Two paper about the growth of firms

Giulio Bottazzi

Scuola Superiore Sant'Anna

June 2019

Theory and Practice

Economics consists of theoretical laws which
nobody has verified

and of empirical laws which
nobody can explain. [Michal Kalecki, 1945]

Two specific questions

The inverse relationship between the size of the firm and the variance of its growth rates. Can we investigate its origin?

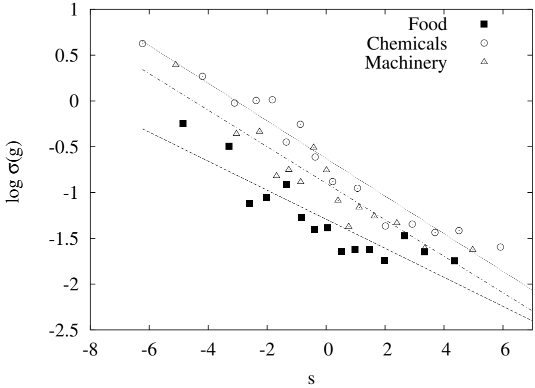
G.Bottazzi and A.Secchi *Gibrat's Law and Diversification* Industrial and Corporate Change, 15, pp. 847-875, 2006

What is the source of the seemingly universal tent shape of the firm growth rate distribution?

G.Bottazzi and A.Secchi *Explaining the Distribution of Firms Growth Rates* Rand Journal of Economics, 37, pp. 234-263, 2006
G.Bottazzi, A.Secchi *Why are distributions of firm growth rates tent-shaped?* Economic Letters vol. 80 pp.415-420, 2003

The evidence

From COMPUSTAT database



typical slope 0.2 is between 0, no scaling, and 0.5, pure “portfolio effect”.

The Data

PHID: developed by EPRIS Program sponsored by Merck Foundation. Sales of a panel of 198 multinational companies in the period 1987 to 1997.

Disaggregated in different micro-classes according to the Anatomical Classification System (ACS). 4-digit disaggregation defines 393 different submarkets.

Definition of the variables

Let $S_i(t)$ be the size (total sales) of firm i at time t. Consider the log size

$$s_i(t) = \log(S_i(t))$$

and the (log) growth rates

$$g_i(t) = s_i(t+1) - s_i(t) \ .$$

One can investigate

$$g_i(t) \sim \alpha + \beta s_i(t) \ .$$

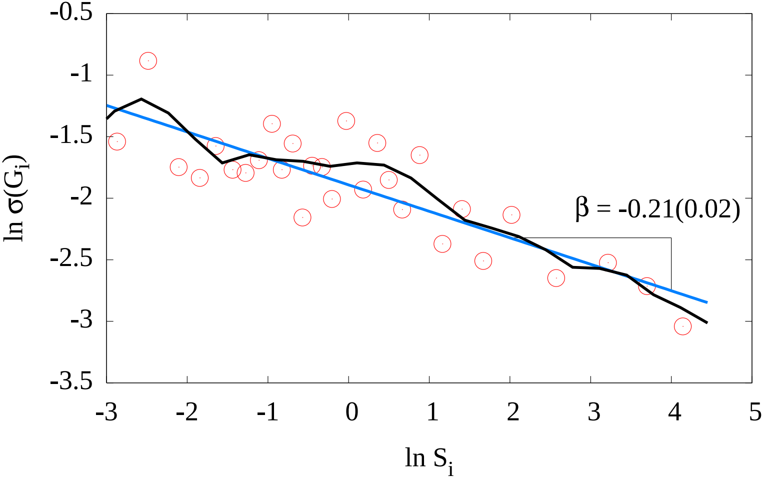
or higher moments.

Binned Regression

$$\Psi(g|s) = \alpha + \beta\, s \;, \quad \Psi = \text{cumulant}$$

Relation	**Slope**	**Std Error**
Size-Mean	-0.11	0.08
Size-Std	-0.21	0.02
Size-Skewness	-0.12	0.09
Size-Kurtosis	-0.03	0.11

Scatter plot Log(Growth Std.Dev.)-Log(Size)



Atomic components

Firms composed by “atomic” units (plants, lines of business, products) of roughly fixed size.

S_u unit fixed size

$N_i(t) = S_i/S_u$ number of units composing firm i at time t.

$G_{i,j}(t)$ growth rate of the j-th unit of firm i.

