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)
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 $S_0$. 

![Graph showing probability density of growth rate](image)
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

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

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

\[
= \frac{1}{\sqrt[3]{\frac{(3/\gamma)}{(1/\beta)}}}
\]

• It includes the Laplace (\( \beta=1 \)) and the Gaussian (\( \beta=2 \))
• 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

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)