Size Distribution of Portuguese Firms between 2006 and 2012

Rui Pascoal, Mário Augusto and A. M. Monteiro

School of Economics University of Coimbra

1st Electronic International Conference MDPI-Entropy
November 2014
Outline

Firm size distributions:
- Lognormal
- Pareto
- Zipf
- Simplified Canonical Law (SCL)

Interpretation of SCL coefficients

Empirical application
The aim of this paper is 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 (Zipf as a particular case) and the Simplified Canonical Law (SCL).
Model formulation for firm sizes: Lognormal law

Lognormal law (Gibrat)

- firm size should follow a lognormal distribution
- assume the Law of Proportional Effect:
  \[ S_t = \eta_t S_{t-1}, \]
  \( S_t \) is the firm size at moment \( t \), \( \eta_t \) the corresponding rate of growth
- we obtain \( S_t = \eta_t \eta_{t-1} \cdots \eta_1 S_0 \), taking logarithms:
  \[ \ln S_t = \ln \eta_t + \ln \eta_{t-1} + \cdots + \ln \eta_1 + \ln S_0 \]
- applying the Central Limit Theorem:
  \( \ln S_t \) follows approximately a Gaussian distribution
  \( S_t \) follows a lognormal distribution
Model formulation for firm sizes: Power law

Power law (Pareto)

- introduced by Pareto
- \( P(S > s) \sim s^{-k}, \ k \in \mathbb{R}, \ S \) in the firm size variable
- alternative form:
  \( S(r) \sim r^{-A}, \ r \) is the rank of the firms, \( A \in \mathbb{R} \)
- we have \( A = \frac{1}{k}, \ k \neq 0 \) since the tail frequency is represented by \( P(S > s) \) and \( r \)
Model formulation for firm sizes: Power law

Power law (Zipf, a particular case of Pareto law with $A = 1$)

- $S(r) \sim \frac{1}{r}$

Theoretical approach

- Zipf law is obtained from:
  Gibrat law + some additional constraint
Model formulation for firm sizes: Simplified Canonical Law (SCL)

- generalization of Power law
- introduced by Mandelbrot

\[ p(r) = P \cdot (r + m)^{-A}, \]

\[ r = 1, \cdots, N, \text{ } r \text{ is the rank of entities} \]

\[ N \text{ is the maximum number of different entities} \]

\[ P \text{ is a normalizing constant,} \]

\[ P = \left[ \sum_{r=1}^{n} (r + m)^{-A} \right]^{-1}, \text{ } p(r) \text{ is the relative frequency corresponding to } r \]
Mandelbrot:
- study, in linguistics, the frequency of words in a text
- coding of a message: word by word versus block of words
- criterium: minimizing the cost difference between the codings
- word by word coding as smaller delay, but has a greater cost

Ramsden and Kiss Haypál:
- apply SCL in order to model firm size
- \( S(r) = P(r + \rho)^{-\frac{1}{\theta}} \)
  obtained by reparametrization: \( \theta = \frac{1}{A} \) and \( \rho = m \)
Remark

since $P$ is a normalizing constant and $S(r)$ are not proportions, $P$ must be given by

$$P = \sum_{K=1}^{N} S(k) \left[ \sum_{r=1}^{N} (r + \rho)^{-\frac{1}{\theta}} \right]^{-1}.$$
Analogy linguistic/firm size:

- letters (resp. economic units) can be aggregated to build words (resp. firms)
- words (resp. firms) can be assembled to build a message (resp. project)

- it’s more time efficient to combine words to code a message (resp. to have a flexible structure of firms to build projects)
- it’s economically more efficient to code the message as a block (resp. to build the whole project from the scratch)
\( \theta \) is identified with the concept of temperature in thermodynamics.

In firm size studies it represents:
- the degree of “economic activity”
- the degree of countries economic development
Interpretation of SCL coefficients: $\theta$ parameter

- $\theta$ is negatively correlated with firm size concentration (i.e. positively correlated with Shannon entropy at firm level)

**Remark**

We claim that there is a time series effect: recession (corresponding to negative evolution in $\theta$) can lead to greater losses in big firms, this may attenuate the effect in firm size concentration.

