=c_1=h=d=0$ and no minimum exit size. Then, $N(s,t)=N(s)=\int_s^\infty g(s')ds'$, where

$$g(s)=\frac{2\nu_0}{\sigma^2(1-\delta)}\,s_0^{1-\delta}\,s^{\delta-2}\,, \qquad s>s_0\,, \qquad \delta=\frac{2\mu}{\sigma^2}<1\,. \tag{31}$$

This expression derives from the general expression (65) for Zipf's law given in appendix A.1 in the limit $s_{\min}\to 0$. This leads to

$$N(s)=N_0\,\left(\frac{s_0}{s}\right)^{1-\delta}\,, \tag{32}$$

where

$$N_0=\int_{s_0}^\infty g(s)ds=\frac{2\nu_0}{\sigma^2(1-\delta)^2} \tag{33}$$

is the mean number of firms, whose sizes, at a given time t, are larger than the initial size s_0. Using (30) for the variance of the relative distance between the number of firms in one realization and its statistical average in Corollary 2, and with (32), we obtain that the variance of the relative distance is given by $\frac{1}{N_0}\,\frac{s}{s_0}$, where we have assumed that $\delta=0$, so that Zipf's law $N(s)\sim s^{-1}$ holds for the mean distribution of firm sizes.

In this illustrative example, the total number of firms is infinite while N_0 remains finite. Let us consider a data set spanning the range $s\in(s_0,s_*)$ where $s_*=0.01\,N_0\,s_0$ is such that the variance of the relative distance between the number of firms in one realization and its statistical average remains smaller than 10^{-2} over the range $s\in(s_0,s_*)$. Suppose that the mean number of firms in the economy, whose sizes are larger than s_0, is equal to $N_0=10^6$. Then $s_*=10^4\,s_0$, showing that Zipf's law should be observed, in a single realization of an economy, with good accuracy over four orders of magnitudes in this example. Figure 3 depicts ten simulation results obtained for such an economy.

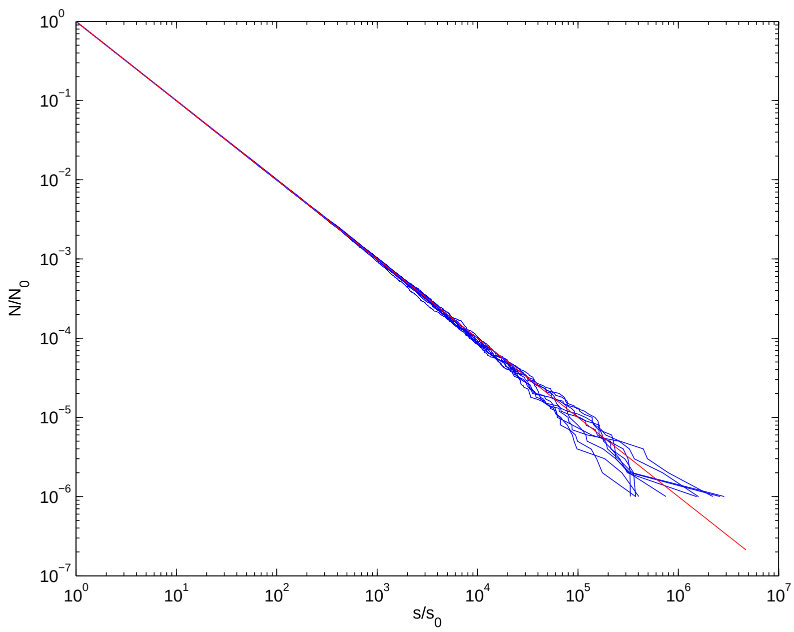

Figure 3: Number of firms whose size is larger than s when $\sigma = 0.01$, $\nu_0 = 50$ and $\delta = 0$ for ten realizations of the economy. The straight red line depicts Zipf's law for the *mean* number of firms.

