SCUOLA NORMALE SUPERIORE PISA

Crescita e dimensione delle aziende: un approccio statistico

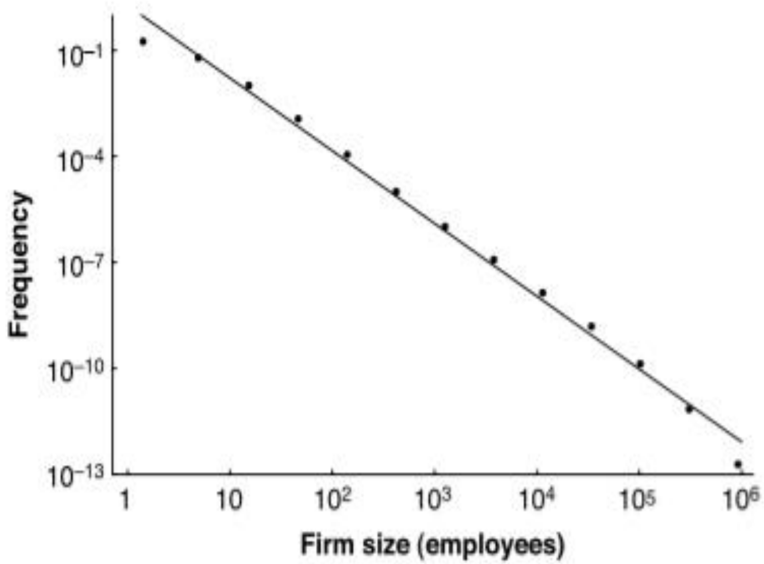
Fabrizio Lillo

Scuola Normale Superiore (Italy)

Measuring size and growth

- First of all firm size can be measured in many different ways: number of employees, revenue/turnover, sales, etc.
- On long time series it is important to normalize monetary size measures (e.g. revenue, sales) controlling for inflation.
- The relevant variable for growth is either
  - [S(t)-S(t-1)]/S(t-1)
  - log [S(t)/S(t-1)]

- The first stylized fact is an extreme heterogeneity of firm size, well described by a power law (Pareto) distribution (see figure below from Axtell, Zipf's distribution for U.S firm size, Science 2001)



REPORTS

Fig. 1. Histogram of U.S. firm sizes, by employees. Data are for 1997 from the U.S. Census Bureau, tabulated in bins having width increasing in powers of three (*30*). The solid line is the OLS regression line through the data, and it has a slope of 2.059 (SE = 0.054; adjusted $R^2 = 0.992$), meaning that $\alpha = 1.059$; maximum likelihood and nonparametric methods yield similar results. The data are slightly concave to the origin in log-log coordinates, reflecting finite size cutoffs at the limits of very small and very large firms.

Gibrat' s law

- Under the assumptions
  - proportionate growth of a firm in a given period is a random variable independent of the initial firm size
  - statistical independence of successive growths

  Gibrat (1931) concluded that after a long period the logarithmic growth rates are Gaussian distributed and independent of the initial firm's size
- Gibrat' s law has been tested empirically, but significant deviations from the normal distribution have been observed

Measuring the firm growth distribution

- Despite the availability of large firm size datasets, single firm models are difficult to test because the number of firms, N, is large, but the number of data points per firm, T, is very small (5-50 points)
- Two approaches used in the literature:
  - assume that the growth time series of each individual firm is a specific realization of the same stochastic process (Model Firm hypothesis)
  - assume that all firms in a balanced panel have the same specific functional form of the growth rate distribution, although the parameters that characterize the distribution may be different from firm to firm (Common distribution hypothesis)

- Gibrat law has been tested empirically (figures taken from Stanley et al., Nature **379**, 804 (1996)) and it has been found that the distribution of firm growth r depends on the size s_0

