report on

Investigating the Exponential Age Distribution of Firms

Economics E-journal discussion paper 2010-12

In the past few years the availability of large commercial databases recording firm level data made it possible to investigate the empirical distributions of the variables found in the databases. The main merit of the paper is that of looking at the yet rarely explored distribution of firm ages by gathering data from a number of freely available datasets.

The author's trust on the reliability of the exponential distribution in providing a benchmark for the shape of the firm ages distribution pervades the paper. However, in my opinion, the results from the actual level of the data processing are not enough to convince readers as well. In fact, the statement that age distribution of firms is exponential is supported by a visual inspection of figures 2-4. Beside the fact the visual inspection is not a so scientific tool, at a closer look, doubts about the absence of systematic deviations from an hypothetical straight line fitting the main portion of the displayed empirical distributions arise.

I think the paper should be enriched by a section where quantitative tools such as hypothesis testing for equality between the empirical and theoretical distributions should be implemented. In this additional work, the author could identify which is the portion of the distribution better approximated by a theoretical benchmark. In addition, the list of potential candidates for the best fitting theoretical distribution could be enlarged. Of course this call the theoretical result of section 2 into question, but after all, the validity of the Zipf, or more generally the power law distribution of firm sizes (which theoretically justify the exponential distribution of ages) on the whole support seems to be still an open question (Growiec et al. (2008) On the size distribution of business firms. *Economics Letters*, 98, pp. 207 - 212).

In summary, I think the topic is rather new and the author succeeded in finding interesting datasets, but his thesis should be supported by quantitative tests. I encourage the author to make further efforts to better underpin his arguments and eventually to enlarge his investigation to other theoretical distributions to find out which one has the best fit.

Minor remarks

- σ is used in equation (5) but not previously defined;

- Figure 1 could be omitted: readers probably know how the exponential function looks like in a semi-log plot;

- is there a reason for the partitioning of the age space used to build figure 6?;

- there is a missed word on page 13, next-to-last line "In this section we ????? the ...";

- contrary to what expected the previous literature is nod discussed in section 3.1;

- the author should take into account the eventuality to remove the word "Exponential" from the title.