5 Conclusion

We have presented a general theoretical derivation of Zipf's law, which states that, for most countries, the size distribution of firms is a power law with a specific exponent: the number of firms with size greater than S is inversely proportional to S. Our framework has taken into account time-varying firm creation, firms' exit resulting from both a lack of sufficient capital and sudden external shocks, and Gibrat's law of proportional growth. We have identified that four key parameters control the tail index m of the power law distribution of firms sizes: the expected growth rate μ of incumbent firms, the hazard rate h of random exits of firms of any size, the growth rate c_0 of the size of entrant firms, and the growth rate d of the number of new firms. We have identified that Zipf's law holds exactly when a balance condition holds, namely when the growth rate $d + c_0$ of investments in new entrant firms is equal to the average growth rate $\mu - h$ of incumbent firms. Thus, Zipf's law can be interpreted as a remarkable statistical signature of the long-term optimal allocation of resources that ensures the maximum sustainable growth rate of an economy. We have also found that Zipf's law is recovered approximately when the volatility of the growth rate of individual firms becomes very large, even when the balance condition does not hold exactly. We have studied the deviations from Zipf's law due to the finite age of the economy and shown that a deviation of the balance condition $d + c_0 = \mu - h$ can be compensated approximately by the effect of the finite age of the economy to give again an approximate Zipf's law. We have also shown that the presence of a minimum size below which firms exit allows us to account for the age-dependent hazard rate documented in the empirical literature. Our results hold not only for statistical averages over ensemble of economies (i.e., in expectations) but also apply to a single typical economy, as the variance of the relative difference between the number of firms in one realization and its statistical average decays as the inverse of the number of firms and thus goes to zero very fast for sufficiently large economies. Therefore, our results can be compared with empirical data which are usually sampled for a single economy. Our theory improves significantly on previous works by getting rid of many constraints and conditions that are found unnecessary or artificial, when taking into account the proper interplay between birth, death and growth.

A Appendix

A.1 Derivation of the distribution of firms' sizes: proof of proposition 1

Consider an economy with many firms born at random times $t_i \geq t_0$, $i \in \mathbb{N}$, where t_0 is the starting time of the economy. We assume that no two firms are born at the same time so that the random sequence $\{t_i\}_{i\in\mathbb{N}}$ defines a *simple point process* (Daley and Vere-Jones, 2007, def. 3.3.II).

Let $S_i(t)$, $i \in \mathbb{N}$, $t \geq t_0$ be a positive real-valued stochastic process representing the size, at time t, of the firm born at t_i. Obviously, $S_i(t) = 0$, $\forall t < t_i$. The sequence $\{t_i, S_i(t)\}_{i\in\mathbb{N}}$ defines a *simple marked point process* (Daley and Vere-Jones, 2007, def. 6.4.I - 6.4.II) with ground process $\{t_i\}_{i\in\mathbb{N}}$ and marks $\{S_i(t)\}_{i\in\mathbb{N}}$. We assume that $\{t_i\}$ and $\{S_i(t)\}$ are mutually independent and such that the distribution of $S_i(t)$ depends only on the corresponding location in time t_i. Consequently, the *mark kernel* $F_{m,i}(s,t) := \Pr[S_i(t) < s]$ simplifies to $F_m(s, t|t_i)$.

For any subset $T \times \Sigma$ of $[t_0, \infty) \times \mathbb{R}_+$, we introduce the *counting measure*

$$\begin{aligned} N_t(T \times \Sigma) &:= \#\{t_i \in T,\ S_i(t) \in \Sigma\}, && (34)\\ &= \sum_{i\in\mathbb{N}:\ t_i\in T} 1_{S_i(t)\in\Sigma}\,. && (35) \end{aligned}$$

The total number of firms whose sizes are larger than s at time t then reads

$$\begin{aligned} \tilde{N}(s,t) &:= N_t([t_0, t) \times [s, \infty)), && (36)\\ &= \int_{[t_0,t)\times[s,\infty)} N_t(du \times ds), && (37)\\ &= \sum_{i\in\mathbb{N}:\ t_i\leq t} 1_{S_i(t)\geq s}\,. && (38) \end{aligned}$$

As a consequence of theorem 6.4.IV.c in Daley and Vere-Jones (2007) we can state that

Lemma 1. *Provided that the ground process $\{t_i\}_{i\in\mathbb{N}}$ admits a first order moment measure with density $\nu(t)$ w.r.t Lebesgue measure, the counting process $\tilde{N}(s,t)$ admits a first moment*

$$\begin{aligned} N(s,t) &:= E\left[\tilde{N}(s,t)\right], && (39)\\ &= \int_{t_0}^{t} [1 - F_m(s, t|u)] \cdot \nu(u)\, du. && (40) \end{aligned}$$

Remark 5. When the ground process is an (inhomogeneous) Poisson process, $\nu(t)$ is nothing but the intensity of the process.

