UNIVERSITEIT GENT

FACULTEIT ECONOMIE
EN BEDRIJFSKUNDE

TWEEKERKENSTRAAT 2
B-9000 GENT
Tel. : 32 - (0)9 – 264.34.61
Fax. : 32 - (0)9 – 264.35.92

WORKING PAPER

Common Correlated Effects Estimation of Dynamic Panels with Cross-Sectional Dependence

Tom De Groote and Gerdie Everaert

SHERPPA, Ghent University

June 2011

2011/723

D/2011/7012/28

Common Correlated Effects Estimation of Dynamic Panels with Cross-Sectional Dependence*

Tom De Groote and Gerdie Everaert†

SHERPPA, Ghent University

June, 2011

Abstract

We study estimation of dynamic panel data models with error cross-sectional dependence generated by an unobserved common factor. We show that for a temporally dependent factor, the standard within groups (WG) estimator is inconsistent even as both N and T tend to infinity. Next we investigate the properties of the common correlated effects pooled (CCEP) estimator of Pesaran [Econometrica, 2006] which eliminates the cross-sectional dependence using cross-sectional averages of the data. In contrast to the static case, the CCEP estimator is only consistent if next to N also T tends to infinity. It is shown that for the most relevant parameter settings, the asymptotic bias of the CCEP estimator is larger than that of the infeasible WG estimator, which includes the common factors as regressors. Restricting the CCEP estimator results in a somewhat smaller asymptotic bias. The small sample properties of the various estimators are analysed using Monte Carlo experiments. The simulation results suggest that the CCEP estimator can be used to estimate dynamic panel data models provided T is not too small. The size of N is of less importance.

JEL Classification: C13, C15, C23

Keywords: Cross-Sectional Dependence; Dynamic Panel; Common Correlated Effects

*We thank Lorenzo Pozzi for helpful suggestions and constructive comments on an earlier version of this paper.

†We acknowledge financial support from the Interuniversity Attraction Poles Program - Belgian Science Policy, contract no. P5/21. Corresponding author: Tom De Groote, Tweekerkenstraat 2, B-9000 Gent, Belgium. Email: tom.degroote@UGent.be. Website: http://www.sherppa.be.

1 Introduction

Over the last decades, estimation of dynamic panel data models has received a lot of attention. Nickell (1981) demonstrated that in dynamic panel data regressions the within groups (WG) estimator is inconsistent for fixed T and $N \to \infty$. Given that the asymptotic bias may be quite sizable in many cases relevant to applied research, various alternative estimators have been suggested ranging from, among others, general method of moments (GMM) estimators (Arellano and Bond, 1991; Blundell and Bond, 1998), over analytical bias corrected WG estimators (Kiviet, 1995; Bun and Carree, 2005) to a bootstrap-based bias corrected WG estimator (Everaert and Pozzi, 2007). However, new challenges arise when it comes to the estimation of dynamic panel data models. The recent panel data literature shifted its attention to the estimation of models with error cross-sectional dependence. A particular form that has become popular is a common factor error structure with a fixed number of unobserved common factors and individual-specific factor loadings (see e.g. Coakley *et al.*, 2002; Phillips and Sul, 2003; Bai and Ng, 2004; Pesaran, 2006).

The most obvious implication of error cross-sectional dependence is that standard panel data estimators are inefficient and estimated standard errors are biased and inconsistent. Phillips and Sul (2003) for instance show that if there is high cross-sectional correlation there may not be much to gain from pooling the data. However, cross-sectional dependence can also induce a bias and even result in inconsistent estimates. In general, inconsistency arises as an omitted variables bias when the observed explanatory variables are correlated with the unobserved common factors (see e.g. Pesaran, 2006). More specifically, Phillips and Sul (2007) show that in dynamic panel data models with fixed T and $N \to \infty$, the unobserved common factors induce additional small sample bias and variability in the inconsistency of the WG estimator. This is true, even under the assumption of temporarily independent factors[1]. This bias disappears as $T \to \infty$. Sarafidis and Robertson (2009) show that also dynamic panel data IV and GMM estimators (either in levels or first-differences) are inconsistent for fixed T and $N \to \infty$ as the moment conditions used by these estimators are invalid under error cross-sectional dependence.

In this paper we further analyze the impact of error cross-sectional dependence in linear dynamic panels. Explicit asymptotic bias formulas are derived for both the standard WG estimator, which ignores error cross-sectional dependence, and the common correlated effects pooled (CCEP) estimator of Pesaran (2006), which is explicitly designed to deal with unobserved common factors in the error term. We first extend the work of Phillips and Sul (2007) by relaxing the assumption of a temporally independent common factor. In line with their results we find that for fixed T and $N \to \infty$ the inconsistency of the WG estimator is a combination of the nonrandom dynamic panel data bias, as obtained by Nickell (1981), and a random component induced by the common factor in the error term. Importantly, the latter component of the inconsistency becomes nonrandom but does not disappear as also $T \to \infty$ since the temporal dependence in the unobserved com-

[1] This implies that the factors are uncorrelated with the lagged dependent variable.

mon factor implies that the error term is correlated with the lagged dependent variable even for $(N,T)_{\text{seq}} \to \infty$[2]. This finding should warn against the use of the WG estimator in cross-sectionally dependent dynamic panels even when T is large. Second, we extend the work of Pesaran (2006) by analyzing the asymptotic behaviour of the CCEP estimator in a dynamic panel data setting. The basic idea of CCEP estimation is to deal with error cross-sectional dependence by filtering out the unobserved common factors using the cross-section averages of both the dependent and the explanatory variables. We show that contrary to the static model, the CCEP estimator is no longer consistent for $N \to \infty$ and fixed T. Similar to the results for the WG estimator, the inconsistency is a combination of the standard nonrandom dynamic panel data bias and a random component which is now induced by orthogonalising on the cross-sectional averages. The main difference is that both components of the inconsistency disappear as we also let $T \to \infty$. As a benchmark, we derive the asymptotic bias of the infeasible WG estimator, which includes the unobserved factor as an observed explanatory variable. This infeasible WG estimator has a random inconsistency for fixed T and $N \to \infty$ which disappears for $(N,T)_{\text{seq}} \to \infty$. However, for the cases most relevant to applied research, the asymptotic bias of the CCEP is bigger than that of the infeasible WG estimator. One possible reason for this is that the CCEP estimator as suggested by Pesaran (2006) ignores the restrictions on the individual-specific factor loadings as implied by the derivation of the cross-sectional averages augmented specification of the model. Imposing these restrictions, the asymptotic bias of the restricted CCEP estimator is closer to that of the infeasible WG estimator.

We next analyse the small sample properties of the WG and CCEP estimators using Monte Carlo experiments. First, the infeasible WG estimator is biased for small T, with the bias increasing in the degree of temporal dependence in both the model and in the common factor. Second, both the unrestricted and the restricted CCEP estimators have a higher bias than the infeasible WG estimator for small values of T but the restricted CCEP estimator outperforms the unrestricted CCEP estimator and is not much worse than the infeasible WG estimator for moderate T. Interestingly, the performance of the CCEP estimators is highly similar comparing $N = 20$ with $N = 50$. This shows that these estimators are not very sensitive to the size of N. Finally, the results illustrate that the standard WG estimator, ignoring cross-sectional dependence, has a persistent (as N and T growing larger) bias for temporally dependent factors, is inefficiency compared to the CCEP estimators and has substantially biased estimated standard errors. Overall, the results suggest that the CCEP estimator is quite useful for estimating cross-sectional dependent dynamic panel data models provided T is not too small.

The remainder of this paper is organized as follows: Section 2 sets out the basic model and its assumptions. Section 3 explores the asymptotic behaviour of the naive WG, the infeasible WG and the unrestricted and restricted CCEP estimators in a dynamic model with error cross-sectional dependence. Section 4 adds exogenous explanatory variables. Section 5 reports the Monte Carlo results. Section 6 concludes.

[2]Denote $(N,T)_{\text{seq}} \to \infty$ as the sequential limit, i.e., $N \to \infty$ first and $T \to \infty$ later.

2 Model and assumptions

Consider the following first-order autoregressive panel data model

$$y_{it} = \alpha_i + \rho y_{i,t-1} + \nu_{it}, \qquad |\rho| < 1, \qquad (i = 1, \ldots, N; \quad t = 1, \ldots, T) \tag{1}$$

with y_{it} being the observation of the dependent variable for the ith cross-sectional unit at time t. For notational convenience we assume y_{i0} is observed. We further assume:

Assumption A1. *(Cross-section dependence) The error term ν_{it} has a single-factor structure*

$$\nu_{it} = \gamma_i F_t + \varepsilon_{it}, \tag{2}$$

$$F_t = \theta F_{t-1} + \mu_t, \qquad |\theta| < 1, \tag{3}$$

where F_t is an individual-invariant time-specific unobserved effect with $\mu_t \sim i.i.d.\left(0, \sigma_\mu^2\right)$. The individual-specific factor loadings γ_i are nonrandom parameters satisfying $\underset{N\to\infty}{plim} \dfrac{1}{N}\sum_{i=1}^{N} \gamma_i^2 = m_\gamma^2$. ε_{it} satisfies A2.

The restriction of a single-factor structure is for expositional purposes only.

Assumption A2. *(Error condition) $\varepsilon_{it} \sim i.i.d.\left(0, \sigma_\varepsilon^2\right)$ across i and t and is independent of F_s for all i, t, s.*

Assumption A3. *(Fixed effect condition) $\alpha_i \sim i.i.d.\left(0, \sigma_\alpha^2\right)$ across i and independent of F_t for all i, t.*

The model in equations (1)-(3) can be written in a convenient component form as

$$y_{it} = y_{it}^+ + \gamma_i F_t^+, \tag{4a}$$

$$y_{it}^+ = \alpha_i + \rho y_{i,t-1}^+ + \varepsilon_{it}, \tag{4b}$$

$$F_t^+ = (1 - \rho L)^{-1} F_t = (\rho + \theta) F_{t-1}^+ - \rho\theta F_{t-2}^+ + \mu_t. \tag{4c}$$

For further discussion, stacking (1)-(2) for each i yields

$$y_i = \alpha_i + \rho y_{i,-1} + \gamma_i F + \varepsilon_i, \tag{5}$$

where $y_i = (y_{i1}, \ldots, y_{iT})'$, $y_{i,-1} = (y_{i0}, \ldots, y_{i,T-1})'$, $F = (F_1, \ldots, F_T)'$ and $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iT})'$.

3 Estimators

In this section we analyse the properties of various estimators for ρ in equation (1) given the error structure in A1-A3. We start with two 'extreme' approaches, i.e. the naive within groups (WGn) estimator which ignores cross-sectional dependence and the infeasible within groups (WGi) estimator which adds the unobserved common factor F_t as an observed explanatory variable to the model. Next, we analyse an unrestricted and a restricted version of the CCEP estimator suggested by Pesaran (2006) and compare their properties to those of the WGn and WGi estimators.