- $\theta$ is negatively correlated with sector unevenness (i.e. negatively correlated with Shannon entropy at the higher sector aggregation level)
- different levels of aggregation are considered in the study of entropy evolution
Interpretation of SCL coefficients

- $\rho$ defines the departure of SCL from Power law
- for higher $\rho$, competition between firms inside the same sector is less tolerated
- as a consequence we have positive correlation between $\rho$ and entropy (sector diversity)
- there is negative correlation between $\rho$ and firm size concentration
Empirical application

- Data source: IBAS (Iberian Balances Analysis System)
- Period: 2006-2012
- Criteria:
  - only firms without accounting information during the review period were eliminated
  - only firms in activity were taken into account (total assets and operating results are both positives)
  - only firms that do not suffer any significant changes such as mergers or total shut-down, in this period were considered
  - only private firms were considered
- Sample: 8,480
## Empirical application: number of firms by main economic activity

<table>
<thead>
<tr>
<th>Economic activity</th>
<th>Number of firms</th>
<th>% of the total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture, animal breeding, hunting and forestry</td>
<td>135</td>
<td>1.59</td>
</tr>
<tr>
<td>Mining and quarrying</td>
<td>31</td>
<td>0.37</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>1,856</td>
<td>21.89</td>
</tr>
<tr>
<td>Electricity, gas, steam, cold and hot water and cold air</td>
<td>38</td>
<td>0.45</td>
</tr>
<tr>
<td>Water collection, treatment and distribution; sewerage, waste management and remediation activities</td>
<td>46</td>
<td>0.54</td>
</tr>
<tr>
<td>Construction</td>
<td>990</td>
<td>11.67</td>
</tr>
<tr>
<td>Wholesale and retail trade; repair of motor vehicles and motorcycles</td>
<td>1,964</td>
<td>23.16</td>
</tr>
</tbody>
</table>
## Empirical application: number of firms by main economic activity

<table>
<thead>
<tr>
<th>Economic activity</th>
<th>Number of firms</th>
<th>% of the total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transportation and storage</td>
<td>357</td>
<td>4.21</td>
</tr>
<tr>
<td>Accommodation and food service activities</td>
<td>482</td>
<td>5.68</td>
</tr>
<tr>
<td>Information and communication activities</td>
<td>276</td>
<td>3.25</td>
</tr>
<tr>
<td>Financial and insurance activities</td>
<td>39</td>
<td>0.46</td>
</tr>
<tr>
<td>Real estate activities</td>
<td>700</td>
<td>8.25</td>
</tr>
<tr>
<td>Consultancy, scientific and technical activities</td>
<td>804</td>
<td>9.48</td>
</tr>
<tr>
<td>Administrative and support service activities</td>
<td>360</td>
<td>4.25</td>
</tr>
<tr>
<td>Public administration and defence; compulsory social security</td>
<td>5</td>
<td>0.06</td>
</tr>
</tbody>
</table>
Empirical application: number of firms by main economic activity

<table>
<thead>
<tr>
<th>Economic activity</th>
<th>Number of firms</th>
<th>% of the total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>23</td>
<td>0.27</td>
</tr>
<tr>
<td>Human health and social work activities</td>
<td>299</td>
<td>3.53</td>
</tr>
<tr>
<td>Arts, entertainment, sports and recreation activities</td>
<td>46</td>
<td>0.54</td>
</tr>
<tr>
<td>Other service activities</td>
<td>29</td>
<td>0.34</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>8,480</strong></td>
<td><strong>100.00</strong></td>
</tr>
</tbody>
</table>

Rui Pascoal, Mário Augusto and A. M. Monteiro
1st Electronic Inter. Conference MDPI-Entropy Nov. 2014
Empirical application: estimation methods

