

Referee Report for the Economics E-Journal

Title: “Investigating the Exponential Age Distribution of Firms” by Alex Coad.

Manuscript ID: DP2010-12.

Referee: Marco Capasso.

Overview:

The paper studies the empirical distribution of firm age, by using the exponential function as a benchmark. The paper bases its theoretical importance in the finding by Coad (2008) that the power law found for the firm size distribution could be related to the exponential distribution of firm age. Such argument finds place within the old debate in the industrial dynamics literature about the causes of a power-law (or the related Zipf distribution) observed for firm size empirical distribution, whereas Gibrat's Law alone would predict a lognormal distribution. Coad (2008) had shown that the exponential distribution of firm age, when paired by a Gibrat law acting for incumbent firms, can lead to a power-law of firm size. The present paper goes along the same line of research, by showing that empirical data resemble an exponential distribution of firm age, and in particular investigates the deviations from the theoretical exponential distribution that are observed in both the lower and upper parts of the dominion of the empirical distribution. The analysis is conducted at different levels of aggregation, and provides important information for the future studies of industrial dynamics.

Comments:

As the author explains in the reply to the first two referees, “the main thrust of the paper is empirical investigation of the aggregate age distribution.” My opinion of the paper is good, in that it provides a good analysis of the characteristics of the empirical distribution of firm age, as observed at different scales. In the same reply, the author says that “criticism of the mathematical model is not crucial for this particular paper”. This is not clear in the way the paper is currently written, as a strong accent is given to the theoretical model by Huberman and Adamic (1999) brought into this context by Coad (2008), which seems to justify the interest in the age distribution. I personally have doubts that such model brings us to a better comprehension of industrial dynamics, as it seems mainly to introduce a new point of view on the previous theoretical models in the literature, whose links with the age distribution have not been properly emphasized.

In the past, explanations of the firm size power law were generally based on entry and exit phenomena (like in Simon and Bonini, 1958) or in the existence of a lower bound of firm size, like in Axtell (2001).

There are two main points I want to stress:

1) We know that there are some violations of Gibrat's Law on incumbents. While old firms seem to behave according to Gibrat's Law, as Lotti, Santarelli and Vivarelli (2009) have shown by following a cohort of firms of the same age over time, when including young firms in the picture (that can still be considered as incumbents) some violations of Gibrat's Law can be found, in particular involving mean and variance heterogeneity. As Axtell (2001) said, the fact that the growth of smaller firms is characterized by higher mean and higher variance should be taken into account by the models of industrial dynamics;

2) Age can easily be linked to size. It is known that new firms have a lower than average size with respect to the whole population. Therefore, considering in the same scenario growth, age and size without investigating the relation between age and size brings to a very partial view of the overall industrial dynamics. The present model is considering entry and exit by means of the hypothesis about the age distribution, i.e. linking younger age to a higher probability of exit, and at the same time keeps the assumption of same growth process for incumbent firms, while attention should be pointed on the fact that for very small incumbent firms (which can be young) it is very difficult to be governed by the same mean and variance of the growth process of large firms (the growth rate distribution of small firms must necessarily be left truncated). In other words, we are assuming that the growth process is the same for all incumbent firms and that young firms have a higher probability to exit; but if we look at the stylized fact that usually young firms are smaller (as suggested by the author citations in the note 3, about the change over time of the size distribution of the same cohort) the two assumptions sound not too much unrelated: can entry and exit be phenomena that are so separated by the growth process of incumbents? Can growth process and age distribution be two separate assumptions?

The two points of my critique can be faced together at the same time, by comparing the paper to the previous literature. The theoretical part of the paper goes more in the tradition of Simon than in the tradition of Axtell (and Gabaix, and the Kesten process) in that it explains the power law by means of entry and exit phenomena rather than of boundaries put on incumbents' size. It does so by showing a stylized fact on the age distribution of firms. However, the exponential age distribution can well be a consequence of a model *à la* Simon rather than an additional assumption. I now show a simple model that can translate the two assumptions given by Coad (2008) into the same language used in the previous literature. I hope this can bring the author and me to a better comparison with the previous literature and to a better comprehension of the economic phenomenon under study.

Suppose that Gibrat's Law holds for all the firms, e.g. that growth (expressed as difference of log size) of all firms at each time period is an event drawn from a normal distribution having mean equal to zero and standard deviation equal to one. Suppose there is a minimum size of the firm, say equal to one, and that whenever an incumbent firm reaches a size lower than one, because of the previously described growth process, it exits the market. Every new time period, ten firms enter the market with the minimum size. Notice that the main assumptions here are: Gibrat's Law, exit when falling below the minimum size, constant entry at the minimum size. Let's now simulate this model for 1000 consecutive time periods (check the appendix of this referee report for the Matlab code I have written). We obtain figure 1, which shows that the age distribution of firms resemble an exponential, with positive deviations at the extremes of the domain and negative deviations in the central part of the domain.