3.1 Naive within groups

Consider the naive within groups (WGn) estimator for ρ in (1)

$$\widehat{\rho}_{\text{WGn}} = \frac{(1/NT)\sum_{i=1}^{N} y_{i,-1}' M_I y_i}{(1/NT)\sum_{i=1}^{N} y_{i,-1}' M_I y_{i,-1}} = \rho + \frac{(1/NT)\sum_{i=1}^{N} y_{i,-1}' M_I (\gamma_i F + \varepsilon_i)}{(1/NT)\sum_{i=1}^{N} y_{i,-1}' M_I y_{i,-1}}, \qquad (6)$$

where $M_I = I_T - \iota(\iota'\iota)^{-1}\iota'$ and ι a (Tx1) vector of ones. The exact form of the inconsistency and its large T expansion are given in the following Proposition. All proofs are in appendix A.

Proposition 1. *In model (1) with errors satisfying A1-A3, the WGn estimator is inconsistent as* $N \to \infty$

$$\underset{N\to\infty}{plim}\,(\widehat{\rho}_{WGn} - \rho) = \frac{-A(\rho,T) + \frac{m_\gamma^2}{\sigma_\varepsilon^2}\sum_{t=1}^{T} F_{t-1}^{+}\widetilde{F}_t}{B(\rho,T) + \frac{m_\gamma^2}{\sigma_\varepsilon^2}\sum_{t=1}^{T}\left(\widetilde{F}_{t-1}^{+}\right)^2}, \qquad (7)$$

where $A(\rho,T) = \frac{1}{1-\rho}\left(1 - \frac{1}{T}\frac{1-\rho^T}{1-\rho}\right)$, $B(\rho,T) = \frac{T-1}{1-\rho^2}\left(1 - \frac{1}{T-1}\frac{2\rho}{1-\rho}\left[1 - \frac{1}{T}\frac{1-\rho^T}{1-\rho}\right]\right)$, $\widetilde{F}_t = F_t - \overline{F}$, $\widetilde{F}_{t-1}^{+} = F_{t-1}^{+} - \overline{F}_{-1}^{+}$, $\overline{F} = \frac{1}{T}\sum_{t=1}^{T} F_t$ *and* $\overline{F}_{-1}^{+} = \frac{1}{T}\sum_{t=1}^{T} F_{t-1}^{+}$.
The inconsistency in (7) has the following asymptotic representation as $(N,T)_{seq} \to \infty$

$$\underset{N\to\infty}{plim}\,(\widehat{\rho}_{WGn} - \rho) = \left(-\frac{1+\rho}{T} + \frac{(1-\rho^2)\,\theta}{(1-\theta\rho)(1-\theta^2)}\frac{m_\gamma^2\sigma_\mu^2}{\sigma_\varepsilon^2}\right)\left(1 + \frac{(1+\theta\rho)}{(1-\theta\rho)(1-\theta^2)}\frac{m_\gamma^2\sigma_\mu^2}{\sigma_\varepsilon^2}\right)^{-1} + o_p(1). \quad (8)$$

Proposition 1 shows that the inconsistency of $\widehat{\rho}_{\text{WGn}}$ for $N \to \infty$ and fixed T is induced by two components. The first is the standard Nickell dynamic panel data bias, which depends on the persistence ρ in y_{it} and on the time dimension T. This can be seen by setting the error cross-sectional dependence to zero ($m_\gamma^2 = 0$) such that equation (7) reduces to the standard Nickell bias formula $-A(\rho,T)/B(\rho,T)$ and equation (8) to its large T approximation $-(1+\rho)/T$. The second component stems from the error cross-sectional dependence. It is apparent from equation (7) that the inconsistency induced by the common factor is random for fixed T as $1/T\sum_{t=1}^{T} F_{t-1}^{+}\widetilde{F}_t$ depends on the particular realisation for F_t. Also letting $T \to \infty$, it can be seen from equation (8) that for a temporally independent factor ($\theta = 0$) this random inconsistency disappears. Inertia in F_t ($\theta \neq 0$) results in an additional random inconsistency that, as $(N,T)_{\text{seq}} \to \infty$, becomes nonrandom, but does not disappear since $1/T\sum_{t=1}^{T}F_{t-1}^{+}\widetilde{F}_t$ does not converge to zero. Essentially, this is an omitted variable bias as $\theta\neq 0$ implies $E\left(y_{i,t-1}F_t\right)\neq 0$ such that omitting F_t from the regression results in an inconsistent estimator for ρ.

3.2 Infeasible within groups

The infeasible within groups (WGi) estimator for ρ, including F_t as an observed regressor, is given by

$$\widehat{\rho}_{\mathrm{WGi}}=\frac{(1/NT)\sum_{i=1}^{N}y_{i,-1}'M_{\dot{F}}y_i}{(1/NT)\sum_{i=1}^{N}y_{i,-1}'M_{\dot{F}}y_{i,-1}}=\rho+\frac{(1/NT)\sum_{i=1}^{N}y_{i,-1}'M_{\dot{F}}\varepsilon_i}{(1/NT)\sum_{i=1}^{N}y_{i,-1}'M_{\dot{F}}y_{i,-1}},\tag{9}$$

where $M_{\dot{F}}=I_T-\dot{F}\left(\dot{F}'\dot{F}\right)^{-1}\dot{F}'$ and $\dot{F}=(\iota,F)$.

Proposition 2. *In model (1) with errors satisfying A1-A3, the WGi estimator is inconsistent as $N\to\infty$*

$$\underset{N\to\infty}{plim}\left(\widehat{\rho}_{WGi}-\rho\right)=\frac{-A(\rho,T)-\sum_{t=1}^{T-1}\rho^{t-1}\widetilde{g}_{F,t}}{B(\rho,T)-\frac{1}{1-\rho^2}\left(1+2\rho\sum_{t=1}^{T-1}\rho^{t-1}\widetilde{g}_{F,t}\right)+T\frac{m_\gamma^2}{\sigma_\varepsilon^2}\widetilde{h}_F},\tag{10}$$

where $\widetilde{g}_{F,t}=\sum_{s=t+1}^{T}\widetilde{\tau}_{s,s-t}$ with $\widetilde{\tau}_{s,s-t}$ being the $(s,s-t)$th element in $\widetilde{F}\left(\widetilde{F}'\widetilde{F}\right)^{-1}\widetilde{F}'$, $\widetilde{F}=\left(\widetilde{F}_1,\ldots,\widetilde{F}_T\right)'$ and $\widetilde{h}_F=\frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{F}_{t-1}^{+}\right)^2\left(1-\frac{\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{F}_t\widetilde{F}_{t-1}^{+}\right)^2}{\frac{1}{T}\sum_{t=1}^{T}\widetilde{F}_t^2\frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{F}_{t-1}^{+}\right)^2}\right)$.

The inconsistency in (10) has the following asymptotic representation as $(N,T)_{seq}\to\infty$

$$\underset{N\to\infty}{plim}\left(\widehat{\rho}_{WGi}-\rho\right)=-\frac{1}{T}\left(1+\rho+\frac{\theta\left(1-\rho^2\right)}{(1-\theta\rho)}\right)\left(1+\frac{m_\gamma^2}{(1-\theta\rho)^2}\frac{\sigma_\mu^2}{\sigma_\varepsilon^2}\right)^{-1}+o_p\left(\frac{1}{T}\right).\tag{11}$$

Proposition 2 shows that the inconsistency of the WGi estimator for $N\to\infty$ is also induced by two components. The first is again the standard Nickell bias. The second part now stems from orthogonalizing on the observed factor F_t. For fixed T, this induces randomness in the inconsistency as the orthogonalisation depends on the particular realization of the process F_t. Moreover, equation (11) shows that temporary dependence in the common factor ($\theta\neq 0$) also induces a nonrandom inconsistency, which disappears as $(N,T)_{\mathrm{seq}}\to\infty$. The intuition for this result is that for fixed T the transformed error term $M_{\dot{F}}\varepsilon_i$ in (9) is, next to being a function of the average error term $\overline{\varepsilon}_i$ due to the within transformation, now also a function of the entire series F (as represented by $\widetilde{g}_{F,t}$) due to the orthogonalisation on F_t which results in correlation with the explanatory variable $y_{i,t-1}$. The denominator of equation (11) further shows that the inconsistency is smaller when the cross-sectional dependence is stronger as this implies more variability in the explanatory variable $y_{i,t-1}$, which is induced by F_{t-1} and is not completely filtered out by including F_t as a control variable in the regression. This additional variability is captured by the term $\widetilde{h}_F$ in the denominator of equation (11).

3.3 CCEP

The CCEP estimator suggested by Pesaran (2006) eliminates the unobserved common factors by including cross-section averages of the dependent and the explanatory variables. Taking cross-sectional averages of equation (1) gives

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N} y_{it} &= \frac{1}{N}\sum_{i=1}^{N}\alpha_i + \rho\frac{1}{N}\sum_{i=1}^{N} y_{i,t-1} + F_t\frac{1}{N}\sum_{i=1}^{N}\gamma_i + \frac{1}{N}\sum_{i=1}^{N}\varepsilon_{it},\\ \overline{y}_t &= \overline{\alpha} + \rho\overline{y}_{t-1} + \overline{\gamma}F_t + \overline{\varepsilon}_t, \qquad (12)\end{aligned}$$

which can be solved for F_t as

$$F_t = \frac{1}{\overline{\gamma}}(\overline{y}_t - \overline{\alpha} - \rho\overline{y}_{t-1} - \overline{\varepsilon}_t). \qquad (13)$$

Note that

$$\operatorname*{plim}_{N\to\infty} \overline{\varepsilon}_t = 0, \qquad (14)$$

$$\operatorname*{plim}_{N\to\infty} \overline{y}_t = \operatorname*{plim}_{N\to\infty} \left(\overline{y}_t^{+} + \overline{\gamma}F_t^{+}\right) = (1-\rho)^{-1}\overline{\alpha} + \overline{\gamma}F_t^{+}. \qquad (15)$$

Inserting (13) in (1) yields the following augmented form[3]

$$\begin{aligned}y_{it} &= \alpha_i + \rho y_{i,t-1} + \frac{\gamma_i}{\overline{\gamma}}(\overline{y}_t - \overline{\alpha} - \rho\overline{y}_{t-1} - \overline{\varepsilon}_t) + \varepsilon_{it},\\ &= \alpha_i^* + \rho y_{i,t-1} + \gamma_{1i}\overline{y}_t + \gamma_{2i}\overline{y}_{t,-1} + \varepsilon_{it}^*, \qquad (16)\end{aligned}$$

with $\gamma_{1i} = \gamma_i / \overline{\gamma}$, $\gamma_{2i} = \rho\gamma_{1i}$, $\alpha_i^* = \alpha_i - \gamma_{1i}\overline{\alpha}$ and $\varepsilon_{it}^* = \varepsilon_{it} - \gamma_{1i}\overline{\varepsilon}_t$.

