

Referee report on Ainara Gonzalez de San Romn and Yolanda F. Rebollo-Sanz (2014), "An Estimation of Worker and Firm Effects with Censored Data". Economics Discussion Papers, No 2014-28, Kiel Institute for the World Economy.

This paper provides an algorithm for the estimation of a linear model with two-sided heterogeneity and a censored outcome variable. The algorithm's properties is illustrated using both simulated and real data. The algorithm outperforms the Tobit model in terms of estimation speed, but at the cost of higher variance and therefore loss of precision.

1 Major Comments

- It seems that a central assumption for the algorithm to work is to assume a particular distribution of the error terms. However, the choice of a normal distribution is not motivate anywhere. Since identification comes from the distribution of the error terms, I think that the authors should spend more time in the paper in order to discuss the choices taken in this regard. Also, I wonder if it would be possible to estimate the model for those workers that are not censored and then use the empirical distribution of the estimated error terms to check if they are normally distributed?
- The authors have chosen a normal distribution, which I am fine with given the caveat above. However, it would be interesting to do a robustness check for different distributions. I.e. if one assumes that the errors terms are normal, but they are really e.g. from a t-distribution, how does that affect the results.
- Given the assumption of normally distributed errors would it not be feasible just to do maximum likelihood in the model without worker or firm effects?
- P. 5: The authors note: "The properties derived also apply to the general case of both-side coding and to more general models.". If this is the case then I think that the authors should provide these proofs. In particular, I do not like the fact that the authors use a lot of space and energy to show consistency, how to compute standard errors, etc., for the model with no heterogeneity. This is not the contribution and is ultimately not that interesting from my point of view.
- P. 18: It is a little hard to find out exactly how the assignment of workers to firms is done. I hope that the authors can elaborate a little on this.

2 Minor Comments

- P. 12: The authors write that the FLS algorithm converges faster than Tobit. What is the gain in CPU time?
- P. 17: The authors note that an assumption behind identification is that some workers move firms. It is more than that. In order to identify the model all workers should be connected in a graph-theoretical sense. I.e. This simply means that all workers and rms have to be connected by other rms or workers. This is first mentioned a lot later in the paper.
- P. 30: The reference to Bagger et al. (2010) cannot be found in the reference list.