Discussion Paper Series

No.212

The Bias-Corrected First-Difference Estimator in Dynamic Panel Data Models with Cross Section Dependence and Heteroscedasticity

Kazuhiko Hayakawa

May 2007

Hitotsubashi University Research Unit
for Statistical Analysis in Social Sciences
A 21st-Century COE Program

Institute of Economic Research
Hitotsubashi University
Kunitachi, Tokyo, 186-8603 Japan
http://hi-stat.ier.hit-u.ac.jp/

The Bias-Corrected First-Difference Estimator in Dynamic Panel Data Models with Cross Section Dependence and Heteroscedasticity

Kazuhiko Hayakawa*†

Department of Economics, Hitotsubashi University

JSPS Research Fellow

May 20, 2007

Abstract

In this paper, we show that the bias-corrected first-difference (BCFD) estimator suggested by Chowdhury (1987) can be applied to the case where the error terms are cross-sectionally dependent and heteroscedastic. By deriving the finite sample bias of the BCFD estimator, we find that the BCFD estimator has small bias when T, the dimension of the time series, is not very large and ρ, the autoregressive parameter, is close to one. Simulation results show that the BCFD estimator performs better than existing estimators, especially when T is not very large.

*E-mail : em031112@ yahoo.co.jp (Remove the space after @.)

†The author is deeply grateful to two referees, Taku Yamamoto, Eiji Kurozumi, Hiroaki Chigira, Hosung Jung, Asuka Takeuchi, and the participants at the 2006 Japanese Joint Statistical Meeting held at Tohoku University for helpful comments. Financial support from the JSPS Fellowship is also gratefully acknowledged. All remaining errors are my own.

1 Introduction

In recent years, a considerable number of studies discussing the estimation and inference of stationary dynamic panel data models have been published.[1] However, most of these studies assume that individuals are cross-sectionally independent, a restriction that is unlikely to hold in many applications. For example, if we use panel data of countries, regions, or industries, it is natural to assume that individuals are correlated.

Representative studies dealing with cross-sectional dependence in the context of stationary dynamic panel data models are those by Phillips and Sul (2003, 2007). The first of these (Phillips and Sul, 2003) investigated the effect of cross section dependence on the performance of the least squares dummy variables (LSDV) estimator,[2] and proposed panel median unbiased estimators as alternatives to the LSDV estimator. Their second study (Phillips and Sul, 2007) derives the asymptotic properties of the LSDV estimator under cross section dependence and heteroscedasticity and proposed a panel feasible generalized mean unbiased estimator to reduce the bias of the LSDV estimator.

The purpose of this paper is to show that the bias-corrected first-difference (BCFD) estimator by Chowdhury (1987) can be applied to the case where errors are cross-sectionally dependent and heteroscedastic.[3]

One of the advantages of the BCFD estimator is that, in the case of AR(1) models, it is very easy to compute. Although the BCFD estimator can be extended to more general AR(p) models, this would require the use of numerical optimization procedures.[4] To utilize the advantage of the BCFD estimator, i.e., its tractability, and to simplify the derivation of the theoretical properties, we mainly consider AR(1) panel data models in this paper. However, it will be argued later that the BCFD estimator can be applied to models with exogenous regressors exactly in the same way as Phillips and Sul's (2007) approach.

Although the use of AR(1) specifications is somewhat limited, a number of empirical

[1]For a recent review, see Arellano (2003).

[2]Phillips and Sul (2003) refer to the LSDV estimator as the pooled panel least squares estimator.

[3]For other studies on the BCFD estimator, see Wansbeek and Knaap (1999), Ramalho (2005), Han and Phillips (2007), Phillips and Han (2006), and Hayakawa (2006, 2007).

[4]See Chowdhury (1987).

studies have used AR(1) panel data models. In the field of macroeconomics, Frankel and Rose (1996), for example, used an AR(1) panel data model to estimate the speed of adjustment toward purchasing power parity, while Shioji (2004) and Ho (2006) used AR(1) panel data models to analyze growth convergence.[5] And in microeconomics, Hirano (2002) used an AR(1) model to examine earning dynamics. Therefore, as these empirical examples show, AR(1) specifications do have their uses.

The remainder of the paper is organized as follows. In Section 2, we define the model, provide some assumptions, and review the BCFD estimator. The main results of this paper are then presented in Section 3, while Section 4 examines the performance of the BCFD estimator using Monte Carlo simulations. Section 5 concludes.

2 Model, assumptions, and the bias-corrected first-difference estimator

We consider an AR(1) panel data model given by

$$y_{it} = a_i + \rho y_{i,t-1} + u_{it}, \qquad \rho \in (-1,1) \quad (i=1,...,N;\ t=1,...,T) \tag{1}$$

$$y_{it} = a_i^0 + y_{i,t}^0, \quad y_{it}^0 = \rho y_{i,t-1}^0 + u_{it}, \qquad \rho = 1 \qquad (i=1,...,N;\ t=1,...,T) \tag{2}$$

where ρ is the parameter of interest and u_{it} has the following factor component structure:

$$u_{it} = \sum_{s=1}^{K} \delta_{si}\theta_{st} + \varepsilon_{it} = \delta_i'\theta_t + \varepsilon_{it} \tag{3}$$

In the unit root case, we assume that y_{i0}^0 is $O_p(1)$. Moreover, following Phillips and Sul (2007), we make the following assumptions:

Assumption 1. *ε_{it} have zero mean, finite $2+2\nu$ moments for some $\nu > 0$, are independent over i and t with $var(\varepsilon_{it}) = \sigma_i^2$ for all t, and $\lim_{N\to\infty} N^{-1}\sum_{i=1}^{N}\sigma_i^2 = \sigma^2$.*

Assumption 2. *The factors θ_t are $iid(0,\Sigma_\theta)$ over t and the factor loadings δ_i are nonrandom parameters satisfying $\lim_{N\to\infty} N^{-1}\sum_{i=1}^{N}\delta_i\delta_i' = M_\delta$. When $K=1$, we set $\Sigma_\theta = \sigma_\theta^2$ and $M_\delta = m_\delta^2$.*

[5] For a recent review of the literature on growth convergence, see Maddala (1999).

