Physica A 512 (2018) 68–73

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Granularity of the top 1,000 Brazilian companies

Sergio Da Silva [a,*], Raul Matsushita [b], Ricardo Giglio [c], Gunther Massena [a]

[a] Department of Economics, Federal University of Santa Catarina, Florianopolis S.C., 88040-900, Brazil
[b] Department of Statistics, University of Brasilia, Brasilia D.F., 70910-900, Brazil
[c] Department of Production Engineering, Federal University of Santa Catarina, Florianopolis S.C., 88040-900, Brazil

HIGHLIGHTS

- We assess the granularity hypothesis by considering data for Brazilian companies.
- We adjust power laws to the data and estimate Pareto and Gumbel exponents.
- We find we cannot dismiss the hypothesis of granularity.
- We also find the Pareto exponent is approximately one, roughly a Zipf's law.
- We find a power-law progress curve fits the data even better.

ARTICLE INFO

Article history:
Received 23 March 2018
Received in revised form 20 June 2018
Available online xxxx

Keywords:
Granularity
Companies
Firm-size distribution
Power law
Zipf's law
Power-law progress curve

ABSTRACT

"Granularity" refers to the fact that economies are populated by a few large companies (the big "grains") that coexist with many smaller companies. Such a distribution of firm sizes is modeled by power laws. This study adds to the international evidence of the granularity hypothesis by considering data for the top 1,000 Brazilian companies. We sort the companies from top to bottom in terms of their net revenues. Then, we adjust power laws to the data and estimate Pareto and Gumbel exponents. We find we cannot dismiss the hypothesis of granularity for the Brazilian companies. We also find the Pareto exponent is approximately one ($1.070 \pm .015$), roughly a Zipf's law. Such a result is in line with the previous one found for American companies where the Pareto exponent = 1.059. We also find a power-law progress curve best fits the data.

1. Introduction

Macroeconomic theory tacitly assumes the activity of one individual company matters very little for the economy as a whole. The activity of millions of firms tends to cancel each other out so that the impact of the individual company on the aggregate is negligible. Thus, the economy is assumed to be populated by a smooth continuum of firms [1]. However, the assumption that the shocks of individual firms average out on the aggregate will break down if the distribution of firm sizes is fat-tailed, as documented empirically [1]. A power law for the size of firms suggests economic activity is concentrated in a small number of firms [2]. In South Korea, for instance, the top 10 business groups accounts for 54 percent of GDP, and Samsung alone accounts for 14 percent of GDP [3]. Shocks occurring in big companies explain about one-third of GDP fluctuation in the American economy [1] and almost half of GDP changes in France [4]. The recent period known as the Great Moderation in the American economy was characterized by a fall in the importance of aggregate shocks, while the volatility

* Corresponding author.
E-mail address: professorsergiodasilva@gmail.com (S. Da Silva).

https://doi.org/10.1016/j.physa.2018.08.027
0378-4371/

of sectoral shocks was unchanged—meaning a substantial role of idiosyncratic shocks after the mid-1980s explains half of the quarterly variation in the industrial production [5].

In light of this piece of evidence, one might wish to consider the alternative assumption of "granular" firms [1]. "Granularity" refers to the fact that economies are populated by a few large companies (the big "grains") that coexist with many smaller companies [1]. Such a distribution of firm sizes is modeled by "power laws". Samsung, for instance, is a big grain whose activity has a huge impact on the aggregate. It is about time to dismiss the assumption of Gaussianity in the distribution of firm sizes implicitly adopted by macroeconomic theory and eventually embrace the evidence of its fat tails. After all, economic fluctuations are not due to small diffuse shocks that directly affect every firm. Instead, they are attributable to the incompressible "grains" of economic activity – that is, the large firms – whose idiosyncratic shocks generate significant aggregate shocks, affecting GDP and therefore all the firms [1].

When the probability of measuring a particular value of some quantity varies inversely as a power of that value, the quantity follows a power law. In general, a power law is a functional relation between two quantities, where a relative change in one quantity yields a proportional relative change in the other, regardless of the initial size of such quantities, that is, one quantity varies as a power of the other. Power laws appear widely in physics, biology, economics and many other disciplines. More than 100 years ago, it was economist Vilfredo Pareto who first found a power law when studying the distribution of wealth. He proposed the number of people y with wealth $\geq x$ is distributed as $y = x^{-\nu}$, where the "Pareto exponent" $\nu \approx 1.5$. Thereafter, power laws have been discovered for events ranging from forest fires, earthquakes, mass extinctions of species and stock markets [6]. In economics, there are several well-established power laws, apart from the distribution of wealth [2]. Hume's quantity theory of money is another. Stock price returns follow an inverse cubic law with $\nu = 3$. City sizes distribute as a power law with Pareto exponent $\nu = 1$. A power law with a unitary Pareto exponent is known as Zipf's law. Of note, Pareto expressed his law in terms of the cumulative distribution function, while Zipf formulated his law in terms of the probability distribution function. This, and more details, can be found in Ref. [7]. The distribution of American firm sizes is claimed to follow a Zipf's law, but $\nu = 1.059$ [8]. When the firm size distribution follows a power law with an exponent close to one, it implies that idiosyncratic shocks to large firms have an important impact on aggregate volatility [3].

This paper considers data for the top 1,000 Brazilian companies, by net revenue. We adjust a classic power law to the data and estimate the Pareto exponent to test for the hypothesis of granularity and Zipf's law. We also show a power law of the log of data adjusts better than Zipf's. Such a "power-law progress curve" is shown in connection with its "Gumbel exponent".

The rest of this paper is organized as follows. Section 2 shows the data and the methods employed in their analysis. Section 3 reports the results, and Section 4 discusses them. Section 5 concludes this study.

2. Materials and methods

We collect the data from the newspaper *Valor Econômico* (available online on http://www.valor.com.br/valor1000/2016/ranking1000maiores). The source of data was the accounting balance sheet of the companies. We consider the rank of net revenues (a company's revenue net of discounts and returns) for the 2015 accounting year. The top 1,000 companies were considered. The dataset is available at Figshare (https://doi.org/10.6084/m9.figshare.5952004.v1).

We let L_r be the net revenue of company r such that $L_r > L_{r+1}$, where $r = 1, \ldots, 1000$. Then, we define the relative fraction of revenue as

$$Y_r = \frac{L_r}{\sum_{r=1}^{1000} L_r}. \quad (1)$$

Therefore, $\sum_{r=1}^{1000} Y_r = 1$ and thus $\{Y_r\}$ is the normalized distribution of revenues of the 1,000 companies. For our purposes, this treatment is more computationally convenient.

A classic power law is written as

$$Y_r = ar^{-\nu} \quad (2)$$

or, in logs,

$$\log Y_r = \log a - \nu \log r, \quad (3)$$

where $0 < Y_r < 1$ due to the normalization (1), a is a proportionality constant, and ν is the Pareto exponent.

Alternatively, we can represent a power law as [9]

$$Y_r = a \cdot b^{(r-1)^c}, \quad (4)$$

which expresses a dampened exponential decay of Y_r. It can be rewritten as

$$\log Y_r = \log a + (\log b)(r-1)^c. \quad (5)$$

Coefficient b is the rate of progress, and c is the Gumbel exponent, an analogy to the extreme-value Gumbel distribution, whose cumulative distribution is described by a double exponential [10]. Eq. (4) is a power-law progress curve [11].

We then fit Eqs. (3) and (5) to our data through linear and nonlinear regression, respectively.