Let's now slightly change the model and include some heterogeneity in the growth process. In particular, we assume that the mean of the growth distribution linearly decreases with size (intercept: 1; slope: -0.01), thus giving a higher probability of growth for smaller firms. We start from period zero, and simulate the model for 1000 consecutive periods. Now the exponential function fits much better the age distribution of firms (figure 2).

One of the many reasons why this simulation cannot be considered as adequate to explain reality lies in the fact that we are not considering the tent shape of growth rates and many other stylized facts. My aim here is just to point out that observing in the real data an exponential age distribution of firms may not be a good reason for assuming it in a theoretical model. In other words, the observation of an exponential age distribution, as such and at this stage of its study, does not imply the need for a new variable (age) to be modelled separately by the traditional ones (size and growth). Exponential age cannot be a primitive assumption, in that firms' death date is not written (in probabilistic terms) in their chromosomes when they are born; instead, it is most often a consequence of the same causes that can in other cases drive a simple decrease in size. However, firm age can well be observed and studied as a separate variable, as its observation can be fruitfully used as a new stylized fact retrieved on empirical data in order to provide a confirmation or a counterfactual for previous models already present in the literature. Only a finer examination of the properties of this distribution will tell us if age is exerting an influence that does not go through size, and the present paper, with its analysis of the departures from the exponential function, provides interesting results in this sense.

My suggestion to the author would then be to emphasize the empirical findings, which represent the original part of the paper, and to make clear the position of the author about the model of Huberman and Adamic (1999). I read at the beginning of Section 2 "We now demonstrate how the age distribution is of interest in theoretical models". My opinion is that the described model is a simple didactic example of how age and size distribution can be linked, but this not the reason why we are interested in the age distribution. I think that the age distribution is interesting because it provides information about entry and exit, and because it connects entry and exit between themselves, which in turn are both connected to the growth process.

I do not have a definitive opinion, and I will respect the opinion of the author if he thinks that the model by Huberman and Adamic provides an alternative explanation to the ones already suggested in the previous literature. If this is not the case, I suggest the author to state it clearly.

A final thing: a hint for future research (read: future papers) can be to go deeper into relating the empirical findings with the theoretical models already in the literature. Simulation exercises, e.g. in the direction of Richiardi (2004), could then be used to check which theoretical models are consistent with an exponential age empirical distribution, both in its fit and in its departures from the theoretical exponential distribution.

References (not already cited in the paper):

Lotti, F., Santarelli, E., and Vivarelli, M. (2009), "Defending Gibrat's Law as a long-run regularity", *Small Business Economics* 32: 31-44.

Richiardi, M. (2004), "Generalizing Gibrat: Reasonable Multiplicative Models of Firm Dynamics", *Journal of Artificial Societies and Social Simulation* 7 (1).

Simon, H. and Bonini, C. (1958), "The Size Distribution of Business Firms" *American Economic Review* 48: 607-617.

Figure 1: Firm age distribution with mean homogeneity.

Firm age distribution (histogram) and exponential fit (line)

