Title

Chloe Segale

January 9, 2018

Use employees as variable to investigfate

**1 Introduction**

I am going to investigate Gibrat’s Law, or the “law of proportionate effect”, which states that the rate of growth of a firm is independent of its size. This suggests that the growth of firms is driven by small idiosyncratic random shocks and hence there is no short-run or long-run convergence of firms to an “optimal size” (Sharfenaker).

If Gibrat’s law holds true, then one can conclude that it is impossible to make stable predictions about firm growth in the long run. Instead, growth is only attributed to small idiosyncratic shocks. This is an important question to answer because it dismisses the idea of market selection and where more efficient firms grow while those that aren’t shrink and eventually leave the market. It also rejects the notion of convergence to an optimal size (Cefis, 1).

To test Gibrat’s law, I will perform a statistical analysis of firm growth relative to firm size on within

Include a brief literature review

Why is this an important question

How do you propose to answer it?

Studying \_\_ sectors. See appendix for those listed.

Perform a statistical analysis of firms growth relative to firm size (in appendix) within \_\_ industries. Posterior distributions

Axtell argues that the ZIPF distribution is a better representation of U.S. firm sizes and identifies the limitations implicit in the COMPUSTAT dataset. In comparison to data pulled from the US Census Bureau, COMPUSTAT filters out many smaller firms despite their significance in the economy and therefore only represents a portion of the population of firms in the U.S., reinforcing the lognormal distribution (1818, Axtell). However, Axtell suggests that the power law distribution may be a better fit to the data and finds that it describes it well, suggesting that “firm growth operates on firms of all sizes, and that the fundamental unit of analysis is the individual employee” (Axtell, 1819). Axtell does produce some significant findings: That is, (i) firm growth rates follow a Laplace distribution, (ii) the standard deviation in growth rates falls with initial firm size according to a power law, and (iii) large firms pay higher wages for the same job according to yet another power law (the so-called wage-size effect) (Axtell, 1820).

Botazzi points out the shortcomings of pooling firms that operate in different sectors. Analyzing heterogeneous firms through risks the scenario in which aggregation procedures that are the sole (or at least main) contributor to any statistical regularity and ultimately “conceal the specific characteristics of the dynamics of firms operating in different sectors”. Instead, Botazzi studies firm growth rates relative to each sector and compares his findings to aggregated results to observe which statistical features at the sectoral level match those at the aggregate level and identify which differ when analyzed at the sectoral level (Botazzi, 218). Neither in the aggregate or sectoral level does Botazzi evidence of a relationship between firm size and average firm growth in aggregate or sectoral level, but does observe in both that “the variance of the growth rates robustly depend on size” (Botazzi, 228).

Mention data source

**2 Model Specification and Inference**

In this section you will detail your model, including your likelihood, priors, and

posterior. You will also detail how will proceed with inference, i.e. the use of a

Gibbs Sampler, STAN, Metropolis-Hastings, etc.

“If the evolution of firm sizes follows Gibrat’s law then firm growth rates are determined by a random walk with an autoregressive specification”:

Gibrat’s law asserts that growth rates are determined, i.e., that growth rates are independent of firm size. In this model, the growth of firm size (growth rate) is defined as . “Dividing by each year’s revenue will control for the increasing trend of total revenue as well as any common shocks experienced by all firms”.

This model is tested by running the below AR(1) model in STAN.

Accordingly, represents firm growth rates. is independently and normally distributed with mean zero and variance Analyzing the posterior of density of will tell us how probable Gibrat’s law is.

­If , Gibrat’s Law is confirmed, i.e. “the growth of the firm is unrelated to it’s current size and only depends on the sum of idiosyncratic shocks” (Sharfenaker).

As Botazzi had emphasized, firms within the same sector do not necessarily share similarities. To account for this heterogeneity, all slope and intercept parameters in the model are firm-specific (Cefis, 5).

Specification is as follows:

for firms (, all the same length (as specified in data section) at time between time and again specified in data.

e.g. i= firm id. i= 1 indicates firm 1, with the first observation occurring at and the last at .

All complete cases.

Since varying intercept and slope, both alpha and rho will have coefficients that vary

Observe initial condition

Varying intercept and slope model:

log(x\_ij)=log(A\_j)+a\_j log(k\_ij)

Lognormal distribution is skewed to the right, where the probability mass is mostly to the right of the mode (1818, Axtell).

The model is tested by running the above AR(1) times series model in STAN.

Panel data hierarchical regression

Hierarchical linear model

Hierarchical model group of J levels

What git equals allows firms within same industry to be able to differ from eachother

Time series hierarchical

“hierarchical Bayesian normal linear model to autoregressive time series panel data

**3 Data**

The source of the data set is from the COMPUSTAT database. The data is appropriate for the problem because it includes all data from\_\_ and is the designated database for this type of data.