3.3.1 Unrestricted CCEP

Ignoring the restrictions on γ_{1i}, γ_{2i} and α_i^*, the unrestricted CCEP estimator for ρ in (16) is given by

$$\widehat{\rho}_{\text{CCEPu}} = \frac{(1/NT)\sum_{i=1}^{N} y_{i,-1}' M_{\dot{G}} y_i}{(1/NT)\sum_{i=1}^{N} y_{i,-1}' M_{\dot{G}} y_{i,-1}} = \rho + \frac{(1/NT)\sum_{i=1}^{N} y_{i,-1}' M_{\dot{G}} \varepsilon_i^*}{(1/NT)\sum_{i=1}^{N} y_{i,-1}' M_{\dot{G}} y_{i,-1}}, \qquad (17)$$

where $M_{\dot{G}} = I_T - \dot{G}(\dot{G}'\dot{G})^{-1}\dot{G}'$ and $\dot{G} = (\iota_t, G)$ with $G = (\overline{y}, \overline{y}_{-1})$, $\overline{y} = (\overline{y}_1, \ldots, \overline{y}_T)'$ and $\overline{y}_{-1} = (\overline{y}_0, \ldots, \overline{y}_{T-1})'$.

Theorem 1. *In model (1) with errors satisfying A1-A3, the CCEPu estimator is inconsistent as*

[3] See Phillips and Sul (2007) for an expression with multiple factors.

$N \to \infty$

$$\underset{N\to\infty}{plim}\left(\widehat{\rho}_{CCEPu}-\rho\right)=\frac{-A(\rho,T)-\sum_{t=1}^{T-1}\rho^{t-1}\widetilde{g}_{F,t}^{+}}{B(\rho,T)-\frac{2}{1-\rho^{2}}\left(1+\rho\sum_{t=1}^{T-1}\rho^{t-1}\widetilde{g}_{F,t}^{+}\right)}, \tag{18}$$

where $\widetilde{g}_{F,t}^{+}=\sum_{s=t+1}^{T}\widetilde{\tau}_{s,s-t}^{+}$ with $\widetilde{\tau}_{s,s-t}^{+}$ being the $(s,s-t)$th element in $\widetilde{H}\left(\widetilde{H}'\widetilde{H}\right)^{-1}\widetilde{H}'$, $\widetilde{H}=\left(\widetilde{F}^{+},\widetilde{F}_{-1}^{+}\right)$, $\widetilde{F}^{+}=\left(\widetilde{F}_{1}^{+},\ldots,\widetilde{F}_{T}^{+}\right)'$ and $\widetilde{F}_{-1}^{+}=\left(\widetilde{F}_{0}^{+},\ldots,\widetilde{F}_{T-1}^{+}\right)'$.
The inconsistency (18) has the following asymptotic representation as $(N,T)_{seq}\to\infty$

$$\underset{N\to\infty}{plim}\left(\widehat{\rho}_{CCEPu}-\rho\right)=-\frac{1}{T}\left(1+2\rho+\frac{\theta(1-\rho^{2})}{1-\theta\rho}\right)+o_{p}\left(\frac{1}{T}\right). \tag{19}$$

The implication of Theorem 1 is that the CCEPu estimator is consistent for $(N,T)_{seq}\to\infty$ but has a different asymptotic bias compared to the WGi estimator for $N\to\infty$ and T fixed. The intuition for this is that the error term ε_{it} is now orthogonalised on a constant, $\overline{y}_t$ and $\overline{y}_{t-1}$ with, as can be seen from (15), the latter two converging to F_t^{+} and F_{t-1}^{+} respectively as $N\to\infty$. For fixed T, this implies two differences compared to the WGi estimator. First, orthogonalising on F_{t-1}^{+}, next to on a constant and on F_t^{+}, results in extra correlation between the orthogonalised error term and the explanatory variable $y_{i,t-1}$ as the latter is by construction a function of F_{t-1}^{+}. As such, comparing (11) and (19), the numerator of the latter contains an extra term in ρ. Second, the extra variability in the explanatory variable $y_{i,t-1}$ induced by F_{t-1} is now completely filtered out by orthogonalising on F_{t-1}^{+}. As such, stronger cross-sectional dependence raises the denominator in (11) but doesn't affect (19). Further comparing (11) and (19), it is clear that both asymptotic biases need not have the same direction and that the absolute value of the asymptotic bias of the CCEPu estimator is not necessarily bigger than that of the WGi estimator. However, for the majority of values for ρ and θ, the absolute value of the asymptotic bias is larger for the CCEPu estimator. For the most relevant case of both $\rho>0$ and $\theta>0$, both the WGi and CCEPu estimator have a downward asymptotic bias. However, the downward bias of the latter is bigger compared to the former. Thus, in these cases, approximating the unobserved F_t using cross-sectional averages of the observed data comes at the cost of a higher asymptotic bias for $N\to\infty$ and T fixed compared to a situation where F_t is observed.

3.3.2 Restricted CCEP

Taking into account the restrictions on γ_{1i}, γ_{2i} and α_i^{*}, the restricted CCEP estimator for ρ in (16) can be obtained by minimizing the objective function

$$S_{NT}(\rho,F)=\frac{1}{NT}\sum_{i=1}^{N}\left(y_i-\rho y_{i,-1}\right)'M_{\dot{F}}\left(y_i-\rho y_{i,-1}\right). \tag{20}$$

Although F is not observed when estimating ρ and similarly, ρ is not observed when estimating F, we can replace the unobserved quantities by initial estimates and iterate until convergence. The continuously-updated estimator for (ρ, F) is defined as

$$\left(\widehat{\rho}_{\text{CCEPr}}, \widehat{F}\right) = \underset{\rho, F}{\operatorname{argmin}}\, S_{NT}(\rho, F). \tag{21}$$

More specifically, $\left(\widehat{\rho}_{\text{CCEPr}}, \widehat{F}\right)$ is the solution to the following two equations

$$\widehat{\rho}_{\text{CCEPr}} = \frac{(1/NT)\sum_{i=1}^{N} y_{i,-1}' M_{\widehat{\dot{F}}} y_i}{(1/NT)\sum_{i=1}^{N} y_{i,-1}' \widehat{\dot{F}} y_{i,-1}} = \rho + \frac{(1/NT)\sum_{i=1}^{N} y_{i,-1}' M_{\widehat{\dot{F}}}(\gamma_i F + \varepsilon_i)}{(1/NT)\sum_{i=1}^{N} y_{i,-1}' M_{\widehat{\dot{F}}} y_{i,-1}}, \tag{22}$$

$$\widehat{F} = \frac{1}{\overline{\gamma}}\left(\overline{y} - \widehat{\rho}_{\text{CCEPr}}\overline{y}_{-1} - \widehat{\overline{\alpha}}\right), \tag{23}$$

where $M_{\widehat{\dot{F}}} = I_T - \widehat{\dot{F}}\left(\widehat{\dot{F}}'\widehat{\dot{F}}\right)^{-1}\widehat{\dot{F}}'$ with $\widehat{\dot{F}} = \left(\iota, \widehat{F}\right)$. Note that the restricted CCEP estimator bears some similarities with the continuously updated (Cup) estimator presented in Bai *et al.* (2009). The difference being that the CCEP estimates the unobserved components via the cross-sectional averages of both dependent and indepent variables, whereas the Cup estimator uses a principal component approach.

Theorem 2. *In model (1) with errors satisfying A1-A3, the CCEPr estimator is consistent as* $(N, T)_{seq} \to \infty$,

$$\widehat{\rho}_{CCEPr} \underset{p}{\to} \rho, \tag{24}$$

Theorem 3. *In model (1) with errors satisfying A1-A3, the CCEPr estimator is inconsistent as* $N \to \infty$

$$\underset{N\to\infty}{plim}\left(\widehat{\rho}_{CCEPr} - \rho\right) = \frac{-A(\rho, T) - \sum_{t=1}^{T-1}\rho^{t-1}\widetilde{g}_{\widehat{F},t}}{B(\rho, T) - \frac{2}{1-\rho^2}\left(1 + \rho\sum_{t=1}^{T-1}\rho^{t-1}\widetilde{g}_{\widehat{F},t}\right)}, \tag{25}$$

where $\widetilde{g}_{\widehat{F},t} = \sum_{s=t+1}^{T}\widehat{\widetilde{\tau}}_{s,s-t}$ *with* $\widehat{\widetilde{\tau}}_{s,s-t}$ *being the* $(s, s-t)$ *th element in* $\widehat{\widetilde{F}}\left(\widehat{\widetilde{F}}'\widehat{\widetilde{F}}\right)^{-1}\widehat{\widetilde{F}}'$. *Using theorem 2, it can be shown that the inconsistency of (25) has the following asymptotic representation as* $(N, T)_{seq} \to \infty$

$$\underset{N\to\infty}{plim}\left(\widehat{\rho}_{CCEPr} - \rho\right) = -\frac{1}{T}\left(1 + \rho + \frac{\theta(1-\rho^2)}{1-\theta\rho}\right) + o_p\left(\frac{1}{T}\right). \tag{26}$$

Comparing (26) and (11), the asymptotic bias of the CCEPr equals the asymptotic bias of the WGi multiplied by a factor $\left(1 + \frac{m_\gamma^2}{(1-\theta\rho)^2}\frac{\sigma_\mu^2}{\sigma_\varepsilon^2}\right) > 1$. This implies that the asymptotic CCEPr bias has the same direction and is bigger than that of the WGi estimator. The intuition for this is that for fixed T, the deviation of $\widehat{F}_t$ from F_t is a function of F^+_{t-1}, as can be seen from (A-27), which induces extra correlation between the error term ε_{it} and $y_{i,t-1}$ after orthogonalisation on $\widehat{F}_t$. Further comparing (26), (19) and (11), the asymptotic bias of the CCEPr estimator is smaller than that of the CCEPu estimator and closer to that of the WGi estimator for the most relevant case of $\rho > 0$ and $\theta > 0$.

4 Including exogenous variables

This section extends the model in (1) by including a vector of exogenous variables, x_{it}[4]. Consider the following autoregressive model

$$y_{it} = \alpha_i + \rho y_{i,t-1} + x'_{it}\beta + \nu_{it}, \qquad |\rho| < 1, \qquad (i = 1, \dots, N; \quad t = 1, \dots, T), \tag{27}$$

with $x_{it} = \left(x^1_{it}, \dots, x^K_{it}\right)'$ a $(K \times 1)$ vector of explanatory variables and β a $(K \times 1)$ coefficient vector. In addition to assumptions A1 - A3, we assume

Assumption A4. *(x_{it} condition) x_{it} is strictly exogenous with respect to the residuals ε_{it} but allowed to be correlated with the individual effects and common components*

$$E(x_{it}\varepsilon_{js}) = 0, \quad E(x_{it}\alpha_i) = \sigma^2_{x\alpha}, \quad E(x_{it}F_t) = \sigma^2_{xF}, \qquad \forall i, t, j, s \tag{28}$$

with $\sigma^2_{x\alpha}$ and σ^2_{xF} being $(K \times 1)$ vectors.

