RESPONSES TO REVIEWER 1

General: We thank the reviewer for many excellent comments. In the space below, we discuss how we will address these comments if given the chance to revise and resubmit our manuscript.

We should mention that one of the authors (Reed) has a long history of researching the PCSE estimator, as indicated by these listings in the references:

- Chen, X., Lin, S., & Reed, W.R. "A Monte Carlo Evaluation of the Efficiency of the PCSE Estimator." *Applied Economics Letters,* 17 (2010):7-10.
- Reed, W.R. & Webb, R. "The PCSE Estimator is Good – Just Not as Good as You Think." *Journal of Time Series Econometrics*, Vol. 2, No. 1 (2010): Article 8.
- Mantobaye, M., Rea, W., and Reed, W.R. "Which Panel Data Estimator Should I Use?: A Corrigendum and Extension." *Economics: The Open-Access, Open-Assessment E-Journal*, 12 (2018-4): 1–31.

We mention this because some of what the reviewer asks us to do has been done in these other papers. These previous papers were written to (i) demonstrate that the Parks estimator had smallest MSE in finite samples relative to a large number of competitor estimators (including OLS and the PCSE estimator), (ii) contest Beck and Katz's (BK's) claim that PCSE was just as efficient as Parks, (iii) point out that many of BK's claims are based on simulations where the DGP has no serial correlation, and (iv) demonstrate that the PCSE's Type I error rates are sometimes substantially larger than alpha.

While these previous papers pointed out deficiencies in the PCSE estimator, they were unable to supply a satisfactory solution to Parks's poor performance in estimating standard errors - *the raison d'être of the PCSE estimator*. That is the purpose of the current paper. We believe the combination of Parks with bootstrapped inference removes any reason to use the PCSE estimator when T>N. Given the widespread adoption of the PCSE estimator, we believe our paper makes a significant contribution and should be published. Specific responses to the reviewers comments follow.

1. **Comment: "References include Rilstone and Veall which is close to what the authors are doing. Rilstone and Veall looked at bootstrapping inferences for the basic SUR model. It should be mentioned early and contextualized by saying that the current paper does a similar analysis as Rilstone and Veall with the complications of serially correlated errors."**

   Response: We propose to include the following response in our revised version: "Rilstone and Veall (1996) showed improved performance for confidence intervals based on a parametric bootstrap in the context of a simple SUR model. Their paper helped to shift the focus of bootstrap work toward test statistics and away from standard errors, based on then-recent theoretical work on the bootstrap. An important contribution of the present paper is the development of a non-parametric bootstrap for the more complicated case of a SUR model with serially correlated errors." We will also explain why the addition of serial correlation to the SUR model poses substantial technical challenges since one can no longer simply block bootstrap on cross-sectional units à la Rilstone and Veall.

2. **Comment: "…in the preliminaries the authors should state the GLS estimator and provide its asymptotic distribution, noting that FGLS, using standard consistent estimates of the $\rho_i's$ and Σ, will be asymptotically equivalent. Also mention variety of estimators would work and discuss why the ones that are popular are used."**

Response: The revised version will state the GLS estimator and provide its asymptotic distribution. It will note that FGLS with standard consistent estimates of the $\rho_i's$ and Σ is asymptotically equivalent.

We will also point the reader to Mantobaye, Rea, & Reed (2018), where the relative performance of a large number of alternative estimators are assessed (see Table 2 below, taken from MR&R). MR&R compared the performance of these estimators with respect to MSE and Coverage Rates in a wide variety of "realistic research environments". The list of estimators they assessed are reported in their Table 2. Note that several OLS estimators with robust standard errors are included on this list.