The date range selected was from 1962 to 2016, which should be an adequate time span to yield meaningful results. The following variables were selected for import (with descriptions on the right):

|  |  |
| --- | --- |
| Company Name |  |
| SIC | Standard Industry Classification Code |
| FYEAR | Fiscal Data Year |
| SALE | Sales/Turnover (net) |
| AT | Total Assets |
| GVKEY | Company Code |

The data was first subset to include only the manufacturing industry (SIC 2000-3900). This step and the following data grooming was repeated \_ times for the \_\_\_ industry, \_\_\_\_ industry, and \_\_\_\_ industry. To ensure statistical reliability, the data was again subset to only include complete cases each individual industry.

Additional variables were added to the data table. A new variable was created to represent the number of years survived by the firm within the selected date range. The data was then again subset to only include firms that had survived the 54 year span of the data set. A numeric identifier for each firm was created for the analysis.

Axtell identifies potential limitations of the COMPUSTAT data set, which does not include data from smaller firms. By filtering out these smaller firms despite their significance in the economy and therefore only representing a portion of the population of firms in the U.S., the model may inappropriately reinforce the lognormal distribution (1818, Axtell).

**4 Results**

This section should describe your model results and posterior inference in detail.

In this section make sure to include plots of posterior distributions, posterior summaries,

HPDs, etc.

Inference for Stan model: 3fba6a6fa9267e7b79c9c48adc31de16.

3 chains, each with iter=10000; warmup=5000; thin=2;

post-warmup draws per chain=2500, total post-warmup draws=7500.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **mean** | **se\_mean** | **sd** | **2.50%** | **25%** | **50%** | **75%** | **97.50%** | **n\_eff** | **Rhat** |
| **alpha** | -0.02 | 0 | 0 | -0.02 | -0.02 | -0.02 | -0.02 | -0.01 | 6951 | 1 |
| **beta** | 0.9 | 0 | 0 | 0.89 | 0.9 | 0.9 | 0.91 | 0.91 | 6631 | 1 |
| **sigma** | 0.25 | 0 | 0 | 0.24 | 0.25 | 0.25 | 0.25 | 0.25 | 7475 | 1 |
| **lp\_\_** | 8988.43 | 0.02 | 1.24 | 8985.2 | 8987.88 | 8988.75 | 8989.33 | 8989.82 | 5491 | 1 |

n\_eff is a crude measure of effective sample size,

and Rhat is the potential scale reduction factor on split chains (at

convergence, Rhat=1).

It appears that Gibrat’s law holds for long-lived manufacturing firms in the United States. , with beta (p) converging at 0.9.

Extracting the MCMC samples and to produce replications from the posterior and inspecting: plotting the replicated posterior mins, maxes, and means against the actual means found by the model shows that:

Histogram of (replicated) posterior minimums vs the actual minimum

Actual minimum of growth, -1.4089, (Black line) vs a histogram of replicated minimums. Centered near -1.40849.



Histogram of (replicated) posterior maximums vs the actual maximum, at 0.7417343. Again, the actual value stands near the observable center of posterior maximums.



Histogram of (replicated) posterior standard deviations vs the actual standard deviation, at 0.5778813. From the graph below, it is clear to see that the actual standard deviation is far from the replicated posterior standard deviation.

Histogram of growth; Histogram with overlapping predicted data:

The model did a decent job compared to the given data

Does it appear Gibrat’s law holds for long-lived manufacturing firms in the US?

Explain in termof the posterior density of rho, i.e. how probable is Gibrat’s law?

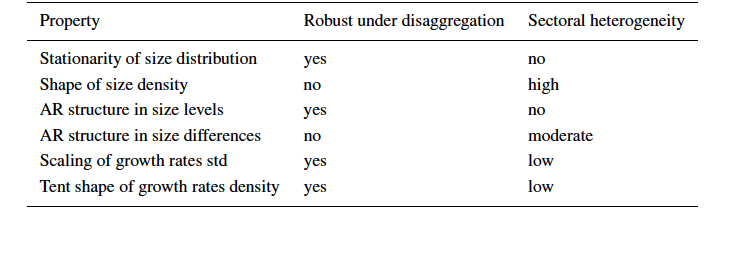
**5 Model Checks**

In this section diagnose your model. Include convergence checks and robustness

checks.

Test that this finding using total sales (SALE) is robust to other measures of firm

size such as total assets (AT).

Make graph similar to botazzi’s with label property and robustness. 

Compare your results to long-lived firms in other sectors of your choice.then do with other sectors.

**6 Conclusion**

What can you conclude based on your model and inference and what are the implications

of your findings?

Botazzi finds that Firm growth is random – meaning average growth rates are independent of firm size.

**7 References**

 Please cite all references appropriately.

ZIPF Distribution of U.S. Firm Sizes. Robert L. Axtell, et al. Science 293, 1818 (2001). DOI: 10.1126/science.1062081

Testing Gibrat's Legacy:

A Bayesian Approach to

Study the Growth of Firms

**8 Appendix**

Please include an appendix for such things as detailed tables, mathematical proofs,

model diagnostics, etc.

Final data table: