Log-growth distributions of US city sizes and non-Lévy processes

Ramos, Arturo

Universidad de Zaragoza

9 September 2015

Log-growth distributions of US city sizes and non-Lévy processes

Arturo Ramos*

September 9, 2015

Abstract

We study whether the hypothesis that the log-population of US cities follows a Lévy process can be rejected or not. The result seems to be rejection.

As a consequence, the cited process seems not to be described by a standard Brownian motion with drift (with a Yule process), thus explaining in another way the rejection of the lognormal and double Pareto lognormal distributions for US city size in recent studies. The datasets employed are that of US incorporated places on the period 1890-2000.

However, we recall a way of obtaining a family of stochastic Itô differential equations whose sample paths are associated to the time-dependent probability density functions for city size that in principle could be observed empirically.

JEL: C46, R11, R12.

Keywords: Lévy process, Brownian motion with drift, Yule process, stochastic Itô differential equation, US city size

*Department of Economic Analysis, Universidad de Zaragoza (SPAIN) aramos@unizar.es. I thank constructive comments from Rafael González-Val, José Olmo and Fernando Sanz-Gracia on a previous version of the manuscript, although all remaining errors are mine. This work is supported by the project ECO2013-45969-P of the Spanish Ministry of Economy and Competitiveness and by the Aragon Government, ADETRE Consolidated Group.

1 Introduction

There is an ample amount of work concerning Zipf's Law and Gribrat's Law in the field of Urban Economics. Two of the main references are Gabaix (1999, 2009), where the author finds an explanation for Zipf's Law assuming that the US urban units follow a geometric Brownian motion process with a lower barrier and a Poisson process of city creation. On its side, Eeckhout (2004) proposes the lognormal distribution for describing US city size, and the generation of this distribution is based on the standard multiplicative Gibrat's process, which is another way of considering a geometric Brownian motion. In the firm size distribution literature, Sutton (1997) and Delli Gatti et al. (2005) postulate that if Gibrat's Law holds, the resulting log-size distribution will be normal, and that the log-growth rates are expected to follow a normal distribution as well.[1] Almost simultaneously, a quite remarkable density function has been proposed for city size (Reed, 2002, 2003; Reed and Jorgensen, 2004), later embraced by Giesen et al. (2010); Giesen and Suedekum (2012, 2014), namely the double Pareto lognormal (dPln). This last distribution can be generated by a variation of the geometric Brownian motion, adding the effects of city age to yield the associated Yule process. Thus, until the year 2015 the dPln offered the best fit for a number of countries in the literature (Giesen et al., 2010; González-Val et al., 2015).

However, the recent article Ramos and Sanz-Gracia (2015) has proposed new parametric models for which the tails are essentially Pareto, and the body is Generalized Beta 2, the tails and the body delineated at two exact population thresholds. These distributions are based on economic models and not so heavily on pure statistical reasoning, and they offer overwhelmingly better fits than the lognormal and dPln in the sense that they are not rejected by standard Kolmogorov–Smirnov (KS) and Crámer–

[1] However, work by Stanley et al. (1996); Amaral et al. (1997) shows that the log-growth rate distribution of firm sizes is described better by a Laplace distribution. See also Toda (2012) for something similar regarding the income distribution. More recently, Ramos (2015) has shown a new parametric density function for US city log-growth rates that is not empirically rejected by the Kolmogorov–Smirnov (KS), Crámer–von Mises (CM) and Anderson–Darling (AD) tests.

von Mises (CM) tests, contrary to the other cases. Apparently, their derivation has no relation with stochastic processes associated to them so the arguments of Gabaix (1999, 2009) may not apply at first sight. Since the observed distributions are clearly not lognormal nor dPln, then the hypothesis of geometric Brownian motion (with a Yule process) may not apply in practice.

The aim of this paper is to reconcile the non appearance of a Lévy process (a generalization of the previously mentioned processes) with the possibility of constructing an associated Itô differential equation by investigating whether the log-city size distribution follows a Lévy process. If it is not the case, then in particular the processes of the lognormal and dPln may not occur.

In Ramos and Sanz-Gracia (2015) it has been checked already that the lognormal and the dPln are rejected for the US city size distribution, but it is our aim to relate these facts to the study of the associated random growth (Lévy processes) and Itô stochastic differential equations to see to what extent the underlying ideas of Gabaix (1999, 2009) can still be preserved.

Even in the case of the log-population process be non-Lévy, we will explore more about the relation of stochastic processes and the associated density functions to show that for any given time-dependent density function it is possible to find a stochastic Itô differential equation describing a process associated to the former.

The rest of the paper is organized as follows. Section 2 reviews Gibrat's process and Lévy processes. Section 3 describes the databases used. Section 4 studies the stationarity and independence of the log-growth rates for the US log-population of incorporated places. Section 5 describes the theoretical procedure of constructing a family of Itô stochastic differential equations associated to any prescribed time-dependent probability density function. Finally, Section 6 offers a discussion and conclusions.