"Portfolio" decomposition

The aggregate growth can be decomposed in its atomic components

$$G_i(t) = \frac{1}{N_i(t)} \sum_{j=1}^{N_i(t)} G_{i,j}(t)$$

if $G_{i,j}(t)$ are independent

$$\sigma(G_i(t)) \sim 1/\sqrt{N_i} \sim 1/\sqrt{S_i} \ .$$

Too much!!

Definition of the Variables

$S_{ij}(t)$ sales of firm i in sector j at time t.

$\mathbf{N}_i(t)$ Set of active submarkets of firm i at time t.

$G_{ij}(t) = S_{ij}(t+1)/S_{ij}(t) - 1$ sectoral growth rates.

$$G_i(t) = \frac{S_i(t+1)}{S_i(t)} - 1 = \sum_{j \in \mathbf{N}_j(t)} \frac{S_{ij}(t+1)}{S_i(t)} - 1 \ .$$

Firm size is previously rescaled such that $\sum_i S_i(t) = 1$ at each t (get rid of nominal or aggregate real effects).

Disentangling different contributions

$$G_i(t) = \sum \frac{S_{ij}(t+1)}{S_i(t)} - 1 = \sum_j \frac{1}{N_i(t)} \; G_{ij}(t) \; \Delta_{ij}(t) \tag{1}$$

where

- $G_{ij}(t) = \dfrac{S_{ij}(t+1)}{S_{ij}(t)} - 1$ Growth in a given sub-market
- $\Delta_{ij}(t) = \dfrac{N_i(t)\; S_{ij}(t)}{S_i(t)}$ Diversification structure
- $N_i(t)$ Number of active submarkets

Sectoral Growth rates

From the data we observe that growth rates in different sub-markets are uncorrelated! ($r \sim 0.03$, not significant), so that

$$\mathrm{var}_{it}[G_i(t)] = \sum_j \mathrm{var}_{it}\left[G_{ij}(t)\Delta_{ij}(t)\frac{1}{N_i(t)}\right] \tag{2}$$

The relation $\mathrm{var}_{it}[G_i(t)] \sim S_i$ may come from:

1. Increasing returns to scale: $G_{ij} \sim S_i$
2. Degree of corporate coherence: $\Delta_{ij} \sim S_i$
3. Diversification effect: $N_i(G) \sim S_i$

First possible source: $G_{ij} \sim S_i$?

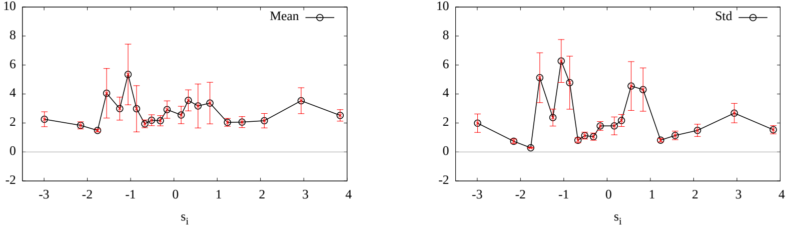


Figure: Average sub-market growth rates vs. firm size. Bars represent two standard errors.

No increasing (decreasing) returns to aggregate scale.

Second possible source: $\Delta_{ij} \sim S_i$?

Diversification heterogeneity index:

$$\tilde{\Delta}_i(t) = \frac{\sqrt{(\sum_{j=1}^{N_i(t)} \Delta_{i,j}(t) - 1)^2}}{\sqrt{N_i(t)(N_i(t) - 1)}}$$

The index range from 0, evenly distributed activity, to 1, complete concentration.

Second possible source: $\Delta_{ij} \sim S_i$?

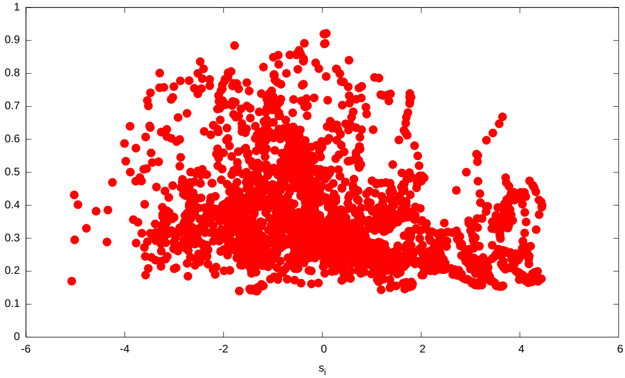


Figure: Scatter plot of $\tilde{\Delta}$ and firm log-size s

No different diversification structure at different sizes.

Third possible source: $N_i \sim S_i$?

