

Carlo Bianchi^a, Pasquale Cirillo^b, Mauro Gallegati^c, Pietro A. Vagliasindi^d

^a DSE, Università di Pisa

^b DEP/IMQ, Università Bocconi

^c DEA/SIEC, Università Politecnica delle Marche

^d DDEFI, Università di Parma

Validation and Calibration in ACE Models:

Some Experiments with the CATS Model

Contents

1. Why validation?
2. The model
3. Validation experiments:
 - Aggregate Output
 - Total Capital and its components .
 - Growth Rates
4. Calibration: a preliminary experiment
5. Conclusions and what comes next

Why do we aim to validate the CATS model?

In the Industrial Literature there are two stylized facts, accepted by most authors (obviously, as far as each fact is concerned there is a little quarrel over details).

1. The distribution of firms' size is right skewed and it presents a fat tails behavior (Axtell, 2001; Gaffeo et al., 2003; Kiss-Haypal, 2000).
2. The distribution of firms's growth rates is tent-shaped (Stanley et al., 1996; Bottazzi et al., 2003).

The Cats model succeeds in replicating such regularities; so it can be useful to validate it.

In reality, there's something more

In fact, Cats model is able to reproduce several secondary stylized facts. In particular, as we will see:

INDUSTRIAL DYNAMICS

- The variance of the aggregate is lower than that of the individual agents (Lee et al., 1998; Gabaix, 2004)
- The average standard deviation of growth rates linearly decreases with firms' size (Stanley et al. 1996; Gabaix, 2004)

FINANCIAL STYLIZED FACT

- The distribution of loans is power law (Fujiwara, 2003)

The model

A new approach to business fluctuations: heterogeneous interacting agents, scaling laws and financial fragility (2003)

Delli Gatti, Gallegati, Di Guilmi, Gaffeo, Giulioni e Palestrini

Introduction

Our model considers a fully supply-determined sequential economy à la Greenwald and Stiglitz (1993), with interacting firms on a discrete time pattern. In each period the market for an homogeneous produced good is opened.

It may constitute a simple version of Gallegati et al. (2003), since the banking sector, originally present in Gallegati et al. (2003), is reduced here to an exogenous set of financial data.

Our choices are mainly due to a lack of reliable information concerning the credit market. Since our aim is to calibrate and validate a HIA (*Heterogeneous Interacting Agent*) world, we decided to only consider the parts of the model for which we had a sufficient amount of data.

1. At any time $t = 1..6$, the economy consists of N_t firms.
2. Every firm has a very simple production function based on capital

$$Y_{it} = \phi_{it} K_{it}$$

3. The capital productivity ϕ_{it} differs among firms and depends on their relative size (inverse relationship).
4. The demand for goods for each firm is affected by an iid idiosyncratic real shock

$$P_{it} = u_{it} P_t \quad E(u_{it}) = \mu \text{ and } \sigma_{u_{it}}^2 < +\infty$$

5. Looking at actual data, we have decided to split the price generator process into two different processes, depending once again on firms' size.

6. In our minds, small firms show a higher average price and a stronger volatility, while big firms face more concentrated prices with a lower mean. We have chosen two uniform distributions.
7. Credit is the only external source of finance for firms. So, every firm can finance its capital simply recurring to net worth (A_{it}) or bank loans (L_{it}), $K_{it} = A_{it} + L_{it}$.
8. Given an exogenous real interest rate as cost for loans, at each time t profits and losses are

$$\pi_{it} = u_{it} Y_{it} - \bar{r} L_{it}$$

9. A firm goes bankrupt if its net worth become negative, that is to say $A_{it} < 0$, where

$$A_{it} = A_{it-1} + \pi_{it}$$

10. As in Greenwald and Stiglitz (1993), we assume that the probability of bankruptcy (Pr^b) is directly incorporated into the firm's profit/loss function. We design a convex bankruptcy cost function, quadratic by supposition:

$$C^b = cY_{it}^2 \quad c > 0$$

11. Then every firm will maximize its objective function, in order to determine its optimal capital stock K_{it}^*

$$\max_{K_{it}} \Gamma_{it} = E(\pi_{it}) - E(C^b)$$

12. At this point, each firm can decide the total amount of bank loans it needs, simply comparing the desired capital with the stock capital inherited from the previous period.

Validation

The dataset

All our validation experiments, together with the subsequent empirical analysis, are based on firm-level observations from the AIDA database, for the period 1996-2001.

To be more exact, the data we have dealt with when writing this paper are a particular subset of AIDA.