By first-differencing model (1), we have

$$\Delta y_{it} = \rho \Delta y_{i,t-1} + \Delta u_{i,t} \qquad\qquad i = 1, ..., N \quad and \quad t = 2, ..., T. \tag{4}$$

The OLS estimator of this model is given by

$$\widehat{\rho}_{fd} \;=\; \frac{\sum_{t=2}^{T}\sum_{i=1}^{N}\Delta y_{i,t-1}\Delta y_{i,t}}{\sum_{t=2}^{T}\sum_{i=1}^{N}\Delta y_{i,t-1}^2} = \rho + \frac{\sum_{t=2}^{T}\sum_{i=1}^{N}\Delta y_{i,t-1}\Delta u_{i,t}}{\sum_{t=2}^{T}\sum_{i=1}^{N}\Delta y_{i,t-1}^2} = \rho + \frac{A}{B}. \tag{5}$$

The BCFD estimator considered by Chowdhury (1987) takes the following form:

$$\widehat{\rho}_{bcfd} = 2\widehat{\rho}_{fd} + 1. \tag{6}$$

The BCFD estimator is closely related to the bias-corrected estimator developed by Bun and Carree (2005a, b, 2006) and Phillips and Sul (2007). Both studies proposed a bias-correction method using the inverse function of $\rho + B(\rho, T)$, where $B(\rho, T)$ is the asymptotic bias of the LSDV estimator when $N \to \infty$. Since $\rho + B(\rho, T)$ is a complex function of ρ and T, it is quite difficult to derive the explicit expression of the bias-corrected estimator and in practice we therefore need to use a numerical optimization procedure. However, when the same bias-correction method is applied to the OLS estimator of the first-differenced model, we can derive the bias-corrected estimator explicitly as $\widehat{\rho}_{bcfd}$. In other words, although the basic idea of bias-correction underlying the BCFD estimator and the bias-corrected estimator by Bun and Carree (2005a,b, 2006) and Phillips and Sul (2007) is the same, they differ in that the BCFD estimator can be expressed explicitly as $\widehat{\rho}_{bcfd}$, while the bias-corrected estimator by Bun and Carree (2005a,b, 2006) and Phillips and Sul (2007) cannot.

This implies that although the analysis in this paper concentrates on AR(1) panel data models, it is also possible to use the BCFD estimator for models with exogenous regressors using the same methodology as Phillips and Sul (2007, pp.176-177), since the basic structure of bias-correction is identical.

In the next section, we consider the asymptotic and finite sample properties of the BCFD estimator.

3 Some properties of the BCFD estimator

In this section, we first derive the asymptotic properties of the BCFD estimator under a large N and fixed T and the large N and large T asymptotics. We then derive the finite sample bias associated with the order of T^{-1}.

The asymptotic properties of $\widehat{\rho}_{fd}$ are established in the following theorem:

Theorem 1. *Let Assumptions 1 and 2 hold. Then, as $N \to \infty$ with T fixed, we have*

$$\operatorname*{plim}_{N\to\infty} \widehat{\rho}_{fd} = \rho + \frac{T_0^{-1}\sum_{t=2}^{T} trace(\Delta F_{\theta,t-1}\Delta\theta_t' M_\delta) - \sigma^2}{T_0^{-1}\sum_{t=2}^{T} trace(\Delta F_{\theta,t-1}\Delta F_{\theta,t-1}' M_\delta) + \frac{2}{1+\rho}\sigma^2}, \quad \rho \in (-1,1) \tag{7}$$

$$\operatorname*{plim}_{N\to\infty} \widehat{\rho}_{fd} = \frac{T_0^{-1}\sum_{t=2}^{T} trace(\theta_{t-1}\theta_t' M_\delta)}{T_0^{-1}\sum_{t=2}^{T} trace(\theta_{t-1}\theta_{t-1}' M_\delta) + \sigma^2} \qquad \rho = 1 \tag{8}$$

where $T_0 = T-1$. When $T \to \infty$ regardless of whether N is fixed or $N \to \infty$, we have

$$\operatorname*{plim}_{T\to\infty} \widehat{\rho}_{fd} = \frac{\rho-1}{2} \qquad \rho \in (-1,1) \tag{9}$$

$$\operatorname*{plim}_{T\to\infty} \widehat{\rho}_{fd} = 0 \qquad \rho = 1 \tag{10}$$

where $F_{\theta,t} = \sum_{j=0}^{\infty} \rho^j \theta_{t-j}$.

Remark 1 When N is large and T is fixed, the probability limit of $\widehat{\rho}_{fd}$ is random, since it depends on the random variable θ_t. Phillips and Sul (2007) obtained a similar result for the LSDV estimator. When T is large regardless of whether N is fixed or large, $\widehat{\rho}_{fd}$ is inconsistent and the probability limit is a linear function of ρ when $|\rho| < 1$.

The asymptotic properties of $\widehat{\rho}_{bcfd}$ are provided in the following theorem:

Theorem 2. *Let Assumptions 1 and 2 hold. Then, as $N \to \infty$ with T fixed, we have,*

$$\operatorname*{plim}_{N\to\infty} \widehat{\rho}_{bcfd} = 2\rho + 1 + \frac{2T^{-1}\sum_{t=1}^{T} trace(\Delta F_{\theta,t-1}\Delta\theta_t' M_\delta) - 2\sigma^2}{T^{-1}\sum_{t=1}^{T} trace(\Delta F_{\theta,t-1}\Delta F_{\theta,t-1}' M_\delta) + \frac{2}{1+\rho}\sigma^2} \quad \rho \in (-1,1) \tag{11}$$

$$\operatorname*{plim}_{N\to\infty} \widehat{\rho}_{bcfd} = 1 + \frac{2T^{-1}\sum_{t=1}^{T} trace(\theta_{t-1}\theta_t' M_\delta)}{T^{-1}\sum_{t=1}^{T} trace(\theta_{t-1}\theta_{t-1}' M_\delta) + \sigma^2} \qquad \rho = 1. \tag{12}$$

When $T \to \infty$ regardless of whether N is fixed or large, we have

$$\operatorname*{plim}_{T\to\infty} \widehat{\rho}_{bcfd} = \rho \qquad \rho \in (-1,1) \tag{13}$$

$$\operatorname*{plim}_{T\to\infty} \widehat{\rho}_{bcfd} = 1 \qquad \rho = 1 \tag{14}$$

Remark 2 We find that $\widehat{\rho}_{bcfd}$ is still inconsistent under large N and fixed T asymptotics since its probability limit depends on the random variable θ_t. However, when T is large regardless of N, $\widehat{\rho}_{bcfd}$ is consistent. This result is different from the independent panel case where either only N or T is required to tend to infinity to be consistent.

We proceed to consider the finite sample properties of the BCFD estimator. The finite sample bias of the BCFD estimator is given in the next lemma.

Lemma 1. *Let Assumptions 1 and 2 hold. Furthermore, let us assume that θ_t is normally distributed. Then, in the single factor case ($K = 1$), the finite sample bias of $\widehat{\rho}_{bcfd}$ with T fixed is given by*

$$E\left(\operatorname*{plim}_{N\to\infty} \widehat{\rho}_{bcfd}\right) = \rho + \frac{1-\rho}{T-1}\left(\frac{m_\delta^2\sigma_\theta^2}{\sigma^2 + m_\delta^2\sigma_\theta^2}\right)^2 + o(T^{-1}) \qquad \rho \in (-1,1) \tag{15}$$

$$E\left(\operatorname*{plim}_{N\to\infty} \widehat{\rho}_{bcfd}\right) = 1 + o(T^{-1}) \qquad \rho = 1. \tag{16}$$

Remark 3 We impose the normality assumption on θ_t to simplify the expression of the finite sample bias. Note that normality is not required to obtain the consistency of $\widehat{\rho}_{bcfd}$.