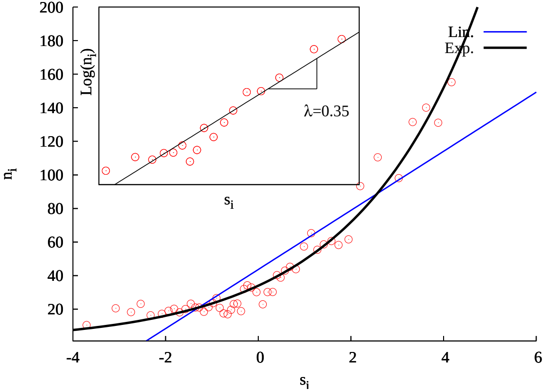


Figure: The number of active submarkets as a function of firm log-size. In the inset: the log of active submarkets.

The sole diversification explains the relation: $N \sim \ S_i^{.35}$ hence $\mathrm{var}_{it}[G_i(t)] \sim N^{1/2} \sim S_i^{.18}$.

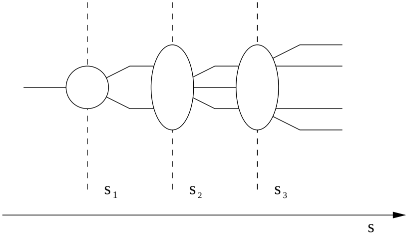
A Stochastic Model of Firm Diversification

What we learned: power-like relation between the number of active sectors and the size of the firm.

Log(size) is the independent variable.

$P(n, s; m, s')$ = probability that a firm which possesses n active markets when its (log) size is s, will possess $m \geq n$ active markets when its (log) size is $s' \geq s$

Random diversification



- Diversification events are seen as mutually independent “events”.
- these events are uncorrelated in time.
- neglects the possibility of “instantaneous” multiple events.

Poisson process (infinitesimal description)

Consider the transition probability $P(n, s; m, s+\delta)$ (joint probability) between two sizes differing by an infinitesimal quantity δ

$$P(n, s; m, s+\delta) = \begin{cases} \lambda\,\delta + o(\delta) & m = n+1 \\ o(\delta) & m > n+1 \end{cases} \quad .$$

Whence the Chapman-Kolmogorov equation

$$\begin{cases} p_{n_0}(s+\delta) = p_{n_0}(s)\,P(n_0, s; n_0, s+\delta), n = n_0 \\ p_n(s+\delta) = p_{n-1}(s)\,P(n-1, s; n, s+\delta) + p_n(s)\,P(n, s; n, s+\delta), n > n_0 \end{cases}$$

where $P_n(s)$ is the probability that a firm of size s is active in n sub-markets.

Poisson process (master equation)

Substituing and takig the limit $\delta \to 0$ one gets the master equation

$$\begin{cases} p'_n(s) &= -\lambda p_n(s) + \lambda p_{n-1}(s) \qquad n > n_0 \\ p'_{n_0}(s) &= -\lambda p_{n_0}(s) \qquad\qquad\qquad n = n_0 \ , \end{cases}$$

with initial conditions

$$p_n(s_0) = \begin{cases} 1, & n = n_0 \\ 0, & n \neq n_0 \ . \end{cases}$$

depending on initial size s_0 and initial number of active sub-markets n_0.

Poisson process (solution)

The solutions foe $s \geq s_0$ and $n \geq n_0$ reads

$$p_n(s) = \frac{\lambda(s - s_0)^{n-n_0}}{(n - n_0)!} \; e^{-\lambda(s-s_0)} \quad n \geq n_0 \quad .$$

It is easy to compute the average

$$m(s) = \sum_{n=0}^{+\infty} n\, p_n(s) = n_0 + \lambda(s - s_0)$$

and the variance

$$\sigma^2(s) = \sum_{n=0}^{+\infty} (n - m(s))^2\, p_n(s) = \lambda(s - s_0)$$

Variance increases proportionally to log(size)

Generalized Poisson process

$\lambda \to \Lambda(s)$ probability to enter in a new market when the size is s.
If $\Lambda'(s) > 0$ scale economies to diversification
If $\Lambda'(s) < 0$ barriers to diversification for larger firms.
Formally for $\delta << 1$ one has

$$P(n, s; m, s + \delta) = \begin{cases} \Lambda(s)\,\delta + o(\delta) & m = n + 1 \\ o(\delta) & m > n + 1 \end{cases} \quad .$$