Stacking the model in (27) and (2) for each i yields

$$y_i = \alpha_i + \rho y_{i,-1} + x_i\beta + \gamma_i F + \varepsilon_i. \tag{29}$$

The WGi and CCEPu estimators[5] for ρ and β in (29) can be written as

$$\widehat{\rho} = \left(\frac{1}{NT}\sum_{i=1}^{N} y^{\perp\prime}_{i,-1} M_{x^\perp} y^\perp_{i,-1}\right)^{-1} \left(\frac{1}{NT}\sum_{i=1}^{N} y^{\perp\prime}_{i,-1} M_{x^\perp} y^\perp_{i}\right), \tag{30a}$$

$$\widehat{\beta} = \left(\frac{1}{NT}\sum_{i=1}^{N} x^{\perp\prime}_{i} x^\perp_{i}\right)^{-1} \left(\frac{1}{NT}\sum_{i=1}^{N} x^{\perp\prime}_{i} \left(y^\perp_i - \widehat{\rho} y^\perp_{i,-1}\right)\right), \tag{30b}$$

where $y^\perp_i = M_{\dot{Q}} y_i$, $y^\perp_{i,-1} = M_{\dot{Q}} y_{i,-1}$, $x^\perp_i = M_{\dot{Q}} x_i$ and $\varepsilon^\perp_i = M_{\dot{Q}}\varepsilon_i$. $M_{\dot{Q}} = I_T - \dot{Q}\left(\dot{Q}'\dot{Q}\right)^{-1}\dot{Q}'$ is a matrix which orthogonalizes out both individual effects and unobserved common factors with $\dot{Q} = (\iota, Q)$. For the WGi estimator $Q = F$ while $Q = \left(\overline{y}, \overline{y}_{-1}, \overline{x}\right)$ for the CCEPu estimator with

[4] For the case of endogenous variables, Harding and Lamarche (2011) and Everaert and Pozzi (2010) respectively describe an IV and GMM approach based on the pooled CCE estimator.

[5] Restricting our attention to these two estimators is for notational purposes only. Derivations for the naive WG are similar to the ones of the WGi and CCEPu. Formulae for CCEPr estimator directly follow from rewriting the proofs of theorems 2 and 3 into matrix notation.

$\overline{x} = (\overline{x}_1, \ldots, \overline{x}_T)'$, $\overline{x}_t = \left(\overline{x}_t^1, \ldots, \overline{x}_t^K\right)'$ and $\overline{x}_t^k = \frac{1}{N}\sum_{i=1}^N x_{it}^k$.

Next, letting $N \to \infty$, (30a) and (30b) are given by

$$\operatorname*{plim}_{N\to\infty}(\widehat{\rho} - \rho) = \left(\operatorname*{plim}_{N\to\infty} \frac{1}{NT}\sum_{i=1}^N y_{i,-1} M_{\dot{Q}} M_{x^\perp} y_{i,-1}\right)^{-1} \left(\operatorname*{plim}_{N\to\infty} \frac{1}{NT}\sum_{i=1}^N y_{i,-1} M_{\dot{Q}} \varepsilon\right), \tag{31a}$$

$$\operatorname*{plim}_{N\to\infty}(\widehat{\beta} - \beta) = -\left(\operatorname*{plim}_{N\to\infty} \left(\frac{1}{NT}\sum_{i=1}^N x_i M_{\dot{Q}} x_i\right)^{-1} \frac{1}{NT}\sum_{i=1}^N x_i M_{\dot{Q}} y_{i,-1}\right) \operatorname*{plim}_{N\to\infty}(\widehat{\rho} - \rho), \tag{31b}$$

where use is made of the fact that (i) $x_i^\perp$ an exogenous variable, (ii) for the WGi estimator $\dot{Q}$ contains F and (iii) for the CCEPu estimator

$$F = \frac{1}{\overline{\gamma}}\left(\overline{y} - \overline{\alpha} - \rho\overline{y}_{-1} - \overline{x}\beta - \overline{\varepsilon}\right),$$

with $\operatorname{plim}_{N\to\infty}\overline{\varepsilon} = 0$ such that for $N \to \infty$ the factor F is eliminated by orthogonalising on $\left(\overline{y}, \overline{y}_{-1}, \overline{x}\right)$ and a constant.

Adding an additional vector of exogenous explanatory variables to our original model (1) has little implications with respect to our earlier findings. For $\operatorname{plim}_{N\to\infty}(\widehat{\rho} - \rho)$, (31a) indicates that an additional, positive term is added to the denominator. This follows directly from rewriting model (27) to its compontent form

$$y_{it} = y_{it}^0 + x_{it}^0\beta, \quad y_{it}^0 = \alpha_i + \rho y_{i,t-1}^0 + \gamma_i F_t + \varepsilon_{it}, \quad x_{it}^0 = (1 - \rho L)^{-1} x_{it}, \tag{32}$$

where y_{it}^0 is equivalent to the model described by (1), and inserting it into the denominator. This wil lessen the degree of the bias. The nominator is almost identical to our earlier findings. The nominator of the CCEPu estimator for (27) differs slightly from the one given by (18) due to the fact that $\overline{x}_t$ is added to the orthogonalisation matrix $M_{\dot{Q}}$. The corresponding expansion for $T \to \infty$ of this matrix will depend on the specification of x_t. Turning to the asymptotic behaviour of $\operatorname{plim}_{N\to\infty}(\widehat{\beta} - \beta)$, equation (31b) reveals that the inconsistency depends (i) on the asymptotic behaviour of $(\widehat{\rho} - \rho)$ and (ii) on the relationship between the exogenous variables x_{it} and $y_{i,t-1}$.

5 Monte Carlo simulation

In this section we investigate the small sample properties of the WGn, WGi, CCEPu and CCEPr estimators under cross-sectional dependence[6]. We are interested in the effects of (i) the extent of cross-sectional dependence, (ii) the degree of inertia in both the dependent variable and the factor and (iii) the relative importance of the variance of the factor loadings and the idiosyncratic errors.

[6]Sarafidis and Robertson (2009) explore the behaviour of GMM estimators under error cross-sectional dependence. They find that (i) the standard moment conditions of GMM estimators are no longer valid and (ii) the WGn often outperforms the naive first-difference GMM in terms of root median squared error. Therefore, we do not include a naive GMM estimator in our simulations.

5.1 Experimental design

The data generating process we consider is given by

$$\begin{aligned} y_{it} &= \alpha_i + \rho y_{i,t-1} + \beta x_{it} + \gamma_i F_t + \varepsilon_{it}, \\ x_{it} &= \phi_i F_t + \omega_{it}, \\ F_t &= \theta F_{t-1} + \mu_t, \end{aligned} \tag{33}$$

with ε_{it}, ω_{it} and μ_t randomly drawn in each replication from $i.i.d.N\left(0, \sigma_\varepsilon\right)$, $i.i.d.N\left(0, \sigma_\omega\right)$ and $i.i.d.N\left(0, \sigma_\mu\right)$ respectively and γ_i and ϕ_i drawn from $i.i.d.U\left[\gamma_L, \gamma_U\right]$ and $i.i.d.U\left[\phi_L, \phi_U\right]$ respectively. We initialise $y_{i,-49}$, $x_{i,-49}$ and F_{-49} at zero and discard the first 50 observations.

When comparing the estimators over different values of the dynamic parameters ρ and θ, we wish to control (i) the signal-to-noise ratio, (ii) the relative importance of the error components (α_i, $\gamma_i F_t$, ε_{it}) in terms of contribution to the variance of y_{it} and (iii) the relative importance of factor loading heterogeneity γ_i in the overall cross-sectional dependence m_γ^2. To this end, similar to Sarafidis *et al.* (2009) we extend the Monte Carlo design of Kiviet (1995) by allowing a factor structure in the error process. First note that the total variance in y_{it} is given by

$$\begin{aligned} Var(y_{it}) &= \frac{1}{(1-\rho)^2}\sigma_\alpha^2 + \frac{\beta^2}{1-\rho^2}\sigma_\omega^2 + \frac{(1+\theta\rho)(\beta^2 m_\phi^2 + m_\gamma^2)}{(1-\rho^2)(1-\theta^2)(1-\rho\theta)}\sigma_\mu^2 + \frac{1}{1-\rho^2}\sigma_\varepsilon^2, \\ &= a_1\sigma_\alpha^2 + a_2\beta^2\sigma_\omega^2 + a_2 a_3\sigma_\mu^2 + a_2\sigma_\varepsilon^2, \end{aligned} \tag{34}$$

where $a_1 = \frac{1}{(1-\rho)^2}$, $a_2 = \frac{1}{1-\rho^2}$ and $a_3 = \frac{(1+\theta\rho)(\beta^2 m_\phi^2 + m_\gamma^2)}{(1-\theta^2)(1-\rho\theta)}$ with $m_\phi^2 = \underset{N\to\infty}{\text{plim}} \dfrac{1}{N}\sum_{i=1}^{N}\phi_i^2$. The terms on the right hand side of (34) capture the contributions to the variance of y_{it} of the individual effects, the exogenous variable, the common factor and the idiosyncratic error term respectively. As such, we can control the relative importance of

- the individual effects and the idiosyncratic error term by setting ψ_1

$$\psi_1 = \frac{a_1\sigma_\alpha^2}{a_2\sigma_\varepsilon^2}, \tag{35}$$

- the common factor and idiosyncratic error term by setting ψ_2

$$\psi_2 = a_3\frac{\sigma_\mu^2}{\sigma_\varepsilon^2}, \tag{36}$$

- the mean $\overline{\gamma}$ and the variance σ_γ^2 of γ_i within the overall degree of cross-sectional dependence

m_γ^2 by setting ψ_3[7]

$$\psi_3 = \frac{\overline{\gamma}^2}{\overline{\gamma}^2 + \sigma_\gamma^2} = \frac{\overline{\gamma}^2}{m_\gamma^2}. \quad (37)$$

After applying the normalisations $\sigma_\varepsilon^2 = \sigma_\mu^2 = 1$, $\beta = 1 - \rho$, $(\phi_L, \phi_U) = (0, 1)$ (implying $m_\phi^2 = 1/3$), for given values of ρ and θ the values for σ_α^2, m_γ^2 and (γ_L, γ_U) follow from (35), (36) and (37) respectively.

We further define the signal, σ_s^2, as the amount of variance in y_{it} induced by information contained in $y_{i,t-1}$ and x_{it}

$$\sigma_s^2 = Var\left(y_{it} - \alpha_i/(1-\rho) - \gamma_i F_t - \varepsilon_{it}\right), \quad (38)$$

$$= \frac{\beta^2}{1-\rho^2}\sigma_\omega^2 + \left(\frac{(1+\theta\rho)(\beta^2 m_\phi^2 + m_\gamma^2)}{(1-\rho^2)(1-\theta^2)(1-\theta\rho)} - \frac{m_\gamma^2}{1-\theta^2}\right)\sigma_\mu^2 + \frac{\rho^2}{1-\rho^2}\sigma_\varepsilon^2. \quad (39)$$

The signal-to-noise ratio is then given by $\xi = \sigma_s^2/\sigma_{\gamma_i F_t + \varepsilon_{it}}^2$. Setting ξ to a specific value, allows us to calculate σ_ω^2.