Table 2: List and Description of Panel Data Estimators to Be Studied

Estimator	*Procedure*	*Assumed Error Structure*
1	OLS-1A	IID
2	OLS-1B	Robust heteroskedasticity
3	OLS-1C	Robust heteroskedasticity + Robust autocorrelation
4	OLS-1D	Robust heteroskedasticity + Robust cross-sectional dependence
5	FGLS-1A	Groupwise heteroskedasticity
6	FGLS-2	Groupwise heteroskedasticity + autocorrelation
7	FGLS-3 (Parks)	Groupwise heteroskedasticity + autocorrelation + cross-sectional dependence
8	FGLS-4 (PCSE)	Groupwise heteroskedasticity + autocorrelation + cross-sectional dependence
9	FGLS-1B	Weight = Groupwise heteroskedasticity Var-Cov = Robust heteroskedasticity + Robust cross-sectional dependence
10	FGLS-1C	Weight = Groupwise heteroscedasticity Var-Cov = Robust heteroskedasticity + Robust autocorrelation
11	FGLS-1D	Weight = Groupwise heteroskedasticity Var-Cov = Robust heteroskedasticity

Source: Mantobaye, M., Rea, W., and Reed, W.R. "Which Panel Data Estimator Should I Use?: A Corrigendum and Extension." Economics: The Open-Access, Open-Assessment E-Journal, 12 (2018-4): 1–31.

3. **Comment: "This is a purely Monte Carlo paper so I would expect that a substantially wider range of results be reported (where viable or at least summarized) including different estimators."**

Response: We agree with the reviewer. The revised version will report additional bootstrapping results to supplement the results of TABLE 3.

4. **Comment: "My reading of the Beck and Katz paper is that their intent was to outline certain shortcomings of the Parks method and propose some alternatives….Their point seems to be to use OLS but with correct standard errors. (Beck and Katz is not**

the paper I'm reviewing but my comments are relevant in this regard). OLS always has its role and reporting it along with corrections for its distribution are important. I'm not sure if Beck and Katz got it right wrt correct standard errors."

Response: Please see our response to Comment #9. As we note there, BK were misleading in their characterization of PCSE as OLS with "corrected standard errors". PCSE only produces estimates equivalent to OLS in the case of zero serial correlation.

We agree with the reviewer that readers may find comparisons with OLS to be of interest. MR&R reports an extensive set of performance results. For example, MR&R's Table 7 compares efficiency of the different estimators with OLS. Values less than 100 indicate that the respective estimators are more efficient than OLS. Note the relative performance of the Parks and PCSE estimators. The revised version of the manuscript will point the reader to these results.

Table 7: Comparison of Estimator *EFFICIENCY*

Estimator	*Average EFFICIENCY*		*Percentage of Times the Estimator Is More Efficient Than OLS*	
	T/N > 1.5 (1)	*T/N ≤ 1.5* (2)	*T/N > 1.5* (3)	*T/N ≤ 1.5* (4)
	REED AND YE (2011) DATASETS			
Estimator 5/9/10/11	96.6	84.2	68.8	78.9
Estimator 6	82.8	74.7	75.0	89.1
Estimator 7 (Parks)	45.5	66.1*	100.0	100.0*
Estimator 8 (PCSE)	86.9	89.1	62.5	72.7
	NEW DATASETS			
Estimator 5/9/10/11	70.8	54.5	88.6	97.9
Estimator 6	61.5	48.0	97.7	99.0
Estimator 7 (Parks)	46.9	80.1*	97.7	96.4*
Estimator 8 (PCSE)	85.1	92.1	95.5	80.2

Source: Mantobaye, M., Rea, W., and Reed, W.R. "Which Panel Data Estimator Should I Use?: A Corrigendum and Extension." Economics: The Open-Access, Open-Assessment E-Journal, 12 (2018-4): 1–31.