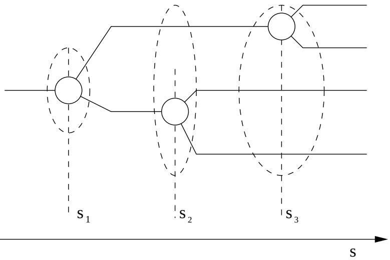
The average and variance

$$m(s) = n_0 + \theta(s; s_0) \qquad \sigma^2(s) = \theta(s; s_0)$$

where

$$\theta(s; s_0) = \int_{s_0}^{s} dx\ \Lambda(x)\ .$$

Possible power-like relation between N and S



Trying to open the diversification “black box”: diversification as a branching process: each opened branch (sub-market) becomes eventually the origin of a new branching (diversification) event.

Yule process

Formally for $\delta << 1$ one has

$$P(n,s;m,s+\delta) = \begin{cases} n\,\lambda\,\delta + o(\delta) & m = n+1 \\ o(\delta) & m > n+1 \end{cases},$$

The average number of active sectors and variance

$$m(s) = n_0\, e^{\lambda(s-s_0)} \qquad\qquad \sigma^2(s) = n_0\, e^{\lambda(s-s_0)}\left[e^{\lambda(s-s_0)} - 1\right] \;.$$

where $n_0 > 0$ is the initial number of active sectors.
Power-like relation between N and S

Validation 1: variance-average relation

Both generalized Poisson and Yule predict power-like relation between N and S.
Look at the relation between conditional mean and conditional variance.
For generalized Poisson

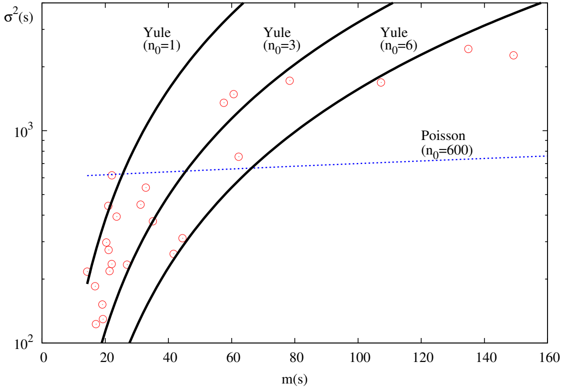
$$\sigma^2(s) = m(s) + n_0 \ ,$$

For Yule

$$\sigma^2(s) = \frac{1}{n_0} m(s)^2 - m(s) \ ,$$

Variance-average binned plot

Figure: Sample variance of the number of active sectors versus its average for firms in different equipopulated size classes.



Validation 2: empirical density of number of sectors

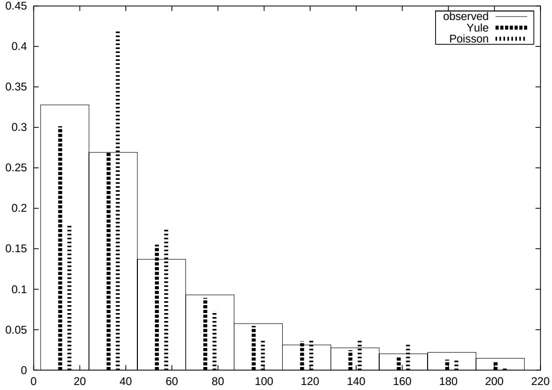


Figure: The binned probability density for the number of sub-market computed directly by the data and theoretically predicted by the Poisson and Yule models.

In summary

- The relation between a size of a firm and the variance of its growth rates has a long story! We provide evidence that in the Pharmaceutical sector this effect is completely due to the diversification dynamics

- The diversification dynamics is explained by a model supporting the dual interpretation of "increasing return to scope economy" and the existence of "limit to diversification" (technological? organizational?)

Firms Size

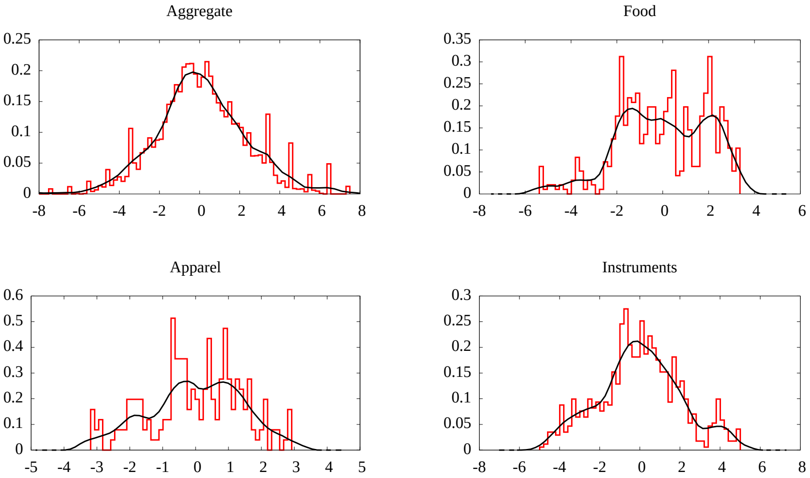
Let $S_i(t)$ be the size of firm i at time t. Consider the normalized (log) size

