Another Look at "Which Panel Data Estimator Should I Use?"

Author: Mantobaye Moundigbaye

Abstract: *Panel data can be characterized by complex error structures. Heteroskedasticity, serial correlation, and cross-sectional dependence are all likely present in many empirical applications. The presence of these nonspherical errors can generate inefficiency in coefficient estimation and bias in the estimation of standard errors. Unfortunately, robust estimators that accommodate all three sources of nonspherical error behaviour do not exist. This creates a confusing situation for researchers using panel data. On the one hand, there is a plethora of panel data estimators available from statistical software packages like EViews, LIMDEP, RATS, SAS, Stata, TSP, and others. On the other hand, the finite sample performances of these estimators are not well known. Thus, it is not clear which estimator one should use in a given research situation. To address this situation, Reed and Ye (2011) performed an extensive set of Monte Carlo analyses to measure the finite-sample performance of a large number of estimators in "realistic data environments." Their analysis compared estimator performance on two dimensions: (i) efficiency and (ii) coverage rates; leading to a series of recommendations regarding the "best" panel data estimator to use in specific data environments. The paper has been relatively well-cited in Web of Science. Unfortunately, there is a flaw in their Monte Carlo design. Our research (i) corrects the flaw in their Monte Carlo simulation design; (ii) replicates their analysis; and (iii) identifies the extent to which this flaw affects their recommendations.*

Table of contents

1.1. Introduction 1

1.2. Experimental design: old vs. new versions 3

1.2.1: Brief review and criticism of the original experimental design 3

a. Construction of data specific inputs 3

b. Experimental parameters of interest 5

c. Criticism of the initial experimental design 8

1.2.2. Modified experimental design 8

1.3. Description of replicated and redesigned experiments' output 9

1.3.1. Relative efficiency of estimators 9

1.3.2. Confidence interval precision 10

1.4. Further re-examination of estimators' performances 11

1.4.1. Assessment of the original recommendations 11

a. Recommendation 1 13

b. Recommendation 2 14

c. Recommendation 3 15

1.4.2. Further implications of the redesign of experiments 16

1.5. Conclusion 18

References 19

1.1. Introduction

Modeling cross-section and serial dependence in panel data is motivated by strong evidence of such complications in the context of panel models. Regular interactions among peer units such as economic agents, unobserved common shocks and factors, and the consistency of individual units' behaviour over time (adaptive decision making process) are examples of potential sources of error dependence – among and within units - in models using panel data. Even though unobserved time-constant influences could be addressed by heterogeneous parameters (intercepts and slopes), the residuals independence hypothesis may still be violated due to time-varying random effects that persist for more than a single period (Skrondal and Rabe-Hesketh, 2008).

These complex error structures, if not properly handled, may result in serious consequences including, but not limited to, biased coefficient estimates[1], inaccurate hypothesis testing, erroneous statistical inference, and *in fine* misleading analyses and policy recommendations. In the literature, discussions about sources and consequences of cross-units and serial correlation of panel model residuals can be found in Kézdi (2004), Hoechle (2007), Petersen (2009), Sarafidis and Wansbeek (2012), and Chudik and Pesaran (2011), to name but a few relatively recent works.

The need for panel estimators that are robust to cross-sectional dependence and/or serial correlation has been met with an abundance of such estimators proposed in the literature and programmed in statistical packages. Estimators robust to both cross-sectional dependence and serial correlation or to either one of the two exist. Among earlier researchers interested in modelling these complications are Zellner (1962, 1963), Zellner and Huang (1963), Zellner and Theil (1962), and Parks (1967). Additionally, there are a whole new set of sophisticated estimators when handling panel data characterized by cross-sectional and serial correlation. More recent treatments building on flaws and limitations of the earlier attempts include Pesaran and Smith (1995), Beck and Katz (1995), Driscoll and Kraay (1998), Cameron et al. (2006, 2009,

[1] Coefficient estimates are biased only if the source of dependence is significantly correlated with at least one covariate. If this is not the case, coefficient estimates will be consistent.

2011), Bond and Eberhardt (2009), or Kapetanios et al. (2011). Various panel estimators robust to serial and cross-sectional correlation are proposed, with different underlying assumptions about the data structure. Though these necessary assumptions are critical for estimating and diagnosing panel models, they may be considered too strong when applied to general forms of relationships in the error terms characterized by serial correlation and cross-section dependence.

As a result, the plethora of robust panel estimators may appear to be good news for researchers using panel models. For those having little knowledge or an unclear idea of the implications of the complicated errors on the output produced by statistical packages, there are many robust estimators to choose from. Yet, for researchers aware of these implications, this situation poses another problem of estimator selection. At least two motives could be invoked as evidence for this problem: (i) simulation work has provided evidence of performance differences in the proposed estimators, and (ii) no straightforward rule to guide users in the estimator selection exists. While the first motive is a direct result of differences in underlying assumptions used for different estimators, the second constitutes a limit for many applied econometric investigations. In many instances, interest is mostly focussed on differences between estimators' relative performances while the concern about the optimal choice criteria is limited.

A recent attempt to compare panel data estimators was made by Reed and Ye (2011). A benefit of their study is that they provided a framework and a point of comparison for evaluating panel data estimators. A deficiency of their study was that they did not include many recent panel data estimators. Another deficiency was that they introduced an error in their data simulation procedures that calls into question their results. I begin by replicating their study. I do this both (i) to gauge the consequences of their mistake, and (ii) to calibrate my own efforts at simulating panel data estimators in the presence of cross-sectional and serial correlation.

The remaining of the paper is organised as follows. Section 2 presents the original and modified experimental designs. The output under both experimental designs is presented in Section 3. I undertake a comparison of the implications of the two experimental designs in Section 4 to attempt to validate Reed and Ye's (2011) recommendations. Section 5 concludes.