We conduct experiments for combinations of the following parameter values: $\rho \in \{0.4; 0.6; 0.8\}$, $\theta \in \{0; 0.4; 0.8\}$, $\psi_2 \in \{0.5; 1; 2\}$, $\psi_3 \in \{0.4; 0.8\}$ and $\psi_3 \in \{3; 10\}$. ψ_1 is set fixed at 1. Experiment 1 serves as a point of reference and has the following settings: $\rho = 0.6$, $\beta = 0.4$, $\theta = 0.4$, $\xi = 3$, $\psi_1 = 1$, $\psi_2 = 1$, $\psi_3 = 0.80$ and $\xi = 3$. The other simulations assess the impact on the small sample properties of the estimators for (i) a higher persistence in both y_{it} and the common factors, (ii) different degrees of error cross section dependence, (iii) the degree of heterogeneity of γ_i and (iv) a change in the signal-to-noise ratio. T and N respectively take the following values:{5;10;20;30;40;50} and {20;50}. The estimators are compared in terms of (i) mean bias (bias), (ii) mean estimated standard error (stde), (iii) standard deviation (stdv) and (iv) root mean squared error (rmse). All experiments are based on 5000 iterations.

5.2 Simulation Results

The simulation results can be found in the tables in appendix B. Table 1 reports the results for the benchmark experiment. With respect to estimating ρ, the following conclusions stand out. First, the WGn estimator is biased for all combinations of N and T, with the bias being negative for small T and positive for larger values of T. This switch in sign is due to the fact as T increases the Nickell part of the bias, which is negative, diminishes whereas the positive bias originating from the unobserved common components does not. Second, for small values of T the bias of the WGi estimator is highly similar to that of the WGn estimator but as T increases, this bias shrinks to zero. Third, the CCEP estimators both have a considerably larger bias compared to the WG estimators for small values of T but CCEPr clearly outperforms CCEPu. For both CCEP

[7]Setting ψ_3 to one results in a homogeneous time-effect.

estimators the bias diminishes as T increases and is more or less comparable to that of the WGi estimator for $T = 50$. Fourth, looking at the standard deviations, the WGi estimator is clearly more efficient than the other estimators. Also note that for small values of T, the WGn estimator has smaller standard deviations compared to the CCEP estimators but the CCEP estimators become more efficient when T increases and even have standard deviations that are not much bigger than the WGi estimator for $T = 50$. Fifth, in terms of estimated standard errors, all estimators substantially underestimate the true standard deviations for small T. For larger values of T, estimated standard errors converge to the true standard deviations for the WGi and CCEP estimators but not for the WGn estimator. Turning to the estimates of β, a small bias is found for $T = 5$ for all estimators. This bias disappears as T increases for the WGi and CCEP estimators but again not for the WGn estimator. Finally, note that the size of the cross-sectional dimension N does not affect the size of the bias for any of the considered estimators, i.e. in all of the experiments there is virtually no difference between setting $N = 20$ or $N = 50$. Therefore, when analysing deviations from the benchmark experiment in Tables 2-6, we only report simulation results setting $N = 50$ [8].

Tables 2 and 3 demonstrate that higher persistence in y_{it}, by increasing either the persistence in the model through the dynamic coefficient ρ or the persistence in the common factor through the coefficient θ, results in a larger downward bias for all estimators. Especially the CCEP estimators are strongly downward biased when $\theta = 0.8$ and T is small. Table 4 shows that a higher degree of error cross-sectional dependence (governed by ψ_2), reduces the small T bias of all estimators. For a larger value of T, the bias of the WGn estimator increases though. Table 5 shows that the heterogeneity in the cross-sectional dependence (keeping overall heterogeneity fixed) does not affect the size of the bias. Table 6 illustrates that the stronger the signal contained in the explanatory variables, the smaller the small sample bias is.

6 Conclusion

This papers examines the effects of error cross-sectional dependence, modelled as an unobserved common factor, on WG and CCEP estimators in a linear dynamic panel data model. In general, the asymptotic behaviour as $N \to \infty$ of each estimator breaks down in two parts: the well known Nickell dynamic panel data bias and a random asymptotic bias which depends on the particular realisation of the unobserved common factor. First, in line with Phillips and Sul (2007), we find that the naive WG estimator is inconsistent for $N \to \infty$ and T fixed. For a cross-sectionally dependent factor, the inconsistency remains even for $(N,T)_{\text{seq}} \to \infty$. Second, the benchmark infeasible WG estimator, including the unobserved common factor as an observed explanatory variable, is also inconsistent for $N \to \infty$ and fixed T but this inconsistency disappears as $(N,T)_{\text{seq}} \to \infty$. Third, contrary to the findings in Pesaran (2006) for a static model the unrestricted CCEP estimator is inconsistent for $N \to \infty$ and fixed T. For a relevant range of

[8] The simulation results for $N = 20$ are available upon request

parameter combinations, the asymptotic bias is larger compared to the infeasible WG estimator. Restricting the CCEP estimator by taking into account the restrictions on the individual-specific factor loadings as implied by the derivation of the specification of the model augmented with cross-sectional averages results in a somewhat smaller asymptotic bias. Letting $(N,T)_{\text{seq}} \to \infty$, both the unrestricted and the restricted CCEP estimators are consistent. The small sample properties of the various estimators are analysed using Monte Carlo experiments. The simulation results confirm the breakdown of the naive WG estimator under error cross-sectional dependence and shows that the performance of the CCEP estimators is not that different from the infeasible WG estimator provided T is not too small.

References

ARELLANO, M. and BOND, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies*, **58** (2), 277–297.

BAI, J., KAO, C. and NG, S. (2009). Panel cointegration with global stochastic trends. *Journal of Econometrics*, **149** (1), 82–99.

BAI, J. S. and NG, S. (2004). A panic attack on unit roots and cointegration. *Econometrica*, **72** (4), 1127–1177.

BLUNDELL, R. and BOND, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, **87** (1), 115–143.

BUN, M. and CARREE, M. (2005). Bias-corrected estimation in dynamic panel data models. *Journal of Business & Economic Statistics*, **23** (2), 200–210.

COAKLEY, J., FUERTES, A.-M. and SMITH, R. (2002). A principal components approach to cross-section dependence in panels. *Unpublished Manuscript, Birkbeck College, University of London*.

EVERAERT, G. and POZZI, L. (2007). Bootstrap-based bias correction for dynamic panels. *Journal of Economic Dynamics & Control*, **31** (4), 1160–1184.

— and — (2010). *The Stickiness of Aggregate Consumption Growth in OECD Countries: A Panel Data Analysis*. Working Papers Series 10/654, Ghent University, Faculty of Economics and Business Administration.

HARDING, M. and LAMARCHE, C. (2011). Least squares estimation of a panel data model with multifactor error structure and endogenous covariates. *Economics Letters*, **111** (3), 197 – 199.

KIVIET, J. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics*, **68** (1), 53–78.

NICKELL, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, **49** (6), 1417–1426.

PESARAN, M. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, **74** (4), 967–1012.

PHILLIPS, P. C. B. and SUL, D. (2003). Dynamic panel estimation and homogeneity testing under cross section dependence. *Econometrics Journal*, **6** (1), 217–259.

— and — (2007). Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence. *Journal Of Econometrics*, **137** (1), 162–188.

SARAFIDIS, V. and ROBERTSON, D. (2009). On the impact of error cross-sectional dependence in short dynamic panel estimation. *Econometric Journal*, **12** (1), 62–81.

—, YAMAGATA, T. and ROBERTSON, D. (2009). A test of cross section dependence for a linear dynamic panel model with regressors. *Journal of Econometrics*, **148** (2), 149–161.

Appendices

Appendix A Proofs

Lemma A-1.

Under Assumption A1 we have from (4c)

$$\begin{aligned}
\lambda_0 &= E\left(F_t^{+}\right)^2 = (\rho+\theta)\,\lambda_1 - \rho\theta\lambda_2 + \sigma_\mu^2, \\
\lambda_1 &= E\left(F_t^{+}F_{t-1}^{+}\right) = (\rho+\theta)\,\lambda_0 - \rho\theta\lambda_1, \\
\lambda_s &= E\left(F_t^{+}F_{t-s}^{+}\right) = (\rho+\theta)\,\lambda_{s-1} - \rho\theta\lambda_{s-2}, \qquad \forall s \geq 2
\end{aligned} \tag{A-1}$$

which can be solved to obtain

$$\begin{aligned}
\lambda_0 &= \frac{1+\theta\rho}{(1-\theta\rho)(1-\theta-\rho+\theta\rho)(1+\theta+\rho+\theta\rho)}\sigma_\mu^2, \\
&= \frac{1+\theta\rho}{(1-\theta\rho)(1-\theta^2)(1-\rho^2)}\sigma_\mu^2, & \text{(A-2)} \\
\lambda_1 &= \frac{\theta+\rho}{(1-\theta\rho)(1-\theta^2)(1-\rho^2)}\sigma_\mu^2. & \text{(A-3)}
\end{aligned}$$

Next, using that $F_t = F_t^{+} - \rho F_{t-1}^{+}$ we have

$$E\left(F_t F_{t-1}^{+}\right) = E\left(F_t^{+}F_{t-1}^{+}\right) - \rho E\left(F_{t-1}^{+}\right)^2 = \lambda_1 - \rho\lambda_0 = \frac{\theta}{(1-\theta\rho)(1-\theta^2)}\sigma_\mu^2. \tag{A-4}$$

Proof of Proposition 1. Suppose Assumptions A1-A2 hold, then:

$$\begin{aligned}
\operatorname*{plim}_{N\to\infty}\left(\widehat{\rho}_{\text{WGn}} - \rho\right) &= \operatorname*{plim}_{N\to\infty} \frac{(1/NT)\sum_{i=1}^N y_{i,-1}' M_I\left(\gamma_i F + \varepsilon_i\right)}{(1/NT)\sum_{i=1}^N y_{i,-1}' M_I y_{i,-1}}, \\
&= \frac{\operatorname*{plim}_{N\to\infty}(1/NT)\sum_{i=1}^N\sum_{t=1}^T \left(y_{i,t-1}^{+} + \gamma_i F_{t-1}^{+}\right)\left[\gamma_i\left(F_t - \frac{1}{T}\sum_{s=1}^T F_s\right) + \left(\varepsilon_{it} - \frac{1}{T}\sum_{s=1}^T \varepsilon_{is}\right)\right]}{\operatorname*{plim}_{N\to\infty}(1/NT)\sum_{i=1}^N\sum_{t=1}^T\left[\left(y_{i,t-1}^{+} - \frac{1}{T}\sum_{s=1}^T y_{i,s-1}^{+}\right) + \gamma_i\left(F_{t-1}^{+} - \frac{1}{T}\sum_{s=1}^T F_{s-1}^{+}\right)\right]^2}, \\
&= \frac{\frac{1}{T}\sum_{t=1}^T \operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^N\left[-y_{i,t-1}^{+}\frac{1}{T}\sum_{s=1}^T\varepsilon_{is} + \gamma_i^2 F_{t-1}^{+}F_t - \gamma_i^2 F_{t-1}^{+}\frac{1}{T}\sum_{s=1}^T F_s\right]}{\frac{1}{T}\sum_{t=1}^T \operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^N\left(y_{i,t-1}^{+} - \frac{1}{T}\sum_{s=1}^T y_{i,s-1}^{+}\right)^2 + \gamma_i^2\left(F_{t-1}^{+} - \frac{1}{T}\sum_{s=1}^T F_{s-1}^{+}\right)^2}, \\
&= \frac{-\frac{1}{1-\rho}\left(1 - \frac{1}{T}\frac{1-\rho^T}{1-\rho}\right) + \frac{m_\gamma^2}{\sigma_\varepsilon^2}\sum_{t=1}^T F_{t-1}^{+}\widetilde{F}_t}{\frac{T-1}{1-\rho^2}\left(1 - \frac{1}{T-1}\frac{2\rho}{1-\rho}\left[1 - \frac{1}{T}\frac{1-\rho^T}{1-\rho}\right]\right) + \frac{m_\gamma^2}{\sigma_\varepsilon^2}\sum_{t=1}^T\left(\widetilde{F}_{t-1}^{+}\right)^2},
\end{aligned} \tag{A-5}$$

with $\widetilde{F}_t = F_t - \overline{F}$, $\widetilde{F}_{t-1}^{+} = F_{t-1}^{+} - \overline{F}_{-1}^{+}$, $\overline{F} = \frac{1}{T}\sum_{t=1}^T F_t$ and $\overline{F}_{-1}^{+} = \frac{1}{T}\sum_{t=1}^T F_{t-1}^{+}$.