Remark 4 From this expression, we find that when ρ is close to unity, the bias of $\widehat{\rho}_{bcfd}$ becomes small. This is in contrast to the LSDV estimator, $\widehat{\rho}_{lsdv}$. Phillips and Sul (2007) derived the following results using an expansion when T is large:

$$E\left(\operatorname*{plim}_{N\to\infty} \widehat{\rho}_{lsdv}\right) = \rho - \frac{1+\rho}{T} - \frac{2\rho}{T}\left(\frac{m_\delta^2\sigma_\theta^2}{\sigma^2 + m_\delta^2\sigma_\theta^2}\right) + o(T^{-1}). \tag{17}$$

From (17), we find that the bias of $\widehat{\rho}_{lsdv}$ increases as ρ approaches unity. Therefore, in terms of the bias, $\widehat{\rho}_{bcfd}$ is preferable to $\widehat{\rho}_{lsdv}$, especially when T is not very large and ρ is close to one.

In terms of empirical applications, and especially the analysis of macro panel data, there are two cases where the property that $\widehat{\rho}_{bcfd}$ has small bias when T is not very large and ρ is close to one is important. First, if we use large T panel data, we can of course capture the dynamics of variables more precisely than in the case where we use small T panel data. However, if we use large T panel data, data are likely to be subject to structural breaks and resulting estimates might not be reliable. Although data are less likely to be subject to structural breaks if we use shorter panel data, the bias of $\widehat{\rho}_{lsdv}$ tends to be large. In a situation such as this, $\widehat{\rho}_{bcfd}$ is useful since the bias of $\widehat{\rho}_{bcfd}$ is small even if T is not very large. The second situation in which $\widehat{\rho}_{bcfd}$ is useful is in the analysis of growth models. In the analysis of growth models, we often use data at, say, five-year intervals, since data then are less likely to be serially correlated and the effects of business cycle fluctuations are mitigated.[6] In this case, the dimension of available time series T becomes small and $\widehat{\rho}_{lsdv}$ tends to be biased, while $\widehat{\rho}_{bcfd}$ has small bias. Thus, in these situations, the BCFD estimator is more useful than the LSDV estimator in terms of bias.

In the next section, we compare the performance of the BCFD estimator with existing estimators via Monte Carlo experiments.

4 Monte Carlo experiments

In this section, we conduct Monte Carlo experiments to examine the performance of the BCFD estimator. We consider the following AR(1) model:

$$y_{i,t} = a_i + \rho y_{i,t-1} + u_{it} \tag{18}$$

$$u_{i,t} = c\delta_i\theta_t + v_{it} \tag{19}$$

where a_i and θ_t are independently generated from $N(0,1)$, δ_i are generated from uniform distribution $U[0,1]$, and $y_{i,0}$ are generated to satisfy stationarity.[7] We consider the following four cases:

(i) u_{it} are iid, $(c = 0, var(v_{it}) = \sigma^2)$.

(ii) u_{it} are cross-sectionally dependent and v_{it} are homoscedastic, $(c = 1, var(v_{it}) = \sigma^2)$.

(iii) u_{it} are cross-sectionally independent and v_{it} are heteroscedastic, $(c = 0, var(v_{it}) = \sigma_i^2)$.

(iv) u_{it} are cross-sectionally dependent and v_{it} are heteroscedastic, $(c = 1, var(v_{it}) = \sigma_i^2)$.

[6] See Islam (1995, p.1140).

[7] δ_i are generated once and used repeatedly in each replication.

For cases (i) and (ii), v_{it} are generated from $v_{it} \sim iidN(0,1)$, while for cases (iii) and (iv), v_{it} are independently generated from $N(0,\sigma_i^2)$ with $\sigma_i^2 \sim \chi^2(1)$. The sample sizes we consider are $N = 10, 25, 50, 100, 200$ and $T = 10, 15, 25, 50$. ρ is set to $\rho = 0.5, 0.9, 0.95$. The number of replications is 5000 for all cases. We computed the BCFD estimator, $\widehat{\rho}_{bcfd}$, the LSDV estimator, $\widehat{\rho}_{lsdv}$, and Phillips and Sul's (2007) feasible generalized mean unbiased estimator (FGMUE) based on the residual variance of mean unbiased estimator with common time effects, $\widehat{\rho}_{ps}$.[8] Tables 1, 2, 3, and 4 correspond to case (i), (ii), (iii), and (iv), respectively.

Looking at Tables 1 to 4, we find that the bias of the BCFD estimator is very small in almost all cases. In case (i), $\widehat{\rho}_{ps}$ has larger bias than $\widehat{\rho}_{bcfd}$ when $T = 10$. In case (ii), the bias of $\widehat{\rho}_{lsdv}$ is substantial when $\rho = 0.95$ even when $T = 50$. $\widehat{\rho}_{ps}$ performs well when T is as large as 25. However, when $T = 10$ and $\rho = 0.95$, $\widehat{\rho}_{ps}$ exceeds one. In case (iii), the performance of $\widehat{\rho}_{lsdv}$ is similar to that in case (ii). However, $\widehat{\rho}_{ps}$ exhibits quite different results from case (ii). The bias of $\widehat{\rho}_{ps}$ is substantial. Especially when $\rho = 0.95$, $\widehat{\rho}_{ps}$ exceeds unity even when $T = 50$. In terms of the bias and the root mean squred error (RMSE), $\widehat{\rho}_{bcfd}$ performs best when $T \leq 25$ and ρ is large. The reason why $\widehat{\rho}_{ps}$ does not work well with heteroscedastic errors is that it requires the estimation of the variance σ_i^2. In fact, estimating σ_i^2 based on the method by Phillips and Sul (2003) is not expected to work well unless T is large. In contrast, since $\widehat{\rho}_{bcfd}$ does not require the estimation of σ_i^2, the bias of $\widehat{\rho}_{ps}$ remains small. In case (iv), similar comments as in case (iii) apply. The bias of $\widehat{\rho}_{lsdv}$ is large, although it has the smallest RMSE among the three estimators in some cases. The bias of $\widehat{\rho}_{ps}$ is also substantial. For example, even when $T = 50$, there remains large bias, especially when ρ is large. However, unlike $\widehat{\rho}_{lsdv}$ and $\widehat{\rho}_{ps}$, $\widehat{\rho}_{bcfd}$ has very small bias.