We have collected a set of 6422 manufacturing firms, all satisfying the following prerequisites:

1. No lack of data or discontinuities for every year or variable;
2. Reliable data as far as capital;
3. At least an employee and relative reliable costs.

Simulation and Results

In $t = 0$, every firm is initialized with its actual data from 1996.

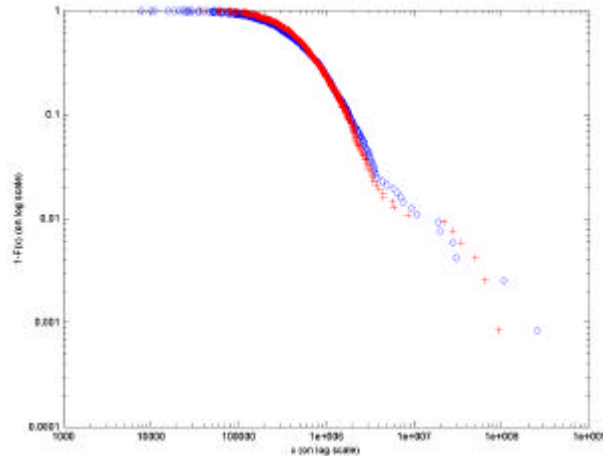
The market interest rate is **exogenous** and equal for all the firms; **it decreases every year** starting from 11.4% (1996) and arriving at 10% (2001). This values reproduce the average interests paid by the actual firms in the database.

The results we achieve are quite promising, since the simulation succeeds in reproducing several empirical evidences.

Whenever our model does not reproduce a particular empirical evidence, we find out that our database behaves in the same way, indicating that, anyway, the simulation achieves its goal: **replicating actual data**.

Total Capital

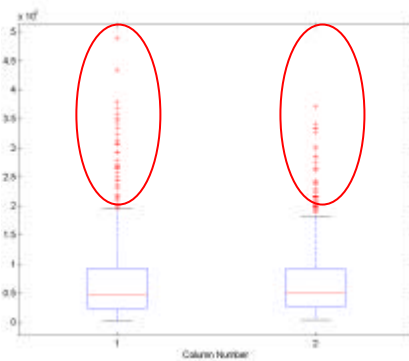
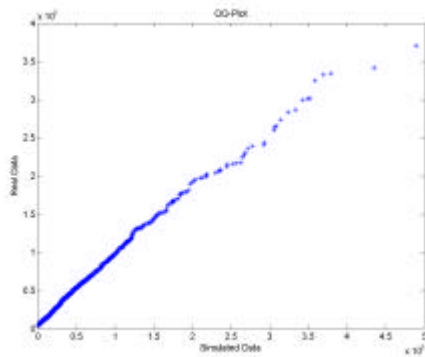
Accepti
simulate
6422 (8)
Both ob
skewed,
This is a
(Axtell,



nd
s over
arly
ature

Zipf's Plot of the total capital distributions: observed (red plus) and simulated (blue diamonds).

We have performed many graphical and analytical tests to discover if our two samples (observed and simulated data) may be considered belonging to the same distribution.



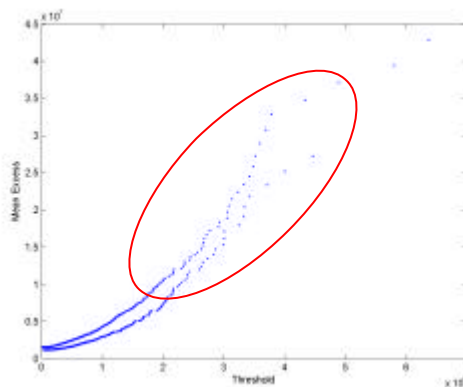
The same results are supported by the Generalised Kolmogorov-Smirnov and Cramér-von Mises tests, both concerning the ECDF of the two samples. Results are achieved with a confidence interval of 95%.

So, assuming the same distribution for both samples, we have found out that such distribution follows the GPD family:

$$H(x) = \begin{cases} 1 - (1 + \xi \frac{x}{\beta})^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ 1 - e^{-\frac{x}{\beta}} & \text{if } \xi = 0 \end{cases}$$

In particular, a Pareto II type ($\xi=0$).

As we can see below, our samples show a clear Paretian Behaviour
This suggests more analysis.

