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, η_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

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

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

Theorethical 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
- $p(r) = P \cdot (r + m)^{-A}$,
  $r = 1, \cdots, N$, r is the rank of entities
  N is the maximum number of different entities
  P is a normalizing constant,
- $P = \left[\sum_{r=1}^{n} (r + m)^{-A} \right]^{-1}$, $p(r)$ is the relative frequency
  corresponding to r

Model formulation for firm sizes: Simplified Canonical Law (SCL)

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 $\rho = m$

Model formulation for firm sizes: Simplified Canonical Law (SCL)

Remark

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

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

Model formulation for firm sizes: Simplified Canonical Law (SCL)

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)

- θ 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

- θ 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 θ) can lead to greater losses in big firms, this may attenuate the effect in firm size concentration

- θ 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

- ρ defines the departure of SCL from Power law
- for higher ρ, competition between firms inside the same sector is less tolerated
- as a consequence we have positive correlation between ρ and entropy (sector diversity)
- there is negative correlation between ρ 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

Economic activity	Number of firms	% of the total
Agriculture, animal breeding, hunting and forestry	135	1.59
Mining and quarrying	31	0.37
Manufacturing	1,856	21.89
Electricity, gas, steam, cold and hot water and cold air	38	0.45
Water collection, treatment and distribution; sewerage, waste management and remediation activities	46	0.54
Construction	990	11.67
Wholesale and retail trade; repair of motor vehicles and motorcycles	1,964	23.16

Empirical application: number of firms by main economic activity

Economic activity	Number of firms	% of the total
Transportation and storage	357	4.21
Accommodation and food service activities	482	5.68
Information and communication activities	276	3.25
Financial and insurance activities	39	0.46
Real estate activities	700	8.25
Consultancy, scientific and technical activities	804	9.48
Administrative and support service activities	360	4.25
Public administration and defence; compulsory social security	5	0.06

Empirical application: number of firms by main economic activity

Economic activity	Number of firms	% of the total
Education	23	0.27
Human health and social work activities	299	3.53
Arts, entertainment, sports and recreation activities	46	0.54
Other service activities	29	0.34
Total	8,480	100.00

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

- Lognormal law and SCL fitting
- Evolution of Shannon entropy
- Different levels of aggregation from firm level (measuring concentration) to great activity (measuring sector diversity)
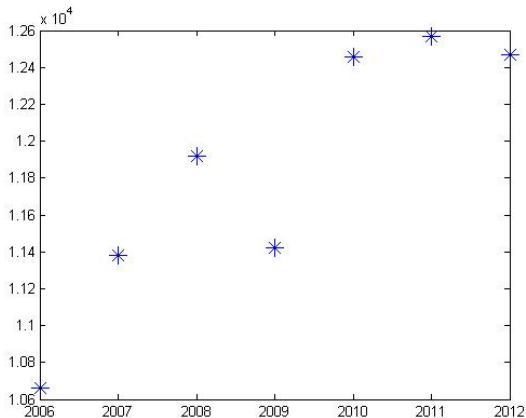
Figure: Lognormal distribution fitting for firm asset mean value
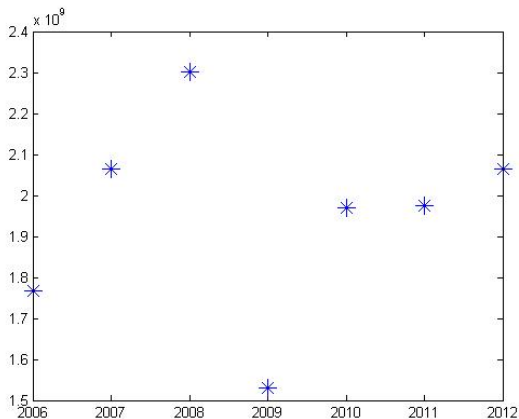
Figure: Lognormal distribution fitting for firm asset variance

Figure: Asset entropy results

- 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
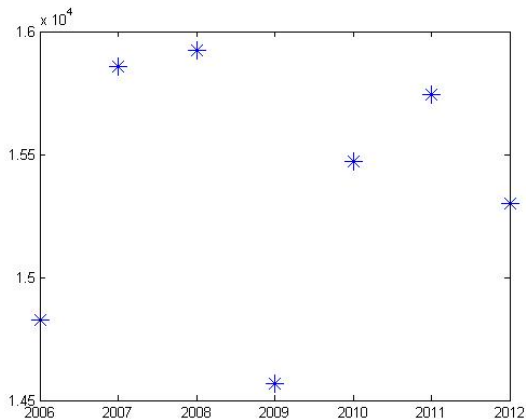
Figure: Lognormal distribution fitting for firm sales mean value
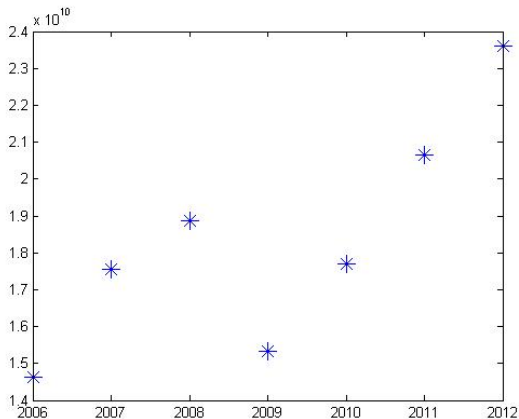
Figure: Lognormal distribution fitting for firm sales variance