As a final assessment of the estimators, we consider the case of small T, say, $T = 5$. In this case, the common factor θ_t, which is assumed to be random (Assumption 2), might be fixed since the choice of the time horizon is not random if long panel data are not available. In this case, of course, Lemma 1 does not hold since it is derived under the assumption that θ_t is random. Thus, we compare the cases when θ_t is random and fixed for $T = 5$. The simulation result is given in Table 5. For the case of random θ_t, a new θ_t is generated for each replication, and for the case when θ_t is fixed, θ_t is generated

[8] $\widehat{\rho}_{ps}$ corresponds to (E) in Table 2 in Phillips and Sul (2007).

once and used repeatedly in each replications. We find from Table 5 that when θ_t is random, the biases of $\widehat{\rho}_{bcfd}$ are quite small, while when θ_t is fixed, all estimators do not work well. This implies that if there is evidence that θ_t cannot be regarded as random, all estimators considered in this paper are inappropriate and a new estimator will be required.

Therefore, we can conclude that, in terms of the bias, the BCFD estimator performs best in almost all the cases, and in terms of the RMSE, the BCFD estimator is superior to the other two estimators when T is moderately large, say $T = 15$, and when the error terms are heteroscedastic. When T is small and when it might be inappropriate to assume that θ_t is random, all the estimators do not work well, resulting in the need to develop a new estimator. This is, however, beyond the scope of the present paper.

5 Conclusion

In this paper, we showed that the bias-corrected first-difference estimator developed by Chowdhury (1987) can be applied to the case where the errors are cross-sectionally dependent and heteroscedastic. The theoretical analysis indicated that the BCFD estimator is consistent when T is large regardless of N, and is inconsistent when N is large and T is fixed. However, using asymptotic expansions, we showed that the BCFD estimator has small bias when ρ is close to one even if T is not very large. The simulation results showed that the BCFD estimator has smaller bias than the two existing estimators, $\widehat{\rho}_{lsdv}$ and $\widehat{\rho}_{ps}$. Therefore, we conclude that the BCFD estimator is useful when T is not very large and when ρ is close to one. Also, if θ_t is considered to be fixed as in the case of small T, we need to note that all the estimators including the BCFD and Phillips and Sul's (2007) estimators do not work well.

A Mathematical Proofs

Proof of Theorem 1 Note that A and B in (5) can be decomposed as follows:

$$\begin{aligned}
\frac{1}{NT_0}A &= \frac{1}{NT_0}\sum_{i=1}^{N}\sum_{t=2}^{T}\left[\delta_i'\Delta F_{\theta,t-1}\Delta\theta_t'\delta_i + \delta_i'\Delta F_{\theta,t-1}\Delta\varepsilon_{it} + \Delta w_{i,t-1}\delta_i'\Delta\theta_{t-1} + \Delta w_{i,t-1}\Delta\varepsilon_{it}\right] \\
&= A_1 + A_2 + A_3 + A_4 \\
\frac{1}{NT_0}B &= \frac{1}{NT_0}\sum_{i=1}^{N}\sum_{t=2}^{T}\left[\delta_i'\Delta F_{\theta,t-1}\Delta F_{\theta,t-1}'\delta_i - 2\delta_i'\Delta F_{\theta,t-1}\Delta w_{i,t-1} + \Delta w_{i,t-1}^2\right] \\
&= B_1 - 2B_2 + B_3
\end{aligned}$$

Let us define $w_{it} = \sum_{j=0}^{\infty}\rho^j\varepsilon_{i,t-j}$. Then, the probability limit of A and B under large N and fixed T asymptotics are given by

$$\begin{aligned}
\operatorname*{plim}_{N\to\infty} A_1 &= \lim_{N\to\infty}\frac{1}{NT_0}\sum_{i=1}^{N}\sum_{t=2}^{T} trace\left[\Delta F_{\theta,t-1}\Delta\theta_t'\delta_i\delta_i'\right] = \frac{1}{T_0}\sum_{t=2}^{T} trace\left[\Delta F_{\theta,t-1}\Delta\theta_t' M_\delta\right] \\
\operatorname*{plim}_{N\to\infty} A_2 &= 0 \\
\operatorname*{plim}_{N\to\infty} A_3 &= 0 \\
\operatorname*{plim}_{N\to\infty} A_4 &= -\sigma^2 \\
\operatorname*{plim}_{N\to\infty} B_1 &= \lim_{N\to\infty}\frac{1}{NT_0}\sum_{i=1}^{N}\sum_{t=2}^{T} trace\left[\Delta F_{\theta,t-1}\Delta F_{\theta,t-1}'\delta_i\delta_i'\right] = \frac{1}{T_0}\sum_{t=2}^{T} trace\left[\Delta F_{\theta,t-1}\Delta F_{\theta,t-1}' M_\delta\right] \\
\operatorname*{plim}_{N\to\infty} B_2 &= 0 \\
\operatorname*{plim}_{N\to\infty} B_3 &= \frac{2}{1+\rho}\sigma^2
\end{aligned}$$

Then the first result follows. The result under large T asymptotics is obtained from the following probability limit:

$$\begin{aligned}
\operatorname*{plim}_{T\to\infty} A_1 &= -\frac{1}{N}\sum_{i=1}^{N}\delta_i'\Sigma_\theta\delta_i \\
\operatorname*{plim}_{T\to\infty} A_2 &= 0 \\
\operatorname*{plim}_{T\to\infty} A_3 &= 0 \\
\operatorname*{plim}_{T\to\infty} A_4 &= -\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 \\
\operatorname*{plim}_{T\to\infty} B_1 &= \left(\frac{2}{1+\rho}\right)\frac{1}{N}\sum_{i=1}^{N}\delta_i'\Sigma_\theta\delta_i \\
\operatorname*{plim}_{T\to\infty} B_2 &= 0 \\
\operatorname*{plim}_{T\to\infty} B_3 &= \left(\frac{2}{1+\rho}\right)\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2
\end{aligned}$$

Note that we allow N to be finite or infinite since the terms associated with a summation over i are cancelled out.

Proof of Theorem 2 The proof of Theorem 2 is straightforward to show using Theorem 1.

Proof of Lemma 1 We first prove the stationary case. After taking the limit $N \to \infty$, we have

$$\operatorname*{plim}_{N\to\infty} \widehat{\rho}_{fd} = \frac{m_\delta^2 \frac{1}{T_0}\sum_{t=2}^{T}(F_{\theta,t-1}-F_{\theta,t-2})(F_{\theta,t}-F_{\theta,t-1}) - \sigma^2\left(\frac{1-\rho}{1+\rho}\right)}{m_\delta^2 \frac{1}{T_0}\sum_{t=2}^{T}(F_{\theta,t-1}-F_{\theta,t})^2 + \frac{2\sigma^2}{1+\rho}} \tag{20}$$

$$= \frac{m_\delta^2 \tilde{X} - \sigma^2\left(\frac{1-\rho}{1+\rho}\right)}{m_\delta^2 \tilde{Y} + \frac{2\sigma^2}{1+\rho}} = \frac{X}{Y} \tag{21}$$

where $T_0 = T-1$, $\tilde{X} = T_0^{-1}\sum_{t=2}^{T}(F_{\theta,t-1}-F_{\theta,t-2})(F_{\theta,t}-F_{\theta,t-1})$, and $\tilde{Y} = T_0^{-1}\sum_{t=2}^{T}(F_{\theta,t-1}-F_{\theta,t-2})^2$.