$$s_i(t) = \log(S_i(t)) - < \log(S_i(t)) >_i \tag{3}$$

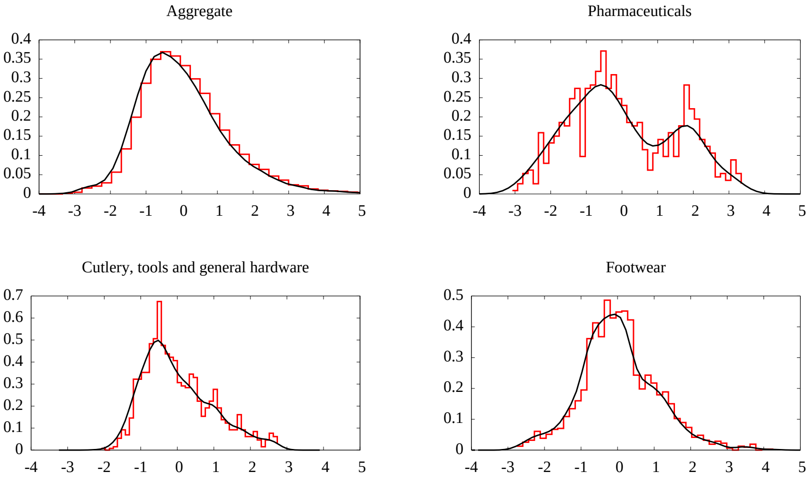
Main results on empirical firms size densities

1. Heterogeneity of shapes across sectors
2. Bimodality and no log-normality
3. Separation core-fringe
4. Paretian upper-tails?

Empirical Size Densities - US



Empirical Size Densities - ITA



Firms Growth Rates

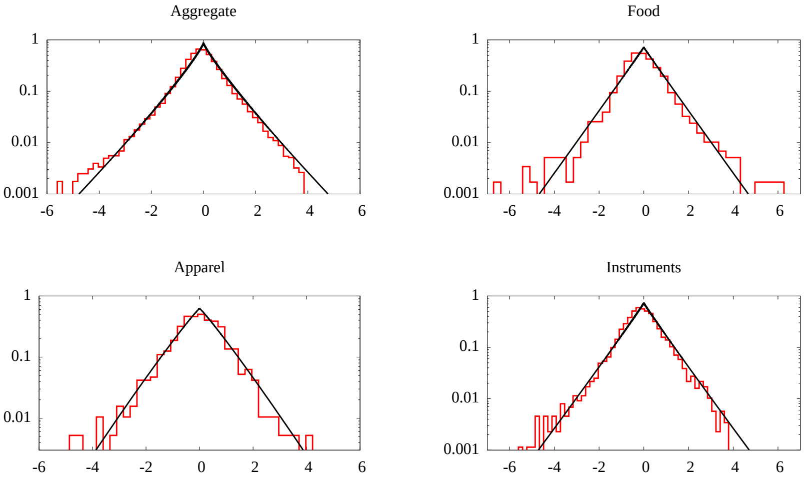
We build firms growth rates as the first difference of S_i