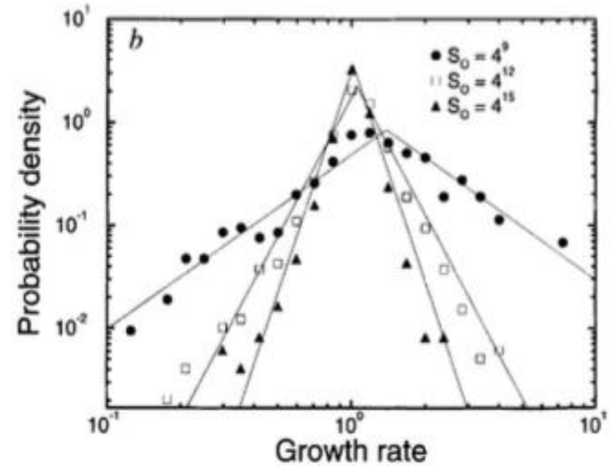


FIG. 1 *a*, Probability density $p(r \mid s_0)$ of the growth rate $r \equiv \ln(S_1/S_0)$ from year 1990 to 1991 for all publicly traded US manufacturing firms in the 1994 Compustat database with standard industrial classification index of 2000–3999. We examine 1991 because between 1992 and 1994 there are several companies with zero sales that either have gone out of business or are 'new technology' companies (developing new products). We show the data for two different bins of initial sales (with sizes increasing by powers of 4): $4^{11.5} < S_0 < 4^{12.5}$ (squares) and $4^{14.5} < S_0 < 4^{15.5}$ (triangles). Within each sales bin, each firm has a different value of *R*, so the abscissa value is obtained by binning these *R* values. The solid lines are fits to equation (1) (in the text) using the mean $\bar{r}(s_0)$ and standard deviation $\sigma(s_0)$ calculated from the data. *b*, Probability density $p(r \mid s_0)$ of the annual growth rate, for three different bins of initial sales: $4^{8.5} < S_0 < 4^{9.5}$ (circles), $4^{11.5} < S_0 < 4^{12.5}$ (squares) and $4^{14.5} < S_0 < 4^{15.5}$ (triangles). The data were averaged over all 16 one-year periods between 1975 and 1991. The solid lines are fits to equation (1) using the mean $\bar{r}(s_0)$ and standard deviation $\sigma(s_0)$ calculated from all data.

Tent shape

- The distribution is double exponential (or Laplace)

$$p(r \mid s_0) = \frac{1}{\sqrt{2}\sigma(s_0)} \exp\left(-\frac{\sqrt{2}|r - \bar{r}(s_0)|}{\sigma(s_0)} \right)$$

Subbotin family of distributions

- As specific distribution to test we consider the Subbotin family

$$p(r) = \frac{1}{2\gamma\beta^{1/\beta}\Gamma(1+1/\beta)} \exp\left(-\frac{1}{\beta} \left| \frac{r-\mu}{\gamma} \right|^{\beta} \right)$$

where μ is the mean, β characterizes the shape (kurtosis decreases with beta) and the standard deviation is

$$s = gb^{1/b} \sqrt{\frac{G(3/b)}{G(1/b)}}$$

- It includes the Laplace (β=1) and the Gaussian (β=2)