Next, letting $T \to \infty$ and using Lemma A-1, we have

$$\frac{1}{T}\sum_{t=1}^{T}\left(F_{t-1}^{+}-\overline{F}_{-1}^{+}\right)^{2}=E\left(F_{t-1}^{+}\right)^{2}+O_{p}\left(\frac{1}{\sqrt{T}}\right)=\frac{1+\theta\rho}{\left(1-\theta\rho\right)\left(1-\theta^{2}\right)\left(1-\rho^{2}\right)}\sigma_{\mu}^{2}+O_{p}\left(\frac{1}{\sqrt{T}}\right),$$

$$\frac{1}{T}\sum_{t=1}^{T}\left(F_{t-1}^{+}F_{t}-F_{t-1}^{+}\overline{F}\right)=E\left(F_{t-1}^{+}F_{t}\right)+O_{p}\left(\frac{1}{\sqrt{T}}\right)=\frac{\theta}{\left(1-\theta\rho\right)\left(1-\theta^{2}\right)}\sigma_{\mu}^{2}+O_{p}\left(\frac{1}{\sqrt{T}}\right),$$

where use is made of $\frac{1}{T}\sum_{t=1}^{T}F_{t-1}^{+}\overline{F}_{-1}^{+}$, $\frac{1}{T}\sum_{t=1}^{T}\left(\overline{F}_{-1}^{+}\right)^{2}$ and $\frac{1}{T}\sum_{t=1}^{T}F_{t-1}^{+}\overline{F}$ all being $O_{p}\left(\frac{1}{T}\right)$.

Thus, the inconsistency in (7) has the following asymptotic representation as $T \to \infty$

$$\begin{aligned}
\operatorname*{plim}_{N\to\infty}\left(\widehat{\rho}_{\mathrm{WGn}}-\rho\right) &= \frac{-\frac{1}{1-\rho}\left(1-\frac{1}{T}\frac{1-\rho^{T}}{1-\rho}\right)+T\frac{m_{\gamma}^{2}}{\sigma_{\varepsilon}^{2}}\left(\frac{\theta}{(1-\theta\rho)(1-\theta^{2})}\sigma_{\mu}^{2}+o_{p}\left(1\right)\right)}{\frac{T-1}{1-\rho^{2}}\left(1-\frac{1}{T-1}\frac{2\rho}{1-\rho}\left[1-\frac{1}{T}\frac{1-\rho^{T}}{1-\rho}\right]\right)+T\frac{m_{\gamma}^{2}}{\sigma_{\varepsilon}^{2}}\left(\frac{1+\theta\rho}{(1-\theta\rho)(1-\theta^{2})(1-\rho^{2})}\sigma_{\mu}^{2}+o_{p}\left(1\right)\right)},\\
&= \left(-\frac{1+\rho}{T}+\frac{\left(1-\rho^{2}\right)\theta}{(1-\theta\rho)(1-\theta^{2})}\frac{m_{\gamma}^{2}\sigma_{\mu}^{2}}{\sigma_{\varepsilon}^{2}}\right)\left(1+\frac{(1+\theta\rho)}{\left(1-\theta\rho\right)\left(1-\theta^{2}\right)}\frac{m_{\gamma}^{2}\sigma_{\mu}^{2}}{\sigma_{\varepsilon}^{2}}\right)^{-1}+o_{p}\left(1\right).
\end{aligned}$$

Proof of Proposition 2. First, the probability limit as $N \to \infty$ of the numerator of equation (9) is given by

$$\begin{aligned}
\operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N}y_{i,-1}'M_{\dot{F}}\varepsilon_{i} &= \operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}y_{i,t-1}\left[\left(\varepsilon_{it}-\overline{\varepsilon}_{i}\right)-\sum_{s=1}^{T}\widetilde{\tau}_{st}\left(\varepsilon_{is}-\overline{\varepsilon}_{i}\right)\right],\\
&= \frac{1}{T}\sum_{t=1}^{T}\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}y_{i,t-1}\varepsilon_{it}-\frac{1}{T}\sum_{t=1}^{T}\left(1-\sum_{s=1}^{T}\widetilde{\tau}_{st}\right)\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}y_{i,t-1}\overline{\varepsilon}_{i}\\
&\qquad -\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{\tau}_{st}\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}y_{i,t-1}\varepsilon_{is}, \qquad \text{(A-6)}
\end{aligned}$$

where $\widetilde{\tau}_{st}=\widetilde{F}_{s}\widetilde{F}_{t}\Big/\sum_{t=1}^{T}\widetilde{F}_{t}^{2}$. Using

$$\begin{aligned}
\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}y_{i,t-1}\varepsilon_{i,t-s} &= \sigma_{\varepsilon}^{2}\rho^{s-1} \qquad \forall s\geq 1, \qquad \text{(A-7)}\\
&= 0 \qquad \forall s<1, \qquad \text{(A-8)}
\end{aligned}$$

the first term of (A-6) drops and the last term rewrites to

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{\tau}_{st}\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}y_{i,t-1}\varepsilon_{is} &= -\sigma_{\varepsilon}^{2}\frac{1}{T}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\widetilde{\tau}_{st}\rho^{s+t-1}=-\sigma_{\varepsilon}^{2}\frac{1}{T}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\widetilde{\tau}_{t,t-s}\rho^{s-1},\\
&= -\sigma_{\varepsilon}^{2}\frac{1}{T}\left(\sum_{t=2}^{T}\widetilde{\tau}_{t,t-1}+\rho\sum_{t=3}^{T}\widetilde{\tau}_{t,t-2}+\rho^{2}\sum_{t=4}^{T}\widetilde{\tau}_{t,t-3}+\ldots+\rho^{T-2}\widetilde{\tau}_{T,1}\right),\\
&= -\sigma_{\varepsilon}^{2}\frac{1}{T}\sum_{t=1}^{T-1}\rho^{t-1}\sum_{s=t+1}^{T}\widetilde{\tau}_{s,s-t}=\sigma_{\varepsilon}^{2}\frac{1}{T}\sum_{t=1}^{T-1}\rho^{t-1}\widetilde{g}_{F,t}, \qquad \text{(A-9)}
\end{aligned}$$

where $\widetilde{g}_{F,t} = \sum_{s=t+1}^{T} \widetilde{\tau}_{s,s-t} = \sum_{s=t+1}^{T} \widetilde{F}_s \widetilde{F}_{s-t} \Big/ \sum_{t=1}^{T} \widetilde{F}_t^2$.
The second term is given by

$$\frac{1}{T}\sum_{t=1}^{T}\left(1-\sum_{s=1}^{T}\tau_{st}\right)\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}y_{i,t-1}\overline{\varepsilon}_i = \frac{1}{T^2}\frac{\sigma_\varepsilon^2}{1-\rho}\sum_{t=1}^{T}(1-\rho^{t-1})\left(1-\frac{\widetilde{F}_t\sum_{s=1}^{T}\widetilde{F}_s}{\sum_{t=1}^{T}\widetilde{F}_t^2}\right)$$
$$= \frac{1}{T^2}\frac{\sigma_\varepsilon^2}{1-\rho}\sum_{t=1}^{T}(1-\rho^{t-1}) = \frac{1}{T}\left[\frac{\sigma_\varepsilon^2}{T}\frac{1}{1-\rho}\left[T-\frac{1-\rho^T}{1-\rho}\right]\right],$$
$$= \frac{1}{T}\sigma_\varepsilon^2 A(\rho,T). \tag{A-10}$$

Using (A-7) - (A-10), equation (A-6) can be written as

$$\operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N}y_{i,-1}'M_{\dot{F}}\varepsilon_i = -\frac{\sigma_\varepsilon^2}{T}A(\rho,T) - \sigma_\varepsilon^2\frac{1}{T}\sum_{t=1}^{T-1}\rho^{t-1}\widetilde{g}_{F,t}. \tag{A-11}$$

Second, the probability limit as $N\to\infty$ of the denominator of equation (9) is given by

$$\operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N}y_{i,-1}'M_{\dot{F}}y_{i,-1} = \operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}y_{i,t-1}\left[\left(y_{i,t-1}-\overline{y}_{i,-1}\right)-\sum_{s=1}^{T}\widetilde{\tau}_{st}\left(y_{i,s-1}-\overline{y}_{i,-1}\right)\right],$$
$$= \operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[y_{i,t-1}^2-y_{i,t-1}\overline{y}_{i,-1}\right] - \operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{\tau}_{st}y_{i,t-1}y_{i,s-1}$$
$$+ \operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{\tau}_{st}y_{i,t-1}\overline{y}_{i,-1},$$
$$= \frac{1}{T}\sum_{t=1}^{T}\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left[y_{i,t-1}^2-y_{i,t-1}\overline{y}_{i,-1}\right] - \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{\tau}_{st}\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}y_{i,t-1}y_{i,s-1}$$
$$+ \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{\tau}_{st}\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}y_{i,t-1}\overline{y}_{i,-1}. \tag{A-12}$$