$$g_i(t) = s_i(t) - s_i(t-1) \tag{4}$$

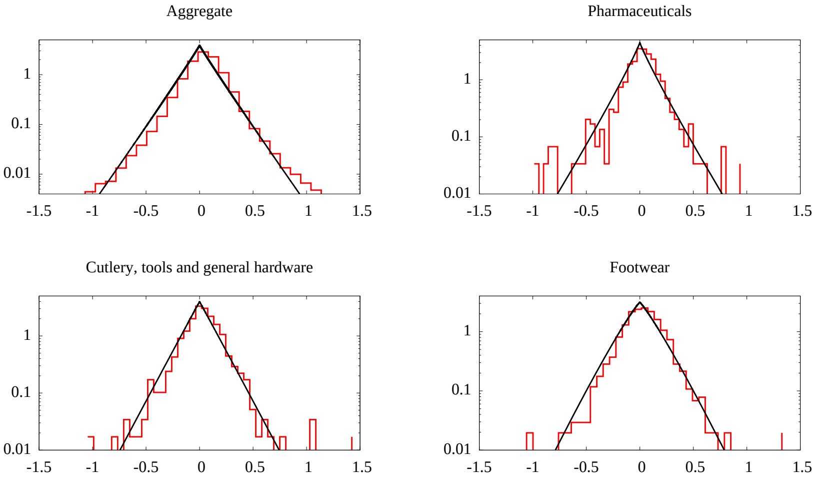
Main results on empirical growth rates densities

1. shape is stable over time
2. display similar shapes across sectors
3. look similar to the Laplace
4. present similar width(?)

Empirical Growth Rates Densities - US



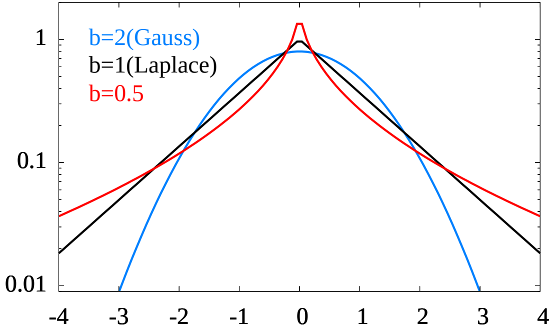
Empirical Growth Rates Densities - ITA



The Subbotin Distribution

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

Mikhail Fyodorovich Subbotin (1883-1966)



ML Estimation Procedure

We consider:

$$-\log(L_{\mathrm{S}}(x;a,b,\mu)) = \\ n\log\left(2b^{1/b}\,a\,\Gamma(1+1/b)\right) + (ba^b)^{-1}\sum_{i=1}^{n}|x_i-\mu|^b \tag{5}$$

and we minimize it with respect to the parameters using a multi-step procedure.

These ML estimators are asymptotically consistent in all the parameter space, asymptotically normal for $b > 1$ and asymptotically efficient for $b > 2$.

Estimates on Italian Sectors

		Parameter b		Parameter a	
Ateco code	**Sector**	**Coef.**	**Std Err.**	**Coef.**	**Std Err.**
151	Production, processing and preserving of meat	0.83	0.05	0.089	0.004
155	Dairy products	0.91	0.07	0.080	0.004
158	Production of other foodstuffs (brad, sugar, etc...)	0.89	0.05	0.097	0.004
159	Production of beverages (alcoholic and not)	0.88	0.06	0.108	0.006
171	Preparation and spinning of textiles	1.19	0.07	0.142	0.005
172	Textiles weaving	1.12	0.06	0.122	0.004
173	Finishing of textiles	1.11	0.06	0.107	0.004
175	Carpets, rugs and other textiles	1.02	0.08	0.118	0.006
177	Knitted and crocheted articles	0.97	0.05	0.124	0.005
182	Wearing apparel	0.92	0.03	0.120	0.003
191	Tanning and dressing of leather	1.12	0.09	0.140	0.007
193	Footwear	1.12	0.05	0.150	0.004
202	Production of plywood and panels	0.98	0.09	0.104	0.007
203	Wood products for construction	0.94	0.08	0.105	0.007
205	Production of other wood products (cork, straw, etc...)	1.31	0.13	0.106	0.006

Estimates on US Sectors

Ateco code	Sector	Parameter b Coef.	Parameter b Std Err.	Parameter a Coef.	Parameter a Std Err.
20	Food and kindred products	0.9888	0.0010	0.7039	0.0005
23	Apparel and other textile products	1.0819	0.0027	0.7664	0.0013
26	Paper and allied products	1.0999	0.0024	0.7663	0.0011
27	Printing and publishing	0.9621	0.0015	0.7115	0.0008
28	Chemicals and allied products	1.0164	0.0004	0.7562	0.0002
29	Petroleum and coal products	1.1841	0.0043	0.8370	0.0019
30	Rubber and miscellaneous plastics products	0.9487	0.0018	0.7148	0.0010
32	Stone, clay, glass, and concrete products	1.1023	0.0039	0.7720	0.0018
33	Primary metal industries	1.1254	0.0015	0.7870	0.0007
34	Fabricated metal products	0.9081	0.0013	0.6639	0.0007
35	Industrial machinery and equipment	0.9466	0.0003	0.6761	0.0002
36	Electrical and electronic equipment	0.8989	0.0003	0.6303	0.0001
37	Transportation equipment	1.0033	0.0011	0.7107	0.0005
38	Instruments and related products	0.9722	0.0004	0.6980	0.0002
39	Miscellaneous manufacturing industries	1.0232	0.0022	0.7447	0.0011

