Report on "An estimation of worker and firm effects with censored data", MS 1247

The authors propose a simple estimator for linear regressions with a censored dependent variable. The method relies on iterations of OLS regressions, during which the data points for the censored observations are imputed by draws from a censored normal distribution, using the coefficient estimated from the previous iteration. The authors call this method fill-in iterated least squares (FILS). They derive and discuss the properties of FILS, implement several Monte Carlo simulations, and provide an empirical application to Spanish matched employer-employee data. They emphasize the usefulness of this method to deal with censored data in applications with two high-dimensional fixed effects, such as for example worker and firm fixed effects.

I think the paper is relevant, because it deals with a dilemma many empirical economists face when using administrative matched employer-employee data, or other panel data: These datasets are very well suited to include high-dimensional fixed effects (such as worker and firm fixed effects), but in many of the available administrative datasets, earnings or wages are often censored. This paper addresses the question of what the consequences of censoring are, if ignored, and compares it with a novel method, FILS. This is very useful, and I believe the analysis in the paper is correct. It was, however, not entirely clear to me whether the proofs of the properties of FILS are all complete, and whether they extend to the model with fixed effects or not. Also, I am wondering how generalizable the results from specific simulation and from an empirical application to real data are to other datasets and data generating processes.

My comments in detail are the following:

1. I was a bit confused about whether the proof of convergence in section 3.1 was complete or not. At the end of that section, the authors say: "The simulation results reported in the next section confirm that the parameterised function is contractive because convergence occurs most of the time." Does this mean that section 3.1 does not contain a full analytical proof of convergence? Instead, convergence is inferred from the fact that in a set of simulations, convergence is almost always achieved? This, then, however does not seem to proof convergence generally, but only for the data generating processes used for the simulations?

2. In Section 3.2, consistency is shown based on an "Identification Assumption", and based on taking sigma as known. It is not clear to me whether this is a general proof of consistency, or whether again this incomplete (for example because sigma will in general not be known)? Also, as an applied empirical economist, I would benefit from an intuitive description of what the identification assumption means (last formula in section 3.2). Given that consistency follows from it, how strong an assumption is it? Are the instances in which it might be likely to be violated?

3. I think a crucial point when applying this estimator to censored data in a worker (or worker and firm) fixed effects model is the incidental parameters problem discussed in the last paragraph of section 5.1. For N going towards infinity with a fixed T, the fixed effects are inconsistent. My understanding is that in many nonlinear models (such as for example the probit and the tobit model), the inconsistency of the fixed effects contaminates also the other parameter estimates, and therefore we usually do not use fixed effects probit or tobit models. So in principle, FILS applied to censored data with high-dimensional fixed effects is inconsistent, right? If that is so, it should be emphasized more clearly throughout the paper. For example, in the before-last paragraph of section 5.2.1 the authors suggest that the estimator is consistent ("we observe that consistency is still guaranteed…").
If it is true that FILS is generally inconsistent in fixed effects models, the paper still makes a valuable contribution. If the choice is between not taking censoring into account at all in a fixed effects model, and employing FILS, then this might be a choice between two inconsistent estimators. So the question is then which of the two is less biased. This is where I see the contribution of the paper and where I find its results relevant and interesting. (And for T towards infinity, the incidental parameters problem goes away, for this case FILS then becomes of course even more appealing.)

4. The results in Table 6 seem to suggest that in a worker and firm fixed effects model with censoring of the dependent variable, FILS performs better than an AKM regression. This is an interesting result. In particular, the AKM regression seems to estimate the correlation between worker and firm fixed effects too negative. The same seems to be the case in the empirical application to real data (Table 9). My comment here is that it is somewhat unclear how generalisable the results from these simulations and from the empirical application are. Will AKM as compared to FILS always bias the correlation downwards? Or can it in other samples and application also lead to an upward bias? What do the direction and the size of the bias depend on?