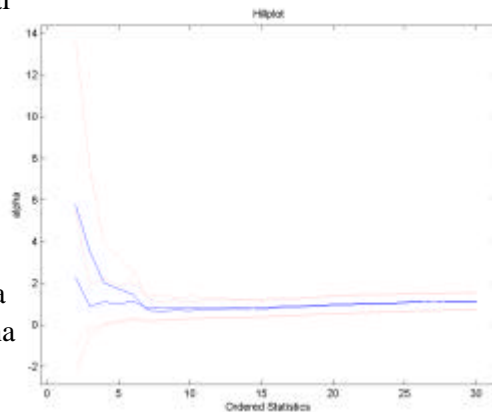


Meplot of actual and simulated data

We have performed Hill's estimates of the shape parameters for simulated ($a = 1,48$) and actual ($a = 1,52$) capital.

So, the two parameters are very similar and belong to the Paretian field ($0,8 < a < 2$), but we cannot state that the two tails behave in the same way.

Simulated capital, in fact, shows a slightly heavier tail (since its alpha is lower), demonstrating that we overestimates a little bit the observed values.



Growth Rates

As far as firms growth rates, several studies (Axtell, 2001; Bottazzi e Secchi, 2002) find a tent-shape behavior. In particular, the Laplace and the Lévy distributions seem to provide the best fitting (Bottazzi e Secchi, 2003; Gabaix, 2004).

In order to have a general approach, we have investigated if the empirical distributions of growth rates (in terms of capital) belong to the well-known Subbotin's family (Subbotin, 1923), that represents a generalization of several particular cases, such as Laplace and Gaussian distributions:

$$f(x, a, b) = \frac{1}{2ab^{\frac{1}{b}} \Gamma(1 + \frac{1}{b})} e^{-\frac{1}{b} \left| \frac{x-\mu}{b} \right|^{\frac{1}{b}}}$$

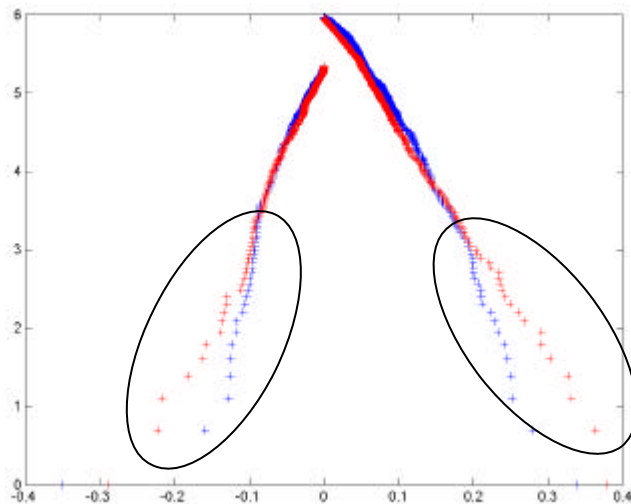
where μ is the mean, a and b two parameters and Γ the standard Gamma. If $b = 1$ the Subbotin distribution becomes a Laplace, a Gaussian for $b = 2$.

Both analytical and graphical tests support the idea of two Laplace distributions.

As a matter of fact, using the maximum likelihood method, we have found out the following values:

	observed data	Simulated data
μ	-0.0030 (0.0017)	0.0048 (0.0021)
a	0.0587 (0.0244)	0.0614 (0.0278)
b	1.0184 (0.3498)	1.0626 (0.3664)
$-\loglik$	1.1528	1.1549

1. The two means are very close and slightly positive \rightarrow asymmetric distribution
2. Since b is very near to 1, both distributions are in the field of attraction of the Laplacian case.
3. The values of a , the Laplacian shape parameter, are not very different in both cases, even if simulated data show slightly fatter tails ($0,061 > 0,058$).



Comparison of the two Laplace distributions: actual (blue) and simulated (red) data.

Wild@ce 2004

Following the procedures used by Gabaix (2004), we have also studied the relationship between firms size and firms growth rates.

Some authors find that bigger firms show a lower volatility as far as their growth rates and that this volatility (σ_{rates}) is linearly decreasing with size (S), that is to say

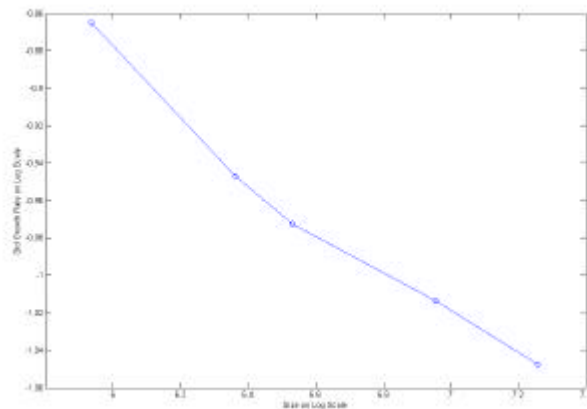
$$\ln \sigma_{rates} = -\alpha \ln S + \beta,$$

with $\alpha \sim 0.15$

We divided firms in four bins with regard to their sizes. Then we calculated the standard deviation of their growth rates. Finally, we plotted a log-log graph of the average standard deviation of growth rates versus the average size in each bin (Gabaix, 2004).