2 Gibrat's process and Lévy processes

Gibrat's process for cities can be understood as follows (we base our development mainly in Sutton (1997) and references therein, Eeckhout (2004) and Delli Gatti et al. (2005)). Let $x_{i,t}$ be the population of city i at time t, and $g_{i,t} = \ln x_{i,t} - \ln x_{i,t-1}$ the log-growth rate of city i between times $t-1$ and t. From the relation

$$\ln x_{i,t} = \ln x_{i,t-1} + g_{i,t}$$

and assuming that t is an integer number, we can iterate the former and arrive to

$$\begin{aligned} \ln x_{i,t} &= \ln x_{i,t-2} + g_{i,t-1} + g_{i,t} \\ &= \ln x_{i,0} + g_{i,1} + \cdots + g_{i,t-1} + g_{i,t} \end{aligned}$$

Then, if the log-growth rates or increments $g_{i,t}$ are independent variables with mean m and variance σ^2 *for all* i, t,[2] by the Central Limit Theorem (see, e.g., Feller (1968)) we have that as $t \to \infty$ the quantity $\ln x_{i,t} - \ln x_{i,0}$ will follow a normal distribution with mean mt and variance $\sigma^2 t$.[3]

In contrast, we have empirically obtained in González-Val et al. (2015) and Ramos and Sanz-Gracia (2015) that the lognormal specification for US city size distribution is strongly rejected by the Kolmogorov–Smirnov (KS) and Crámer–von Mises (CM) tests. In the second of these references, we obtain an excellent model that is non-rejected by the same tests and is the so-called "threshold double Pareto Generalized Beta 2" (tdPGB2) for incorporated places.[4] Thus the key assumption in obtaining the normal distribution for the log-populations in the previous paragraph, namely that the

[2] And therefore the increments are clearly *stationary* and *independent*, see below for a rigourous definition.

[3] Kalecki (1945) modifies this derivation so as to obtain a lognormal distribution for the size with constant variance, by allowing a negative correlation between the log-growth rates and log-size.

[4] See Section 3 for an explanation of the urban units and datasets used in this paper.

increments $g_{i,t}$ are stationary and independent, deserves a reconsideration.[5]

Thus it is one of our main interests in this paper to study empirically, in the most general standard framework, the question of whether the previous log-growth rates $g_{i,t}$ are stationary and independent, based on our relatively ample database. There exists a well established theory of the stochastic processes with stationary and independent increments, also known as Lévy processes. For this topic, we will mainly follow Kyprianou (2006), see also Sato (1999) and Lukacs (1970). We simply recall the definition of this kind of processes, to be used below:

Definition 1 (Lévy process)**.** *A process $Y = \{Y_t : t \geq 0\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$[6] is said to be a Lévy process if it possesses the following properties:*

(i) The paths of Y are $\mathbb{P}$-almost surely right continuous with left limits.

(ii) $\mathbb{P}(Y_0 = 0) = 1$.

(iii) For $0 \leq s \leq t$, $Y_t - Y_s$ is equal in distribution to Y_{t-s}.

(iv) For $0 \leq s \leq t$, $Y_t - Y_s$ is independent of $\{Y_u : u \leq s\}$.

It can be shown, see again Kyprianou (2006) and Sato (1999), that variables that follow Lévy processes can be associated to probability laws that are *infinitely divisible* and reciprocally. Particular and paradigmatic cases of Lévy processes are Poisson processes and the standard Brownian motion (with drift). Also, the standard Brownian motion with drift (with a Yule process) that can be used to generate the asymmetric double Laplace-normal for the log-population (double Pareto lognormal for the population) (Reed, 2002, 2003; Reed and Jorgensen, 2004) is a Lévy process since the characteristic function of the distribution of the log-population $y = \ln x$ in this case

[5] One could argue that for US cities the current t is not big enough to give sense to the previous limit. The convergence is known to be of the order $O(t^{-1/2})$ (see, e.g, Feller (1968)) and we will assume that the limiting distribution should have approximately been reached already.

[6] Ω denotes the sample space, i.e. the set of all possible outcomes, $\mathcal{F}$ denotes the σ-algebra of the set of events, and $\mathbb{P}$ is a function from events to probabilities.

takes the form

$$\phi_y(\theta) = \exp\left(iA_0\theta - \frac{1}{2}B_0^2\theta^2\right)\frac{1}{(1 - i\theta/\alpha)\,(1 + i\theta/\beta)}$$

where A_0, B_0, α, β are real constants (and here, i is the imaginary unit). In fact, this characteristic function is the product of the characteristic functions of a normal distribution and of two Gamma distributions, each of them being infinitely divisible. According to Theorem 5.3.2 in Lukacs (1970) the product is infinitely divisible as well and the underlying process of the asymmetric double Laplace-normal distribution (double Pareto lognormal distribution for the population) can be associated to a Lévy process. This straightforward result shows the relation between Lévy processes and the double Pareto lognormal distribution.

It is our interest here to test whether the hypothesis that the log-population follows a Lévy process can be rejected or not. Once the log-growth rates are computed, we can proceed to test whether the conditions for a Lévy process in Definition 1 hold.