5. **Comment: "The simulations from the Grunfeld data should be rerun making sure to include the following.**
   - **inferences based on both OLS and Parks estimators**
   - **inferences based on asymptotic standard errors under correct specification**
   - **inferences based on bootstrapped standard errors**
   - **inferences based on HAC consistent standard errors**
   - **inferences based on critical points using the χ^2 distribution**
   - **inferences based on bootstrapped critical points"**

Response: Much of this work has already been done in previous research. For example, see Table 8 in MR&R (reproduced below). The revised version will alert the reader to these results and reiterate why we focus on the pairwise comparison of Parks with PCSE: Parks is generally more efficient than all other estimators in situations where it can be estimated. However, it suffers from poor standard error estimation. The PCSE estimator greatly improves on Parks, producing more reliable inference, albeit at the cost of efficiency. Our paper makes clear that one can have the best of all worlds, efficiency via Parks and reliable inference via bootstrapping. In other words, whenever Parks can be estimated, there is no reason to use PCSE. Given the widespread use of PCSE, we believe this result will be of great interest to researchers.

Table 8: Comparison of Estimator Coverage Rates

	Autocorrelation < 0.30		*Autocorrelation ≥ 0.30*	
	Coverage (1)	*|95 – Coverage| (2)*	*Coverage (3)*	*|95 – Coverage| (4)*
REED AND YE (2011) DATASETS (T/N ≥ 1)				
Estimator 1	65.7	29.3	90.9	6.1
Estimator 2	64.1	30.9	91.0	4.9
Estimator 3	86.5	8.9	88.6	6.5
Estimator 4	60.1	34.9	91.5	3.8
Estimator 5	59.8	35.2	88.6	6.4
Estimator 6	88.0	7.1	90.9	4.4
Estimator 7 (Parks)	42.9	52.1	45.6	49.4
Estimator 8 (PCSE)	89.5	5.5	92.7	2.3
Estimator 9	51.9	43.1	85.6	9.4
Estimator 10	70.5	24.5	77.3	17.7
Estimator 11	58.5	36.5	88.2	6.8
NEW DATASETS (T/N ≥ 1)				
Estimator 1	87.9	9.3	74.6	21.4
Estimator 2	83.1	11.9	73.2	22.0
Estimator 3	88.4	6.6	90.2	6.7
Estimator 4	83.6	11.4	74.7	20.3
Estimator 5	85.8	9.2	73.4	22.9
Estimator 6	90.9	4.1	90.9	5.7
Estimator 7 (Parks)	38.1	56.9	42.0	53.0
Estimator 8 (PCSE)	91.4	3.6	92.1	3.3
Estimator 9	75.5	19.5	64.7	30.3
Estimator 10	68.9	26.1	72.7	22.4
Estimator 11	80.6	14.4	68.6	26.4

Source: Mantobaye, M., Rea, W., and Reed, W.R. "Which Panel Data Estimator Should I Use?: A Corrigendum and Extension." Economics: The Open-Access, Open-Assessment E-Journal, 12 (2018-4): 1–31.

6. **Comment: "report true GLS results as benchmark."**

   Response: This is an excellent suggestion. The revised version will report GLS results as a benchmark.

7. **Comment: It would be interesting to report the MSEs of OLS and Parks from the Monte Carlo. If Parks is the correct specification it may have greater variance but better MSE.**

   Response: This has been done in previous research. Please see the response to Comment #4 above. The revised version will point readers to these results.

8. **Comment: Also report average of estimated standard errors to the standard deviation of the estimates to show how wrong the estimated standard errors are.**

Response: This has been done in previous research, albeit with a focus on coverage rates. Please see Table 8 in the response to Comment #5 above. The revised version will point readers to these results.

9. **Comment: I'm not certain that equation (7) (and hence (8)) is correct. From my reading, I believe Beck and Katz suggest using OLS along with "correct" standard errors for OLS. I may be wrong.**

Response: This is an astute observation by the reviewer, because BK are, at best, vague about the relationship between PCSE and OLS. The key Monte Carlo experiments that BK rely on to highlight the efficiency properties of their PCSE estimator are reported in their Table 5. While not obvious, these assume zero serial correlation.