1.2. Experimental design: old vs. new versions

1.2.1: Brief review and criticism of the original experimental design

The experimental design in Reed and Ye (2011) partially followed the traditional practice of simulation in empirical econometrics. Altogether, the experiments were implemented in the following three steps.

(i) The first step consisted in generating a series of dependent data using a pre-determined model specification - or in other terms the data generating process (DGP), exogenous covariates, and a priori coefficient and disturbance parameters generally provided for by the researcher. This step guaranteed that the experiments are controlled.

(ii) In the second step, the same covariates were regressed on the generated dependent variable, using the same model specification.

(iii) These two first steps were iterated a large enough number of time so to allow for a realistic interpretation of estimated parameters' finite properties in the third step.

In Reed and Ye (2011), the DGP was a simple static panel model with single covariate. The originality of their approach was that the model independent variable, intercept and residuals parameters were dataset-specific rather than guesses by the researchers as we usually come across in empirical research. For this purpose, a total of four different macro-economic datasets of various characteristics (level and growth rates) and geographic coverage (US datasets at State level, and worldwide datasets at country level) are used, suggesting a rich assortment of associations among data points across and within individuals in time. This forms the ground for the authors' claim that their generated datasets looked like a "real-world" ones. Key technical details of Reed and Ye (2011) are presented in the next sub-sections, followed by the major criticism of the approached adopted by the authors.

a. Construction of data specific inputs

As noted in the paragraph above, the regressor, the intercept and error variance covariance matrix used to run simulations in Reed and Ye (2011) were dataset-specific. Different data sets are used to construct these simulation inputs following an identical procedure. For a given data set and panel dimension N (number of units) and T (number of time periods) with N and T less than or equal to the total number of individual and time periods in that data set, multiple static panel fixed effects regressions with a single exogenous covariate were run using successive data slices or windows of length T for each individual[2]. The residuals from these regressions were collected and used to form the simulation disturbance parameters. The dependent variables on the one hand and the exogenous covariates on the other were averaged across the slices (see equations (2) and (3) below) to compute the DGP intercept parameter.

Model (1) describes each window's treatment at this preliminary stage of the experimental design.

$$
\begin{aligned}
Y_{it} &= \boldsymbol{\beta} X_{it} + u_{it} \\
u_{it} &= \begin{cases} \mu_i + \varepsilon_{it} \\ \text{or} \\ \mu_i + \eta_t + \varepsilon_{it} \end{cases} \qquad (1) \\
\varepsilon_{it} &= \boldsymbol{\rho} \varepsilon_{i,t-1} + r_{it}
\end{aligned}
$$

where Y_{it} is the NTx1 vector of the dependent variable observed on i = 1, 2, …, N individuals over time periods t = 1, 2, …, T; X_{it} is a NTx1 vector of observed deterministic regressor values; $\boldsymbol{\beta}$ is the slope coefficient, u_{it} is the error term including unobserved fixed or random effects; μ_i is the individual fixed or random effects; η_t is the time specific effects; ε_{it} is the error term assumed autocorrelated of first order; $\boldsymbol{\rho}$ is the common autocorrelation coefficient of ε_{it}, and r_{it} is a white noise.

Both specifications of the error decomposition above were adopted as separate experiments with each data set and a given pair of panel individual and time dimensions.

Averages of the regressands and the regressors were calculated for each individual i over the different windows as:

$$\bar{Y}_i = \frac{1}{N}\sum_{k=1}^{N} y_{ik},\ y_{ik} = (Y_{i,k}, Y_{i,k+1}, \ldots, Y_{i,k+T-1}) \qquad (2)$$

and

$$\bar{X}_i = \frac{1}{N}\sum_{k=1}^{N} x_{ik},\ x_{ik} = (X_{i,k}, X_{i,k+1}, \ldots, X_{i,k+T-1}) \qquad (3)$$

[2] A total maximum of 31 regressions for each set of N and T are run with each data set at this step.

The simulation intercept is the difference $\overline{Y_\iota} - b_x.\overline{X_\iota} = b_0$ where b_x is the only parameter created by the researchers. $\overline{X_\iota}$ is used as regressor for the simulated data sets.

The simulation disturbance variance covariance matrix Ω_{NT} construction follows the Park's (1967) structure as the direct product of averaged matrices of cross-sectional and serial correlations, respectively denoted by $\bar{\Sigma}$ and $\bar{\Pi}$, and computed as below.

$$\Omega_{NT} = \bar{\Sigma} \otimes \bar{\Pi} \qquad (4)$$

where

$$\bar{\Sigma} = \begin{bmatrix} \bar{\sigma}_{\varepsilon,11} & \bar{\sigma}_{\varepsilon,12} & \cdots & \bar{\sigma}_{\varepsilon,1\mathrm{N}} \\ \bar{\sigma}_{\varepsilon,21} & \bar{\sigma}_{\varepsilon,22} & \cdots & \bar{\sigma}_{\varepsilon,2\mathrm{N}} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\sigma}_{\varepsilon,\mathrm{N}1} & \bar{\sigma}_{\varepsilon,\mathrm{N}2} & \cdots & \bar{\sigma}_{\varepsilon,\mathrm{NN}} \end{bmatrix}, \; \bar{\Pi} = \begin{bmatrix} 1 & \bar{\rho} & \bar{\rho}^2 & \cdots & \bar{\rho}^{\mathrm{T}-1} \\ \bar{\rho} & 1 & \bar{\rho} & \cdots & \bar{\rho}^{\mathrm{T}-2} \\ \bar{\rho}^2 & \bar{\rho} & 1 & \cdots & \bar{\rho}^{\mathrm{T}-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \bar{\rho}^{\mathrm{T}-1} & \bar{\rho}^{\mathrm{T}-2} & \bar{\rho}^{\mathrm{T}-3} & \cdots & 1 \end{bmatrix},$$

and

$$\bar{\sigma}_{\varepsilon,\mathrm{ij}} = \frac{\bar{\sigma}_{\mathrm{r,ij}}}{1-\bar{\rho}^2}.$$