The first condition cannot be checked by means of our empirical data, since it comprises only cross-sections on discrete time and continuity would require the knowledge of population at *all* times. The second condition states that the initial population of all cities is one (the log-population is zero) with probability one. Assuming that all cities start their existence having the same population the previous requirement could be met by an appropriate shift of the population values. It is however difficult, if not impossible, to check empirically if the assumption holds. The third and fourth requirements are thus the essential ones to be tested. The third expresses the stationarity of the log-growth rates and the fourth the independence of the log-growth rates on past values of the log-population (in particular, this includes Gibrat's Law for cities).

3 The databases

We have used in this article data about US urban centers from the decennial data of the US Census Bureau of "incorporated places" without any size restriction, for the period 1890-2000. These include governmental units classified under state laws as cities, towns, boroughs or villages. Alaska, Hawaii and Puerto Rico have not been considered due to data limitations. The data have been collected from the original documents of the annual census published by the US Census Bureau.[7] These data sets were first introduced in González-Val (2010), see therein for details, and later used in other works like González-Val et al. (2013, 2015); Ramos and Sanz-Gracia (2015).[8]

[Table 1 near here]

We offer in Table 1 the descriptive statistics of the used data for the US.

4 Non-stationarity and dependence of the log-growth rates of US incorporated places

In this Section we analyze whether the decennial log-growth rates for the data of US incorporated places are equal in distribution, namely whether requirement (iii) in Definition 1 holds for the process followed by $Y_t = \ln x_t$.

Also, we will comment about the dependence of log-growth rates on the initial sizes so that requirement (iv) in Definition 1 may not occur.

[7] `http://www.census.gov/prod/www/decennial.html` Last accessed: September 8^{th}, 2015.

[8] We have not used the dataset of all US urban places, unincorporated and incorporated, and without size restrictions, also provided by the US Census Bureau for the years 2000 and 2010, because of consistency of the definition of the urban units. This dataset for the year 2000 was first used in Eeckhout (2004) and later in Levy (2009), Eeckhout (2009), Giesen et al. (2010), Ioannides and Skouras (2013) and Giesen and Suedekum (2014). The two samples were also used in González-Val et al. (2015); Ramos and Sanz-Gracia (2015). Likewise, the datasets of "City Clustering Algorithm" (CCA) (Rozenfeld et al., 2008, 2011) have not been considered because their temporal span is short (1991-2000) in order to consider a long-term perspective.

If the cited condition (iii) holds, it should happen that

$$\ln x_{i,t} - \ln x_{i,t-1} = g_{i,t}$$

is equal in distribution to $\ln x_{i,1}$ for all i. Thus, all $g_{i,t}$ should be equal in distribution for all t.

We have available eleven samples of decennial intervals. We will test whether the corresponding log-growth rates come from the same distribution.

For that, we simply perform the Kolmogorov–Smirnov (KS) and Crámer–von Mises (CM) tests to the empirical log-growth rates of each period compared to all other periods' samples. The null hypothesis in all cases is that the empirical log-growth distributions come from the same distribution.

[Table 3 near here]

[Table 4 near here]

The results of the KS test are shown in Table 3 and of the CM test in Table 4. In them, it is seen that the null is (strongly) rejected in *all* cases.[9]

Thus we have that the increments of the log-population of US incorporated places seem to be *not* stationary, and requirement (iii) in Definition 1 seems to be not fulfilled.

With regards to independence, we can resort to previously published work with the same database for US incorporated places, namely González-Val et al. (2013). It is shown in it that Gibrat's Law is rejected sometimes for the US incorporated places data and that also sometimes there exists a threshold value for the sample size above which Gibrat's Law is rejected, so therefore the independence of log-growth rates on previous values of log-sizes is rejected in a number of cases. Thus, requirement (iv) in Definition 1 seems to be not fulfilled always.

[9] We have checked that a similar result holds for the decennials in the period 1951-2011 for Italy and 1900-2010 for Spain, where the sample sizes are almost constant. These last results are available from the author upon request.

5 Construction of Itô stochastic differential equations

We have seen in the previous Section that the population process of US cities does not seem to qualify as a Lévy process, which is a generalization of the processes leading to the lognormal (Brownian motion) or the dPln (Brownian motion with a Yule process). These last two distributions are empirically rejected for US cities (Ramos and Sanz-Gracia, 2015) and instead, alternative new distributions are observed, starting from an economic model evolved from one of Parker (1999).

Thus, the framework of Lévy processes seems to be not general enough to cover the empirical processes occurring in the description of city sizes. But even in this case, it is still theoretically possible to associate the time-dependent observed city size distribution to an Itô stochastic differential equation, and in this way the fundamental idea of Gabaix (1999, 2009) of associating to the city size distribution a random growth, is preserved.