Proof. By theorem 6.4.IV.c in Daley and Vere-Jones (2007), the first-moment measure $M_1(\cdot) := \mathrm{E}[N_t(\cdot)]$ of the marked point process $\{t_i, S_i(t)\}_{i\in\mathbb{N}}$ exists since the corresponding moment measure exists for the ground process $\{t_i\}_{i\in\mathbb{N}}$. It reads

$$M_1(du \times ds) = \nu(u)du \cdot F_m(ds, t|u). \qquad (41)$$

As a consequence

$$N(s,t) = \mathrm{E}\left[\int_{[t_0,t)\times[s,\infty)} N_t(du \times ds)\right] = \int_{[t_0,t)\times[s,\infty)} M_1\,(du \times ds)\,, \tag{42}$$

$$= \int_{[t_0,t)\times[s,\infty)} \nu(u)du \cdot F_m(ds,t|u) = \int_{t_0}^{t} [1 - F_m\,(s,t|u)] \cdot \nu(u)\,du. \tag{43}$$

$\square$

As an immediate consequence, provided that $S(t)$ admits a density $f_m(s,t|u)$ with respect to Lebesgue measure, the counting process $\tilde{N}(s,t)$ admits a first-moment density

$$g(s,t) := \int_{t_0}^{t} f_m\,(s,t|u) \cdot \nu(u)\,du. \tag{44}$$

This first-moment density does not sum up to one but to a value $N_\infty(t) = \lim_{s\to 0} N(s,t)$, which remains finite for all finite t. A sufficient condition is that the growths of the number of firms and of their sizes are not faster than exponential in time, in agreement with condition ($\imath\imath$) in proposition 1. Many faster-than-exponential growth processes of the number of firms and of their sizes are also permitted, as long as they do not lead to finite-time singularities.

Lemma 2. *Under the assumptions 1, 2 and 5, the first-moment density of sizes of all the firms existing at the current time t reads*

$$g(s,t) = \int_{t_0}^{t} \nu(u) e^{-h\cdot(t-u)} f(s,t|u) du\,, \qquad t > t_0\,, \tag{45}$$

where $t_0 (> t_*)$ is the starting time of the economy (with t_* given by (2)) and $f(s,t|u)$ is the probability density function of a firm's size at time t and born at time u.

Proof. Assumptions 1 and 2 are enough for lemma 1 to hold. Besides, by assumption 5, the exit rate of a firm is independent from its size so that $f_m(s,t|u) = e^{-h(t-u)} \cdot f(s,t|u)$, where $f(s,t|u)$ denotes the probability density function of a firm's size at time t and born at time u. $\square$

Lemmas 1 and 2 show that, in order to derive proposition 1, we just need to consider the law of a single firm's size, given that it has not yet crossed the level $s_{\min}(t)$. The density of a single firm's size, that is solution to equation (1) embodying Gibrat's law, for a firm born at time $t_i = t - \theta_i$ and given the condition that the firm's size $S_i(t,\theta_i)$ is larger than $s_{\min}(t)$, $\forall \theta_i \geq 0$, is given by the following result.

Lemma 3. *Under the assumptions 2, 3 and 4, the probability density function $f\,(s,t|t-\theta,\tilde{s}_0 = s_0)$ of a firm's size at time t and aged θ conditional on $\tilde{s}_0 = s_0$, taking into account the condition that the firm would die if its size would reach the exit level $s_{\min}(t)$, is*

$$f\,(s,t|t-\theta,\tilde{s}_0 = s_0) = \frac{1}{2\sqrt{\pi\tau}s}\left[\exp\left(-\frac{1}{4\tau}\left(\ln\left(\frac{s}{s_{\min}(t)}\right) - \ln\left(\frac{s_0(t)}{s_{\min}(t)}\right) - (\delta - 1 - \delta_0)\tau\right)^2\right) - \left(\frac{s_0(t)}{s_{\min}(t)}\right)^{-(\delta-1-\delta_0)}\left(\frac{s}{s_{\min}(t)}\right)^{\delta_0-\delta_1}\exp\left(-\frac{1}{4\tau}\left(\ln\left(\frac{s}{s_{\min}(t)}\right) + \ln\left(\frac{s_0(t)}{s_{\min}(t)}\right) - (\delta - 1 - \delta_0)\tau\right)^2\right)\right]\,, \tag{46}$$