The Theoretical Framework

Observed growth as the cumulative effect of diverse "events"

$$g(t;T) = s(t+T) - s(t) = \epsilon_1(t) + \epsilon_2(t) + \ldots = \sum_{j=1}^{G(t;T)} \epsilon_j(t)$$

- The Gibrat Tradition: ϵ_j are r.v. independent from size s (strong form: ϵ_j are i.i.d.) Limitation: No interaction among firms
- Simon's model introduces a finite number of M opportunities progressively captured by N firms. $G(t;T)$ becomes a r.v. Limitation: Equipartition of opportunities among firms $\rightarrow$ Gaussian growth rates

The Model

Multi-step simulation model

Business Events $\rightarrow$ Micro-Shocks $\rightarrow$ Growth

Self-reinforcing effect in events assignment. Idea of "competition among objects whose *market success*...[is] cumulative or self-reinforcing" (B.W. Arthur)

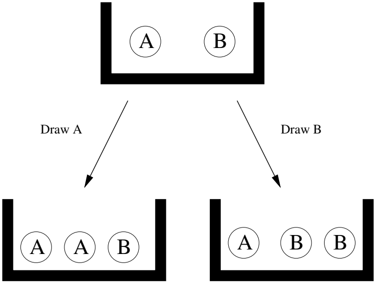
Discrete time stochastic growth process; at each round a two steps procedure is implemented:

- determine the number of events captured by a firm, $G(t;T)$
- disclose $\epsilon_j \quad j = \{1, \ldots, G(t;T)\}$, i.e. the effect of these events on firm size

STEP 1 - The Assignment of Business Events

1. Consider an urn with N different balls, each representing a firm

2. Draw a ball and replace with TWO of the same kind. (Here the first draw from an urn with two types of ball)



3. Repeat this procedure M times

RESULT: partition of M events on N firms.

STEP 2 - The Generation of Shocks

From the previous assignment procedure

$$m_i(t) = \text{\# of opportunity given to firm i at time t}$$

A very simple relation between “opportunities” and growth:

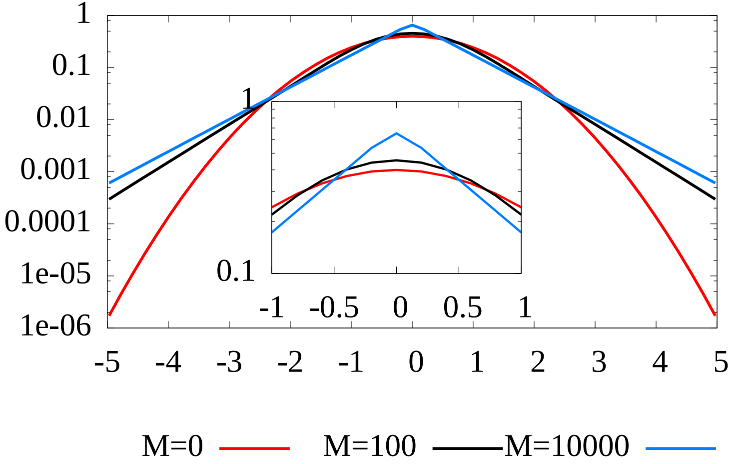
$$s_i(t+T) - s_i(t) = \sum_{j=1}^{m_i(t)+1} \epsilon_j(t) \tag{6}$$

ϵ are i.i.d. with a common distribution $f(\epsilon)$.

Run the simulation and collect statistics.

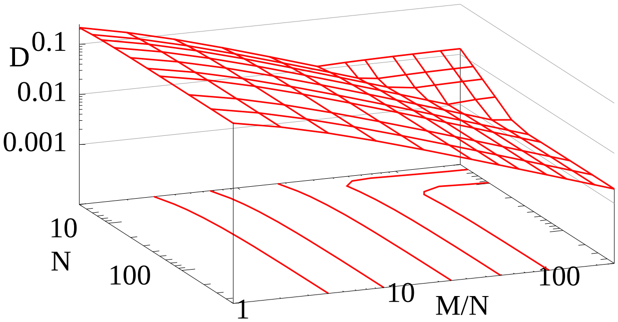
Simulation Results

Growth rates densities for $N = 100$ and different values of M.