Note that $E(\widehat{\rho}_{fd}) = E(X/Y)$ can be expanded up to $O(T^{-1})$ as follows:

$$E\left(\frac{X}{Y}\right) = \frac{E(X)}{E(Y)}\left[1 - \frac{cov(X,Y)}{E(X)E(Y)} + \frac{var(Y)}{[E(Y)]^2}\right] + o(T^{-1}). \tag{22}$$

First, after a simple manipulation, we have

$$E(X) = -\left(\frac{1-\rho}{1+\rho}\right)(\sigma_\theta^2 m_\delta^2 + \sigma^2) \tag{23}$$

$$E(Y) = \left(\frac{2}{1+\rho}\right)(\sigma_\theta^2 m_\delta^2 + \sigma^2). \tag{24}$$

Next, we consider $cov(X,Y)$ and $var(Y)$. Note that $cov(X,Y) = m_\delta^4 cov(\tilde{X},\tilde{Y})$ and $var(Y) = m_\delta^4 var(\tilde{Y})$. Hayakawa (2006) shows that

$$cov(\tilde{X},\tilde{Y}) = \sigma_\theta^4\left[\frac{-4}{T_0}\frac{(1-\rho)(2+\rho)}{(1+\rho)^3}\right] \tag{25}$$

$$var(\tilde{Y}) = \sigma_\theta^4\left[\frac{1}{T_0}\frac{\rho+3}{(1+\rho)^3}\right]. \tag{26}$$

Using these results, it follows that

$$cov(X,Y) = m_\delta^4\sigma_\theta^4\left[\frac{-4}{T_0}\frac{(1-\rho)(2+\rho)}{(1+\rho)^3}\right] \tag{27}$$

$$var(Y) = m_\delta^4\sigma_\theta^4\left[\frac{1}{T_0}\frac{\rho+3}{(1+\rho)^3}\right]. \tag{28}$$

Substituting (23), (24), (27) and (28) into (22), we obtain

$$E\left(\operatorname*{plim}_{N\to\infty}\widehat{\rho}_{fd}\right) = \frac{\rho-1}{2}\left[1 - \frac{1}{T_0}\left(\frac{\sigma_\theta^2 m_\delta^2}{\sigma_\theta^2 m_\delta^2 + \sigma_v^2}\right)^2\right] + o(T^{-1}). \tag{29}$$

Then, it follows that

$$E\left(\underset{N\to\infty}{\text{plim}}\,\widehat{\rho}_{bcfd}\right) = \rho + \frac{1-\rho}{T_0}\left(\frac{\sigma_\theta^2 m_\delta^2}{\sigma_\theta^2 m_\delta^2 + \sigma_v^2}\right)^2 + o(T^{-1}). \tag{30}$$

In the nonstationary case, note that $X = m_\delta^2 T_0^{-1}\sum_{t=2}^T \theta_{t-1}\theta_t$ and $Y = m_\delta^2 T_0^{-1}\sum_{t=2}^T \theta_{t-1}^2 + \sigma^2$. Since $E(X) = 0$ and $E(Y) = m_\delta^2\sigma_\theta^2 + \sigma^2$, the result is readily obtained in view of (22).

References

[1] Arellano (2003) *Panel Data Econometrics*, Oxford University Press, New York.

[2] Bun, M.J.G. and M.A. Carree (2005a) "Bias-Corrected Estimation in Dynamic Panel Data Models," *Journal of Business and Economic Statistics*, 23, 200-210.

[3] Bun, M.J.G., and M.A. Carree (2005b) "Correction: Bias-corrected estimation in dynamic panel data models," *Journal of Business and Economic Statistics*, 23, 363.

[4] Bun, M.J.G. and M.A. Carree (2006) "Bias-Corrected Estimation in Dynamic Panel Data Models with Heteroscedasticity," *Economics Letters*, 92, 220-227.

[5] Chowdhury, G. (1987) "A Note on Correcting Biases in Dynamic Panel Models," *Applied Economics*, 19, 31-37.

[6] Frankel, J.A. and A.K. Rose (1996) "A Panel Project on Purchasing Power Parity: Mean Reversion Within and Between Countries," *Journal of International Economics*, 40, 209-224.

[7] Han, C. and P. C. B. Phillips (2007) "GMM Estimation for Dynamic Panels with Fixed Effects and Strong Instruments at Unity," Cowles Foundation Discussion Paper No.1599, Yale University.

[8] Hayakawa, K. (2006) "A Note on the Bias in First Differenced AR(1) Models," *Economics Bulletin*, Vol. 3, No.27 pp. 1-10.

[9] Hayakawa, K. (2007) "Consistent OLS Estimation of AR(1) Dynamic Panel Data Models with Short Time Series," forthcoming in *Applied Economics Letters*.

[10] Hirano, K. (2002) "Semiparametric Bayesian Inference in Autoregressive Panel Data Models," *Econometrica*, 70, 781-799.

[11] Ho, T.W. (2006) "Income Thresholds and Growth Convergence: A Panel Data Approach," *Manchester School*, 74, 170-189.

[12] Islam, N. (1995) "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, 110, 1127-1170.

[13] Maddala, G.S. (1999) "On the Use of Panel Data Methods with Cross-Country Data," *Annales D'Économie et de Statistique*, 55-56, 429-448.

[14] Phillips, P.C.B. and C. Han (2006) "Gaussian Inference in AR(1) Time Series with or without a Unit Root," Cowles Foundation Discussion Paper, No.1546, Yale University.

[15] Phillips, P.C.B. and D. Sul (2003) "Dynamic Panel Estimation and Homogeneity Testing under Cross Section Dependence," *Econometrics Journal*, 6, 217-259.

[16] Phillips, P.C.B. and D. Sul (2007) "Bias in Dynamic Panel Estimation with Fixed Effects, Incidental Trends and Cross Section Dependence," *Journal of Econometrics*, 137, 162-188.

[17] Ramalho, J.J.S. (2005) "Feasible Bias-Corrected OLS, Within-Groups, and First-Differences Estimators for Typical Micro and Macro AR(1) Panel Data Models," *Empirical Economics*, 30, 735-748.

[18] Shioji, E. (2004) "Initial Values and Income Convergence: Do 'The Poor Stay Poor'?," *Review of Economics and Statistics*, 86, 444-446.

[19] Wansbeek, T.L. and T. Knaap (1999) "Estimating a Dynamic Panel Data Model with Heterogeneous Trends," *Annales D'Économie et de Statistique*, 55-56, 331-349.