The first and second term can be written as

$$\frac{1}{T}\sum_{t=1}^{T}\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left[y_{i,t-1}^2-y_{i,t-1}\overline{y}_{i,-1}\right] = \frac{1}{T}\sum_{t=1}^{T}\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left[y_{i,t-1}^+-\overline{y}_{i,-1}^+\right]^2+\gamma_i^2\left[F_{t-1}^+-\overline{F}_{-1}^+\right]^2,$$
$$= \frac{\sigma_\varepsilon^2}{T}\frac{T-1}{1-\rho^2}\left[1-\frac{1}{T-1}\frac{2\rho}{1-\rho}\left[1-\frac{1}{T}\frac{1-\rho^T}{1-\rho}\right]\right] + m_\gamma^2\frac{1}{T}\sum_{t=1}^{T}\left[F_{t-1}^+-\overline{F}_{-1}^+\right]^2,$$
$$= \frac{\sigma_\varepsilon^2}{T}B(\rho,T) + m_\gamma^2\frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{F}_{t-1}^+\right)^2, \tag{A-13}$$

and

$$\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}y_{i,t-1}y_{i,s-1} = \operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left(y_{i,t-1}^+ + \gamma_i F_{t-1}^+\right)\left(y_{i,s-1}^+ + \gamma_i F_{s-1}^+\right),$$

$$
\begin{aligned}
&= \operatorname*{plim}_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} y^{+}_{i,t-1}y^{+}_{i,s-1} + \operatorname*{plim}_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \gamma_i^2 F^{+}_{t-1}F^{+}_{s-1}, \\
&= \frac{\rho^{|t-s|}}{1-\rho^2}\sigma_\varepsilon^2 + m_\gamma^2 F^{+}_{t-1}F^{+}_{s-1}. \qquad \text{(A-14)}
\end{aligned}
$$

The last term in (A-12) can be dropped since

$$
\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{\tau}_{st} \operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} y_{i,t-1}\overline{y}_{i,-1} = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\widetilde{F}_t\sum_{s=1}^{T}\widetilde{F}_s}{\sum_{t=1}^{T}\widetilde{F}_t^2}\right)\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} y_{i,t-1}\overline{y}_{i,-1} = 0.
$$

Thus, equation (A-12) can be written as

$$
\begin{aligned}
&\operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N} y'_{i,-1}M_{\dot{F}}y_{i,-1} \\
&\quad= \frac{\sigma_\varepsilon^2}{T}B(\rho,T) + m_\gamma^2\frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{F}^{+}_{t-1}\right)^2 - \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{\tau}_{ts}\left(\frac{\rho^{|t-s|}}{1-\rho^2}\sigma_\varepsilon^2 + m_\gamma^2F^{+}_{t-1}F^{+}_{s-1}\right), \\
&\quad= \frac{\sigma_\varepsilon^2}{1-\rho^2}\left(\frac{1-\rho^2}{T}B(\rho,T) - \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{\tau}_{ts}\rho^{|t-s|}\right) + m_\gamma^2\left(\frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{F}^{+}_{t-1}\right)^2 - \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\widetilde{\tau}_{st}F^{+}_{t-1}F^{+}_{s-1}\right), \\
&\quad= \frac{\sigma_\varepsilon^2}{1-\rho^2}\left(\frac{1-\rho^2}{T}B(\rho,T) - \left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\tau}_{tt} + 2\rho\frac{1}{T}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\widetilde{\tau}_{t,t-s}\rho^{s-1}\right)\right) + \\
&\qquad\quad m_\gamma^2\frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{F}^{+}_{t-1}\right)^2\left(1 - \frac{\sum_{t=1}^{T}\widetilde{F}_tF^{+}_{t-1}\sum_{s=1}^{T}\widetilde{F}_s\left(F^{+}_{s-1}\right)}{\sum_{t=1}^{T}\widetilde{F}_t^2\sum_{t=1}^{T}\left(\widetilde{F}^{+}_{t-1}\right)^2}\right), \\
&\quad= \frac{\sigma_\varepsilon^2}{1-\rho^2}\left(\frac{1-\rho^2}{T}B(\rho,T) - \frac{1}{T}\left(1 + 2\rho\sum_{t=1}^{T-1}\rho^{t-1}\widetilde{g}_{F,t}\right)\right) + m_\gamma^2\widetilde{h}_F, \qquad \text{(A-15)}
\end{aligned}
$$

where

$$
\widetilde{h}_F = \frac{1}{T}\sum_{t=1}^{T}\left(\widetilde{F}^{+}_{t-1}\right)^2\left(1 - \frac{\left(\sum_{t=1}^{T}\widetilde{F}_tF^{+}_{t-1}\right)^2}{\sum_{t=1}^{T}\widetilde{F}_t^2\sum_{t=1}^{T}\left(\widetilde{F}^{+}_{t-1}\right)^2}\right).
$$

Dividing (A-11) by (A-15) yields the result in equation (10).

Next, letting $T\to\infty$

$$
\begin{aligned}
\widetilde{g}_{F,t} &= \frac{\frac{T-t}{T}\left(E\left(F_sF_{s-t}\right) + O_p\left(\frac{1}{\sqrt{T}}\right)\right)}{E\left(F_t^2\right) + O_p\left(\frac{1}{\sqrt{T}}\right)} = \frac{T-t}{T}\theta^t + O_p\left(\frac{1}{\sqrt{T}}\right), \\
\widetilde{h}_F &= E\left(\left(F^{+}_{t-1}\right)^2\right)\left(1 - \frac{E\left(F_tF^{+}_{t-1}\right)^2}{E\left(F_t^2\right)E\left(\left(F^{+}_{t-1}\right)^2\right)}\right) + O_p\left(\frac{1}{\sqrt{T}}\right), \\
&= \frac{1}{\left(1-\theta\rho\right)^2\left(1-\rho^2\right)}\sigma_\mu^2 + O_p\left(\frac{1}{\sqrt{T}}\right),
\end{aligned}
$$

where use is made of lemma A-1. Hence, as $T \to \infty$,

$$\sum_{t=1}^{T-1} \rho^{t-1} \widetilde{g}_{F,t} = \sum_{t=1}^{T-1} \rho^{t-1} \left(\frac{T-t}{T} \theta^t + O_p \left(\frac{1}{\sqrt{T}} \right) \right) = \theta \sum_{t=1}^{T-1} (\rho\theta)^{t-1} - \frac{1}{\rho} \frac{1}{T} \sum_{t=1}^{T-1} t (\rho\theta)^t + O_p \left(\frac{1}{\sqrt{T}} \right),$$

$$= \theta \frac{1-(\theta\rho)^{T-1}}{1-\theta\rho} - \frac{1}{\rho} \frac{1}{T} \left(\theta\rho \frac{1-(\theta\rho)^{T-1}}{(1-\theta\rho)^2} - (T-1) \frac{(\theta\rho)^T}{1-\theta\rho} \right) + O_p \left(\frac{1}{\sqrt{T}} \right),$$

$$= \frac{\theta}{1-\theta\rho} \left[1 - \frac{1}{T} \frac{1-\theta^T \rho^T}{1-\theta\rho} \right] + O_p \left(\frac{1}{\sqrt{T}} \right) = \frac{\theta}{1-\theta\rho} + O_p \left(\frac{1}{\sqrt{T}} \right). \tag{A-16}$$

Thus, taking limits as $N \to \infty$ followed by an expansion as $T \to \infty$, equation (10) is given by

$$\operatorname*{plim}_{N\to\infty} (\widehat{\rho}_{\text{WGi}} - \rho) = \frac{1}{T} \frac{-A(\rho,T) - \frac{\theta}{1-\theta\rho} + o_p(1)}{\frac{1}{1-\rho^2} - \frac{1}{T}\frac{1}{1-\rho^2} \left(1 + 2\rho \frac{\theta}{1-\theta\rho} \right) + \frac{1}{(1-\theta\rho)^2 (1-\rho^2)} \frac{m_\gamma^2 \sigma_\mu^2}{\sigma_\varepsilon^2} + o_p(1)},$$

$$= -\frac{1}{T} \left(1 + \rho + \frac{\theta (1-\rho^2)}{1-\theta\rho} \right) \left(1 + \frac{m_\gamma^2}{(1-\theta\rho)^2} \frac{\sigma_\mu^2}{\sigma_\varepsilon^2} \right)^{-1} + o_p \left(\frac{1}{T} \right). \tag{A-17}$$

Proof of Theorem 1. Using (14) and (15), the probability limit as $N \to \infty$ of the numerator of equation (17) equals

$$\operatorname*{plim}_{N\to\infty} \frac{1}{NT} \sum_{i=1}^{N} y_{i,-1}' M_{\dot{G}} \varepsilon_i^* = \operatorname*{plim}_{N\to\infty} \frac{1}{NT} \sum_{i=1}^{N} y_{i,-1}' M_{\dot{H}} \varepsilon_i, \tag{A-18}$$

with $M_{\dot{H}} = I_T - \dot{H} \left(\dot{H}' \dot{H} \right)^{-1} \dot{H}'$, $\dot{H} = \left(\iota, F^+, F_{-1}^+ \right)$. Similar to the derivation of (A-11), (A-18) rewrites to

$$\operatorname*{plim}_{N\to\infty} \frac{1}{NT} \sum_{i=1}^{N} y_{i,-1}' M_{\dot{H}} \varepsilon_i = \operatorname*{plim}_{N\to\infty} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} y_{i,t-1} \left[(\varepsilon_{it} - \overline{\varepsilon}_i) - \sum_{s=1}^{T} \widetilde{\tau}_{st}^+ (\varepsilon_{is} - \overline{\varepsilon}_i) \right],$$

$$= -\frac{1}{T} \sum_{t=1}^{T} \operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} y_{i,t-1} \overline{\varepsilon}_i - \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} \widetilde{\tau}_{st}^+ \operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} y_{i,t-1} \varepsilon_{is},$$

$$= -\frac{\sigma_\varepsilon^2}{T} A(\rho,T) - \sigma_\varepsilon^2 \frac{1}{T} \sum_{t=1}^{T-1} \rho^{t-1} \widetilde{g}_{F,t}^+, \tag{A-19}$$

where $\widetilde{g}_{F,t}^+ = \sum_{s=t+1}^{T} \widetilde{\tau}_{s,s-t}^+$ and

$$\widetilde{\tau}_{st}^+ = \frac{1}{\widetilde{\alpha}_0} \left(\widetilde{\alpha}_1 \widetilde{F}_t^+ \widetilde{F}_s^+ - \widetilde{\alpha}_2 \widetilde{F}_t^+ \widetilde{F}_{s-1}^+ + \widetilde{\alpha}_3 \widetilde{F}_{t-1}^+ \widetilde{F}_{s-1}^+ - \widetilde{\alpha}_2 \widetilde{F}_{t-1}^+ \widetilde{F}_s^+ \right), \tag{A-20}$$

with $\widetilde{\alpha}_0 = \widetilde{\alpha}_1 \widetilde{\alpha}_3 - \widetilde{\alpha}_2^2$, $\widetilde{\alpha}_1 = \sum_{t=1}^{T} \left(\widetilde{F}_{t-1}^+ \right)^2$, $\widetilde{\alpha}_2 = \sum_{t=1}^{T} \widetilde{F}_t^+ \widetilde{F}_{t-1}^+$ and $\widetilde{\alpha}_3 = \sum_{t=1}^{T} \left(\widetilde{F}_t^+ \right)^2$.