All parameters in $\bar{\Sigma}$ and $\bar{\Pi}$ were estimated using the residuals of OLS regressions of model (1) on data slices mentioned above assuming first order serially correlated residuals. Specifically, a consistent way to estimate ρ coefficients averaged to get $\bar{\rho}$ suggested in Greene (2003, p 326) uses the following expression with OLS innovations $\hat{e}_{it}$.

$$\rho = \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{e}_{it}\hat{e}_{i,t-1}}{\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{e}_{it}^{\,2}}$$

Furthermore, $\bar{\sigma}_{\mathrm{r,ij}}$ was calculated as by averaging the sample correlations between individuals i and j's residuals $\hat{r}_{it}$ and $\hat{r}_{jt}$ arrived at by Prais-transforming model (1).

b. Experimental parameters of interest

The central parameter of interest for analysis purposes in Reed and Ye (2011) was the slope coefficient b_x of $\overline{X_\iota}$ in the model described below.

$$\bar{Y}_{it} = \boldsymbol{b_0} + \boldsymbol{b_x}.\bar{X}_{it} + e_{it} \tag{5}$$

and

$e_{it} = \boldsymbol{\rho} e_{i,t-1} + v_{it}$

where e_{it} is the first order auto correlated residuals with Ω_{NT} described above as variance covariance matrix, and v_{it} is assumed to be a white noise. .

The interest lied in the precision with which $\boldsymbol{b_x}$ could be estimated, and the relative efficiency of its estimation with regards to a number of characteristics of the experimental disturbance term. The experiments covered eleven estimators of $\boldsymbol{b_x}$. The precision of a given estimator was defined as the 95 per cent coverage rate measured by the percentage of times the hypothesis test of the equality between the true and experiment-based values of $\boldsymbol{b_x}$ failed to reject the null hypothesis out of a given number of trials. Denoting by L the number of trials in a given experiment, efficiency is computed using the formula below.

$$efficiency = 100\frac{\sqrt{\sum_{l=1}^{L}\left[\hat{b}_k^{(l)} - b_k\right]^2}}{\sqrt{\sum_{l=1}^{L}\left[\tilde{b}_k^{(l)} - b_k\right]^2}}$$

where $\tilde{b}_k^{(l)}$ is the value of b_k for the reference estimator (here OLS) at the l^{th} trial, $\hat{b}_k^{(l)}$ is the value of b_k for the estimator that is being compared to the reference.

A value of efficiency greater (less) than 100 characterizes an estimator that is less (more) efficient than the reference estimator (see, for example, Beck and Katz 1995). Likewise, a value of coverage significantly below 95 is an indication of poor confidence interval construction.

Three main features of the generated data sets were used to analyse and interpret the experimental outcomes. These were the N-T ratio, the common first order autocorrelation coefficient and the degree of heteroscedasticity (HETCOEF) of the OLS disturbances term from model (5). The computation of a consistent estimate of ρ, ρ (RHOHAT) in the experiments followed the formula below proposed in Greene (2003, p 326).

$$\rho = \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{e}_{it}\hat{e}_{i,t-1}}{\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{e}_{it}^{\,2}}$$

where $\hat{e}_{it}$ is the residuals series from the OLS re-estimation of the DGP equation with simulated data.

The HETCOEF parameter was determined as the ratio of the first and third quintiles of the estimate of population variances of the residuals e_{it} from model (5). This indicator captures the degree of heteroscedasticity in the error term of the simulated data. Higher values of HETCOEF might be associated with higher distortions in the hypothesis tests conclusions and the coefficients confidence intervals due to inaccurate estimates of coefficients' standard errors.

Table 1: Features of estimators' residuals modeled.

No	**Estimator**	**Features of the residuals modeled**
	From Stata	
Estimator 1	OLS	Independent
Estimator 2	OLS	Heteroscedasticity
Estimator 3	OLS	Heteroscedasticity, serial correlation
Estimator 4	OLS	Heteroscedasticity, cross-sectional dependence
Estimator 5	FGLS	Groupwise heteroscedasticity
Estimator 6	FGLS	Groupwise heteroscedasticity, serial correlation
Estimator 7	FGLS (Parks)	
Estimator 8	PCSE (Parks)	
	From EViews	
Estimator 9	FGLS	Weight = Groupwise heteroscedasticity; Covariance = Heteroscedasticity, Cross-sectional dependence
Estimator 10	FGLS	Weight = Groupwise heteroscedasticity; Covariance = Heteroscedasticity, Serial correlation
Estimator 11	FGLS	Weight = Groupwise heteroscedasticity; Covariance = Heteroscedasticity

Source: Reed and Ye (2011)

Experiments in Reed and Ye (2011) were conducted using 11 different static panel estimators. The standard OLS estimator was taken as the reference against which the performance (in efficiency terms) of the remaining ten other estimators were compared.

All the estimators assessed are incorporated in Stata or Eviews among other popular statistical packages. However, specific treatments of these estimators by Stata and Eviews procedures were adopted in the experiments. Details of the residuals' features accounted for by each estimator assessed are described in Table 1.

c. Criticism of the initial experimental design

Reed and Ye (2011)'s investigation resulted in interesting findings as we will discuss below. Nonetheless, we find that their experimental design might be subject to a well-grounded criticism in connection with their construction of the independent variable used for simulation purposes. Equation (3) systematically exaggerates the degree of serial correlation in the regressor. When the error terms are serially correlated, the serial correlation in the regressor affects the variance of its OLS coefficient estimator variance. The following relationship connects the variance of OLS slope estimator characterised by first order serial correlation of both the error term and the regressor, $Var\left(\hat{\beta}_{AR(1)}\right)$, on the one hand, and that of the usual OLS slope estimator, $Var\left(\hat{\beta}_{OLS}\right)$, on the other (see Gujarati 2004, p 452).

$$Var\left(\hat{\beta}_{AR(1)}\right) = Var\left(\hat{\beta}_{OLS}\right)\left(\frac{1 + r\rho}{1 - r\rho}\right)$$

where r and ρ denote the first order serial correlation coefficients of the regressor and the error term respectively.

Therefore, exaggerating the serial correlation in the regressor would worsen the bias in the estimated slope standard error, thereby making the conclusions of analyses based on the accuracy of the confidence interval construction related to the rejection rates and coverage levels misleading.

1.2.2. Modified experimental design

To address the main experimental design flaw identified above, two alternatives exist, of which one is preferred over the other and employed in this paper. One possibility is to keep the constructed independent variable and address this complication in subsequent steps of the experiments. However, we choose to maintain the focus on the structure of the error term, and would prefer avoiding further complications whenever possible. Therefore, we favour the other approach consisting in constructing a regressor without exaggerating its degree of serial correlation. Practically, we perform this method by randomly selecting one of the data windows formed in the data specific parameters generation step to substitute the original version of the regressor. This process does not add further correlation in the regressor used for simulation purposes.

1.3. Description of replicated and redesigned experiments' output

In Reed and Ye (2011), a total of 144 experiments were implemented with both formulations of model (1). Of these experiments, 80 used data sets characterized by N≤T, and 64 were conducted with data sets where N>T. Experiments were carried out with eight different datasets, different N-T combinations formed with 6 different individual dimension (N) values (5, 10, 20, 48, 50 and 77) and 4 different time dimension (T) values (10, 15, 20 and 25). We have fully replicated the experiments twice; once in the same conditions, namely the model specifications, the experimental design, the data sets and the panel individual and time dimensions; and once with the sole difference in the construction of the regressor used for simulation purposes that avoids exaggerating serial correlation. The results of these replications are presented below in light of the original output.

1.3.1. Relative efficiency of estimators

Table 2 contains statistics about estimators' performances on efficiency grounds in three panels, corresponding to the original performances (panel a) as reported by Reed and Ye (2011), the replicated performances (panel b) with unchanged experimental design, and the redesigned performances (panel c) after reconstructing the dependent variable.

The exact replications of efficiencies are very close to the original experiments results when considering the actual efficiency figures, or the number of experiments where OLS is less efficient. The only differences observed relate to estimators 6 (when N>T) and 7 (when N≤T) for which very slight efficiency gains are observed on the one side, and for estimators 8 and 6 (N>T) that dominate more often OLS on efficiency grounds in the replicated experiments.

Table 2: Original, replicated and redesigned relative efficiency statistics.

	Average Efficiency		Percentage of experiments where estimator is more efficient than OLS	
	N<=T	**N>T**	**N<=T**	**N>T**
a. Original output				
Estimator 5/9/10/11	95.2	82.9	58.8	84.4
Estimator 6	95.1	83.1	71.3	79.7
Estimator 7	73.9	--	96.3	--
Estimator 8	100.8	101.0	62.5	51.6
b. Replicated experiments				
Estimator 5/9/10/11	95.2	82.9	58.8	84.4
Estimator 6	95.1	82.6	71.3	81.3
Estimator 7	73.7	--	96.3	--
Estimator 8	100.8	101.0	63.8	57.8
c. Redesigned experiments				
Estimator 5/9/10/11	92.5	76.9	68.8	89.1
Estimator 6	79.5	70.7	85.0	90.6
Estimator 7	62.0	--	100.0	--
Estimator 8	86.4	91.8	71.3	71.9

Source: Reed and Ye (2011) and experiments replications

However, after the regressor is reconstructed, considerable efficiency gains are recorded for all estimators irrespective of whether N is larger or less than T. And the frequencies of trials where the estimators are more efficient than OLS substantially increase for all estimators. This result reveals the significance of the sensitivity of the efficiency indicator - and indirectly the slope coefficient estimate - to the degree of serial correlation in the regressor.

1.3.2. Confidence interval precision

Table 3 summarises precision indices for all estimator under the original (panel a.), exact replication (panel b.) and adjusted replication (panel c.) experimental designs. Our replications of the confidence interval precisions with the unchanged experimental design perfectly match the original results when the individual dimension of simulated panel data sets is not greater than the time dimension. Coverage rates and the absolute gap with the 95 percent theoretical threshold are correctly reproduced for all estimators. Furthermore, when the time dimension is dominated by the individual dimension, differences exist between replicated and original confidence internal precisions but they are within acceptable ranges. Overall, rejection rates decrease by .9 (estimator 3) to 1.8 (estimator 4) percentage points while increases in the gap with the theoretical confidence level by generally the same order for a given estimator are observed.

The impact of the experimental design adjustment for the earlier discussed flaw on the confidence internal precision indicators is substantial. Subsequent to the adjustment, rejection rates decrease for all estimators. When N is less than or equal to T, the highest improvements are associated with estimator 10, OLS, estimator 5 and estimator 3 (gains range from 4.0 to 5.7 percentage points for these estimators) and the lowest precision improvement (.9 percentage point) is observed for FGLS (Parks). When N is larger than T, OLS stands out with the precision gain, followed by estimators 4, 5 and 8.

These differences in the coverage rates improvements indicate that on average, the impact of the regressor serial correlation exaggeration has been especially significant on those estimators that gain more under the experimental redesign.

1.4. Further re-examination of estimators' performances

1.4.1. Assessment of the original recommendations

A major question we investigate in this paper is whether the readjustment of the experimental design has altered the desirability of some estimators over others. That is, we want to examine the validity of recommendations formulated by Reed and Ye (2011) after replacing their regressor contaminated with some extra serial correlation following construction. This is what we do in this section that re-assesses the initial recommendations resulting from the original experiments.

<u>Table 3:</u> Replicated and redesigned confidence interval precision.

Estimator	N<=T		N>T	
	Coverage	|95-Coverage|	Coverage	|95-Coverage|
	a. Original output			
Estimator 1	73.6	21.9	74.2	21.9
Estimator 2	73.7	21.8	77.9	18.8
Estimator 3	83.5	11.6	91.8	3.9
Estimator 4	72.7	22.5	74	21.3
Estimator 5	69.8	25.6	72.6	22.9
Estimator 6	86.4	9.3	88.8	7.2
Estimator 7	43.3	51.7	--	--
Estimator 8	87.8	7.2	88.1	6.9
Estimator 9	66.1	28.9	65.4	29.6
Estimator 10	68.1	26.9	80.1	14.9
Estimator 11	69.5	25.9	72.4	23.2
	b. Replicated experiments			
Estimator 1	73.6	21.9	75.7	20.5
Estimator 2	73.7	21.8	79.3	17.5
Estimator 3	83.5	11.6	92.7	3.0
Estimator 4	72.7	22.5	75.8	19.6
Estimator 5	69.8	25.6	74.1	21.4
Estimator 6	86.4	9.3	90.2	5.5
Estimator 7	43.3	51.7	--	--
Estimator 8	87.8	7.2	89.1	5.9
Estimator 9	66.1	28.9	66.7	28.3
Estimator 10	68.1	26.9	81.5	13.5
Estimator 11	69.5	25.9	73.9	21.7
	c. Extended experiments			
Estimator 1	78.0	18.0	82.1	15.2
Estimator 2	77.2	18.2	80.4	16.0
Estimator 3	87.6	7.7	92.1	5.1
Estimator 4	75.4	19.8	78.0	17.2
Estimator 5	73.8	21.2	76.5	19.2
Estimator 6	89.4	5.8	90.0	5.8
Estimator 7	44.2	50.8	--	--
Estimator 8	91.0	4.0	92.0	3.1
Estimator 9	68.3	26.7	68.7	26.3
Estimator 10	73.8	21.2	81.4	13.6
Estimator 11	73.0	22.0	75.7	19.8

a. Recommendation 1

According to the first recommendation with the original experimental design, the FGLS (Parks) procedure is preferable when the primary concern is efficiency and T/N≥1.5. Figure 1 plots average efficiencies of estimators on the vertical axis against the T/N ratio on the horizontal axis under the old (panel a.) and the new (panel b.) experimental designs.

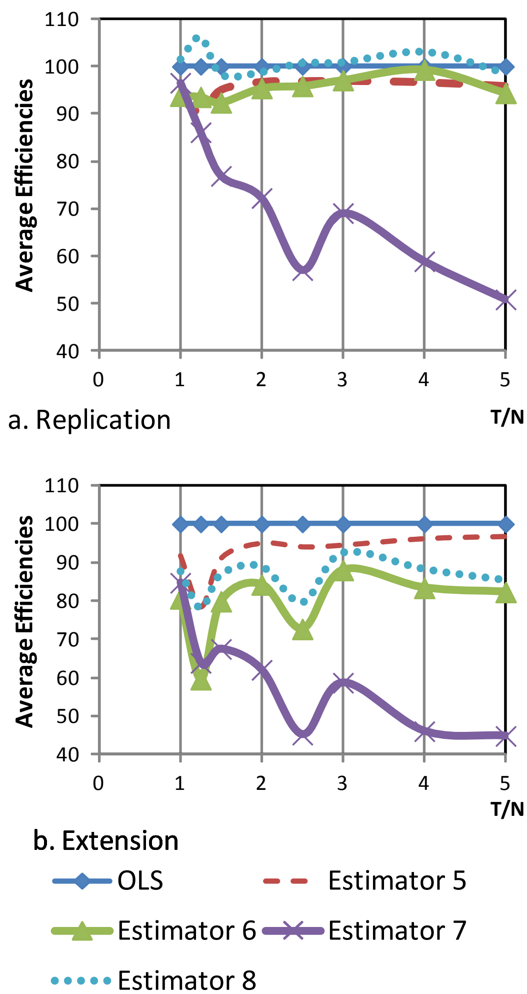


Figure 1: Average efficiencies of estimators when N≤T (vertical axis) against T/N ratio (horizontal axis).

It clearly appears that recommendation 1 holds with both versions of the experimental design. Furthermore, our exact replications indicate that there is a cut point of T/N ratio of 1.25 from which estimator 7 outperforms others; this rate was 1.5 according to the original results.

Additionally, under the new experimental design, when T/N<1.5, estimator 6 appears dominant, leading to a complete classification of estimators when N≤T. Another interesting result established with the redesigned experiments is that from T/N value of 1.5, there is a total order among the estimators with respect to the efficiency indicator as follows: OLS is the least efficient estimator, followed successively by the group of estimators 5/9/10/11, estimator 8, estimator 6, and lastly estimator 7.

b. Recommendation 2

In their second recommendation, Reed and Ye (2011) advise using estimators 5/9/10/11 or estimator 6 to best optimize efficiency when N>T and HETCOEF>1.67. We are able to confirm this recommendation with our exact replications. Panel a. of Figure 2 provides evidence for these preferences in absolute terms. The two conditions hold for a total of 46 experiments of which roughly 2/3 indicate that estimator 6 is more efficient than the group of estimators 5/9/10/11, but these estimators are all strictly preferred to estimator 8 and OLS.

However, subsequent to the correction of the experimental design, no strict preference for a given estimator is revealed over all the experiments with N>T and RHOHAT>1.67. The numbers of experiments that meet these requirements remains the same, and are split the following way with respect to the efficiency performance criteria: estimator 8 outperforms all others in 4 cases; it dominates the group of estimators 5/9/10/11 in 13 cases and estimator 6 in 4 cases. And while the graphical representation reveals a close proximity between estimators 5/9/10/11 and estimator 6, the performance of the group of estimators at least dominates that of estimator 6 in only 15.2 percent of the cases (7 experiments out of 46). Therefore, only estimator 6 stands out as the best. The elimination of extra serial correlation introduced in the simulations through the regressor affects this recommendation by making less recommendable estimators 5/9/10/11 as best estimators on efficiency grounds when N>T. This conclusion aligns with the summaries of Table 2 showing the smaller average efficiency coupled with the number of times the estimator 6 is preferred over OLS. It is worth noting that the order for both indicators are reversed for the group of estimators 5/9/10/11 and estimator 6 post-correction for regressor serial correlation.

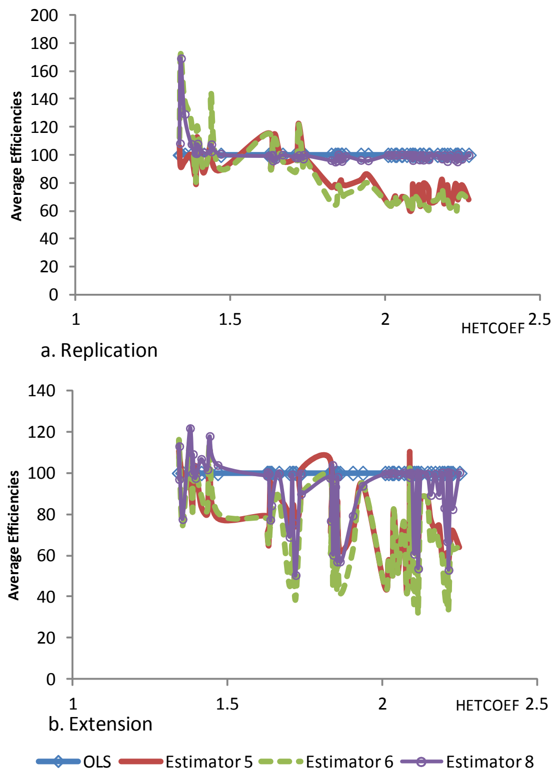


Figure 2: Average efficiencies of estimators when N≤T (vertical axis) against N/T ratio (horizontal axis).

c. Recommendation 3

Recommendation 3 of the original research chooses estimator 4 and estimator 8 as best performers for constructing confidence intervals when RHOHAT<0.30. Two indicators are used to measure the performance of estimators with respect to the confidence interval. These are the coverage and the absolute coverage gap with the 95 percent threshold discussed earlier. The second indicator fills a gap characterizing the first due to the coverage level averaging by capturing the mixed effects of over-rejection and under-rejection in experiments for a given estimator.

According to Table 4 and Figure 3, recommendation 3 keeps the same estimators under both replications experimental designs with some variations. For RHOHAT<0.30, there is a substantial gain in the confidence interval precision for estimator 8 with the experimental design change making it the best estimator to recommend when N≤T and the closest substitute for estimator 4 when N>T. In the meantime, the confidence interval precision of estimators 2 and 3 which were the closest to that of estimator 4 under the initial experimental design significantly deteriorates subsequent to the amendment.

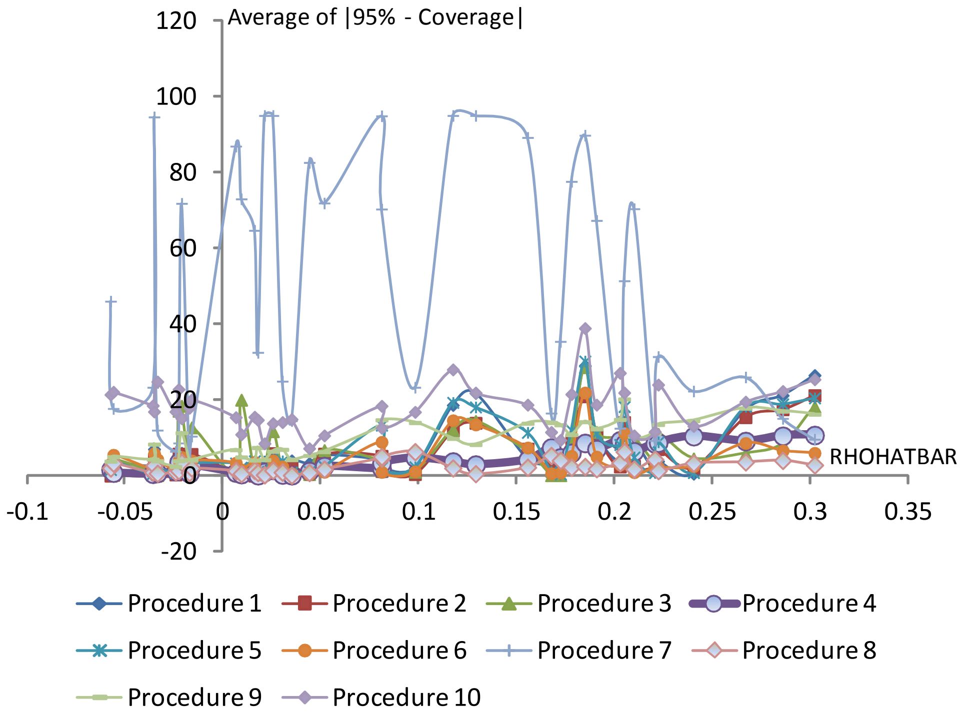
Table 4: Absolute coverage gap with 95 percent for initial and replicated results when RHOHAT<0.30