- The standard deviation $s(s_0)$ depends on the initial size

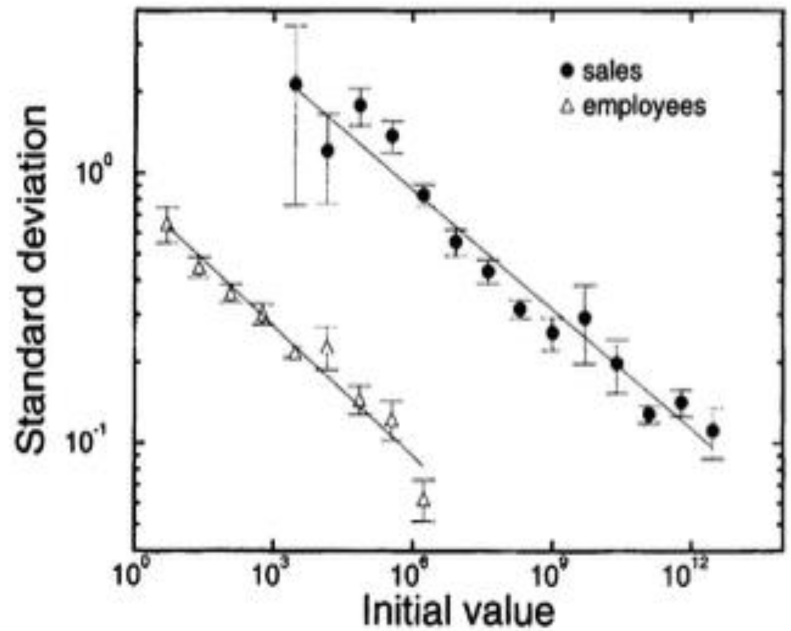


FIG. 2 Standard deviation of the one-year growth rates of the sales (circles) and of the one-year growth rates of the number of employees (triangles) as a function of the initial values. The solid lines are least-square fits to the data with slopes $\beta = 0.15 \pm 0.03$ for the sales and $\beta = 0.16 \pm 0.03$ for the number of employees. We also show error bars of one standard deviation about each data point. These error bars appear asymmetric as the ordinate is a log scale.

More recent empirical works have shown that

- (i) the distribution is slightly asymmetric,
- (ii) the extreme tails are fatter than exponential
- (iii) successive growth rates are slightly correlated,
- (iv) different sectors and subsectors of the economy can have different growth rates and therefore some of the above results might be driven by heterogeneity.

Subsectors

- We considered panels of firms which are homogeneous at the subsector level.
  - For the European Union (Amadeus)
    - Chemical Manufacturing (code 325)
    - Computer and Electronic Product Manufacturing (code 334)
    - Food Manufacturing (code 311).
  - For the US (Compustat)
    - Chemical Manufacturing (code 325)
    - Computer and Electronic Product Manufacturing (code 334)
    - Machinery Manufacturing (code 333).

Bibliography

- R. Axtell, 2001. Zipf Distribution of U.S. Firm Sizes, Science 293, 1818
- Bottazzi, G., Secchi, A., 2006. Explaining the distribution of firm growth rates. RAND J. Econ. 37, 235.
- Fu, D., Pammolli, F., Buldyrev, S.V., Riccaboni, M., Matia, K., Yamasaki, K., Stanley, H.E., 2006. The growth of business firms: theoretical framework and empirical evidence. Proc. Natl. Acad. Sci. USA 102, 18801.
- Lunardi J.T, Miccichè S., Lillo, F.,. Mantegna R.N., Gallegati, M, 2014. Do firms share the same functional form of their growth rate distribution? A new statistical test, Journal of Economic Dynamics and Control 39, 140
- Riccaboni, M., Pammolli, F., Buldyrev, S., Ponta, L., Stanley, H.E., 2008. The size variance relationship of business firm growth rates. Proc. Natl. Acad. Sci. USA 105, 19595.
- Stanley, M.H.R., Amaral, L.A.N., Buldyrev, S.V., Havlin, S., Leschorn, H., Maass, P., Salinger, M.A., Stanley, H.E., 1996. Scaling behaviour in the growth of companies. Nature 379, 804.

Qualche domanda sulla size (1)

- Quantificare la relazione (correlazione) tra le diverse possibili misure di size
- Quale famiglia descrive meglio la distribuzione della size delle aziende italiane? Usare stime non parametriche (istogrammi e kernel) e parametriche (con maximum likelihood)
- Test di ipotesi per forme alternative
- La distribuzione della size in un certo anno delle aziende italiane è con coda a legge di potenza? L’esponente è cambiato col tempo?
- Le risposte alle domande sopra dipendono dall’eterogeneità settoriale e/o geografica?

Qualche domanda sulla crescita (2)

- La legge di Gibrat vale per le aziende italiane?
- Quale famiglia descrive meglio la distribuzione della crescita delle aziende italiane? Usare stime non parametriche (istogrammi e kernel) e parametriche (con maximum likelihood)
- La crescita media è statisticamente diversa da zero in ciascun anno?
- La distribuzione è asimmetrica?
- Esiste una relazione tra varianza della crescita e size?
- Test di ipotesi per forme alternative
- Le risposte alle domande sopra dipendono dall'eterogeneità settoriale e/o geografica?

Qualche domanda sulla crescita nel tempo (3)

- La crescita media è cambiata col tempo? (test di ipotesi)
- La distribuzione della crescita è cambiata nel tempo? (test di ipotesi)
- Come si confrontano crescite annuali con crescite su periodi più lunghi (ad esempio biennali o quinquennali)?
- Le risposte alle domande sopra dipendono dall'eterogeneità settoriale e/o geografica?
- Si può misurare una dipendenza tra la crescita in anni successivi? (predicibilità della crescita)