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
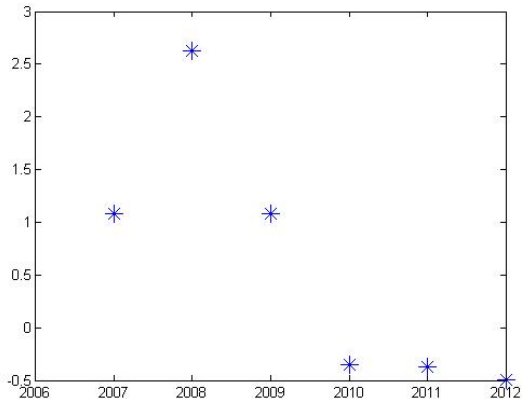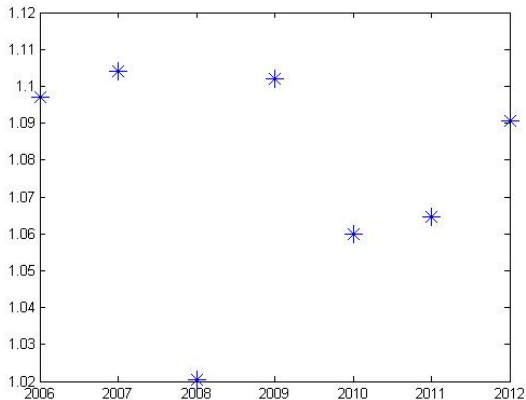
Figure: Asset SCL parameter estimation for ρ

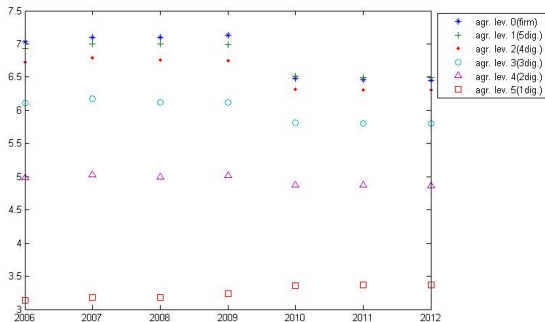Figure: Asset SCL parameter estimation for θ

Figure: Asset entropy results

- Comments:
  - clear dropdown in ρ 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 θ
    in tune with these evolutions
- Explanation:
  - countercyclical behavior of assets level
    reaction by firms to adverse evolution
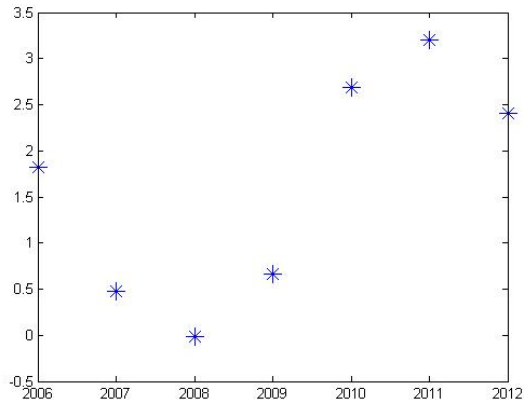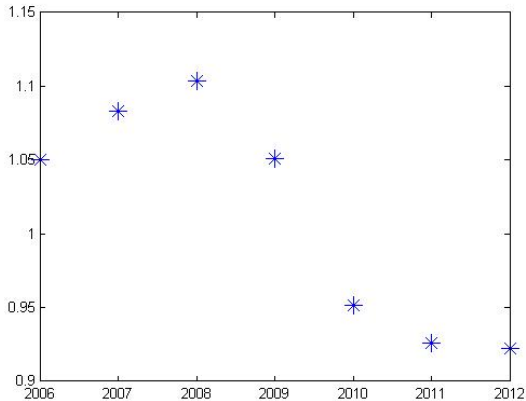
Figure: Sales SCL parameter estimation for ρ
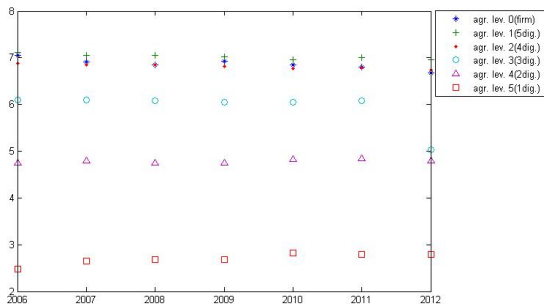
Figure: Sales SCL parameter estimation for θ

Figure: Sales entropy results

Comments:

- evolution of ρ and θ 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

Estimation values for parameter A and corresponding standard deviation:

Year	assets	sales
2006	-1.477	-1.695
	(0.09)	(0.013)
2008	-1.489	-1.751
	(0.08)	(0.013)
2012	-1.425	-1.821
	(0.06)	(0.012)

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