In this Section we will follow mainly Gardiner (2004) and references therein in the presentation of Itô stochastic differential equations. We think of the variable $y_t = \ln x_t$ where x_t is the population of our samples of cities. We establish a standard Itô stochastic differential equation in the form

$$dy_t = m(y_t, t)dt + \sqrt{2s(y_t, t)}\, dB_t \tag{1}$$

where $m(y_t, t)$ models the *drift* term, the $\sqrt{2s(y_t, t)}$ models the *diffusion* term and B_t is a standard Brownian motion (Wiener process) (see, e.g., Itô and McKean Jr. (1996); Kyprianou (2006) and references therein). This process can be associated to the *forward Kolmogorov equation* or *Fokker-Plank equation* for the time-dependent probability density function (conditional on the initial data) $f(y, t)$ (see also Gabaix

(1999, 2009); Toda (2012)):

$$\frac{\partial f(y,t)}{\partial t} = -\frac{\partial}{\partial y}\left(m(y,t)f(y,t)\right) + \frac{\partial^2}{\partial y^2}\left(s(y,t)f(y,t)\right) \tag{2}$$

Given arbitrary $m(y,t)$ and $s(y,t)$ (but subject to the regularity conditions of, e.g., Karatzas and Shreve (1991) and references therein), to solve (2) for $f(y,t)$ is in general a hard problem and several techniques have been developed to deal with it (see, e.g., Gardiner (2004) and references therein). But now we face the *inverse problem*, namely to find suitable functions $m(y,t)$ and $s(y,t)$ starting from a given time-dependent probability density function $f(y,t)$ such that (2) holds. This inverse problem is much easier and it is (formally) solved as follows (Dupire, 1993, 1994):

From (2) we can write[10]

$$\begin{aligned} & f(y,t)s_{yy}(y,t) + 2f_y(y,t)s_y(y,t) + f_{yy}(y,t)s(y,t) \\ & = f_t(y,t) + \left(m(y,t)f(y,t)\right)_y \end{aligned} \tag{3}$$

This differential equation, for a given $f(y,t)$ (and we suppose that also $m(y,t)$ is given), can be regarded as a ordinary linear second-order differential equation as there is no derivative of $s(y,t)$ with respect to t. It is moreover easily integrable, and the explicit general solution can be given as

$$s(y,t) = \frac{1}{f(y,t)}\left(c_1(t) + c_2(t)y + \int_{-\infty}^{y} m(z,t)f(z,t)\,dz + \frac{\partial}{\partial t}\int_{-\infty}^{y} \mathrm{cdf}(z,t)\,dz\right) \tag{4}$$

as can be easily checked, where $c_1(t), c_2(t)$ are arbitrary functions of the variable t, and $\mathrm{cdf}(y,t) = \displaystyle\int_{-\infty}^{y} f(z,t)\,dz$. Inserting this expression into (1) yields a Itô stochastic differential equation which describes a process that has as associated $f(y,t)$ the one

[10] We will denote as usual $s_y(y,t) = \dfrac{\partial s(y,t)}{\partial y}$ and so on.

we have started from.[11]

Note that there appears two undetermined functions $c_1(t), c_2(t)$ of the variable t. In Dupire (1993) it is argued that these two arbitary functions should be zero if the function $f(y,t)$ has a finite expectation. We could consider also the case of distributions with undefined mean, like α-stable distributions with $\alpha \in (0,1]$. The case of $\alpha = 1$ is the Cauchy distribution (see, e.g, Zolotarev (1986); Uchaikin and Zolotarev (1999) and references therein).

The integral version of (1) is (see, e.g., Definition 5.2.1 of Karatzas and Shreve (1991))

$$y_t = y_0 + \int_0^t m(y_u,u)\,du + \int_0^t \sqrt{2s(y_u,u)}\,dB_u$$

so that the log-growth rates $g_{t,\delta} = y_t - y_{t-\delta}$, where $\delta \in [0,t]$ are given in terms of the stochastic process by

$$g_{t,\delta} = \int_{t-\delta}^t m(y_u,u)\,du + \int_{t-\delta}^t \sqrt{2s(y_u,u)}\,dB_u \tag{5}$$

These quantities may be statistically dependent on t and/or $y_u, u \in [t-\delta,t]$ so the conditions (iii) and/or (iv) of Definition 1, respectively, may not hold, and the process $\{y_t : t \geq 0\}$ may not qualify as a Lévy process by these reasons.

We provide an example next to show in an explicit way that even in the case of a normal distribution, the time-dependence might cause that the underlying process is not Lévy.