where

$$s_0(t) := s_0 e^{c_0 \cdot t}\,, \qquad \tau := \frac{\sigma^2}{2}\,\theta\,, \qquad \delta_0 := \frac{2c_0}{\sigma^2}\,, \qquad \delta_1 := \frac{2c_1}{\sigma^2}\,. \tag{47}$$

Proof. Let us consider a firm born at time $u = t - \theta$, where t denotes the current time and $\theta \geq 0$ is the age of the firm. The firm's size $S(\theta, u)$ is given by the following stochastic process

$$S(\theta, u) = s_0(u) e^{c \cdot \theta + \sigma W(\theta)} , \tag{48}$$

where $\theta = t - u$, $W(\theta)$ is a standard Wiener process, while $s_0(u)$ is the initial size of the firm, given $\tilde{s}_0 = s_0$, and $c := \mu - \frac{\sigma^2}{2}$. The process (48) with the initial and boundary conditions in assumptions 2 and 4 can be reformulated as

$$S(\theta, u) = s_{\min}(u + \theta) e^{\mathcal{Z}(\theta, u)} , \tag{49}$$

where

$$\mathcal{Z}(\theta, u) = \ln \rho(u + \theta) + (c - c_1)\theta + \sigma W(\theta), \qquad \rho(t) := \frac{s_0(t)}{s_{\min}(t)} . \tag{50}$$

As a consequence,

$$f\left(s, t | u, \tilde{s}_0 = s_0\right) = \frac{1}{s} \varphi \left[\ln \left(\frac{s}{s_{\min}(t)} \right), \theta; u \right] , \tag{51}$$

where $\varphi(z; \theta, u)$ denotes the density of $\mathcal{Z}(\theta, u)$ which is solution to

$$\begin{gathered} \frac{\partial \varphi(z; \theta, u)}{\partial \theta} + (c - c_1) \frac{\partial \varphi(z; \theta, u)}{\partial z} = \frac{\sigma^2}{2} \frac{\partial^2 \varphi(z; \theta, u)}{\partial z^2} , \\ \varphi(z; \theta = 0, u) = \delta(z - \ln \rho(u)) , \\ \varphi(z = 0; \theta, u) = 0 , \qquad \theta > 0 . \end{gathered} \tag{52}$$

These initial and boundary conditions are equivalent to the initial and boundary conditions in assumptions 2 and 4. Using any textbook on stochastic processes (Redner, 2001, for instance), we get

$$\begin{gathered} \varphi(z; \theta, u) = \frac{1}{2\sqrt{\pi \tau}} \exp\left(-\frac{(z - \ln \rho(u) - (\delta - 1 - \delta_1)\tau)^2}{4\tau} \right) - \\ \frac{[\rho(u)]^{\delta_1 - \delta + 1}}{2\sqrt{\pi\tau}} \exp\left(-\frac{(z + \ln \rho(u) - (\delta - 1 - \delta_1)\tau)^2}{4\tau} \right) , \end{gathered} \tag{53}$$

where δ is defined in (6) and τ in (47). Taking into account the relation

$$\rho(u) = \rho(t) e^{(\delta_1 - \delta_0)\tau} , \tag{54}$$

we rewrite expression (53) as

$$\begin{gathered} \varphi(z; \theta, u) = \frac{1}{2\sqrt{\pi \tau}} \exp\left(-\frac{(z - \ln \rho(t) - (\delta - 1 - \delta_0)\tau)^2}{4\tau} \right) - \\ \frac{[\rho(t)]^{\delta_0 - \delta + 1}}{2\sqrt{\pi\tau}} \exp\left(-\frac{(z + \ln \rho(t) - (\delta - 1 - \delta_0)\tau)^2}{4\tau} + (\delta_0 - \delta_1) z \right) , \end{gathered} \tag{55}$$