Simulation Results - Cont'd

We define $D = |F_{\text{model}}(x; M, N) - F_{\text{L}}(x)|$ the absolute deviation between the empirical growth rates distribution (as approximated by the Laplace) and the distribution predicted by the model. Here D as a function of the number of firms N and the average number of micro-shocks per firm M/N.



Why does the Model work?

The unconditional growth rates distribution implied by this model is given by

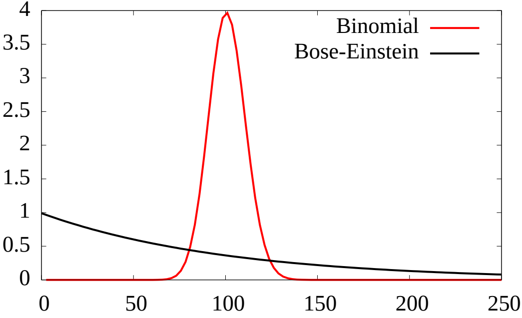
$$\sum_{h=0}^{M} \underbrace{P(h; N, M)}_{\text{Events Distribution}} \quad \underbrace{F(x; v_0)^{\star(h+1)}}_{\text{Distribution of the sum of h micro-shocks}} \quad .$$

In the assignment procedure above P follows a Bose-Einstein

$$P(h; N, M) = \frac{P(X)}{P(X|m_1 = h)} = \frac{\binom{N+M-h-2}{N-2}}{\binom{N+M-1}{N-1}}$$

while follows a Binomial in the Simon tradition.

Occupancy Statistics



Bose-Einstein and binomial with $N = 100$ and $M = 10,000$.

“Large Industry” Limit

Theorem

Suppose that the micro-shocks distribution possesses the second central moment $\sigma_\epsilon^2 < \infty$. Under the Polya opportunities assignment procedure the firms growth rates distribution converges in the limit for $N, M \to \infty$ to a Laplace distribution with parameter $\sqrt{v/2}$, i.e.

$$\lim_{M,N\to\infty} f_{\text{model}} = f_L(x; \sqrt{v/2}) = \frac{1}{\sqrt{2v}}\, e^{-\sqrt{2/v}\,|x|}$$

where $v = \sigma_\epsilon^2 M/N$.

Concluding Remarks on the Model

- A new stylized fact has been presented
- We show its robustness under disaggregation
- Our original explanation is based on a general mechanism of short-horizon “dynamic increasing returns” in a competitive environment
- We provide a “Large Industry” Limit Theorem
- Simulations show that “Large” is not so large

References

Selected reference:

- Bottazzi, G. and Secchi A. *Explaining the Distribution of Firms Growth Rates* Rand Journal of Economics, 2006.
- Gibrat, R. *Les Inégalités Économiques*. Paris: Librairie du Recueil Sirey, 1931.
- Ijiri, Y. and Simon, H.A. *Skew Distributions and the Sizes of Business Firms*. Amsterdam: North Holland, 1977.
- Kalecki, M. "On the Gibrat Distribution." *Econometrica*, Vol. 13(1945), pp. 161-170.
- Steindl, J. *Random Processes and the Growth of Firms*. London: Griffin, 1965.
- Sutton, J. "Gibrat's Legacy" *Journal of Economic Literature*, Vol. 35(1997), pp. 40-59.
- Sutton, J. *Technology and Market Structure, Theory and History*. Cambridge, Mass.: MIT Press, 1998.

Databases

FORTUNE 500 Annual ranking of America's largest public corporations as measured by their gross revenue compiled by *Fortune* magazine.

COMPUSTAT U.S. publicly traded firms in the Manufacturing Industry (SIC code ranges between 2000-3999) in the time window 1982-2001. We have 1025 firms in 15 different two digit sectors.

MICRO.1 Developed by the Italian Statistical Office(ISTAT). More than 8000 firms with 20 or more employees in 97 sectors (3-digit ATECO) in the time window 1989-1996. We use 55 sectors with > 44 firms.

Empirically based Industrial Dynamics

The Law Finding Process (i.e. “Retroduction”)

1. Looking for facts
2. Finding simple generalizations that describe the facts to some degree of approximation
3. Finding Limiting conditions under which the deviations of facts from generalization might be expected to decrease
4. Explaining why the generalization “should” fit the facts
5. The explanatory theories generally make predictions that go beyond the simple generalizations and hence suggest new empirical tests.