Validation in ACE Models

3 December 2004



$\alpha \sim 0,1643$

Growth Rates Std vs Size (observed data)

Calibration

Since our model shows to be very sensitive to the price generator processes, we have decided to perform a first very simple calibration about them.

In particular, we have tried to find the best uniform supports in order to minimize the difference between the actual and the simulated distributions of total capital.

To be more exact, we aim to minimize the distance between the two distributions (Segers, 2004).

Since our distributions belong to the extreme type, we can find a first estimate of the distance studying the differences between 1) the different shape parameters (Beirland et al., 2004; Castillo et al., 2003) and 2) the two quantile distributions (Segers, 2004).

We have estimated the shape parameters using both simple Hill's estimates (elemental percentile method; Castillo et al., 2003)

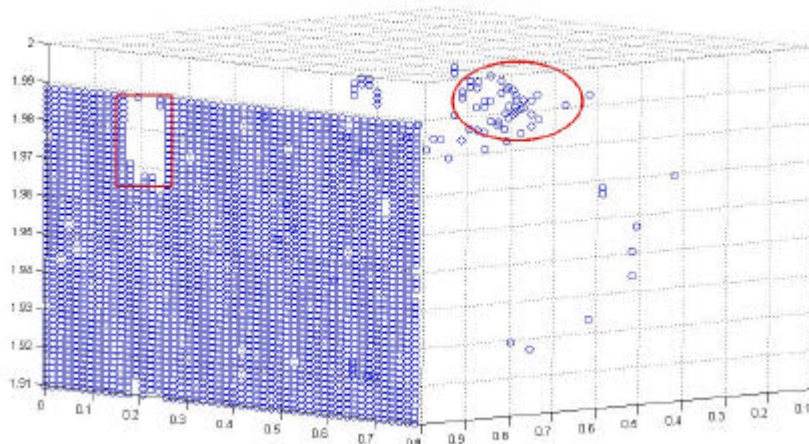
$$\xi = \frac{1}{k-1} \sum_{i=1}^{k-1} \ln x_{i,N} - \ln x_{k,N} \quad \text{for } k \geq 2,$$

and MLE (Balakrishnan et al., 2004).

For the sake of simplicity and for this first experiment, we have decided to use a simple grid method in order to find out the best combination of inferior and superior limits of the uniform supports.

In our plans, we aim to use the GRM algorithm (Gourieroux et al., 1996) in order to get a more precise analysis.

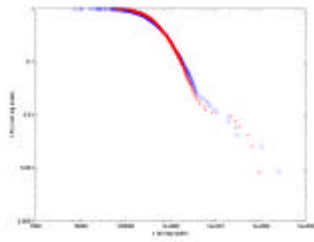
Example: price generator process of small firms



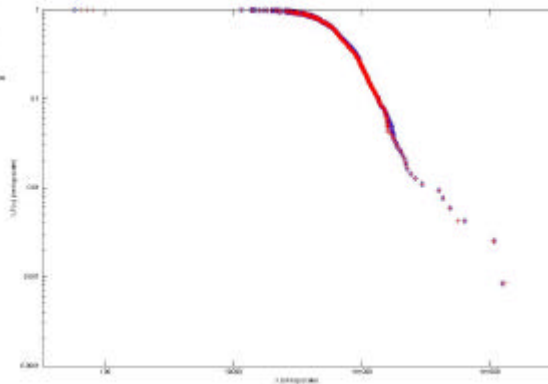
Scatter plot of the grid: distances greater than 0.1 are excluded (set to 1).

Best values: U(0.21, 1.98)

Growth Rates



Before



After Calibration

Conclusions and what comes next

The simple CATS model, first introduced in Gallegati et al. (2003) and here slightly modified to fit data, shows good capabilities in reproducing several empirical evidences.

The results we get in fitting actual data obviously need to be ameliorated, but they can be already considered satisfying.

Most problems occur when analyzing extreme-type events, such as the smallest and the biggest firms, indicating two different things: (i) our database needs to be better investigated and adjusted in order to make it very similar to the universe, (ii) the model must be integrated in some of its parts.