By substitution in (51), this concludes the proof of Lemma 3. □

Performing the change of variable from birthdate u to age $\theta = t - u$ in (45), and accounting for assumption 1, i.e. the fact that $\nu(t) = \nu_0 \, e^{d \cdot t}$, leads to

$$g(s, t) = \nu(t) \int_0^{\theta_0} e^{-(d+h)\theta} \mathrm{E} \left[f\left(s, t | t - \theta, \tilde{s}_0\right) \right] d\theta , \tag{56}$$

where $\theta_0 = t - t_0$ is the age of the given economy. $\mathrm{E}\left[f\left(s; t, |t-\theta, \tilde{s}_0\right)\right]$ denotes the statistical average of $f\left(s; t, \theta|\tilde{s}_0\right)$ over the random variable $\tilde{s}_0$. Inasmuch as t_0 should not be smaller than t_* given by (2), we should thus have $\theta_0 < \theta_* := \frac{\ln \rho(t)}{c_0 - c_1}$.

As a byproduct, the mean density of firm sizes, conditional on $\tilde{s}_0 = s_0$ is

$$g\left(s, t|\tilde{s}_0 = s_0\right) = \nu(t) \int_0^{\theta_0} e^{-(d+h)\theta} f\left(s, t|t-\theta, \tilde{s}_0 = s_0\right) d\theta \ . \tag{57}$$

Thus, substituting (46) into (57) yields

$$g(s, t|\tilde{s}_0 = s_0) = \frac{\tilde{\nu}(t)}{s} G\left(\ln\left(\frac{s}{s_{\min}(t)}\right); t, \tau_0\right) \ , \qquad \tilde{\nu}(t) = \frac{2\nu(t)}{\sigma^2} \ , \tag{58}$$

with

$$G(z; t, \tau_0) := \int_0^{\tau_0} e^{-\eta\tau} \varphi(z; t, \tau) d\tau \ , \tag{59}$$

where $\varphi(z; t, \theta)$ is given by (53) while

$$\tau_0 := \frac{\sigma^2}{2}\theta_0 \qquad (\tau_0 < \tau_*) \ , \qquad \eta := \frac{2}{\sigma^2}(d+h) \ . \tag{60}$$

The substitution of $\varphi(z; t, \theta)$ from (53) into the integral (59) leads to two integrals, which can be reduced to

$$\mathcal{I}(z, \theta, \alpha, \beta) := \int_0^{\theta} \exp\left(-\frac{(z-\alpha\tau)^2}{4\tau} - \beta\tau\right) \frac{d\tau}{2\sqrt{\pi\tau}} \ , \tag{61}$$

whose expression can be obtained by the tabulated integral (7.4.33) in Abramowitz and Stegun (1965) by the change of variable $u = \sqrt{\tau}$. This leads to

$$\begin{aligned} G(z; t, \tau_0) = \frac{1}{2\alpha(\eta)} \times \\ \Bigg\{ e^{\frac{1}{2}(\alpha z_- - \alpha(\eta)|z_-|)} \mathrm{erfc}\left(\frac{|z_-| - \tau_0\alpha(\eta)}{2\sqrt{\tau_0}}\right) - e^{\frac{1}{2}(\alpha z_- + \alpha(\eta)|z_-|)} \mathrm{erfc}\left(\frac{|z_-| + \tau_0\alpha(\eta)}{2\sqrt{\tau_0}}\right) - \\ \rho(t)^{-\alpha}\Big[e^{\frac{1}{2}(\alpha z_+ - \alpha(\eta)|z_+|)} \mathrm{erfc}\left(\frac{|z_+| - \tau_0\alpha(\eta)}{2\sqrt{\tau_0}}\right) - e^{\frac{1}{2}(\alpha z_+ + \alpha(\eta)|z_+|)} \mathrm{erfc}\left(\frac{|z_+| + \tau_0\alpha(\eta)}{2\sqrt{\tau_0}}\right)\Big]\Bigg\} \ , \end{aligned} \tag{62}$$

with

$$\alpha := \delta - 1 - \delta_0 \ , \qquad \alpha(\eta) := \sqrt{\alpha^2 + 4\eta} \ , \qquad z_- := \ln\frac{s}{s_0(t)} \ , \qquad z_+ := \ln\frac{s \cdot s_0(t)}{s_{\min}(t)^2} \ . \tag{63}$$

