Size distribution of Portuguese firms between 2006 and 2012

Rui Pascoal [a,*], Mário Augusto [a,b], A.M. Monteiro [a,c]

[a] School of Economics, University of Coimbra, Avenida Doutor Dias da Silva 165, 3004-512, Coimbra, Portugal
[b] Institute of Systems and Robotics, Coimbra
[c] GEMF, Group for Monetary and Financial Studies, University of Coimbra, Portugal

HIGHLIGHTS

- Size distribution of Portuguese firms as measured by annual sales and total assets.
- Fitting of lognormal, Pareto (in particular Zipf), SCL distributions and the Shannon entropy determined at different levels of aggregation.
- Time evolution for the Portuguese firm size distribution.
- Analysis of the evolution of economic conjuncture, namely economic activity rhythm, diversity of activity sectors and competition.

ARTICLE INFO

ABSTRACT

This study aims to describe the size distribution of Portuguese firms, as measured by annual sales and total assets, between 2006 and 2012, giving an economic interpretation for the evolution of the distribution along the time. Three distributions are fitted to data: the lognormal, the Pareto (and as a particular case Zipf) and the Simplified Canonical Law (SCL). We present the main arguments found in literature to justify the use of distributions and emphasize the interpretation of SCL coefficients. Methods of estimation include Maximum Likelihood, modified Ordinary Least Squares in log–log scale and Nonlinear Least Squares considering the Levenberg–Marquardt algorithm. When applying these approaches to Portuguese's firms data, we analyze if the evolution of estimated parameters in both lognormal power and SCL is in accordance with the known existence of a recession period after 2008. This is confirmed for sales but not for assets, leading to the conclusion that the first variable is a best proxy for firm size.

© 2016 Published by Elsevier B.V.

1. Introduction

Firm size distributions have been widely studied for a long time. Among the efforts in this area are those related to several distributions: lognormal law, power law (in particular Zipf) and Simplified Canonical Law, based in classical studies by, for instance, Pareto [1], Zipf [2], Gibrat [3], and Mandelbrot [4]. Lognormal and power distributions can be obtained from the Gibrat's law and in the power law a constraint must be added, for instance, on the minimum size of firms. As for SCL, firstly presented by Mandelbrot [4] in the context of linguistics, the argument used to deduce the distribution of words frequency in texts can be adapted to the distribution of firms size. This is done by defining production units, which can be aggregated to constitute firms so that its network can be used flexibly to satisfy the most diverse needs in the economy in the less costly way, and reducing as much as possible the delay in achieving these goals.

In this paper we attempt to describe the size distribution of Portuguese firms between 2006 and 2012. Using the interpretations of the several distributions referred above, we analyze the time evolution for the Portuguese firm size distribution, relating it to the evolution of economic conjuncture, namely economic activity rhythm, diversity of activity sectors and competition. Relatively to the economic activity rhythm and competition it is expected their decrease when there is a recession. The diversity of activity sectors is captured by an entropy, measured at the higher aggregation levels. The stylized facts concerning the interrelations between economic rhythm, diversity of activity sectors, competition and others such as the level of concentration of firm size distribution, are gathered from literature. As for firm size concentration, entropy at lower levels of aggregation is used to measure it. The analysis of entropy at different levels of aggregation allows to verify how activity rhythm can have a distinct impact on diversity at those different levels.

This paper departs from traditional literature in several ways: (i) in the SCL model applied to firms, we enhance the role of the project execution's delay and not only its costs, in the decisions concerning the economic projects pursued to fullfil the needs in the economy; (ii) we give a correction in the normalization parameter's formula related to SCL model; when reparameterizing the SCL parameters to the case of firm size distribution, we emphasize that the normalization parameter has to be defined taking into account that firm sizes are not proportions; (iii) we consider an analogy between molecules behavior and firm size referred in literature, which is detailed in order to clarify the relation between economic activity and a central parameter of the SCL model; (iv) we conduct an empirical analysis for time evolution of firm size distribution using several distributions; different interpretations of the estimated parameters collected from the literature are discussed in the light of our data; and (v) we point out a time series effect of recession on firm size concentration.

The rest of this paper is organized as follows. In Section 2, we expose some existing approaches for firm size model formulation. In Section 3, we discuss the interpretation of the coefficients of the Simplified Canonical Law. An empirical application on Portuguese firms distribution over the period 2006 to 2012 is presented and discussed in Section 4. In Section 5, the main conclusions and topics for future research are presented.

2. Firm size model formulation and existing approaches

Gibrat [3] argued, in a first model formulation for firm size, that it is reasonable to assume the Law of Proportional Effect, that is, firm growth is at each moment given by a random value (with finite variance) which is independent of the firm size at that moment. Hereafter we will denote firm size variable by S. As a consequence of this assumption, he concludes that firm size should follow a lognormal distribution. In fact, denoting by S_t the firm size at moment t and by σ_t the corresponding rate of growth:

$$S_t = \eta_t S_{t-1},$$

where

$$\eta_t = 1 + \sigma_t,$$

then, we have

$$S_t = \eta_t S_{t-1} = \eta_t \eta_{t-1} S_{t-2} = \cdots = \eta_t \eta_{t-1} \cdots \eta_1 S_0,$$

and taking logarithms,

$$\ln S_t = \ln \eta_t + \ln \eta_{t-1} + \cdots + \ln \eta_1 + \ln S_0.$$

Since the terms $\ln \eta_i$, $i = 1, \ldots, t$ are independent and identically distributed, by applying the Central Limit Theorem $\ln S_t$ follows approximately a Gaussian distribution and S_t a lognormal one (see Sutton [5] and Steindl [6]).

In this context, Cabral and Mata [7] analyze Portuguese manufacturing firms size, as measured by the number of employees. They find a departure from lognormal law by using nonparametric estimation and parametric estimation in the class of extended generalized gamma distribution. Considering a cohort of firms, born in 1984, they find that this departure is greater at the beginning and tends to vanish as the firms get older. According to these authors, a first argument for this behavior is selection, that is, smaller firms are more likely to disappear. Nonetheless, they find that this is not sufficient to explain that behavior and they consider a model that introduces financing constraints in young firms. Also, Barbosa and Eiriz [8] use nonparametric estimations for firm size distributions in Portuguese districts in 2000 and 2004 and compare their output with the lognormal law. They study whether firms localization influences the relation between firm size and its growth.

Mata and Portugal [9] estimate duration models for a group of Portuguese manufacturing firms (born in 1983), namely the Proportional Hazards Model and the Accelerated Failure Time Model. Introducing a set of explanatory variables, they study the occurrence of the selection effect referred in Cabral and Mata [7]. The concentration measure used here is Herfindahl Index, which is different from the one we use in this paper. They point out the direct effect of firms start-up size and the sectoral determinants of survival rates. Mata et al. [10] study the duration of new firms by opposition to the duration of established firms. The current size as well as the evolution is more important for predicting the failure rate than the initial size. The economic conjuncture in the industry where the firm is inserted, is also important to explain the likelihood of failure. Macroeconomic dummy explanatory variables are also used.