This is implied by the fact that the table only reports values for cross-sectional correlation, not serial correlation. The assumption is made explicit in a footnote. In the first paragraph under the subheading "Ordinary Least Squares with Panel-corrected Standard Errors", Footnote #19 states, "We deal only with the performance of these estimators assuming that serial correlation has already been eliminated."

In other words, when BK talk about OLS with PCSE, they are assuming there is no serial correlation. In that case, the Prais-Winsten transformation matrix ***P*** simplifies to the identify matrix, so that Equation (7) becomes,

$$(7) \qquad \widehat{\boldsymbol{\beta}}_{\boldsymbol{PCSE}} = \left(\boldsymbol{X}'\widehat{\boldsymbol{P}}'\widehat{\boldsymbol{P}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\widehat{\boldsymbol{P}}'\widehat{\boldsymbol{P}}\boldsymbol{y} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = \widehat{\boldsymbol{\beta}}_{\boldsymbol{OLS}}$$

This is why BK use OLS and PCSE synonymously. However, the equivalence only holds in the special case when there is no serial correlation. The general formulation of the PCSE estimator incorporates the Prais-Winsten transformation.

We are certain that our formulation of the PCSE estimator is correct. In the work underlying MR&R, we compared the estimates from our hand-written program of the PCSE estimator with those using Stata's xtpcse procedure applied to the same data. We obtained identical results.

10. **Comment: It would be interesting to see what evidence that the actual data set of Grunfeld used as the basis for the Monte Carlos has the SUR(1) structure. Did the authors (or others who worked on this data) do any testing for these features?**

Response: We didn't test the Grunfeld dataset. However, in MR&R, we reported estimated AR(1) parameters and mean cross-sectional correlations for all the datasets underlying that paper's experiments, which consisted of an eclectic mix of time-series, cross-sectional datasets. These can be found in Table 6 of MR&R (see below).

We found that SUR(1)-type behavior was pervasive across datasets, as evidenced by the substantial first-order serial correlation and cross-sectional dependence in those datasets. The revised version will reference these findings in motivating the SUR(1) model.

Table 6: Description of Simulated Datasets Used in the Experiments

		Heteroskedasticity	***Autocorrelation***	***Cross-sectional Dependence***
		REED AND YE (2011) DATASETS		
$N \leq T$ (80 datasets; 880 observations)	***Minimum***	1.21	-0.06	0.20
	Mean	1.68	0.36	0.44
	Maximum	2.35	0.78	0.90
$N > T$ (64 datasets; 640 observations)	***Minimum***	1.34	-0.04	0.22
	Mean	1.76	0.34	0.43
	Maximum	2.25	0.79	0.79
		NEW DATASETS		
$N \leq T$ (72 datasets; 792 observations)	***Minimum***	1.26	0.08	0.22
	Mean	4.47	0.47	0.35
	Maximum	40.21	0.73	0.52
$N > T$ (68 datasets; 680 observations)	***Minimum***	1.47	0.16	0.23
	Mean	6.93	0.46	0.34
	Maximum	34.91	0.73	0.49

Note: For more details on the construction of the simulated datasets, see Section 2 in the text.

Source: Mantobaye, M., Rea, W., and Reed, W.R. "Which Panel Data Estimator Should I Use?: A Corrigendum and Extension." Economics: The Open-Access, Open-Assessment E-Journal, 12 (2018-4): 1–31.

11. **Comment: Only need to mention "2400 Web of Science references" once.**

Response: The revised manuscript will only mention this once. But note that the number of Web of Science references is now 2,555.

12. **Comment: Typo in Bruckner reference.**

Response: We will fix the typo in the Bruckner reference.

13. **Comment: Remove "innovative" in Abstract. Lots of Monte Carlo experiments are designed using moments from published data sets as basis…Remove the term "Pareto-improving" in Abstract. It's inappropriate here."**

Response: We will remove "innovative" and "Pareto-improving" from the Abstract.