For an old enough economy, i.e., when $\sqrt{\tau_0} \gg 1/\alpha(\eta)$, we can expand expression (62) to obtain

$$G_\infty(z; t) = \frac{1}{\alpha(\eta)}\left[e^{\frac{1}{2}(\alpha z_- - \alpha(\eta)|z_-|)} - \rho(t)^{-\alpha} e^{\frac{1}{2}(\alpha z_+ - \alpha(\eta)|z_+|)}\right] \ . \tag{64}$$

Substituting this last expression into equation (58) for the mean density of firms sizes, and after making explicit the s-dependence of the variable z, we finally get

$$g(s, t|\tilde{s}_0 = s_0) = \frac{\tilde{\nu}(t)}{s\alpha(\eta)} \begin{cases} \left(\frac{s}{s_0(t)}\right)^{\frac{1}{2}(\alpha - \alpha(\eta))}\left(1 - \left(\frac{s_0(t)}{s_{\min}(t)}\right)^{-\alpha(\eta)}\right) \ , & s > s_0(t) \ , \\ \left(\frac{s}{s_0(t)}\right)^{\frac{1}{2}(\alpha + \alpha(\eta))} - \left(\frac{s_0(t)}{s_{\min}(t)}\right)^{-\alpha(\eta)}\left(\frac{s}{s_0(t)}\right)^{\frac{1}{2}(\alpha - \alpha(\eta))} \ , & s_0(t) > s > s_{\min}(t) \ . \end{cases} \tag{65}$$

for large $\tau_0 \gg \alpha(\eta)^{-1}$, with $s_0(t) = s_0 e^{c_0 \cdot t}$, as defined by (47).

According to assumption 2, the expectation of $g(s,t|\tilde{s}_0)$ with respect to $\tilde{s}_0$ provides us with the unconditional mean density of firm sizes

$$g(s,t) \approx \frac{\tilde{\nu}(t)}{s\alpha(\eta)} \cdot \left(\frac{\mathrm{E}\left[\tilde{s}_0^m\right]^{1/m} e^{c_0 \cdot t}}{s} \right)^m , \qquad \text{as } s \to \infty \text{ and } t \to \infty \; . \tag{66}$$

This expression justifies the statement of proposition 1 and concludes the proof.

A.2 Growth rate of the overall economy: Proof of proposition 2

Using the same machinery as in appendix A.1, we define the total size of the economy at time t as

$$\begin{aligned} \tilde{\Omega}(t) &:= \sum_{i \in \mathbb{N}:\, t_i \le t} S_i(t) \\ &= \int_{t_0}^{t} s \cdot N_t(du \times ds) \; . \end{aligned} \tag{67}$$

Under the assumptions of proposition 1, by theorem 6.4.V.iii in Daley and Vere-Jones (2007), we get

$$\begin{aligned} \Omega(t) &:= \mathrm{E}\left[\tilde{\Omega}(t)\right] \\ &= \nu(t) \int_0^{\tau_0} e^{-\eta \cdot \tau} \mathrm{E}\left[S(t,\tau)\right] d\tau \; . \end{aligned} \tag{68}$$

For simplicity, let us consider the case where $s_{\min} = 0$. This assumption is not necessary, but greatly simplifies the calculation. Under this assumption, the size of an incumbent firm follows a geometric Brownian motion so that

$$\mathrm{E}\left[S(t,\tau)\right] = s_0(t) e^{(\delta - \delta_0)\tau} \; , \tag{69}$$

where δ and δ_0 are defined by (6) and $s_0(t)$ by (47). Substituting (69) into (68) gives

$$\Omega(t) = \nu(t) \cdot s_0(t) \int_0^t e^{(\mu - c_0 - h - d)u} du \; , \tag{70}$$

$$= \int_0^t e^{(\mu - h)\cdot(t-u)} \nu(u) \cdot s_0(u) du \; , \tag{71}$$

$$= \int_0^t e^{(\mu - h)\cdot(t-u)} dI(u) \; . \tag{72}$$

This last equation shows that $\mu - h$ is the return on investment of the economy. By integration, we get the limit growth rate of the economy

$$\lim_{t \to \infty} \frac{d \ln \Omega(t)}{dt} = \begin{cases} \mu - h \; , & \mu - h > d + c_0 \\ d + c_0 \; , & \mu - h \le d + c_0 \end{cases} \; . \tag{73}$$

This concludes the proof of proposition 2 when $s_{\min} = 0$. When $s_{\min} \neq 0$ and grows at the rate $c_1 \ge 0$, the result can still be proved along the same lines but at the price of more tedious calculations since the expectation in (69) involves eight error functions.