- Lognormal law: maximum likelihood approach
- Power law: OLS approach in log-log scale
- Simplified Canonical Law: nonlinear least-squares (Levenberg-Marquardt method)
Empirical application

- Lognormal law and SCL fitting
- Evolution of Shannon entropy
- Different levels of aggregation from firm level (measuring concentration) to great activity (measuring sector diversity)
Empirical application: results for lognormal fitting in assets

Figure: Lognormal distribution fitting for firm asset mean value
Empirical application: results for lognormal fitting in assets

Figure: Lognormal distribution fitting for firm asset variance
Empirical application: results for lognormal fitting in assets

Figure: Asset entropy results
Empirical application: results for lognormal fitting in assets

Comments:
- Positive trend in mean
- Less smooth evolution in dispersion but without significant change

Explanation:
- Reinforcement of assets value as a mechanism to ensure survival in time of instability
- Not a good proxy from firm size: doesn’t express effect of economic context in the health of firms
Empirical application: results for lognormal fitting in sales

**Figure:** Lognormal distribution fitting for firm sales mean value
Empirical application: results for lognormal fitting in sales

Figure: Lognormal distribution fitting for firm sales variance
Empirical application: results for lognormal fitting in sales

Figure: Sales entropy results
Comments:

- Decreasing based in mean:
  it confirms freezing in economy

- Dispersion much lower after 2008:
  confirmed in decrease of entropy at firm level
Empirical application: results for SCL fitting in assets

Figure: Asset SCL parameter estimation for $\rho$
Empirical application: results for SCL fitting in assets

Figure: Asset SCL parameter estimation for $\theta$
Empirical application: results for SCL fitting in assets

Figure: Asset entropy results
Empirical application: results for SCL fitting in assets

Comments:
- clear dropdown in $\rho$ after 2008 approaching values near 0
- evolution to a more competitive economy compatible with a down word tendency at all levels except the last one
- positive evolution in $\theta$
  in tune with these evolutions

Explanation:
- countercyclical behavior of assets level
  reaction by firms to adverse evolution
Empirical application: results for SCL fitting in sales

Figure: Sales SCL parameter estimation for $\rho$
Empirical application: results for SCL fitting in sales

Figure: Sales SCL parameter estimation for $\theta$
Empirical application: results for SCL fitting in sales

Figure: Sales entropy results
Empirical application: results for SCL fitting in sales

Comments:

- evolution of $\rho$ and $\theta$ compatible with symptoms of recession
  - less competitive economy
  - slower activity rhythm

- evolution of entropies is in accordance with these findings
  - pronounced decreasing evolution at firm level
    suggests increasing discrepancy between firm sizes
  - this evolution evanesce as aggregation level grows and
    reverts this tendency in the last two levels
    where accrued sector diversity prevails
Empirical application: results for power law fitting

Estimation values for parameter $A$ and corresponding standard deviation:

<table>
<thead>
<tr>
<th>Year</th>
<th>assets</th>
<th>sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>-1.477 (0.09)</td>
<td>-1.695 (0.013)</td>
</tr>
<tr>
<td>2008</td>
<td>-1.489 (0.08)</td>
<td>-1.751 (0.013)</td>
</tr>
<tr>
<td>2012</td>
<td>-1.425 (0.06)</td>
<td>-1.821 (0.012)</td>
</tr>
</tbody>
</table>

**Table:** Power law fitting: parameter $A$

the estimated exponent is statistically greater than one; the evolution for parameter $A$, after 2008, sustains the interpretation given for the SCL fitting
Concluding remarks

- Sales are a good proxy for firm size given its relation to the evolution of activity rhythm, it is consistent with the known fact that there was a pronounced recession after 2008.
- Assets show a countercyclical evolution with the economic tendency, which may be explained by the reaction of firms to an adverse environment.
- Shannon entropy was determined at different levels of aggregation. In the sales case, the decreasing path at firm level reveals an increasing concentration while the inversion of this path at higher levels of integration show an increase in sector diversity. It is significant that this inversion is only achieved on the last two levels.