Example (Normal distribution)**.** *In order to see how the previous construction works*

[11] There exists the problem about whether the $f(y,t)$ is a solution unique or not of each of the processes just constructed. We do not worry about it here since for our purposes it is enough to have $f(y,t)$ as a solution, something which is obtained by construction. Also, the solution sample path of (1) so constructed there exists and is unique when additional conditions on $m(y,t)$ and $s(y,t)$ are imposed. It might be the case of having solutions of (1) in the weak sense if sample paths differ but associated to the same probability distribution. See, e.g., Arnold (1974), Karatzas and Shreve (1991) and Allen (2007) for details.

in a simple example, we consider the time-dependent normal distribution

$$f_{\mathrm{n}}(y, \mu(t), \sigma(t)) = \frac{1}{\sqrt{2\pi}\sigma(t)} \exp\left(-\frac{(y-\mu(t))^2}{2(\sigma(t))^2}\right)$$

The equation (3) can be written in this case (assuming that m depends only on t, and where all the dependences on t are not explicitly shown for the sake of brevity):

$$\begin{array}{l} s_{\mathrm{n}yy} - 2\dfrac{(y-\mu)}{\sigma^2} s_{\mathrm{n}y} + \dfrac{(y-\mu)^2 - \sigma^2}{\sigma^4} s_{\mathrm{n}} = \\ \quad \dfrac{1}{\sigma^3}\left((y-\mu)\sigma\mu' + ((y-\mu)^2 - \sigma^2)\sigma'\right) - m\dfrac{(y-\mu)}{\sigma^2} \end{array}$$

The corresponding $s_{\mathrm{n}}(y, \mu(t), \sigma(t))$ reads, using (4):

$$\begin{array}{l} s_{\mathrm{n}}(y, \mu(t), \sigma(t)) = \exp\left(\dfrac{(y-\mu(t))^2}{2(\sigma(t))^2}\right)\sqrt{\dfrac{\pi}{2}}\sigma(t) \\ \quad \times \left(2c_1(t) + 2c_2(t)y + \left(1 + \mathrm{erf}\left(\dfrac{y-\mu(t)}{\sqrt{2}\sigma(t)}\right)\right)(m(t) - \mu'(t))\right) \\ \quad + \sigma(t)\sigma'(t) \end{array}$$

where erf denotes the error function associated to the normal distribution. With the well-known choice (ordinary Gibrat's process reviewed in Section 2) $m(t) = m$, $\mu(t) = mt + \mu_0$ and $\sigma(t) = \sigma_0\sqrt{t}$, and moreover if $c_1(t) = c_2(t) = 0$ the previous expression reduces simply to

$$s_{\mathrm{n}}(y, \mu(t), \sigma(t)) = \frac{1}{2}\sigma_0^2 \tag{6}$$

Thus, (5) reduces to

$$\begin{aligned} g_{t,\delta} &= \int_{t-\delta}^{t} m\, du + \int_{t-\delta}^{t} \sigma_0\, dB_u \\ &\sim m\delta + \sigma_0 B_\delta \end{aligned}$$

($\sim$ *means here equality in distribution) and thus the* $g_{t,\delta}$ *are normally distributed with mean* $m\delta$ *and variance* $\delta\sigma_0^2$ *(see, e.g., Durrett (1996)). They are obviously stationary (they do not depend on* t*) and independent (they do not depend on* y*). The associated process can be taken as being Lévy, as it corresponds to the standard Gibrat's process reviewed in Section 2.*

However, with the choice $c_1(t) = c_2(t) = 0$*,* $m(t) = mt$*,* $\mu(t) = \frac{1}{2}mt^2 + \mu_0$ *and* $\sigma(t) = \sigma_0\sqrt{t}$ *we obtain as well the expression (6), but then the associated process has log-growth increments*

$$\begin{aligned} g_{t,\delta} &= \int_{t-\delta}^{t} mu\,du + \int_{t-\delta}^{t} \sigma_0\,dB_u \\ &\sim m\left(t\delta - \frac{\delta^2}{2}\right) + \sigma_0 B_\delta \end{aligned}$$

and thus the $g_{t,\delta}$ *are now normally distributed with mean* $m\left(t\delta - \frac{\delta^2}{2}\right)$ *and variance* $\delta\sigma_0^2$*. They are not stationary (they do depend on* t*) and independent (they do not depend on* y*) and thus it is not Lévy.*

Another choice could be $c_1(t) = c_2(t) = 0$*,* $\mu(t) = \mu_0$*,* $m(t) = \mu_0 t$ *and* $\sigma(t) = \sigma_0$*. Then, calling for brevity*

$$h(y) = \exp\left(\frac{(y-\mu_0)^2}{2\sigma_0}\right)\sqrt{\frac{\pi}{2}}\sigma_0\left(1 + \mathrm{erf}\left(\frac{y-\mu_0}{\sqrt{2}\sigma_0}\right)\right)$$

we have that

$$s_n(y, \mu(t), \sigma_0) = h(y)\mu_0 t$$

and then the log-growth increments take the form

$$\begin{aligned} g_{t,\delta} &= \int_{t-\delta}^{t} mu\,du + \int_{t-\delta}^{t} \sqrt{2h(y_u)\mu_0 u}\,dB_u \\ &= m\left(t\delta - \frac{\delta^2}{2}\right) + \int_{t-\delta}^{t} \sqrt{2h(y_u)\mu_0 u}\,dB_u \end{aligned}$$

We see that these increments depend explicitly on t (they are not stationary) and they are expected to depend on $y_u, u \in [t-\delta, t]$ by means of the stochastic integral; the process will be also non-Lévy.

6 Discussion and conclusions

We have analyzed whether the incorporated places' US city size follows a Lévy process, with the following results:

i) The log-growth rates or increments of the US log-population seem to be *strongly not* stationary. Thus condition (iii) in Definition 1 seems to be *not* satisfied in the case under study.

ii) The log-growth rates or increments of the US log-population are *not always* independent of initial log-sizes (rejection of Gibrat's Law) (González-Val et al., 2013).

The important consequence of this analysis is that the log-population process for US incorporated places seems to have not stationary nor independent (in general) increments, so it seems not to qualify as a Lévy process.

Therefore, the cited process seems not to be a standard Brownian motion with drift (eventually, with a Poisson process added[12]), something which is assumed in current theories of city growth (Gabaix, 1999, 2009). This last assumption has been introduced to give an explanation to Zipf's Law. Also, the cited process seems not to be a standard Brownian motion with drift and a Yule process like the one that can be used to generate the asymmetric double Laplace-normal for the log-population (double Pareto lognormal for the population) (Reed, 2002, 2003; Reed and Jorgensen, 2004) which, on its side, can be taken as a Lévy process as we have shown before.

[12]Poisson processes can be used to model entrant cities in the sample, see, e.g., Gabaix (1999, 2009).

The lognormal distribution arises from an economic model of an equilibrium theory of local externalities by Eeckhout (2004), which leads to the Gibrat's process reviewed in Section 2, being a Lévy process. In Ramos and Sanz-Gracia (2015) it is shown that the lognormal specification is always empirically rejected for US incorporated places.

The double Pareto lognormal arises also from an endogenous city creation into a dynamic economic model by Giesen and Suedekum (2014) in which is important the exponential distribution of entrant cities combined with a Gibrat's process like the one for the lognormal. The corresponding process can be taken as a Lévy one as we have shown before. Again, in Ramos and Sanz-Gracia (2015) it is shown that the dPln specification is almost always empirically rejected for US incorporated places.

In this paper we have shown another reason for the rejection of the lognormal and dPln density functions, namely that the population process of US incorporated places is non Lévy: the stationarity of the log-growth rates is strongly rejected and the independence is rejected sometimes (rejection of Gibrat's Law). This means that in particular, the log-population process is not a Brownian motion with or without a Yule process, giving a reason for the non-appearance of the lognormal nor the dPln in empirical terms. This seems to contradict in a first instance the results of Gabaix (1999, 2009) regarding the standard random growth of city sizes. But a closer look yields that the essential ideas of these last two articles remain.

That is, it has been shown in Ramos and Sanz-Gracia (2015) that a very appropriate statistical density function (”tdPGB2” for short) cannot be empirically rejected always, includes pure Pareto tails delineated by exact threshold values of the population, and that this model is derived exactly from a purely economic model in which population self-organizes in city sizes so as to maximize the net output of the overall system of cities in a country. People react therefore to the elasticities of the production function with respect to population, of the production function with respect to the number of cities of each value of population, and of the congestion costs with respect to the population variable, thus increasing or decreasing the number of cities of a given population. This self-organization is a newly observed behaviorial characteristic of the people of US incorporated places. Since the changes in the distribution depend (slowly) on time-dependent elasticities and those ultimately depend on ambient economic conditions, it is not likely that they change always in a time-invariant and size-independent manner, thus it is not to be expected that the population process to be Lévy.

But one of the main ideas of Gabaix (1999, 2009) is preserved, namely, that it is possible to construct *ex post* a family of stochastic Itô differential equations associated to the empirically observed time-dependent density functions for city size (Dupire, 1993, 1994).

From this point of view, the fundamental object is the city-size distribution, empirically observed and with solid economic grounds, and afterwards one can reconstruct a stochastic process, in general non-Lévy, with the associated density function the time-dependent observed one.

References

Allen, E. (2007). *Modeling with Itô stochastic differential equations*. Springer.

Amaral, L. A. N., Buldyrev, S., Havlin, S., Leschhorn, H., Maas, P., Salinger, M., Stanley, E., and Stanley, M. (1997). Scaling behavior in Economics: I. Empirical results for company growth. *Journal de Physique*, 7:621–633.

Arnold, L. (1974). *Stochastic differential equations: Theory and applications*. John Wiley & Sons.

Delli Gatti, D., Di Guilmi, C., Gaffeo, E., Giulioni, G., Gallegati, M., and Palestrini, A. (2005). A new approach to business fluctuations: heterogeneous interacting agents, scaling laws and financial fragility. *Journal of Economic Behavior & Organization*, 56:489–512.

Dupire, B. (1993). Pricing and hedging with smiles. In *Proceedings of AFFI Conference, La Baule, June*.

Dupire, B. (1994). Pricing with a smile. *Risk*, 7(1):271–275.

Durrett, R. (1996). *Stochastic Calculus. A practical introduction*. CRC Press.

Eeckhout, J. (2004). Gibrat's law for (all) cities. *American Economic Review*, 94(5):1429–1451.

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

Feller, W. (1968). *An introduction to probability theory and its applications (third edition). Volume I.* John Wiley & Sons.