Table 1: (i) *iid* errors

Case			Mean			Std. Dev.			RMSE		
ρ	T	N	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$
0.5	10	50	0.504	0.314	0.481	0.088	0.048	0.084	0.088	0.192	0.086
0.5	10	100	0.498	0.317	0.478	0.061	0.033	0.072	0.061	0.186	0.075
0.5	10	200	0.501	0.318	0.480	0.043	0.024	0.125	0.043	0.183	0.126
0.5	15	25	0.504	0.386	0.495	0.097	0.052	0.070	0.097	0.125	0.071
0.5	15	50	0.499	0.384	0.490	0.069	0.038	0.054	0.069	0.122	0.055
0.5	15	100	0.497	0.384	0.490	0.047	0.026	0.041	0.047	0.119	0.042
0.5	25	25	0.504	0.434	0.496	0.073	0.038	0.046	0.073	0.076	0.047
0.5	25	50	0.501	0.434	0.497	0.052	0.027	0.034	0.052	0.071	0.034
0.5	25	100	0.499	0.435	0.498	0.036	0.019	0.025	0.036	0.068	0.025
0.5	50	10	0.502	0.467	0.498	0.080	0.041	0.044	0.080	0.053	0.044
0.5	50	25	0.500	0.467	0.498	0.050	0.026	0.029	0.050	0.042	0.029
0.5	50	50	0.501	0.468	0.499	0.037	0.018	0.020	0.036	0.037	0.020
0.9	10	50	0.901	0.624	0.922	0.098	0.044	0.085	0.098	0.280	0.088
0.9	10	100	0.898	0.628	0.926	0.067	0.031	0.065	0.067	0.274	0.070
0.9	10	200	0.901	0.630	0.928	0.048	0.022	0.101	0.048	0.271	0.105
0.9	15	25	0.904	0.726	0.908	0.109	0.043	0.070	0.109	0.179	0.070
0.9	15	50	0.899	0.724	0.903	0.078	0.031	0.050	0.078	0.179	0.050
0.9	15	100	0.897	0.724	0.903	0.053	0.022	0.037	0.053	0.177	0.037
0.9	25	25	0.903	0.798	0.897	0.081	0.030	0.040	0.081	0.106	0.041
0.9	25	50	0.900	0.800	0.899	0.058	0.021	0.028	0.058	0.102	0.028
0.9	25	100	0.899	0.801	0.899	0.040	0.014	0.021	0.040	0.100	0.021
0.9	50	10	0.901	0.852	0.896	0.090	0.027	0.031	0.090	0.056	0.031
0.9	50	25	0.899	0.853	0.897	0.057	0.017	0.019	0.057	0.050	0.019
0.9	50	50	0.900	0.854	0.899	0.041	0.012	0.014	0.041	0.048	0.014
0.95	10	50	0.950	0.659	0.985	0.100	0.043	0.086	0.100	0.294	0.093
0.95	10	100	0.948	0.663	0.993	0.069	0.031	0.086	0.069	0.288	0.096
0.95	10	200	0.951	0.665	0.999	0.049	0.021	0.148	0.049	0.286	0.155
0.95	15	25	0.955	0.764	0.965	0.110	0.042	0.072	0.110	0.191	0.073
0.95	15	50	0.949	0.761	0.960	0.079	0.029	0.050	0.079	0.191	0.051
0.95	15	100	0.947	0.762	0.960	0.054	0.022	0.036	0.054	0.189	0.038
0.95	25	25	0.953	0.839	0.949	0.082	0.029	0.040	0.082	0.115	0.040
0.95	25	50	0.950	0.841	0.951	0.059	0.020	0.028	0.059	0.111	0.028
0.95	25	100	0.949	0.842	0.952	0.040	0.013	0.020	0.040	0.109	0.020
0.95	50	10	0.951	0.897	0.947	0.091	0.024	0.030	0.091	0.059	0.030
0.95	50	25	0.949	0.897	0.948	0.057	0.015	0.018	0.057	0.055	0.018
0.95	50	50	0.950	0.899	0.949	0.041	0.011	0.013	0.041	0.052	0.013

Table 2: (ii) Cross-sectionally dependent errors

Case			Mean			Std. Dev.			RMSE		
ρ	T	N	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$
0.5	10	50	0.507	0.311	0.493	0.178	0.094	0.075	0.178	0.211	0.075
0.5	10	100	0.504	0.311	0.494	0.155	0.082	0.082	0.155	0.206	0.082
0.5	10	200	0.505	0.313	0.499	0.143	0.075	0.165	0.143	0.201	0.165
0.5	15	25	0.505	0.379	0.492	0.158	0.083	0.066	0.158	0.147	0.066
0.5	15	50	0.503	0.380	0.493	0.140	0.075	0.048	0.140	0.141	0.049
0.5	15	100	0.503	0.382	0.493	0.125	0.065	0.036	0.125	0.135	0.036
0.5	25	25	0.504	0.430	0.496	0.119	0.063	0.044	0.119	0.094	0.044
0.5	25	50	0.503	0.431	0.496	0.108	0.055	0.031	0.108	0.088	0.031
0.5	25	100	0.500	0.433	0.497	0.096	0.049	0.022	0.096	0.083	0.023
0.5	50	10	0.503	0.466	0.498	0.099	0.050	0.043	0.099	0.061	0.043
0.5	50	25	0.502	0.467	0.499	0.084	0.043	0.027	0.084	0.054	0.027
0.5	50	50	0.502	0.467	0.499	0.077	0.039	0.019	0.077	0.051	0.019
0.9	10	50	0.903	0.618	0.937	0.198	0.084	0.089	0.198	0.294	0.096
0.9	10	100	0.900	0.620	0.938	0.173	0.072	0.098	0.173	0.289	0.105
0.9	10	200	0.902	0.622	0.937	0.159	0.065	0.159	0.159	0.286	0.163
0.9	15	25	0.901	0.715	0.903	0.177	0.067	0.069	0.177	0.197	0.069
0.9	15	50	0.900	0.717	0.906	0.157	0.060	0.051	0.157	0.192	0.052
0.9	15	100	0.901	0.719	0.907	0.139	0.051	0.037	0.139	0.188	0.038
0.9	25	25	0.901	0.793	0.897	0.134	0.044	0.037	0.134	0.116	0.037
0.9	25	50	0.901	0.795	0.899	0.120	0.039	0.027	0.120	0.112	0.027
0.9	25	100	0.900	0.798	0.900	0.107	0.034	0.019	0.107	0.108	0.019
0.9	50	10	0.902	0.849	0.896	0.111	0.032	0.030	0.111	0.060	0.030
0.9	50	25	0.901	0.851	0.898	0.095	0.026	0.019	0.095	0.056	0.019
0.9	50	50	0.901	0.852	0.899	0.086	0.023	0.013	0.086	0.053	0.013
0.95	10	50	0.952	0.653	1.000	0.202	0.082	0.092	0.202	0.308	0.105
0.95	10	100	0.950	0.655	1.005	0.175	0.071	0.120	0.175	0.303	0.132
0.95	10	200	0.951	0.657	1.008	0.161	0.064	0.184	0.161	0.300	0.193
0.95	15	25	0.951	0.752	0.959	0.179	0.064	0.072	0.179	0.208	0.072
0.95	15	50	0.949	0.755	0.963	0.159	0.058	0.053	0.159	0.203	0.055
0.95	15	100	0.951	0.757	0.965	0.141	0.049	0.049	0.141	0.199	0.051
0.95	25	25	0.951	0.834	0.949	0.136	0.041	0.038	0.136	0.123	0.038
0.95	25	50	0.951	0.836	0.951	0.122	0.036	0.027	0.122	0.120	0.027
0.95	25	100	0.950	0.838	0.952	0.108	0.032	0.019	0.108	0.116	0.020
0.95	50	10	0.952	0.894	0.946	0.112	0.029	0.029	0.112	0.063	0.029
0.95	50	25	0.951	0.896	0.948	0.096	0.023	0.018	0.096	0.059	0.018
0.95	50	50	0.951	0.897	0.949	0.087	0.020	0.013	0.087	0.057	0.013

