**Responses to Referees' Comments**

1. **Comment: On page 2, the authors introduce the FGLS/Parks estimator. I think the section should contain more details concerning common applications of the Parks estimator.**

   Response. The first paragraph now states the following: "*The Parks model (Parks, 1967) remains the most commonly used estimation procedure for simultaneously handling cross-sectional and serial correlation. For example, the options available with the Stata command "xtgls" are all variations of the Parks model. Recent applications include Congleton and Bose (2010); Stallman and Deller (2010); Kebede, Kagochi, and Jolly (2010); and Roll, Schwartz, and Subrahmanyam (2009). A quick search of papers in Web of Science that cite Parks (1967) produces hundreds more*."

2. **Comment. On page 3, the authors motivate the use of the jackknife by stating that there are no bootstrapping procedures that are valid for the simultaneous occurrence of cross-sectional correlation. I think it may be worthwhile to explain why and to motivate this statement in detail.**

   Response. This was an overstatement. There are no published bootstrapped procedures for the Parks model. There are many types of bootstrapping procedures, some of which are feasible. However, many standard procedures, such as the block bootstrap are not. The last paragraph on page 3 now states: "*While jackknifing with panel data characterized by both serial and cross-sectional correlation is not without its challenges (as we discuss below), it stands in contrast with bootstrapping. To date, no successful bootstrapping procedures have been developed for the Parks model. For example, block bootstrapping techniques have been developed for one-way clustering such as serial correlation or cross-sectional correlation (e.g., Cameron, Gelbach, and Miller, 2008). However, there are no block bootstrapping procedures that are valid for the simultaneous occurrence of both of these. One can resample "blocks" of observations, where the blocks are clusters based on groups or clusters based time, but one cannot do both. Relatedly, newly developed techniques exist for calculating robust standard errors with multi-way clustering such as group and time (Cameron, Gelbach, and Miller, 2006), but these procedures do not allow cross-sectional and serial correlation to interact, as in the Parks model.*" The last sentence is explained in Footnote 3 on page 5.

3. **Comment. Messemer and Parks (2004) propose using the bootstrap to attenuate level distortion in the estimated covariance matrix of the Parks estimator. How does this technique relate to the previous statement and more generally with the findings of the paper?**

   Response. Several years ago I tried to replicate the results of Messemer and Parks (2004). I was unsuccessful. The problem comes in Steps (2b) and (2c) on page 11 of their work. One has to estimate the covariance matrix, then use this as the population error-variance covariance matrix for the DGP, and then generate corresponding

residuals. Iterative resampling from this estimated error variance-covariance matrix is at the heart of the bootstrap procedure. I could never get results that came close to the performance of Messemer and Parks. I contacted both Messemer and Parks (Parks was Messemer's Ph.D. supervisor) for help with the code that they used, and they were unable to provide me their code. We note that their work has never been published.

4. **Comment. On pages 6-7 the authors illustrate their Monte Carlo experiment. Although they quote a previous study (Reed and Ye, 2010) using the same Monte Carlo design, the source of the datasets they use in the simulation is not specified. More generally, I think this section needs to be expanded in order to provided a clearer description of their Monte Carlo design.**

Response. Section IV of the paper has been completely rewritten and expanded according to the referee's comments.

5. **Comment. The Monte Carlo experiment does not compare the performance of the jackknifing estimator with the Panel-Corrected Standard Error (PCSE) estimator of Beck and Katz (1995). Would it be possible to make such a comparison?**

Response. We have done that in other work. Please see TABLE 5 in "The PCSE Estimator is Good – Just Not as Good As You Think," *Journal of Time Series Econometrics* Vol. 2, No. 1 (2010): Article 8. The revised manuscript now cites that work (see Footnote 7).

6. **Comment. If the jackknife procedure is somehow related to the bootstrap methodology of Messemer and Parks (2204), it would be interesting to replicate the Monte Carlo experiment including the bootstrap estimator and using the same, easily accessible, data of Messemer and Parks (2004).**

Response. Please see the response to Comment 3 above. Given our difficulty in replicating their results, we did not think it was advisable to do this.

7. **Comment. On page 9, the third reference should be "Cameron A.C. …"**

Response. The revised manuscript corrects this mistake.

8. **Comment. Instead of reporting the differences between the coverage rates I would just report the two coverage rates side by side in one table.**

Response. We prefer to report first differences, then jackknifing coverage rates (see TABLES 1 and 2) as each table corresponds to the two main results we emphasize on page 10. The FGLS(Parks) coverage rates are easily backed out from the differences. If the referee feels strongly about this, we are open to changing as he/she suggests.