DEPARTMENT OF ECONOMICS

COLLEGE OF BUSINESS AND ECONOMICS

UNIVERSITY OF CANTERBURY

CHRISTCHURCH, NEW ZEALAND

ANOTHER LOOK AT WHAT TO DO WITH TIME-SERIES CROSS-SECTION DATA

by

Xiujian Chen, Shu Lin, and W. Robert Reed

WORKING PAPER

No. 04/2006

Department of Economics, College of Business and Economics
University of Canterbury
Private Bag 4800, Christchurch
New Zealand

WORKING PAPER No. 04/2006

ANOTHER LOOK AT WHAT TO DO WITH

TIME-SERIES CROSS-SECTION DATA

by

Xiujian Chen, Shu Lin, and W. Robert Reed*

Revised, March 31, 2006

Abstract

Our study revisits Beck and Katz' (1995) comparison of the Parks and PCSE estimators using time-series, cross-sectional data (TSCS). Our innovation is that we construct simulated statistical environments that are designed to approximate actual TSCS data. We pattern our statistical environments after income and tax data on U.S. states from 1960-1999. While PCSE generally does a better job than Parks in estimating standard errors/confidence intervals, it too can be unreliable, sometimes producing standard errors/confidence intervals that are substantially off the mark. Further, we find that the benefits of PCSE can come at a large cost in estimator efficiency.

*The contact author is W. Robert Reed, Professor, Department of Economics, University of Canterbury, Private Bag 4800, Christchurch, New Zealand, Phone 64-3-3642846, email: bobreednz@yahoo.com.

Keywords: Panel data, Parks model, PCSE estimator, Monte Carlo methods

JEL Classifications: C23, C15

Acknowledgments: An earlier draft of this paper was presented at the University of Oklahoma and the 11th International Conference on Panel Data. We acknowledge helpful comments from Kevin Grier, Cynthia Rogers, and Aaron Smallwood.

I. INTRODUCTION

Empirical studies frequently employ data consisting of repeated time-series observations on fixed, cross-sectional units. While providing a rich amount of information, time-series cross-sectional (TSCS) data are likely to be characterized by complex error structures. The application of OLS to data with nonspherical errors produces inefficient coefficient estimates, and the corresponding standard error estimates are biased. In contrast, GLS produces coefficient and standard error estimates that are efficient and unbiased, respectively, given certain assumptions. Two such assumptions are (i) the error covariance structure is correctly specified, and (ii) the elements of the error covariance matrix are known. Feasible GLS (FGLS) is used when the structure of the error covariance matrix is known, but its elements are not. The finite sample properties of FGLS are analytically indeterminate.

Beck and Katz (1995) (henceforth, BK) use Monte Carlo methods to study the performance of FGLS in a statistical environment characterized by (i) groupwise heteroscedasticity, (ii) first-order serial correlation, and (iii) contemporaneous cross-sectional correlation. They dub the corresponding FGLS estimator "Parks" (after Parks [1967]). BK report three major findings:

1. Parks produces dramatically inaccurate standard errors.

2. An alternative method, based on OLS but using "panel-corrected standard errors," (henceforth, PCSE) produces accurate standard errors.

3. The efficiency advantage of Parks over PCSE is at best slight, except in extreme cases of cross-sectional correlation, and then only when the number of time periods (T) is at least twice the number of cross-section units (N).

Consequently, BK prescribe that researchers use the PCSE procedure when working with TSCS data.[1]

BK has been very influential. A recent count identified approximately 450 citations (e.g., Nunziata, 2005; Jönsson, 2005; Dejuan and Luengo-Prado, 2006).[2] Their PCSE estimator has been widely applied to both U.S. (e.g., Kacperczyk et al., 2005; Engstrom and Kernell, 2005) and international panel data (e.g., Lee and Roemer, 2005; Soo, 2005; Brülhart and Trionfetti, 2004). It is available as a standard procedure in many statistical software packages, including STATA, Shazam, GAUSS, and RATS.

Our paper constructs a statistical environment patterned after actual TSCS data and revisits BK's analysis of the Parks and PCSE estimators. We construct a "Parks-type" statistical environment, and then attempt to replicate BK's findings using similar Monte Carlo techniques. We confirm BK's result that Parks consistently underestimates coefficient standard errors, with corresponding confidence intervals that are too narrow ("overconfident"). However, we find that PCSE can also substantially underestimate coefficient standard errors/confidence intervals. Further, we find that PCSE is much less efficient than reported by BK.

We next construct a more general statistical environment and repeat our analysis. We once again obtain the result that PCSE generally does a better job than FGLS when estimating standard errors/confidence intervals. However, the standard error benefits of PCSE over Parks are less, and the costs in terms of diminished efficiency are greater.

[1] A further advantage of PCSE is that it is able to incorporate cross-sectional correlation when the number of time series observations (T) is less than the number of cross-sectional observations (N), whereas standard FGLS cannot.
[2] Cf. Web of Science, www.isinet.com/products/citation/wos.

Our results suggest that PCSE is superior to Parks when the researcher's main focus is hypothesis testing. However, even PCSE estimates of standard errors can be misleading. Further, Parks is superior to PCSE when the main concern is obtaining accurate coefficient estimates. Given this tradeoff, we conclude that researchers should use both procedures, relying on the PCSE estimates for hypothesis testing, and Parks for coefficient estimates.

Our paper proceeds as follows. Section II re-evaluates BK's Monte Carlo analysis within a "Parks-type" statistical environment. We set the values of the elements of the population error covariance matrix so that they approximate values found in actual TSCS data. Section III repeats this analysis within a statistical environment that generalizes the Parks model. Section IV concludes.

II. RE-EVALUATING BK WITHIN A "PARKS-TYPE" STATISTICAL ENVIRONMENT

IIA. Methodology for producing a "Parks-type" statistical environment

BK build their Monte Carlo analysis around the following TSCS model:

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_N \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \\ \vdots \\ \boldsymbol{X}_N \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_N \end{bmatrix}, \text{ or } \quad \boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon};$$

where $\boldsymbol{y}_i$ is a $T \times 1$ vector of observations on the dependent variable in the i^{th} group, $i = 1,2,...,N$; $\boldsymbol{X}_i$ is a $T \times K$ matrix of exogenous variables; $\boldsymbol{\beta}$ is a $K \times 1$ vector of coefficients; $\boldsymbol{\varepsilon}_i$ is a $T \times 1$ vector of error terms; and $\boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \boldsymbol{\Omega}_{NT})$.

Following Parks (1967), they allow $\boldsymbol{\Omega}_{NT}$ to consist of (i) groupwise heteroscedasticity; (ii) groupwise, first-order serial correlation; and (iii) cross-sectional (spatial) correlation. Specifically,

$$\boldsymbol{\Omega}_{NT} = \begin{bmatrix} \sigma_{\varepsilon,11}\boldsymbol{\Sigma}_{11} & \sigma_{\varepsilon,12}\boldsymbol{\Sigma}_{12} & \cdots & \sigma_{\varepsilon,1N}\boldsymbol{\Sigma}_{1N} \\ \sigma_{\varepsilon,21}\boldsymbol{\Sigma}_{21} & \sigma_{\varepsilon,22}\boldsymbol{\Sigma}_{22} & \cdots & \sigma_{\varepsilon,2N}\boldsymbol{\Sigma}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{\varepsilon,N1}\boldsymbol{\Sigma}_{N1} & \sigma_{\varepsilon,N2}\boldsymbol{\Sigma}_{N2} & \cdots & \sigma_{\varepsilon,NN}\boldsymbol{\Sigma}_{NN} \end{bmatrix},$$

where $\sigma_{\varepsilon,ij} = \dfrac{\sigma_{u,ij}}{1-\rho_i\rho_j}$ and $\boldsymbol{\Sigma}_{ij} = \begin{bmatrix} 1 & \rho_j & \rho_j^2 & \cdots & \rho_j^{T-1} \\ \rho_i & 1 & \rho_j & \cdots & \rho_j^{T-2} \\ \rho_i^2 & \rho_i & 1 & \cdots & \rho_j^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_i^{T-1} & \rho_i^{T-2} & \rho_i^{T-3} & \cdots & 1 \end{bmatrix}$.

They proceed by selecting various combinations of N and T, specifying the elements of the respective error covariance matrices (the $\boldsymbol{\Omega}_{NT}$'s) by positing values for the population parameters $\sigma_{u,ij}$, ρ_i, and ρ_j, $i, j = 1,2,\ldots,N$.

Given $\boldsymbol{\Omega}_{NT}$, experimental observations are generated in the usual manner. Define $\boldsymbol{u}$ as a vector of standard normal random variables. Define $\boldsymbol{Q}$ such that $\boldsymbol{Q}'\boldsymbol{Q} = \boldsymbol{\Omega}_{NT}$. Error terms are created by $\boldsymbol{\varepsilon} = \boldsymbol{Q}'\boldsymbol{u}$. These simulated errors are added to a deterministic component, $\beta_0 + \beta_x x_i$, to calculate stochastic observations of y_i, where $y_i = \beta_0 + \beta_x x_i + \varepsilon_i$, $i=1,2,\ldots,NT$. BK create the x_i's from a zero-mean normal distribution (fixed in all replications), and set β_0 and β_x equal to 10 in all experiments. They perform 1000 replications for each experiment.

BK compare the (i) Parks and (ii) PCSE estimates of β_x. They employ several performance measures, including "Level" and "Efficiency." "Level" calculates the percent of estimated 95% confidence intervals that include the true value of β_x. "Efficiency" measures the relative efficiency of PCSE to Parks and is defined by

$$Efficiency = 100 \cdot \frac{\sqrt{\sum_{r=1}^{1000}\left(\hat{\beta}^{(r)}_{Parks} - \beta_x\right)^2}}{\sqrt{\sum_{r=1}^{1000}\left(\hat{\beta}^{(r)}_{PCSE} - \beta_x\right)^2}}.$$

An "Efficiency" value less than 100 indicates that PCSE is less efficient than Parks.

IIB. Constructing a "Parks-type" statistical environment based on actual TSCS data

Our experiments follow BK's methodology with one major exception: We create simulated statistical environments to look like actual TSCS data according to the following two-stage procedure: In the first stage, we estimate the parameters $\sigma_{u,ij}$, ρ_i, and ρ_j, $i,j = 1,2,...,N$ from "real-world" TSCS data. In the second stage, we use these estimated values as population parameters in the subsequent Monte Carlo experiments. In this manner, our simulated TSCS data is made to approximate the kind of data that researchers would encounter in actual TSCS data.[3]

For our "real-world" TSCS data, we use two data sets. The first data set consists of annual, U.S. observations on state-level incomes (specifically, the log of real Per Capita Personal Income). The second data set consists of annual, U.S. observations on state-level taxes (specifically, Tax Burden, defined as the ratio of total state and local taxes over Personal Income). Both data series have been the subject of much previous research and continue to be actively researched.

[3] BK recommend that empirical estimation of PCSE's restrict the autocorrelation parameters to be the same across groups (i.e., $\rho_i = \rho_j = \rho$ for all $i,j = 1,2,...,N$). Accordingly, we directly impose this on the simulated statistical environment and then look to the TSCS data to provide a "realistic" value for ρ.

Most advantageous for our approach is that both data series have long time series. We employ 40 years of data on 48 states (omitting Alaska and Hawaii), covering the period 1960-1999. A long time series is crucial for our approach. Most studies use time series where T is between 10 and 25 years (cf. Table 1 in BK). By having a data series substantially longer than that, we can sample multiple T-year, TSCS data sets in order to construct a "representative" error structure for a T-year, TSCS data set. We then use this representative error structure to generate experimental observations through the standard Monte Carlo methodology.

Our approach works like this: Suppose we want to construct a Parks-type, population error covariance structure ($\boldsymbol{\Omega}_{NT}$) in the context of a regression model where the dependent variable is either U.S. state-level incomes or state-level taxes, and the (balanced) TSCS data consist of T annual observations on each of N states. We begin by choosing the first N states in our data set. Next, we choose the T-year period, 1960 to (1960+T-1). We then estimate a fixed effects model relating the respective dependent variable (Y) to a set of state fixed effects (D^j), and an explanatory variable X (more on X below):

$$Y_{it} = \sum_{j=1}^{N} \alpha_j D_{it}^j + \alpha_{N+1} X_{it} + error\ term_{it},$$

where $i=1,2, \ldots ,N;\ t=1960,1961,\ldots,1960+T-1;$ and D^j is a state dummy variable that takes the value 1 for state j. We refer to this equation as the "residual generating function."

The residuals from this estimated equation are used to estimate the "Parks-method" error covariance matrix, $\hat{\boldsymbol{\Omega}}_{NT}$, in the standard manner. Our innovation is that we do this for every possible, T-contiguous year subsample contained within the 40 years of data from 1960-1999 [i.e., 1960-(1960+T-1), 1961-(1961+T-1), 1962-(1962+T-1), …, (1999-T+1)-1999]. This produces a total of $40–T+1$ estimated error covariance matrices, $\hat{\boldsymbol{\Omega}}_{NT}$, one for each possible T-contiguous year subsample. We then average these estimated covariance matrices to obtain a representative error covariance matrix, $\overline{\boldsymbol{\Omega}}$. This becomes the population error covariance matrix for the subsequent Monte Carlo experiments. Note that every element of $\overline{\boldsymbol{\Omega}}$ represents an average value across actual, estimated covariance matrices. In this sense, $\overline{\boldsymbol{\Omega}}$ can be said to be "representative" of the kinds of error structures one encounters in actual TSCS data.

We proceed by generating experimental observations of y_i, where $y_i = \beta_0 + \beta_x x_i + \varepsilon_i$, $i=1,2,...,NT$, and the errors are simulated from the population error covariance matrix, $\overline{\boldsymbol{\Omega}}$. We set the values of β_0 and the x_i's to be representative of their respective data sets, and fix the population value of β_x consistent with the empirical literature on income/taxes.[4]

Given an experimental data set of NT observations of (y_i, x_i), we estimate β_x using the Parks and PCSE estimators, respectively. We perform 1000 replications of this experiment, generating 1000 estimates of β_x for both the Parks and PCSE estimators. These 1000 estimates are then analyzed to compare the performance of the two

[4] For the income equations, we use Tax Burden as the explanatory variable and set $\beta_x = -0.01$ (see, for example, Helms [1985] and Wasylenko [1997]). For the tax equations, we use the log of real Per Capita Personal Income and set $\beta_x = -1.0$ (see, for example, Reed [2006]). The fact that each of the variables appears in the other residual generating function as an explanatory variable may raise concerns. With respect to the literature, these are common specifications. As a practical matter, the inclusion/exclusion of these explanatory variables in the residual-generating functions has a negligible effect on the results. Our only motivation for including them is to address potential concerns that the resulting error structure be independent of the explanatory variable in the simulated data.

estimators. This same procedure can be modified in a straightforward manner to conduct Monte Carlo experiments for alternative N and T values.

At this point it bears revisiting the claim that our simulated statistical environments approximate the error structures of actual TSCS data. Admittedly, the residual-generating function specified above represents a stripped down version of the specifications usually employed by researchers. Other variables typically would be included in the specification.

Unfortunately, there is no single specification that dominates the empirical literature on U. S. state incomes/taxes. As a result, we experimented with alternative residual generating functions that added a lagged dependent variable and/or time fixed effects. We found that our main results were qualitatively unaffected by these more elaborate specifications. Accordingly, we only report results based on the residual-generating function with state fixed effects.

Our study conducts experiments for a wide range of "sizes" of TSCS data sets: We set values for N equal to *5, 10*, *20*, and *48*; and values for T equal to *10, 15, 20,* and *25* -- a total of sixteen N and T combinations. This range encompasses most of the data sets reported in BK's Table 1 (page 635).

The first column of TABLE 1 summarizes salient characteristics of the data for the "Parks-type" statistical environment. The top part of TABLE 1 reports on the income data, the bottom part on the tax data. "*Mean* R^2" refers to the average R^2 for the respective residual generating functions in the first stage of the data-generating process. In other words, a typical fixed-effects regression equation "explained" approximately

73% of the variation in the (actual) TSCS income data, and *70%* of the variation in the (actual) TSCS tax data.

The subsequent rows characterize the serial correlation, heteroscedasticity, and cross-sectional correlation behavior of the *artificial* TSCS data produced in the second-stage of the data-generating process. These data comprised the simulated observations which were used to estimate β_x with the Parks and PCSE procedures, respectively.

Both the simulated income and tax data evidenced substantial degrees of serial correlation. The average of the estimated $\hat{\rho}$ values, "*Mean* $\hat{\rho}$" (averaged over all replications and experiments) was *0.61* for the income data, and *0.58* for the tax data.

As a measure of groupwise heteroscedasticity, we estimated group-specific standard errors ($\hat{\sigma}_i$, $i=1,\ldots,N$) for each replication and rank-ordered them from smallest to largest. We then calculated a "heteroscedasticity coefficient" (*h*), defined as the ratio of the upper quartile value of $\hat{\sigma}_i$ over its lower quartile value, again averaged over all replications and experiments. The "heteroscedasticity coefficient" value for the income data was *1.24*, and the corresponding value for the tax data was *1.59*.

Finally, both the simulated income and simulated tax data were characterized by substantial cross-sectional correlation. "*Mean* r_{ij}" is defined as the mean (absolute) value of the contemporaneous correlation between errors from groups *i* and *j*, averaged over all possible cross-sectional correlations, and over all replications and experiments. "*Mean* r_{ij}" for the income data was *0.74*, and *0.36* for the tax data. Note that the income data displayed a much greater degree of cross-sectional correlation than the tax data.

In summary, our simulated data were characterized by precisely the kinds of statistical problems (i.e., serial correlation, groupwise heteroscedasticity, and cross-sectional correlation) that the Parks and PCSE procedures were designed to handle.

IIC. Monte Carlo experiments assuming a "Parks-type" error structure where the elements of the population error covariance matrix are known

TABLE 2 reports the results of Monte Carlo experiments in which (i) the population error covariance matrix, $\overline{\boldsymbol{\Omega}}$, is constructed to have a "Parks-type" structure; (ii) the elements of $\overline{\boldsymbol{\Omega}}$ are patterned after actual TSCS data on U.S. state-level observations of income; and (iii) $\overline{\boldsymbol{\Omega}}$ is assumed known and available for use in the respective estimation procedures (á la GLS). The Parks and PCSE estimators of β_x are $\hat{\boldsymbol{\beta}}_{Parks} = \left(\boldsymbol{X}'\overline{\boldsymbol{\Omega}}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\overline{\boldsymbol{\Omega}}^{-1}\boldsymbol{y}$ and $\hat{\boldsymbol{\beta}}_{PCSE} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{y}$; and the corresponding estimators of the variance-covariance matrices are $Cov\left(\hat{\boldsymbol{\beta}}_{Parks}\right) = \left(\boldsymbol{X}'\overline{\boldsymbol{\Omega}}^{-1}\boldsymbol{X}\right)^{-1}$ and $Cov\left(\hat{\boldsymbol{\beta}}_{PCSE}\right) = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\overline{\boldsymbol{\Omega}}\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}$.[5]

The purpose of this section is to establish an empirical benchmark to which we can compare subsequent experiments when the elements of $\overline{\boldsymbol{\Omega}}$ will be assumed to be unknown to the researcher and will thus need to be estimated. We expect the estimated confidence intervals, and corresponding "Level" calculations, to be accurate. This follows from the fact that the experimental observations are generated from $\overline{\boldsymbol{\Omega}}$, and this same $\overline{\boldsymbol{\Omega}}$ is used to calculate $s.e.\left(\hat{\beta}_x\right)$ for the subsequent confidence interval calculations.

[5] The PCSE procedure requires that autocorrelated data first be transformed using the Prais-Winsten transformation. In order to facilitate comparison of the two procedures, we assume no autocorrelation and set the AR(1) parameter equal to zero in the Parks-type, population error covariance matrix (cf. $\boldsymbol{\Omega}_{NT}$ above).

In addition, these experiments will illustrate the potential efficiency advantages that drive the use of the Parks estimator.

Each experiment consisted of a 1000 replications of simulated TSCS income data of size *NT*. Separate experiments were conducted for all sixteen *NT* combinations. We note that the Parks method cannot be applied when $N > T$, which is probably why BK do not report Monte Carlo results for these cases. However, as BK's Table 1 shows, $N > T$ for many TSCS data sets, and thus we think researchers will be interested to know how PCSE fares (in an absolute sense) in these environments.

The left hand side of TABLE 2 summarizes the performance of the Parks and PCSE estimators with respect to "Level." The first four rows report the results of the 16 individual experiments. The last row summarizes these results by reporting the average values over all experiments.

For example, when $N = 5$ and $T = 10$, 95.5 percent of the 95% confidence intervals constructed using the Parks estimators of β_x and $\sigma_{\hat{\beta}_x}$ contain the true value of β_x. The corresponding result for the 95% confidence intervals based on the PCSE estimators is 96.4 percent. Across all experiments, 95.1 percent of both the Parks and PCSE 95% confidence intervals contain the true value of β_x. These are, of course, the results one would expect when the error covariance matrix is known, since in this case both Parks and PCSE have accurate finite-sample properties.

Turning now to the right hand side of TABLE 2, we see that Parks is generally much more efficient than PCSE when the error covariance matrix is known. For example, when $N = 10$ and $T = 10$, $Efficiency = 91.5$, which indicates that the sample standard deviation of the Parks estimates of β_x is about a tenth smaller in this experiment than the sample standard deviation of the PCSE estimates. The relative efficiency of Parks versus PCSE improves as both N and T get larger. Once N gets larger than *10*, or T gets larger than *20*, the efficiency gains of Parks over PCSE become quite substantial, with the Parks estimates having a standard deviation 25% or less than the corresponding PCSE estimates.

In summary, this section demonstrates both (i) the accuracy of our Monte Carlo "Level" experiments when $\overline{\boldsymbol{\Omega}}$ is known and available for use in the respective estimation procedures, and (ii) the potential efficiency gains that Parks offers over PCSE. The next section examines the performance of these two estimators when the elements of the error covariance matrix are unknown and must be estimated.

IID. Monte Carlo experiments assuming a "Parks-type" error structure where the elements of the population error covariance matrix must be estimated

This section generally repeats the analyses of the previous section, except this time the elements of the population error covariance matrix, $\overline{\boldsymbol{\Omega}}$, are assumed unknown and must be estimated. In other words, we replace $\overline{\boldsymbol{\Omega}}$ with $\hat{\boldsymbol{\Omega}}$ in the respective formulae for $\hat{\boldsymbol{\beta}}$ and $Cov\left(\hat{\boldsymbol{\beta}}\right)$. The experimental setup is very similar to the one employed by BK, and thus provides an opportunity to confirm their experimental results.

TABLE 3 reports the results. As before, the left hand side of the table summarizes the performances of Parks and PCSE with respect to "Level." The top panel (Panel A) reports the results for the experimental observations patterned after TSCS income data; the lower panel (Panel B) reports results for the simulated TSCS tax data.

Centering our attention first on the Parks results in Panel A, we find -- consistent with BK -- that Parks substantially, in some cases dramatically, underestimates coefficient standard errors, resulting in confidence intervals that are too narrow (i.e., "overconfident"). The Parks "Level" values range from a high of *67.2* percent for *N=5, T=25*; to an abysmally low *8.8* percent for *N=20, T=20*. In other words, when *N=20* and *T=20*, less than 10 percent of the 95% confidence intervals include the true value of β_x, causing the null hypothesis to be rejected much too frequently. The results for the simulated tax data of Panel B are similar. A major contribution of BK's research is that it established the degree to which Parks underestimates standard errors. Our research confirms this finding of theirs.

Turning now to the PCSE "Level" values, we come across our first surprising finding – surprising in the sense that one would not have expected this result from a reading of BK. While PCSE always does a better job than Parks when estimating confidence intervals, it also is guilty of underestimating standard errors. For the simulated income data in Panel A, the PCSE "Level" values range from a high of *88.0* percent (*N=48, T=25)*, to a low of *72.0* percent (*N=5, T=10)*. Across all sixteen *NT* experiments, the mean "Level" value for PCSE is *79.3*, substantially less than its expected value of 95. Again, the results for the tax data of Panel B are similar.

The explanation for this result is straightforward. The analytic expressions for the PCSE standard errors, like those for the Parks standard errors, assume that the elements of the population covariance matrix are known. In reality, they are unknown and must be estimated. Estimation of these parameters introduces an additional degree of uncertainty that is not incorporated in the standard error formulae.[6] Thus, the standard error formulae for both Parks and PCSE are biased downwards.

The right hand side of TABLE 3 reports the efficiency of PCSE relative to Parks. Values less than 100 indicate that PCSE is less efficient than Parks. The right hand side of TABLE 3 makes clear that the improvement of PCSE with respect to standard errors comes at a cost of lower efficiency. For example, for the simulated income data in Panel A, when $N = 10$ and $T = 20$, the standard deviation of the Parks-generated $\hat{\beta}_x$ values are approximately 85% of the size of the corresponding standard deviation of the PCSE estimates. Across all experiments, the mean "Efficiency" value for the simulated income data is *85.9*. The results for the simulated tax data of Panel B are, yet again, quite similar.

This is our second "surprising" result. BK claim that the improvement in standard error estimation provided by PCSE comes at negligible cost in terms of efficiency -- except in extreme cases that researchers are unlikely to encounter in actual TSCS data.[7] In contrast, both of our data sets evidence the existence of a tradeoff between efficiency and accurate standard error estimation. We note that this efficiency tradeoff is not driven

---

[6] Both the Parks and PCSE procedures use the same error covariance matrix, $\hat{\boldsymbol{\Omega}}$, to calculate their respective estimators. Thus, both need to estimate the exact same elements.

[7] We note that BK would have predicted the "Efficiency" results of Panel A "in theory." They write: "[PCSE] is … more efficient than Parks when the errors are uncorrelated (spherical). But even when the average correlation of the errors rises to .25, [PCSE] remains slightly more efficient than Parks. Parks becomes more efficient than [PCSE] when average contemporaneous correlations rise to .50, but this advantage is noticeable only when the number of time points is at least double the number of units. Even here, the efficiency advantage of Parks over [PCSE] is under 20%. Only when the average contemporaneous correlation of the errors rises to .75 is the advantage of Parks marked, and then only when T is twice N (BK, page 642)." Referring back to TABLE 1 we see that the simulated income data sets are indeed characterized by a high degree of contemporaneous correlation (the average contemporaneous correlation across all the data sets used in Panel (A) of TABLE 3 is *0.74*). Large efficiency costs occur only when T is more than twice the size of N. However, BK make the statement that, *in practice*, they have never seen actual TSCS data characterized by such "extreme" values of contemporaneous correlation: "We have done this calculation for a variety of TSCS data sets that were sent to us, and none of them met this condition" (BK, page 642).

by bias in the coefficient estimates, as both procedures produce unbiased estimates. Rather, it is a direct result of the greater precision of the Parks estimates, as foreshadowed by the "Efficiency" results of TABLE 2. While the efficiency advantage of Parks over PCSE is smaller when the population error covariance is unknown and must be estimated (compare TABLE 3 with TABLE 2), it is not negligible.

One potential concern with our efficiency results is that the income data is characterized by substantial contemporaneous, cross-sectional correlation. As noted above, while BK claim that use of their PCSE estimator generally entails negligible efficiency costs, they acknowledge that Parks can dominate in rare cases of severe cross-sectional correlation: "Only when the average contemporaneous correlation of the errors rises to 0.75 is the [efficiency] advantage of Parks marked, and then only when *T* is twice *N*" (BK, page 642). While the tax data fall well within the range for which BK assert there should be no marked advantage for Parks, this is not the case for the income data. As a result, we simulated new income data with reduced cross-sectional correlation to see if the efficiency advantages of Parks would persist.

To reduce cross-sectional correlation, we added time fixed effects to the respective residual generating functions, and used the corresponding population covariance matrix to generate experimental income data. The resulting artificial TSCS data evidenced much lower cross-sectional correlation: "*Mean* r_{ij} " was *0.39*, compared to *0.79* for the experimental income data of TABLE 3. Rather than being diminished, however, we found that the efficiency advantages of Parks were even more pronounced. "*Mean Efficiency*" fell from *85.9* to *64.3*. The full results for these experiments are reported in the Appendix.

The following summarizes our experimental findings when the Monte Carlo statistical environment is characterized by a "Parks-type" error structure and the elements of the error covariance are unknown and must be estimated:

1. Parks substantially underestimates coefficient standard errors, resulting in confidence intervals that are much too narrow.

2. PCSE produces more reliable standard error estimates than Parks. However, PCSE also underestimates coefficient standard errors, producing overly narrow confidence intervals.

3. The improvement in standard error estimates provided by PCSE comes at the cost of decreased efficiency.

Whether the tradeoff in improved standard error estimation associated with PCSE is worth the cost in diminished efficiency is, of course, a subjective evaluation that researchers must make for themselves.

We conclude this section by reporting that we obtained very similar results using several different residual generating functions, all within the "Parks-type" statistical environment studied by BK. Of course, in real life, there is no guarantee that the statistical environment falls within the "Parks-type" category. The next section extends the analysis to a statistical environment that is even more general than the Parks model.

III. EXTENDING BK'S ANALYSIS TO A MORE GENERAL STATISTICAL ENVIRONMENT

While a "Parks-type" statistical environment is generally viewed as being very general, it should be noted that it, in fact, imposes substantial limitations on $\boldsymbol{\Omega}$. Given that $\boldsymbol{\varepsilon}$ is $NT \times 1$, there are $\frac{NT(NT+1)}{2}$ unique parameters in the unrestricted version of $\boldsymbol{\Omega}$. In contrast, there are $\frac{N^2+3N}{2}$ unique parameters in the Parks specification of $\boldsymbol{\Omega}$ (counting the group-specific AR[1] parameters). In other words, the Parks model scales down the number of unique parameters in $\boldsymbol{\Omega}$ by approximately $\frac{1}{T^2}$. As it is common in empirical studies using TSCS data for T to range between 10 and 25 years of data (or more), this constitutes a substantial restriction on $\boldsymbol{\Omega}$. Since both FGLS (Parks) and BK's PCSE procedures are not designed to be applied outside the "Parks-type" statistical environment, it is unclear how they will behave, both absolutely and relatively, in a more general and, perhaps, realistic statistical environment.

This section addresses the following questions: Suppose one uses TSCS data which has a more general error structure than that assumed by the Park model. Will Parks and PCSE still underestimate coefficient standard errors? Will PCSE still do a better job than Parks of estimating standard errors? And will PCSE still be less efficient than Parks? While BK never compare Parks and PCSE outside a "Parks-type" statistical environment, we think that researchers will find our results of interest given the widespread popularity of the PCSE methodology.

To construct our more general statistical environment, we repeat the (first-stage) process described in Section (IIB) up to the point where the residuals from the "residual generating function" are used to construct an error covariance matrix for the respective subsample. Rather than constructing a "Parks-type" error covariance matrix, we construct the generalized error covariance matrix, $\hat{\boldsymbol{\Omega}} = \boldsymbol{ee}'$ (similar to how "robust" covariance matrices are calculated). As before, these sample covariance matrices are then averaged to obtain the "representative" error covariance matrix, $\overline{\boldsymbol{\Omega}}$. $\overline{\boldsymbol{\Omega}}$ becomes the "population" covariance matrix for the subsequent NT Monte Carlo experiment.

The right-hand side column of TABLE 1 reports the salient characteristics of the data for this generalized statistical environment. Of course, "*Mean* R^2" is the same as in the left-hand side column, since the first-stage of the data-generating process – which produces the residuals used to construct the sample covariance matrices – is identical (same original data, same residual generating functions). While the specific values differ, it is clear that the simulated data in this "generalized" statistical environment are likewise characterized by substantial degrees of serial correlation, groupwise heteroscedasticity, and cross-sectional correlation.

TABLE 4 reports the results from these Monte Carlo experiments. These are somewhat different from those of TABLE 3. For example, it is no longer true that Parks and PCSE always underestimate coefficient standard errors. When these procedures are applied in a "generalized" statistical environment, they can either under- or over-estimate coefficient standard errors. For example, for the simulated income data, the "Level' values for Parks range from a low of *39.9* (*N=20, T=20*) to a high of *100* (several experiments). For the simulated tax data, the corresponding range is *48.6* to *100*. The same is true for the PCSE estimates: For both income and tax data, the corresponding "Level" values lie on both sides of 95.

Nor is it necessarily the case that PCSE always produces more accurate hypothesis tests than Parks. For example, for the simulated TSCS income data when *N=10* and *T=20*, Parks produces a marginally more accurate "Level" result than PCSE (*96.1* versus *97.6*). Similar examples can be found for the tax data. Indeed, were it not for a couple of egregious exceptions (*N=10/T=10* and *N=20/T=20*), one might even be led to conclude that Parks was as good, if not slightly better, than PCSE for hypothesis testing. That being said, PCSE appears, overall, to estimate coefficient standard errors more accurately than Parks. However, PCSE can also be severely inaccurate. For example, our experiments identify cases where PCSE estimated coefficient standard errors are twice, or more, their true size.[8]

Turning now to "Efficiency," we see that it is still true that there are efficiency costs in using PCSE rather than Parks to estimate β_x . If anything, the efficiency costs are greater in the "generalized" statistical environment. The average value of "Efficiency" over all sixteen experiments was *51.8* for the simulated income data, and *76.3* for the tax data. The latter value would have been considerably lower were it not for one outlier case where PCSE was substantially more efficient than Parks (*N=5, T=25*). Both values are lower than their counterparts in TABLE 3. Further, it is no longer true that PCSE compares well with Parks on efficiency grounds when N and T are approximately equal. This is evidenced by both the simulated income and tax data (cf. *N=10/T=10* and *N=20/T=20)*.

The following summarizes our main findings from this analysis of the Parks and PCSE estimators within a "generalized" statistical environment:

1. In a "generalized" statistical environment, both Parks and PCSE can either under- or overestimate coefficient standard errors, so that we cannot sign the direction of the bias associated with using these techniques for hypothesis testing.

2. PCSE usually, but not always, produces more reliable standard error estimates than Parks. However, PCSE estimates can sometimes be highly unreliable.

3. Whenever PCSE provides a benefit in the form of more accurate standard error estimates, it comes at a cost of reduced efficiency.

---

[8] This was true for the income data when *N=5* and *T=10*; and true for the tax data in the following cases: *N=5/T=15; N=10/T=20; N=10/T=25;* and *N=20/T=25*. While not reported in the text, we calculated a "Standard Error Ratio" consisting of the ratio of (i) the average estimated standard error based on the associated covariance formula, over (ii) the sample standard deviation calculated from the 1000 values of $\hat{\beta}_x$. This is essentially the inverse of BK's "Overconfidence" measure.

We note that these findings remained valid when alternative, more fully specified residual-generating functions were used to construct our statistical environments.

IV. CONCLUSION

Time-series, cross-sectional (TSCS) data are extremely useful to researchers and have been widely employed in published research. However, the complex nature of the associated error structure can cause inaccurate estimates of coefficients and their standard errors. Beck and Katz (1995) study the properties of FGLS (Parks) and "OLS with Panel-Corrected Standard Errors" (PCSE) within a simulated statistical environment characterized by serial correlation, groupwise heteroscedasticity, and cross-sectional correlation. They find that Parks produces estimates of coefficient standard errors that are too small, and that the extent of this bias can be substantial. In contrast, they claim that PCSE produces accurate estimates of standard errors, at little to no cost in efficiency, except in extreme cases. Consequently, BK prescribe that researchers use the PCSE procedure when working with TSCS data

Our study revisits BK's comparison of the Parks and PCSE estimators. Our innovation is that we construct simulated statistical environments that are patterned after actual TSCS data. We model experimental statistical environments after income and tax data on U.S. states from 1960-1999. For these data, we find that the benefits of PCSE are smaller, and the costs greater, than a reading of BK would suggest: While PCSE generally does a better job than Parks in estimating standard errors, it too can be unreliable, sometimes producing standard errors that are substantially off the mark.

Further, we find that the benefits of PCSE can come at a substantial cost in estimator efficiency.

Based on our study, we would give the following advice to researchers using TSCS data: Given a choice between Parks and PCSE, we recommend that researchers use PCSE for hypothesis testing, and Parks if their primary interest is accurate coefficient estimates. We caution that our advice is predicated on the assumption that researchers' TSCS data resemble our simulated income and tax data. It would be valuable to supplement our findings with results from other simulated statistical environments patterned after actual TSCS data. That is a topic for future research.

REFERENCES

Beck, Nathaniel and Jonathan N. Katz. "What To Do (and Not To Do) With Time-Series Cross-Section Data." *American Political Science Review* Vol. 89, No. 3 (1995): 634-647.

Brülhart, Marius and Federico Trionfetti. "Public Expenditure, International Specialisation and Agglomeration." *European Economic Review* Vol. 48 (2004): 851-881.

Dejuan, Joseph P. and Maria Jose Luengo-Prado. "Consumption and Aggregate Constraints: International Evidence." *Oxford Bulletin of Economics and Statistics* Vol. 68, No. 1 (2006): 81-99.

Engstrom, Erik and Samuel Kernell. "Manufactured Responsiveness: The Impact of Electoral Laws on Unified Party Control of the Presidency and House of Representatives, 1840-1940." *American Journal of Political Science* Vol. 49, No. 3 (July 2005): 531-549.

Helms, L. Jay. "The Effect of State and Local Taxes on Economic Growth: A Time Series-Cross Section Approach." *The Review of Economics and Statistics* (1985): 574-82.

Jönsson, Kristian. "Cross-sectional Dependency and Size Distortion in a Small-sample Homogeneous Panel Data Unit Root Test." *Oxford Bulletin of Economics and Statistics* Vol. 67, No. 3 (2005): 369-392.

Kacperczyk, Marcin, Clemens, Sialm, and Lu Zheng. "On the Industry Concentration of Actively Managed Equity Mutual Funds." *Journal of Finance* Vol. 60, No. 4 (August 2005): 1983-2011.

Lee, Woojin and John E. Roemer. "The Rise and Fall of Unionised Labour Markets: A Political Economy Approach." *The Economic Journal* Vol. 115 (January, 2005): 28-67.

Nunziata, Luca. "Institutions and Wage Determination: A Multi-country Approach." *Oxford Bulletin of Economics and Statistics* Vol. 67, No. 4 (2005): 435-466.

Reed, W. Robert. "Democrats, Republicans, and Taxes: Evidence that Political Parties Matter." *Journal of Public Economics* (2006): forthcoming.

Soo, Kwok Tong. "Zipf's Law for Cities: A Cross-country Investigation." *Regional Science and Urban Economics* Vol. 35, No. 3 (May 2005): 239-263.

Parks, Richard. “Efficient Estimation of a System of Regression Equations When Disturbances Are Both Serially and Contemporaneously Correlated.” *Journal of the American Statistical Association* 62 (1967): 500-509.

StataCorp. *Stata Statistical Software: Release 7.0.* College Station, TX: Stata Corporation, 2001.

Wasylenko, Michael. "Taxation and Economic Development: The State of the Economic Literature." New England Economic Review (March/April 1997): 37-52.

Summary of Diagnostics

	"PARKS-TYPE" STATISTICAL ENVIRONMENT	*GENERALIZED STATISTICAL ENVIRONMENT*
	Income Data	
Mean R^2	0.728	0.728
Mean $\hat{\rho}$	0.61	0.82
Mean $h = \left(\frac{\hat{\sigma}_{75th\ percentile}}{\hat{\sigma}_{25th\ percentile}} \right)$	1.24	1.31
Mean $r_{ij} = \left(\frac{\hat{\sigma}_{\varepsilon,ij}}{\sqrt{\hat{\sigma}_{\varepsilon,ii}}\sqrt{\hat{\sigma}_{\varepsilon,jj}}} \right)$	0.74	0.64
	Tax Data	
Mean R^2	0.701	0.701
Mean $\hat{\rho}$	0.58	0.64
Mean $h = \left(\frac{\hat{\sigma}_{75th\ percentile}}{\hat{\sigma}_{25th\ percentile}} \right)$	1.59	1.55
Mean $r_{ij} = \left(\frac{\hat{\sigma}_{\varepsilon,ij}}{\sqrt{\hat{\sigma}_{\varepsilon,ii}}\sqrt{\hat{\sigma}_{\varepsilon,jj}}} \right)$	0.36	0.41

NOTE: Means are calculated over all replications (1000 replications per experiment) and experiments (a total of 16 experiments based on 16 possible *N* and *T* combinations).

TABLE 2
Performance of Parks and PCSE Estimators in a "Parks-type" Statistical Environment where Elements of the Error Covariance Matrix Are Known (Using Simulated TSCS Income Data)

		Level				*Efficiency*			
		T=10	*T=15*	*T=20*	*T=25*	*T=10*	*T=15*	*T=20*	*T=25*
N=5	*Parks*	95.5	95.2	95.3	94.8	91.5	86.0	70.8	47.0
	PCSE	96.4	95.2	94.6	94.2				
N=10	*Parks*	94.6	96.1	95.4	95.2	76.6	67.1	56.9	24.8
	PCSE	95.2	95.2	95.0	93.8				
N=20	*Parks*	----	----	94.4	94.9	----	----	25.4	13.0
	PCSE	96.4	94.4	94.2	94.6				
N=48	*Parks*	----	----	----	----	----	----	----	----
	PCSE	96.3	95.5	94.7	95.7				
MEAN	*Parks*	95.1				55.9			
	PCSE	95.1							

NOTE: "Level" and "Efficiency" are defined in the text (cf. Section IIA). "Mean" refers to the average value over all replications (1000 replications per experiment) and all experiments. For the PCSE "Level" estimates, there are a total of 16 experiments. For the Parks "Level" estimates and the "Efficiency" estimates, there are only 10 experiments, because Parks cannot be calculated when $N < T$.

TABLE 3
Performance of Parks and PCSE Estimators in a "Parks-type" Statistical Environment where Elements of the Population Error Covariance Matrix Must Be Estimated

		A. Income Data							
		Level				*Efficiency*			
		T=10	***T=15***	***T=20***	***T=25***	***T=10***	***T=15***	***T=20***	***T=25***
N=5	***Parks***	62.1	65.2	67.1	67.2	101.4	98.9	79.6	58.8
	PCSE	72.0	72.6	77.5	76.9				
N=10	***Parks***	29.2	49.3	51.1	49.4	98.2	94.1	84.8	61.6
	PCSE	79.6	75.3	75.3	81.6				
N=20	***Parks***	----	----	8.8	11.2	----	----	98.0	83.6
	PCSE	82.5	79.3	78.8	86.9				
N=48	***Parks***	----	----	----	----	----	----	----	----
	PCSE	82.1	78.7	81.9	88.0				
MEAN	***Parks***	46.1				85.9			
	PCSE	79.3							

TABLE 3 (Continued)
Performance of Parks and PCSE Estimators in a "Parks-type" Statistical Environment where Elements of the Population Error Covariance Matrix Must Be Estimated

		B. Tax Data							
		Level				*Efficiency*			
		T=10	*T=15*	*T=20*	*T=25*	*T=10*	*T=15*	*T=20*	*T=25*
N=5	***Parks***	66.1	70.2	76.5	77.6	84.9	77.8	72.2	71.7
	PCSE	83.3	82.0	84.6	86.1				
N=10	***Parks***	23.4	52.3	61.0	65.3	96.8	84.0	76.0	72.8
	PCSE	86.1	85.3	87.1	86.0				
N=20	***Parks***	----	----	8.4	23.2	----	----	97.4	85.7
	PCSE	87.6	89.9	87.4	86.6				
N=48	***Parks***	----	----	----	----	----	----	----	----
	PCSE	86.5	88.0	90.0	89.8				
MEAN	***Parks***	52.4				81.9			
	PCSE	86.6							

NOTE: "Level" and "Efficiency" are defined in the text (cf. Section IIA). "Mean" refers to the average value over all replications (1000 replications per experiment) and all experiments. For the PCSE "Level" estimates, there are a total of 16 experiments. For the Parks "Level" estimates and the "Efficiency" estimates, there are only 10 experiments, because Parks cannot be calculated when $N < T$.

TABLE 4
Performance of Parks and PCSE Estimators in a "Generalized" Statistical Environment

<table>
<tr><th colspan="2" rowspan="2"></th><th colspan="8">A. Income Data</th></tr>
<tr><th colspan="4">Level</th><th colspan="4">Efficiency</th></tr>
<tr><th></th><th></th><th>T=10</th><th>T=15</th><th>T=20</th><th>T=25</th><th>T=10</th><th>T=15</th><th>T=20</th><th>T=25</th></tr>
<tr><td rowspan="2">N=5</td><td>Parks</td><td>100</td><td>100</td><td>99.9</td><td>99.8</td><td rowspan="2">95.2</td><td rowspan="2">48.5</td><td rowspan="2">30.6</td><td rowspan="2">21.6</td></tr>
<tr><td>PCSE</td><td>100</td><td>100</td><td>83.7</td><td>98.8</td></tr>
<tr><td rowspan="2">N=10</td><td>Parks</td><td>81.7</td><td>100</td><td>96.1</td><td>93.0</td><td rowspan="2">49.8</td><td rowspan="2">28.6</td><td rowspan="2">49.0</td><td rowspan="2">52.9</td></tr>
<tr><td>PCSE</td><td>100</td><td>99.4</td><td>97.6</td><td>97.5</td></tr>
<tr><td rowspan="2">N=20</td><td>Parks</td><td>----</td><td>----</td><td>39.9</td><td>59.7</td><td rowspan="2">----</td><td rowspan="2">----</td><td rowspan="2">77.6</td><td rowspan="2">64.7</td></tr>
<tr><td>PCSE</td><td>96.0</td><td>94.2</td><td>98.9</td><td>96.2</td></tr>
<tr><td rowspan="2">N=48</td><td>Parks</td><td>----</td><td>----</td><td>----</td><td>----</td><td rowspan="2">----</td><td rowspan="2">----</td><td rowspan="2">----</td><td rowspan="2">----</td></tr>
<tr><td>PCSE</td><td>94.2</td><td>91.0</td><td>99.7</td><td>87.5</td></tr>
<tr><td rowspan="2">MEAN</td><td>Parks</td><td colspan="4">87.0</td><td colspan="4" rowspan="2">51.8</td></tr>
<tr><td>PCSE</td><td colspan="4">95.9</td></tr>
</table>

TABLE 4: Continued
Performance of Parks and PCSE Estimators in a "Generalized" Statistical Environment

		B. Tax Data							
		Level				*Efficiency*			
		T=10	*T=15*	*T=20*	*T=25*	*T=10*	*T=15*	*T=20*	*T=25*
N=5	***Parks***	100	100	99.6	99.7	38.5	74.2	100.3	176.4
	PCSE	100	100	100	100				
N=10	***Parks***	72.9	99.4	99.8	99.7	60.4	37.5	65.9	81.2
	PCSE	100	99.9	100	100				
N=20	***Parks***	----	----	48.6	90.9	----	----	72.3	56.4
	PCSE	98.9	99.8	100	100				
N=48	***Parks***	----	----	----	----	----	----	----	----
	PCSE	98.8	100	100	100				
MEAN	***Parks***	91.1				76.3			
	PCSE	99.8							

NOTE: "Level" and "Efficiency" are defined in the text (cf. Section IIA). "Mean" refers to the average value over all replications (1000 replications per experiment) and all experiments. For the PCSE "Level" estimates, there are a total of 16 experiments. For the Parks "Level" estimates and the "Efficiency" estimates, there are only 10 experiments, because Parks cannot be calculated when $N < T$.

APPENDIX:

Replication of TABLE 3/Panel A When the Artificial Income Data Have Lower Cross-sectional Correlation (*Mean* $r_{ij} = 0.39$)

		A. Income Data							
		Level				*Efficiency*			
		T=10	*T=15*	*T=20*	*T=25*	*T=10*	*T=15*	*T=20*	*T=25*
N=5	***Parks***	34.3	52.8	64.4	72.6	61.8	51.0	41.2	34.5
	PCSE	82.1	82.0	84.8	83.5				
N=10	***Parks***	0.2	4.6	7.5	20.6	94.1	71.5	59.3	49.6
	PCSE	85.2	86.7	85.4	86.9				
N=20	***Parks***	----	----	0.0	0.0	----	----	97.1	82.8
	PCSE	84.8	87.7	87.2	89.0				
N=48	***Parks***	----	----	----	----	----	----	----	----
	PCSE	87.8	87.3	88.9	90.6				
MEAN	***Parks***	25.7				64.3			
	PCSE	86.2							

NOTE: "Level" and "Efficiency" are defined in the text (cf. Section IIA). "Mean" refers to the average value over all replications (1000 replications per experiment) and all experiments. For the PCSE "Level" estimates, there are a total of 16 experiments. For the Parks "Level" estimates and the "Efficiency" estimates, there are only 10 experiments, because Parks cannot be calculated when $N < T$.