Procedure	Initial results		Exact replications		Replications with extension	
	N<=T	**N>T**	**N<=T**	**N>T**	**N<=T**	**N>T**
Estimator 1	5.2	4.0	5.1	3.9	6.1	5.8
Estimator 2	4.5	1.8	4.5	1.8	4.9	4.0
Estimator 3	9.9	1.5	9.9	1.5	6.5	4.8
Estimator 4	3.7	1.4	3.7	1.4	3.8	1.3
Estimator 5	6.3	2.1	6.3	2.1	6.4	3.7
Estimator 6	4.9	2.0	4.9	2.6	4.4	3.4
Estimator 7	47.9	--	47.9	--	49.4	--
Estimator 8	3.1	2.4	3.1	2.4	2.3	1.8
Estimator 9	8.6	6.9	8.6	6.9	9.4	7.8
Estimator 10	19.9	6.4	19.9	6.4	17.7	8.8
Estimator 11	6.4	2.1	6.4	2.1	6.8	3.8

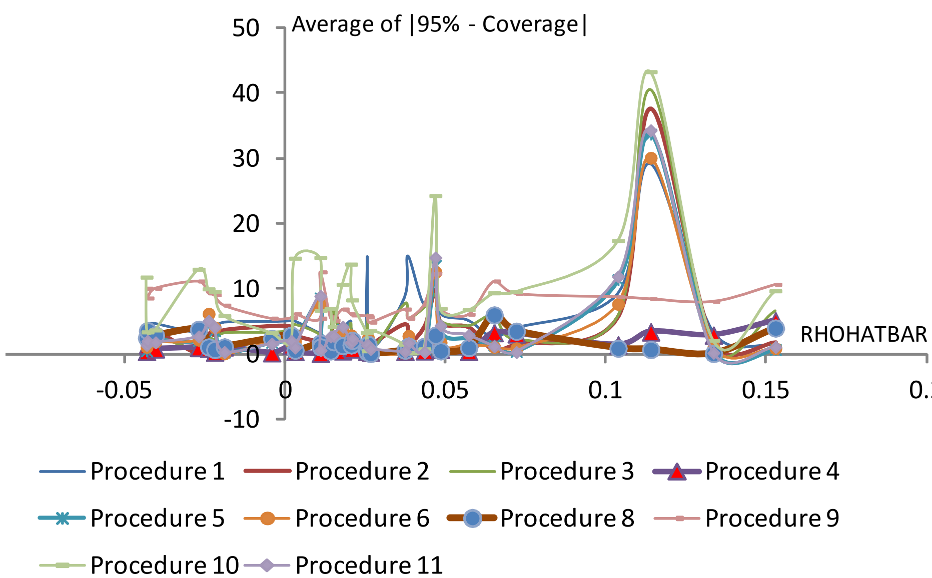
1.4.2. Further implications of the redesign of experiments

We found for the panel individual dimension not greater than the time dimension (N≤T), estimator 6 and estimator 7 are the best performers respectively when T/N<1.5 and when T/N≥1.5. More interestingly, we also found that based on the T-N ratio, estimator 6's efficiency performance when N>T is outstanding as shown on Figure 4. This implies that taking out the HETCOEF indicator would allow formulating a more compelling recommendation on the selection of the most efficient estimator for panel datasets characterized by N>T.

The above findings about the first two recommendations imply that a single recommendation solely based on the T/N ratio would more efficiently combine these two as follows: When the primary concern is efficiency, (i) choose estimator 6 when T/N<1.5 or N>T, and (ii) choose estimator 7 when T/N≥1.5.



a. N≤T



b. N>T

Figure 3: Efficiencies vs. error term serial correlation indicator (RHOHATBAR) under the extended experiments.

1.5. Conclusion

Researchers using econometric models implicitly need to support their theoretical analyses by summarizing relationships among series of relevant data. One important concern they have would be to correctly summarize these relationships. This requires accounting for the characteristics of their data sets in light of underlying assumptions of econometric models that guarantee their goodness of fit along with the hypothesis tests and inferences they may allow. Many estimators are made available and incorporated in statistical packages, but their performances on different data characteristics differ according to the treatment of data by underlying procedures and the types of data relationships they accommodate. This complicates an a priori selection of the right estimator to match specific data sets a given user of econometrics is analyzing. This difficulty is accentuated in the context of panel data models characterized by a much larger number of relationships among variables. Estimator selection appears thus to be one of the key determinants of econometric models performance.

Reed and Ye (2011) attempted to empirically provide researchers with recommendations in choosing the right estimator among a set of commonly used estimators based on the data sets they had available by deploying a Monte Carlo simulation method. Their research presented a set of important recommendations and is being well cited among econometric practitioners.