Gabaix, X. (1999). Zipf's law for cities: An explanation. *Quarterly Journal of Economics*, 114:739–767.

Gabaix, X. (2009). Power laws in Economics and finance. *Annu. Rev. Econ.*, 2009:255–293.

Gardiner, C. W. (2004). *Handbook of stochastic methods for Physics, Chemistry and the Natural Sciences (third edition)*. Springer.

Giesen, K. and Suedekum, J. (2012). The French overall city size distribution. *Région et Développement*, 36:107–126.

Giesen, K. and Suedekum, J. (2014). City age and city size. *European Economic Review*, 71:193–208.

Giesen, K., Zimmermann, A., and Suedekum, J. (2010). The size distribution across all cities-double Pareto lognormal strikes. *Journal of Urban Economics*, 68(2):129–137.

González-Val, R. (2010). The evolution of US city size distribution from a long term perspective (1900–2000). *Journal of Regional Science*, 50:952–972.

González-Val, R., Lanaspa, L., and Sanz-Gracia, F. (2013). Gibrat's law for cities, growth regressions and sample size. *Economics Letters*, 118:367–369.

González-Val, R., Ramos, A., Sanz-Gracia, F., and Vera-Cabello, M. (2015). Size distribution for all cities: Which one is best? *Papers in Regional Science*, 94(1):177–197.

Ioannides, Y. M. and Skouras, S. (2013). US city size distribution: Robustly Pareto, but only in the tail. *Journal of Urban Economics*, 73:18–29.

Itô, K. and McKean Jr., H. P. (1996). *Diffusion processes and their sample paths (reprint of the 1974 edition)*. Springer.

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

Karatzas, I. and Shreve, S. E. (1991). *Brownian motion and stochastic calculus. Second edition*. Springer-Verlag.

Kyprianou, A. E. (2006). *Introductory lectures on fluctuations of Lévy processes with applications*. Springer-Verlag.

Levy, M. (2009). Gibrat's law for (all) cities: Comment. *American Economic Review*, 99:1672–1675.

Lukacs, E. (1970). *Characteristic functions*. Griffin.

Parker, S. C. (1999). The generalised beta as a model for the distribution of earnings. *Economics Letters*, 62:197–200.

Ramos, A. (2015). Are the log-growth rates of city sizes normally distributed? Empirical evidence for the US. *Working paper, available at Munich RePec* `http://mpra.ub.uni-muenchen.de/65584/`.

Ramos, A. and Sanz-Gracia, F. (2015). US city size distribution revisited: Theory and empirical evidence. *Working Paper, available at Munich RePec* `http://mpra.ub.uni-muenchen.de/64051/`.

Reed, W. J. (2002). On the rank-size distribution for human settlements. *Journal of Regional Science*, 42:1–17.

Reed, W. J. (2003). The Pareto law of incomes–an explanation and an extension. *Physica A*, 319:469–486.

Reed, W. J. and Jorgensen, M. (2004). The double Pareto-lognormal distribution–a new parametric model for size distributions. *Communications in Statistics-Theory and Methods*, 33(8):1733–1753.

Rozenfeld, H. D., Rybski, D., Andrade, J. S., Batty, M., Stanley, H. E., and Makse, H. A. (2008). Laws of population growth. *Proceedings of the National Academy of Sciences*, 105(48):18702–18707.

Rozenfeld, H. D., Rybski, D., Gabaix, X., and Makse, H. A. (2011). The area and population of cities: new insights from a different perspective on cities. *American Economic Review*, 101:2205–2225.

Sato, K. I. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge University Press.

Stanley, M. H. R., Amaral, L. A. N., Buldyrev, S. V., Havlin, S., Leschhorn, H., Maass, P., Salinger, M. A., and Stanley, H. E. (1996). Scaling behaviour in the growth of companies. *Nature*, 379:804–806.

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

Toda, A. A. (2012). The double power law in income distribution: Explanations and evidence. *Journal of Economic Behavior & Organization*, 84:364–381.

Uchaikin, V. V. and Zolotarev, V. M. (1999). *Chance and stability. Stable distributions and their applications*. VSP.

Zolotarev, V. M. (1986). *One-dimensional stable distributions*. American Mathematical Society.

Table 1: Descriptive statistics of the data samples used

Sample	Obs.	Mean	SD	Min.	Max.
Inc. places 1890	7,531	3,348	28,730	19	1,515,301
Inc. places 1900	10,596	3,376	42,324	7	3,437,202
Inc. places 1910	14,135	3,561	49,351	4	4,766,883
Inc. places 1920	15,481	4,015	56,782	3	5,620,048
Inc. places 1930	16,475	4,642	67,854	1	6,930,446
Inc. places 1940	16,729	4,976	71,299	1	7,454,995
Inc. places 1950	17,113	5,613	76,064	1	7,891,957
Inc. places 1960	18,051	6,409	74,738	1	7,781,984
Inc. places 1970	18,488	7,094	75,320	3	7,894,862
Inc. places 1980	18,923	7,396	69,168	2	7,071,639
Inc. places 1990	19,120	7,978	71,874	2	7,322,564
Inc. places 2000	19,296	8,968	78,015	1	8,008,278