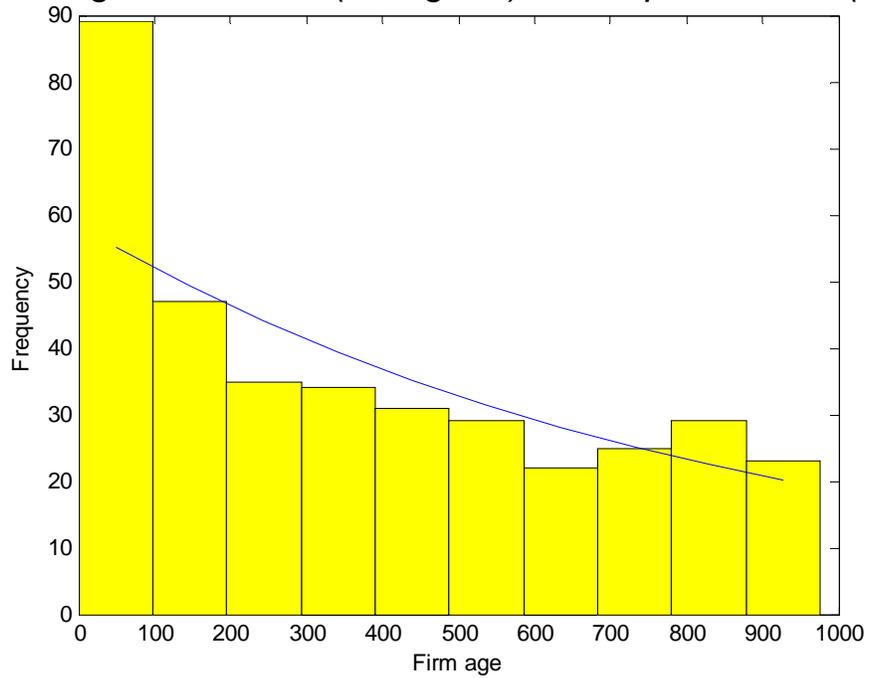
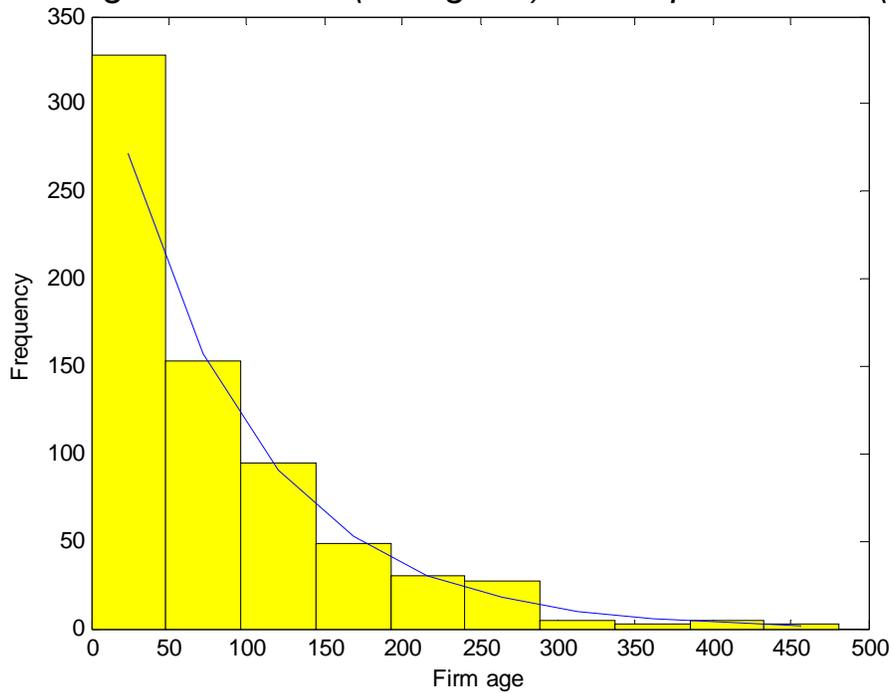


Figure 2: Firm age distribution with mean heterogeneity.

Firm age distribution (histogram) and exponential fit (line)



Appendix (Matlab software code)

```
clear all

mean_homo=input('Do we want Gibrat's Law (type "1") or mean heterogeneity (type "2")? ');

% initialization
number_new_firms=10;
minimum_size_entry=1;
minimum_size=1;
size_vector=zeros(1,number_new_firms)+minimum_size_entry;
age_vector=zeros(1,number_new_firms);

for ttt=1:1000 % we consider a time span of 1000 periods

% new age at each period
age_vector=age_vector+1;

% entry
size_vector(length(size_vector)+1:length(size_vector)+number_new_firms)=minimum_size_entry;
age_vector(length(size_vector)+1:length(size_vector)+number_new_firms)=0;

% growth process
std1=1;

if mean_homo==1
mean1=0; % mean homogeneity (Gibrat's Law)
elseif mean_homo==2
slope1=-0.01; % mean heterogeneity (smaller firms have higher probability to grow)
interc1=1;
mean1=slope1.*size_vector+interc1;
end

growth_rates=randn(1,length(size_vector)).*std1+mean1;
size_vector=size_vector.*exp(growth_rates);

% exit
NOTexits=find(size_vector>=minimum_size);
size_vector=size_vector(NOTexits);
age_vector=age_vector(NOTexits);

end

hist(age_vector)
h = findobj(gca,'Type','patch');
set(h,'FaceColor','y')
[frequencies,bincenters]=hist(age_vector);
logfrequencies=zeros(1,length(bincenters));
logfrequencies(find(frequencies>=1))=log(frequencies(find(frequencies>=1)));
```

```
fit = polyfit(bincenters, logfrequencies, 1);  
line(bincenters, exp(fit(2)).*exp(fit(1)*bincenters));  
  
title('\it{Firm age distribution (histogram) and exponential fit (line)}', 'FontSize', 14)  
xlabel('Firm age')  
ylabel('Frequency')
```