A.3 Representativeness of the mean-distribution of firm size: Proof of proposition 3

The proof of proposition 3 follows from lemma 6.4.VI in Daley and Vere-Jones (2007) which states that a marked point process that has mark kernel $F_m(s,t|u)$, and for which the Poisson ground process has intensity measure $\nu(u)du$, is equivalent to a Poisson process on the product space with intensity measure $\Lambda(du \times ds) = F_m(ds,t|u) \cdot \nu(u)du$.

Thus, under the assumptions 1 and 2, using the notations of appendix A.1, the marked point process $\{t_i, S_i(t)\}_{i\in\mathbb{N}}$ is a compound Poisson process with intensity measure $\Lambda(du \times ds)$. Consequently

$$\Pr\left[\tilde{N}(s,t) = n\right] = \Pr\left[N_t\left([t_0,t) \times [s,\infty)\right) = n\right] \;, \tag{74}$$

$$= \frac{\left(\int_{[t_0,t)\times[s,\infty)} F_m(ds,t|u) \cdot \nu(u)du\right)^n}{n!} \exp\left(-\int_{[t_0,t)\times[s,\infty)} F_m(ds,t|u) \cdot \nu(u)du\right) \tag{75}$$

$$= \frac{N(t,s)^n}{n!} \cdot e^{-N(s,t)}, \tag{76}$$

by lemma 1.

References

Abramowitz, M. and I. Stegun (1965) *Handbook of Mathematical Functions*, Dover.

Acs, Z.J., R. Mork and B. Yeung (1999) Productivity Growth and Firm Size Distribution, in Acs, Carlsson and Karlsson (eds), *Entrepreneurship, Small & Medium-Sized Enterprises and the Macroeconomy*, Cambdrige University Press.

Agarwal, R. and D.B. Audretsch (2001) Does Entry Size Matter? The Impact of the Life Cycle and Technology on Firm Survival, *Journal of Industrial Economics* 49, 21-43.

Axtell, R.L. (2001) Zipf Distribution of U.S. Firm Sizes, *Science* 293, 1818-1820.

Baily, M.N., R.E. Litan, and M.S. Johnson (2008) The origin of the financial crisis, Fixing Finance Series, Paper 3, Initiative on Business and Public Policy at Brookings, November, 4-47.

Bartelsman, E., Scarpetta, S., Schivardi, F. (2005) Comparative Analysis of Firm Demographics and Survival: Micro-level Evidence for the OECD Countries, *Industrial and Corporate Changes* 14, 365-391.

Blank, A. and S. Solomon (2000) Power laws in cities population, financial markets and internet sites (scaling in systems with a variable number of components), Physica A 287 (1-2), 279-288.

Bonaccorsi Di Patti, E. and G. Dell'Ariccia (2004) Bank competition and firm creation. *Journal of Money Credit and Banking* 36, 225-251.

Brixy, U. and R. Grotz (2007) Regional patterns and determinants of birth and survival of new firms in Western Germany, *Entrepreneurship & Regional Development* 19 (4), 293-312.

Bruderl, J., P. Preisendorfer, and R. Ziegler (1992) Survival Chances of Newly Founded Business Organizations, *American Sociological Review* 57, 227-242.

Cabral, L. M. B. and J. Mata (2003) On the Evolution of the Firm Size Distribution: Facts and Theory, *American Economic Review* 93, 1075-1090.

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

Daley, D.J. and D. Vere-Jones (2007) An Introduction to the Theory of Point Processes (Springer).

Davis, S.J., J. Haltiwanger and S. Schu (1996) Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts, *Small Business Economics* 8, 297-315.

de Wit, G. (2005) Firm size distributions, An overview of steady-state distributions resulting from firm dynamics models, *International Journal of Industrial Organization* 23, 423-450.

Dunne, T., Roberts, M.J. and Samuelson, L. (1988) Patterns of Firm Entry and Exit in U.S. Manufacturing Industries, *RAND Journal of Economics* 19 (4), 495-515.