Table 2: Descriptive statistics of the log-growth rates for the consecutive samples used

Sample	Obs	Mean	SD	Min	Max
Ip 1890-1900	7,531	0.198	0.360	-2.381	3.218
Ip 1900-1910	10,503	0.185	0.374	-3.714	2.664
Ip 1910-1920	13,543	0.113	0.322	-3.036	3.723
Ip 1920-1930	15,085	0.068	0.346	-5.053	3.393
Ip 1930-1940	16,199	0.069	0.229	-5.849	3.570
Ip 1940-1950	16,416	0.088	0.293	-5.187	5.645
Ip 1950-1960	16,943	0.099	0.347	-3.235	4.810
Ip 1960-1970	17,826	0.084	0.329	-5.499	8.716
Ip 1970-1980	18,321	0.109	0.294	-2.354	4.166
Ip 1980-1990	18,810	-0.020	0.269	-2.735	2.770
Ip 1990-2000	19,048	0.075	0.262	-4.467	3.581

Table 3: p-values (statistics) of the Kolmogorov–Smirnov (KS) test of the null hypothesis that the decennial log-growth rates or ten-years increments come from the same distribution. The null is rejected in all cases

	KS				
	Ip 1900-1910	Ip 1910-1920	Ip 1920-1930	Ip 1930-1940	Ip 1940-1950
Ip 1890-1900	0 (0.06)	0 (0.13)	0 (0.20)	0 (0.25)	0 (0.19)
Ip 1900-1910		0 (0.10)	0 (0.16)	0 (0.21)	0 (0.15)
Ip 1910-1920			0 (0.09)	0 (0.12)	0 (0.05)
Ip 1920-1930				0 (0.11)	0 (0.07)
Ip 1930-1940					0 (0.07)
	Ip 1950-1960	Ip 1960-1970	Ip 1970-1980	Ip 1980-1990	Ip 1990-2000
Ip 1890-1900	0 (0.19)	0 (0.21)	0 (0.19)	0 (0.39)	0 (0.24)
Ip 1900-1910	0 (0.14)	0 (0.16)	0 (0.15)	0 (0.33)	0 (0.20)
Ip 1910-1920	0 (0.06)	0 (0.09)	0 (0.06)	0 (0.28)	0 (0.12)
Ip 1920-1930	0 (0.03)	0 (0.04)	0 (0.08)	0 (0.20)	0 (0.09)
Ip 1930-1940	0 (0.08)	0 (0.07)	0 (0.07)	0 (0.28)	0 (0.06)
Ip 1940-1950	0 (0.04)	0 (0.05)	0 (0.04)	0 (0.26)	0 (0.08)
Ip 1950-1960		0 (0.03)	0 (0.06)	0 (0.22)	0 (0.06)
Ip 1960-1970			0 (0.04)	0 (0.21)	0 (0.04)
Ip 1970-1980				0 (0.25)	0 (0.06)
Ip 1980-1990					0 (0.22)

Table 4: p-values (statistics) of the Crámer–von Mises (CM) test of the null hypothesis that the log-growth rates or ten-years increments come from the same distribution. The null is rejected in all cases

	CM				
	Ip 1900-1910	Ip 1910-1920	Ip 1920-1930	Ip 1930-1940	Ip 1940-1950
Ip 1890-1900	0 (5.30)	0 (41.11)	0 (105.07)	0 (129.06)	0 (78.47)
Ip 1900-1910		0 (21.09)	0 (74.36)	0 (98.07)	0 (50.96)
Ip 1910-1920			0 (26.80)	0 (35.49)	0 (7.14)
Ip 1920-1930				0 (32.31)	0 (13.36)
Ip 1930-1940					0 (15.25)
	Ip 1950-1960	Ip 1960-1970	Ip 1970-1980	Ip 1980-1990	Ip 1990-2000
Ip 1890-1900	0 (87.08)	0 (109.86)	0 (84.47)	0 (366.36)	0 (139.92)
Ip 1900-1910	0 (57.27)	0 (77.21)	0 (55.48)	0 (345.08)	0 (104.99)
Ip 1910-1920	0 (13.46)	0 (23.77)	0 (10.75)	0 (280.74)	0 (41.65)
Ip 1920-1930	0 (3.56)	0 (3.79)	0 (18.54)	0 (135.53)	0 (17.21)
Ip 1930-1940	0 (23.38)	0 (17.69)	0 (11.59)	0 (288.26)	0 (10.53)
Ip 1940-1950	0 (3.97)	0 (7.88)	0 (2.72)	0 (253.37)	0 (19.12)
Ip 1950-1960		0 (2.50)	0 (7.26)	0 (196.91)	0 (15.80)
Ip 1960-1970			0 (7.93)	0 (189.17)	0 (6.53)
Ip 1970-1980				0 (282.54)	0 (12.90)
Ip 1980-1990					0 (224.52)