Table 3: (iii) Heteroscedastic errors

Case			Mean			Std. Dev.			RMSE		
ρ	T	N	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$
0.5	10	50	0.509	0.316	0.648	0.143	0.073	0.082	0.144	0.198	0.169
0.5	10	100	0.501	0.317	0.659	0.104	0.052	0.039	0.104	0.191	0.164
0.5	10	200	0.503	0.318	0.668	0.072	0.037	0.027	0.073	0.185	0.171
0.5	15	25	0.508	0.380	0.574	0.155	0.080	0.075	0.155	0.144	0.105
0.5	15	50	0.504	0.382	0.589	0.113	0.059	0.069	0.113	0.131	0.112
0.5	15	100	0.498	0.383	0.596	0.081	0.042	0.023	0.081	0.124	0.099
0.5	25	25	0.504	0.431	0.538	0.117	0.060	0.051	0.117	0.091	0.064
0.5	25	50	0.502	0.433	0.546	0.085	0.044	0.044	0.085	0.080	0.064
0.5	25	100	0.502	0.434	0.552	0.061	0.031	0.016	0.061	0.073	0.054
0.5	50	10	0.503	0.464	0.510	0.122	0.063	0.036	0.122	0.073	0.037
0.5	50	25	0.503	0.468	0.515	0.083	0.041	0.022	0.083	0.053	0.027
0.5	50	50	0.503	0.468	0.519	0.060	0.030	0.017	0.060	0.044	0.026
0.9	10	50	0.906	0.624	1.170	0.159	0.067	0.112	0.159	0.284	0.292
0.9	10	100	0.900	0.628	1.193	0.115	0.047	0.070	0.115	0.276	0.302
0.9	10	200	0.902	0.629	1.210	0.080	0.034	0.054	0.080	0.273	0.315
0.9	15	25	0.904	0.718	1.059	0.174	0.063	0.098	0.174	0.193	0.186
0.9	15	50	0.902	0.722	1.079	0.126	0.046	0.067	0.126	0.184	0.191
0.9	15	100	0.897	0.724	1.094	0.089	0.033	0.035	0.090	0.179	0.197
0.9	25	25	0.902	0.797	0.984	0.130	0.042	0.036	0.130	0.111	0.091
0.9	25	50	0.901	0.799	0.996	0.095	0.030	0.024	0.095	0.105	0.099
0.9	25	100	0.901	0.800	1.004	0.068	0.022	0.016	0.068	0.103	0.105
0.9	50	10	0.901	0.847	0.924	0.138	0.039	0.026	0.138	0.066	0.035
0.9	50	25	0.903	0.852	0.933	0.093	0.025	0.017	0.093	0.054	0.037
0.9	50	50	0.902	0.853	0.938	0.067	0.019	0.015	0.067	0.050	0.041
0.95	10	50	0.954	0.658	1.225	0.163	0.068	0.130	0.163	0.300	0.305
0.95	10	100	0.949	0.662	1.257	0.117	0.048	0.083	0.117	0.292	0.318
0.95	10	200	0.952	0.664	1.279	0.081	0.034	0.067	0.081	0.288	0.336
0.95	15	25	0.953	0.754	1.122	0.176	0.063	0.109	0.176	0.205	0.204
0.95	15	50	0.951	0.759	1.147	0.128	0.046	0.063	0.128	0.196	0.207
0.95	15	100	0.947	0.761	1.168	0.090	0.032	0.045	0.090	0.191	0.222
0.95	25	25	0.952	0.837	1.052	0.132	0.039	0.061	0.132	0.119	0.119
0.95	25	50	0.951	0.840	1.068	0.096	0.028	0.042	0.096	0.114	0.126
0.95	25	100	0.950	0.840	1.078	0.069	0.020	0.021	0.069	0.112	0.130
0.95	50	10	0.951	0.892	0.983	0.140	0.033	0.028	0.140	0.067	0.043
0.95	50	25	0.953	0.897	0.994	0.094	0.021	0.018	0.094	0.057	0.047
0.95	50	50	0.952	0.898	1.000	0.068	0.016	0.012	0.068	0.054	0.052

