

Referee Report on MS 419

“Investigating the Exponential Age Distribution of Firms”

This paper is an empirical exploration into the aggregate age distribution of firms. The implicit goal is to provide support for an assumption that is an essential part of the theoretical model that the author has presented elsewhere [Coad (2010)]. The model is capable of generating an aggregate firm size distribution that is consistent with empirical findings.

The topic is of substantial interest to theorists studying industry dynamics. In particular, I view its main contribution as identifying certain patterns in the shape of the aggregate age distribution of firms, independently of whether or not such patterns support the assumptions made in the existing models. Empirical investigations such as this provide the base materials for further theoretical explorations. In that sense, the paper makes a useful contribution. On the other hand, as a stand-alone paper, it falls short of delivering what is needed to make it publishable. First of all, the theoretical model that motivates the paper (the first 7 pages or so) has already been presented in a separate paper by the author [Coad (2010)]. The contribution of the current paper is then on the empirical analysis provided in Sections 3, 4, and 5. While the findings reported in these sections are interesting, the analysis is not sufficiently thorough to justify its publication. I provide below a brief summary of the paper and the results, along with my thoughts on where it falls short.

The theoretical model that the author bases his work on falls in the category of stochastic models of industry dynamics. There are many variants in this category, but the particular version that the author works with has the following features:

- 1) Firm growth follows a Gibrat process so that, in a fixed population of firms with heterogeneous size, each firm faces the same distribution of growth rates, independent of their size. This process gives us log-normal distribution of firm sizes for a *given cohort of firms* – i.e., the firms of the same age.
- 2) Given the size distribution for a given cohort, the author then combines it with the assumption on the distribution of firm ages. In particular, the author assumes that firm age is distributed exponentially. When this assumption is integrated into the Gibrat process of firm growth, the resulting distribution of firm size obeys the power law (Zipf's law) which is a frequently observed phenomenon [Axtell (2001)].

Two assumptions are then crucial for the final result that links the firm growth process with the firm size distribution: 1) A firm's growth rate is independent of its size (Gibrat's assumption); 2) a firm age is distributed exponentially. The author appears to take the first assumption as given and focus on the second one – the exponentially distributed firm age. The author shows some empirical evidence that the firm age may indeed be exponentially distributed – see Figures 2-4. Visually examined, the empirically constructed age distribution does appear to be exponential for a wide range of firm ages in the middle. [A more rigorous statistical testing may be necessary to prove this, however.] This is nice, since the use of exponentially distributed firm age in the author's model generates the power law distribution of firm sizes when combined with the Gibrat's assumption on firm growth. However, the

empirically constructed age distribution violates the exponential distribution at the front- and the back-end of the distribution. In other words, the exponential distribution does not work for the very young firms and the very old firms. The author explores the reasons why these violations may occur.

Section 3 looks at the young firms. Assuming constant entry rates, an exponential age distribution in a cross-section of firms implies a constant survival rate. The author makes the point that the constant survival rate is rejected empirically for young firms. The data, in fact, shows that the survival rate increases with age for young firms [Figure 8]. Hence, the author attributes the failure of the exponential age distribution for young firms to the *non-constant* survival rate for that age group.

I have two comments about the materials presented in Section 3. Firstly, the survival probabilities captured in Figure 8 are only for relatively young cohorts. It certainly seems that the survival probability increases in age for these cohorts. But, for older cohorts, we should see *constant* survival probabilities – this is implied by the assumption of constant entry rates as the author states. Why not present (visually) the survival probabilities for older cohorts to provide support for this implied property?

I also worry about the assumption of “constant entry rate.” This may hold at the aggregate level, but we know that they do not at the industry-level. Note that many of the relatively young industries go through a shakeout phase where both the entry and exit rates are quite high. Given this, most of the young firms at any given point in time may be concentrated in the relatively new industries where the shakeouts are in progress. They face a very different entry rate from the older firms that are distributed in more established industries. It is clear that the assumption of constant entry rates is unlikely to hold for younger firms. This may be another reason that the exponential distribution fails for younger firms.

Section 4 deals with the older firms. Here again, the exponential age distribution fails. The world’s oldest firms appear to be much older than the exponential age distribution would suggest [Figure 9]. The author does not have a good explanation for this. He simply waves his hand by simply resorting to a behavioral explanation – “Some old firms, such as family firms, go to great lengths to continue their operations ...” He further states “... these extremely old firms are small in number, and that even though the exponential benchmark for the empirical age distribution is not verified exactly, it remains a useful approximate benchmark in practical terms.” Perhaps true, but not very satisfying.

Final comment: Ultimately, these observed patterns – e.g., Zipf distribution for firm size and exponential distribution for firm age – need to be explained in terms of the relevant economic primitives. I am not confident that the stochastic models such as the one presented in the paper is capable of doing that. Of course, this criticism is not limited to just this paper. It applies to all papers that belong to this category.

References

Axtell, R. L. (2001). Zipf distribution of US firm sizes. *Science*, 293: 1818-1820.

Coad, A. (2010). The exponential age distribution and the pareto firm size distribution. *Journal of Industry Competition and Trade*, forthcoming.