Dunne, T., Roberts, M.J. and Samuelson, L. (1989) The Growth and Failure of U.S. Manufacturing Plants, *The Quarterly Journal of Economics* 104 (4), 671-698.

Eeckhout, J. (2004) Gibrat's Law for (all) Cities, *American Economic Review* 94, 1429-1451.

Eckhout, J. (2009) Gibrat's law for (all) cities: Reply, *American Economic Review* 99, 1676-1683.

Eurostat (1998) Enterprises in Europe, Data 1994-95, Fifth Report European Commission.

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

Galor, O. (2005) From stagnation to growth: Unified growth theory. In P. Aghion & S. Durlauf (Eds.), Handbook of Economic Growth (pp. 171-293), Elsevier.

Gertler, M. and S. Gilchrist (1994) Monetary Policy, Business Cycles, and the Behavior of Small manufacturing Firms, *Quarterly Journal of Economics* 109, 309-340.

Geroski, P. (2000) The Growth of Firms in Theory and in Practice, in N. Foss and V. Malinke (eds.), *New Directions in Economic Strategy Research* (Oxford University Press).

Gibrat, R. (1931) *Les Inégalités Economiques; Applications aux inégalitiés des richesses, à la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc., d'une loi nouvelle, la loi de l'effet proportionnel* (Librarie du Recueil Sirey, Paris).

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

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

Knaup, A.E. (2005) Survival and longevity in the Business Employment Dynamics data, *Monthly Labor Review*, May, 50-56.

Krugman, P. (1996) *The Self-Organizing Economy* (Blackwell).

Levy, M. (2009) Gibrat's Law for (all) Cities: Comment, *American Economic Review* 99, 1672-1675

Lucas, R. E. (1978) On the size distribution of business firms, *Bell Journal of Economics* 9, 508-523.

Luttmer, E.G.J. (2007) Selection, Growth, and the Size Distribution of Firms, *Quarterly Journal of Economics* 122, 1103-1144.

Luttmer, E.G.J. (2008) On the Mechanics of Firm Growth, *Federal Reserve Bank of Minneapolis – Research Department*, Working Paper 657.

Malevergne, Y., V. Pisarenko and D. Sornette (2009) Gibrat's law for cities: uniformly most powerful unbiased test of the Pareto against the lognormal, `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1479481`)

Pagano, P. and F. Schivardi (2003) Firm Size Distribution and Growth, *The Scandinavian Journal of Economics* 105, 255-274.

Peretto P.F. (1999) Firm Size, Rivalry and the Extent of the Market in Endogenous Technological Change, *European Economic Review* 43, 1747-1773.

Redner, S. (2001) *A Guide to First-Passage Processes*, Cambridge University Press.

Reynolds, P.D., D.J. Storey and P. Weshead (1994) Cross-national comparisons of the variation in new firm formation rates. *Regional Studies* 28, 443-456.

Rossi-Hansberg, E. and M. L. J. Wright (2007a) Establishment Size Dynamics in the Aggregate Economy, *American Economic Review* 97, 1639-1666.

Rossi-Hansberg, E. and M. L. J. Wright (2007b) Urban Structure and Growth, *Review of Economic Studies* 74, 597 - 624.

Saichev, A., Y. Malevergne and D. Sornette (2009) Zipf's law and beyond, *Lecture Notes in Economics and Mathematical Systems* 632 (Springer).

Schumpeter, J. (1934) *Theory of Economic Development* (Harvard University Press, Cambridge).

Simon, H. A. (1955) On a class of skew distribution functions, *Biometrika* 42, 425-440.

Simon, H. A. (1960) Some further notes on a class of skew distribution functions, *Information and Control* 3, 80-88.

Buldyrev, S. V., H. Leschhorn, P. Maass, H. E. Stanley, M. H. R. Stanley, L. A. N. Amaral, S. Havlin, and M. A. Salinger (1997) Scaling Behavior in Economics: Empirical Results and Modeling of Company Growth, in The Physics of Complex Systems, edited by F. Mallamace and H. E. Stanley (IOS Press, Amsterdam), pp. 145-174.

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

Sutton, J. (1997) Gibrat's legacy, *Journal of Economic Literature* 35, 40-59.

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