Table 4: (iv) Cross-sectionally dependent and heteroscedastic errors

Case			Mean			Std. Dev.			RMSE		
ρ	T	N	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$
0.5	10	50	0.506	0.309	0.618	0.198	0.104	0.242	0.198	0.218	0.269
0.5	10	100	0.502	0.311	0.627	0.168	0.089	0.235	0.168	0.209	0.267
0.5	10	200	0.500	0.312	0.630	0.147	0.079	0.262	0.147	0.204	0.292
0.5	15	25	0.507	0.379	0.550	0.183	0.097	0.218	0.183	0.155	0.223
0.5	15	50	0.506	0.380	0.571	0.158	0.086	0.214	0.159	0.148	0.225
0.5	15	100	0.504	0.383	0.576	0.138	0.073	0.245	0.138	0.138	0.256
0.5	25	25	0.505	0.428	0.523	0.140	0.073	0.135	0.140	0.102	0.137
0.5	25	50	0.503	0.431	0.535	0.122	0.064	0.168	0.122	0.094	0.171
0.5	25	100	0.501	0.431	0.544	0.105	0.054	0.207	0.105	0.088	0.211
0.5	50	10	0.502	0.464	0.504	0.128	0.064	0.132	0.128	0.074	0.132
0.5	50	25	0.504	0.465	0.510	0.099	0.051	0.126	0.099	0.061	0.126
0.5	50	50	0.503	0.467	0.519	0.086	0.044	0.161	0.086	0.055	0.162
0.9	10	50	0.901	0.617	1.118	0.221	0.092	0.273	0.221	0.298	0.350
0.9	10	100	0.898	0.620	1.136	0.189	0.078	0.272	0.189	0.291	0.361
0.9	10	200	0.898	0.621	1.146	0.167	0.070	0.273	0.167	0.288	0.367
0.9	15	25	0.905	0.714	1.021	0.206	0.078	0.216	0.206	0.201	0.248
0.9	15	50	0.904	0.716	1.058	0.180	0.068	0.204	0.181	0.196	0.258
0.9	15	100	0.903	0.719	1.068	0.155	0.056	0.223	0.155	0.189	0.279
0.9	25	25	0.902	0.791	0.965	0.157	0.051	0.144	0.157	0.120	0.157
0.9	25	50	0.902	0.795	0.985	0.138	0.044	0.140	0.138	0.114	0.163
0.9	25	100	0.900	0.796	0.990	0.118	0.037	0.199	0.118	0.110	0.219
0.9	50	10	0.901	0.847	0.917	0.143	0.040	0.124	0.143	0.066	0.126
0.9	50	25	0.902	0.849	0.926	0.111	0.031	0.106	0.111	0.060	0.109
0.9	50	50	0.902	0.851	0.939	0.097	0.027	0.135	0.097	0.056	0.140
0.95	10	50	0.950	0.651	1.162	0.225	0.092	0.277	0.225	0.313	0.349
0.95	10	100	0.947	0.654	1.187	0.192	0.077	0.290	0.192	0.306	0.374
0.95	10	200	0.947	0.656	1.188	0.170	0.069	0.317	0.170	0.302	0.396
0.95	15	25	0.955	0.752	1.080	0.209	0.076	0.233	0.209	0.213	0.267
0.95	15	50	0.954	0.753	1.115	0.184	0.066	0.238	0.184	0.208	0.290
0.95	15	100	0.953	0.756	1.123	0.157	0.054	0.261	0.157	0.201	0.313
0.95	25	25	0.952	0.832	1.031	0.159	0.047	0.154	0.159	0.127	0.174
0.95	25	50	0.951	0.835	1.053	0.140	0.041	0.148	0.140	0.122	0.180
0.95	25	100	0.949	0.837	1.057	0.120	0.035	0.203	0.120	0.118	0.229
0.95	50	10	0.951	0.892	0.972	0.144	0.034	0.117	0.144	0.067	0.119
0.95	50	25	0.952	0.894	0.992	0.113	0.027	0.141	0.113	0.062	0.147
0.95	50	50	0.952	0.896	0.996	0.098	0.023	0.125	0.098	0.059	0.133

Table 5: Simulation results for random and fixed θ_t, $(T = 5)$

Case			Mean			Std. Dev.			RMSE		
ρ	T	N	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$	$\widehat{\rho}_{bcfd}$	$\widehat{\rho}_{lsdv}$	$\widehat{\rho}_{ps}$
Cross-sectionally dependent errors with random θ_t											
0.5	5	50	0.511	0.080	0.706	0.265	0.141	0.471	0.266	0.443	0.514
0.5	5	100	0.503	0.084	0.652	0.234	0.129	0.613	0.234	0.435	0.631
0.5	5	200	0.506	0.076	0.679	0.233	0.122	0.788	0.233	0.441	0.808
0.9	5	50	0.902	0.322	1.271	0.294	0.145	0.480	0.294	0.596	0.606
0.9	5	100	0.896	0.329	1.235	0.268	0.135	0.621	0.268	0.587	0.705
0.9	5	200	0.900	0.332	1.222	0.237	0.119	0.798	0.236	0.581	0.860
0.95	5	50	0.939	0.351	1.322	0.312	0.152	0.449	0.312	0.618	0.583
0.95	5	100	0.947	0.355	1.324	0.264	0.130	0.608	0.264	0.609	0.713
0.95	5	200	0.950	0.362	1.280	0.238	0.118	0.798	0.238	0.600	0.863
Cross-sectionally dependent errors with fixed θ_t											
0.5	5	50	0.642	0.166	0.683	0.156	0.112	0.440	0.211	0.352	0.476
0.5	5	100	0.424	0.028	0.641	0.108	0.071	0.499	0.132	0.478	0.519
0.5	5	200	0.439	0.103	0.698	0.148	0.107	0.707	0.160	0.411	0.734
0.9	5	50	1.030	0.388	1.269	0.164	0.103	0.462	0.209	0.522	0.591
0.9	5	100	0.827	0.279	1.215	0.124	0.062	0.520	0.143	0.624	0.607
0.9	5	200	0.831	0.381	1.244	0.130	0.084	0.853	0.147	0.526	0.919
0.95	5	50	1.070	0.410	1.337	0.160	0.096	0.449	0.200	0.548	0.593
0.95	5	100	0.883	0.313	1.340	0.122	0.062	0.541	0.139	0.640	0.666
0.95	5	200	0.878	0.413	1.242	0.113	0.070	0.860	0.134	0.541	0.907
Cross-sectionally dependent and heteroscedastic errors with random θ_t											
0.5	5	50	0.525	0.087	0.805	0.299	0.157	0.448	0.300	0.441	0.542
0.5	5	100	0.513	0.082	0.822	0.255	0.140	0.533	0.255	0.441	0.623
0.5	5	200	0.515	0.086	0.850	0.245	0.130	0.627	0.245	0.434	0.718
0.9	5	50	0.911	0.326	1.323	0.331	0.168	0.540	0.331	0.598	0.685
0.9	5	100	0.905	0.326	1.310	0.288	0.149	0.566	0.288	0.593	0.698
0.9	5	200	0.910	0.333	1.344	0.270	0.137	0.745	0.270	0.583	0.867
0.95	5	50	0.957	0.353	1.341	0.338	0.171	0.601	0.338	0.621	0.717
0.95	5	100	0.954	0.356	1.369	0.293	0.150	0.623	0.293	0.613	0.750
0.95	5	200	0.959	0.363	1.355	0.274	0.137	0.766	0.274	0.603	0.867
Cross-sectionally dependent and heteroscedastic errors with fixed θ_t											
0.5	5	50	0.643	0.106	0.780	0.219	0.138	0.380	0.262	0.417	0.472
0.5	5	100	0.520	0.089	0.827	0.164	0.081	0.367	0.165	0.419	0.491
0.5	5	200	0.195	-0.110	0.729	0.083	0.107	0.799	0.316	0.620	0.831
0.9	5	50	0.991	0.279	1.244	0.235	0.121	0.500	0.251	0.633	0.607
0.9	5	100	0.867	0.321	1.325	0.173	0.090	0.429	0.176	0.585	0.604
0.9	5	200	0.469	0.067	1.124	0.079	0.066	0.827	0.438	0.836	0.856
0.95	5	50	1.028	0.297	1.244	0.237	0.120	0.536	0.250	0.664	0.611
0.95	5	100	0.901	0.343	1.354	0.176	0.092	0.476	0.183	0.614	0.624
0.95	5	200	0.504	0.088	1.179	0.081	0.064	0.864	0.453	0.864	0.893