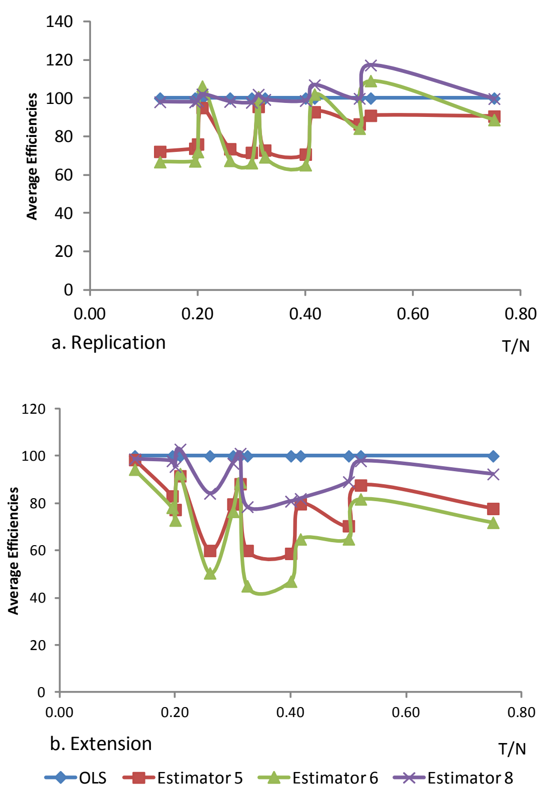


Figure 4: Comparison of efficiencies using the T/N ratio under when N>T.

They use three data sets characteristics to form the recommendations, namely the ratio T/N, the degrees of heteroskedasticity (HETCOEF) and serial correlation (RHOHAT) in the standard OLS residuals.

However, it appears that their experimental design contained a flaw whose implications for the recommendations are investigated in this paper. The serial correlation in the regressor they used in experiments was exaggerated by construction. After adjusting for this flaw, we found the following:

(i) HETCOEF which is one of the data characteristics used for recommendations in Reed and Ye (2011) appears to be irrelevant for this purpose;

(ii) A single complete recommendation could be formulated from the panel data sets dimensions ratio (T/N) in place of two incomplete recommendations based on T/N ratio for the one while the other combines the HETCOEF indicator to the T/N ratio;

(iii) While recommendation 3 still holds post experimental design adjustment, the precision in the recommendation is improved, making it more straightforward compared to its initial version.

References

[1]. Beck, N. & Katz, J. N. (1995). What to do (and not to do) with Time-Series Cross-Section Data. The American Journal of Science Review, 83, 634-647.

[2]. Bond, S. R. & Eberhardt, M. (2009). Cross-section dependence in nonstationary panel models: a novel estimator. Paper presented at the Nordic Econometrics Meeting in Lund, Sweden, October 29-31.

[3]. Cameron, A. C.; Gelbach B. & Miller, D. L. (2006), Robust Inference with Multi-way Clustering, Technical Working Paper 327, available at http://www.nber.org/papers/T0327.

[4]. Cameron, A. C.; Gelbach B. & Miller, D. L. (2009). Robust inference with multi-way clustering, Working Papers, University of California, Department of, Economics, No. 09,9, available at http://hdl.handle.net/10419/58397.

[5]. Cameron, A. C.; Gelbach B. & Miller, D. L. (2011), Robust Inference with Multi-way Clustering. Journal of Business and Economic Statistics, 29(2), 238-249, DOI: 10.1198/jbes.2010.07136.

[6]. Chudick, A, & Pesaran M. H. (2011). Weak and strong cross-section dependence and estimation of large panels. Econometrics Journal, 14, C45-C90. doi: 10.1111/j.1368-423X.2010.00330.x.

[7]. Driscoll, J. C. & Kraay, A. C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. The Review of Economics and Statistics, 80, 549-560.

[8]. Greene, W. H. (2003). Econometric Analysis, 5th Edition, Pearson Education.

[9]. Gujarati, D. N. (2004). Basic Econometrics, 4th Edition, The McGraw−Hill Companies.

[10]. Hoechle, D. (2007). Robust standard errors for panel regressions with cross-sectional dependence. The Stata Journal,7, 281-312.

[11]. Kapetanios, G., Pesaran, M. H. & Yamagata, T. (2011). Panels with non-stationary multifactor error structures. Journal of Econometrics, 160, 326–348.

[12]. Kézdi, G. (2004). Robust Standard Error Estimation in Fixed-Effects Panel Models. Hungarian Statistical Review, Special 9. 2004, 95-116.

[13]. Parks, R. W. (1967). Efficient Estimation of a System of Regression Equations when Disturbances are Both Serially and Contemporaneously Correlated. Journal of the American Statistical Association, 62, 500-509.

[14]. Pesaran, M.H. & Smith, R. P. (1995). Estimating long-run relationships from dynamic heterogeneous panels. Journal of Econometrics 68, 79–113.

[15]. Petersen, M. A. (2009). Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. The Review of Financial Studies, 22(1), 435-80.

[16]. Reed, R. W. & Ye, H. (2011). Which panel data estimator should I use? Applied Economics. 43(8), 985–1000.

[17]. Sarafidis, V. & Wansbeek, T. (2012). Cross-sectional dependence in panel data analysis. Econometric Reviews, 31(5), 483–531.

[18]. Skrondal, A. & Rabe-Hesketh, S. (2008). Multilevel and Related Models for Longitudinal Data. In J. de Leeuw, E. Meijer (eds.), Handbook of Multilevel Analysis, Springer 2008, 275-99.

[19]. Zellner, A. & Huang, D. S. (1962). Further Properties of Efficient Estimators for Seemingly Unrelated Regression Equations. International Economic Review, 3, 300-13.

[20]. Zellner, A. & Theil, H. (1962). Three State Least Squares: Simultaneous Estimation of Simultaneous Equations. Econometrica, 30, 54-78.

[21]. Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. Journal of the American Statistical Association, 57, 348-368.

[22]. Zellner, A. (1963). Estimators of Seemingly Unrelated Regressions: Some Exact Finite Sample Results. Journal of the American Statistical Association, 58, 977-92.