Second, as (15) implies that for $N \to \infty$ $M_{\dot{G}} F^+ = 0$ such that $M_{\dot{G}} y_{i,-1} = M_{\dot{G}} y_{i,-1}^+$, the probability limit as $N \to \infty$ of the denominator of equation (17) is given by

$$
\begin{aligned}
\operatorname*{plim}_{N\to\infty} \frac{1}{NT} \sum_{i=1}^{N} y_{i,-1}' M_{\dot{G}} y_{i,-1} &= \operatorname*{plim}_{N\to\infty} \frac{1}{NT} \sum_{i=1}^{N} y_{i,-1}^{+'} M_{\dot{G}} y_{i,-1}^{+}, \\
&= \operatorname*{plim}_{N\to\infty} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} y_{i,t-1}^{+} \left((y_{i,t-1}^{+} - \overline{y}_{i,-1}^{+}) - \sum_{s=1}^{T} \widetilde{\tau}_{st}^{+} (y_{i,s-1}^{+} - \overline{y}_{i,-1}^{+}) \right), \\
&= \operatorname*{plim}_{N\to\infty} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(y_{i,t-1}^{+} - \overline{y}_{i,-1}^{+} \right)^2 - \operatorname*{plim}_{N\to\infty} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \widetilde{\tau}_{st}^{+} y_{i,t-1}^{+} y_{i,s-1}^{+} \\
&\qquad + \operatorname*{plim}_{N\to\infty} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \widetilde{\tau}_{st}^{+} y_{i,t-1}^{+} \overline{y}_{i,-1}^{+}, \\
&= \frac{1}{T} \sum_{t=1}^{T} \operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \left(y_{i,t-1}^{+} - \overline{y}_{i,-1}^{+} \right)^2 - \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} \widetilde{\tau}_{st}^{+} \operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} y_{i,t-1}^{+} y_{i,s-1}^{+}, \\
&= \frac{\sigma_{\varepsilon}^2}{T} \frac{T-1}{1-\rho^2} \left[1 - \frac{1}{T-1} \frac{2\rho}{1-\rho} \left[1 - \frac{1}{T} \frac{1-\rho^T}{1-\rho} \right] \right] - \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} \widetilde{\tau}_{ts}^{+} \left(\frac{\rho^{|t-s|}}{1-\rho^2} \sigma_{\varepsilon}^2 \right), \\
&= \frac{\sigma_{\epsilon}^2}{1-\rho^2} \left(\frac{1-\rho^2}{T} B(\rho, T) - \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} \widetilde{\tau}_{ts}^{+} \rho^{|t-s|} \right), \\
&= \frac{\sigma_{\epsilon}^2}{1-\rho^2} \left(\frac{1-\rho^2}{T} B(\rho, T) - \left(\frac{1}{T} \sum_{t=1}^{T} \widetilde{\tau}_{tt}^{+} + 2\rho \frac{1}{T} \sum_{t=2}^{T} \sum_{s=1}^{t-1} \widetilde{\tau}_{t,t-s}^{+} \rho^{s-1} \right) \right), \\
&= \frac{\sigma_{\epsilon}^2}{1-\rho^2} \left(\frac{1-\rho^2}{T} B(\rho, T) - \frac{2}{T} \left(1 + \rho \sum_{t=1}^{T-1} \rho^{t-1} \widetilde{g}_{F,t}^{+} \right) \right). \qquad \text{(A-21)}
\end{aligned}
$$

Dividing (A-19) by (A-21) yields the result in equation (18).

Turning to the expression for $T \to \infty$, first note that using Lemma A-1

$$
\frac{1}{T} \sum_{s=t+1}^{T} \widetilde{F}_s^{+} \widetilde{F}_{s-t}^{+} = \frac{T-t}{T} E(F_s^{+} F_{s-t}^{+}) + O_p\left(\frac{1}{\sqrt{T}} \right) = \frac{T-t}{T} \lambda_t + O_p\left(\frac{1}{\sqrt{T}} \right),
$$

such that

$$
\begin{aligned}
\widetilde{g}_{F,t}^{+} &= \frac{1}{\widetilde{\alpha}_0} \left(\widetilde{\alpha}_1 \sum_{s=t+1}^{T} \widetilde{F}_s^{+} \widetilde{F}_{s-t}^{+} - \widetilde{\alpha}_2 \sum_{s=t+1}^{T} \widetilde{F}_s^{+} \widetilde{F}_{s-t-1}^{+} + \widetilde{\alpha}_3 \sum_{s=t+1}^{T} \widetilde{F}_{s-1}^{+} \widetilde{F}_{s-t-1}^{+} - \widetilde{\alpha}_2 \sum_{s=t+1}^{T} \widetilde{F}_{s-1}^{+} \widetilde{F}_{s-t}^{+} \right), \\
&= \frac{T-t}{T} \frac{2\omega_t - \omega_1 \omega_{t+1} - \omega_1 \omega_{t-1}}{1-\omega_1^2} + O_p\left(\frac{1}{\sqrt{T}} \right),
\end{aligned}
$$

with $\omega_t = \lambda_t / \lambda_0$. Next

$$
\begin{aligned}
-\frac{1}{T} \sum_{t=1}^{T-1} \rho^{t-1} \widetilde{g}_{F,t}^{+} &= -\frac{1}{T} \sum_{t=1}^{T-1} \rho^{t-1} \left(\frac{T-t}{T} \frac{2\omega_t - \omega_1 \omega_{t+1} - \omega_1 \omega_{t-1}}{1-\omega_1^2} + O_p\left(\frac{1}{\sqrt{T}} \right) \right), \\
&= -\frac{(1+\theta\rho)^2}{(1-\theta^2)(1-\rho^2)} \frac{1}{T} \sum_{t=1}^{T-1} \rho^{t-1} \left(1 - \frac{t}{T} \right) \left((2 - (\theta+\rho)\,\omega_1)\,\omega_t - (1-\theta\rho)\,\omega_1 \omega_{t-1} \right) + o_p\left(\frac{1}{T} \right),
\end{aligned}
$$

$$= -\left(1+\theta\rho\right)\frac{1}{T}\sum_{t=1}^{T-1}\rho^{t-1}\left(1-\frac{t}{T}\right)\left(\frac{\omega_t-\rho\omega_{t-1}}{1-\rho^2}+\frac{\omega_t-\theta\omega_{t-1}}{1-\theta^2}\right)+o_p\left(\frac{1}{T}\right),$$

$$= -\frac{1}{T}\left(\frac{\theta}{1-\theta\rho}+\frac{\rho}{1-\rho^2}\right)+o_p\left(\frac{1}{T}\right), \tag{A-22}$$

where use is made of $\omega_t=(\theta+\rho)\,\omega_{t-1}-\theta\rho\omega_{t-2}\ \forall t\geq 2$ and

$$\begin{aligned}
\sum_{t=2}^{T-1}\left(1-\frac{t}{T}\right)\rho^{t-1}\omega_t &= (\theta+\rho)\sum_{t=2}^{T-1}\left(1-\frac{t}{T}\right)\rho^{t-1}\omega_{t-1}-\theta\rho\sum_{t=2}^{T-1}\left(1-\frac{t}{T}\right)\rho^{t-1}\omega_{t-2},\\
&= (\theta+\rho)\left(\left(1-\frac{2}{T}\right)\rho\omega_1+\rho\left(\sum_{t=2}^{T-1}\left(1-\frac{t}{T}\right)\rho^{t-1}\omega_t-\frac{1}{T}\sum_{t=2}^{T-1}\rho^{t-1}\omega_t\right)\right),\\
&\quad -\theta\rho\left(\left(1-\frac{2}{T}\right)\rho+\left(1-\frac{3}{T}\right)\rho^2\omega_1+\rho^2\left(\sum_{t=2}^{T-1}\left(1-\frac{t}{T}\right)\rho^{t-1}\omega_t-\frac{2}{T}\sum_{t=2}^{T-1}\rho^{t-1}\omega_t\right)\right),\\
&= (\theta+\rho)\left(\rho\omega_1+\rho\sum_{t=2}^{T-1}\left(1-\frac{t}{T}\right)\rho^{t-1}\omega_t\right),\\
&\qquad -\theta\rho\left(\rho+\rho^2\omega_1+\rho^2\sum_{t=2}^{T-1}\left(1-\frac{t}{T}\right)\rho^{t-1}\omega_t\right)+O\left(\frac{1}{T}\right),\\
&= \frac{(\theta+\rho)\,\rho\omega_1-\theta\rho\left(\rho+\rho^2\omega_1\right)}{1-(\theta+\rho)\,\rho+\theta\rho^3}+O\left(\frac{1}{T}\right),\\
&= \rho\frac{\theta\left(1-\rho^2\right)(\theta+\rho)+\rho^2\left(1-\theta^2\right)}{\left(1-\rho^2\right)\left(1-\theta^2\rho^2\right)}+O\left(\frac{1}{T}\right),
\end{aligned}$$

such that

$$\sum_{t=1}^{T-1}\left(1-\frac{t}{T}\right)\rho^{t-1}\omega_t=\omega_1+\sum_{t=2}^{T-1}\left(1-\frac{t}{T}\right)\rho^{t-1}\omega_t=\frac{\theta\left(1-\rho^2\right)+\rho\left(1-\theta^2\rho^2\right)}{\left(1-\rho^2\right)\left(1-\theta^2\rho^2\right)}+O\left(\frac{1}{T}\right),$$

and

$$\begin{aligned}
\sum_{t=1}^{T-1}\left(1-\frac{t}{T}\right)\rho^{t-1}\omega_{t-1} &= 1+\rho\sum_{t=1}^{T-1}\left(1-\frac{t}{T}\right)\rho^{t-1}\omega_t-\frac{1}{T}\sum_{t=1}^{T}\rho^{t-1}\omega_{t-1},\\
&= 1+\rho\frac{\omega_1-\theta\rho^2}{\left(1-\rho^2\right)(1-\theta\rho)}+O\left(\frac{1}{T}\right)=\frac{\left(1-\theta^2\rho^2\right)+\theta\rho\left(1-\rho^2\right)}{\left(1-\rho^2\right)\left(1-\theta^2\rho^2\right)}+O\left(\frac{1}{T}\right).
\end{aligned}$$

Using (A-22), as $T\to\infty$ (A-19) and (A-21) are given by

$$-\frac{\sigma_\varepsilon^2}{T}A(\rho,T)-\sigma_\varepsilon^2\frac{1}{T}\sum_{t=1}^{T-1}\rho^{t-1}\widetilde{g}_{F,t}^{+}=-\frac{\sigma_\varepsilon^2}{T}\left(A(\rho,T)+\left(\frac{\theta}{1-\theta\rho}+\frac{\rho}{1-\rho^2}\right)\right)+o_p\left(\frac{1}{T}\right), \tag{A-23}$$

$$\frac{\sigma_\epsilon^2}{1-\rho^2}\left(\frac{1-\rho^2}{T}B(\rho,T)-\frac{2}{T}\left(1+\rho\sum_{t=1}^{T-1}\rho^{t-1}\widetilde{g}_{F,t}^{+}\right)\right)=\frac{\sigma_\epsilon^2}{T}B(\rho,T)+O\left(\frac{1}{T}\right). \tag{A-24}$$

Thus taking limits letting $N \to \infty$, followed by $T \to \infty$, equation (19) is given by

$$
\begin{aligned}
\operatorname*{plim}_{N\to\infty} \left(\widehat{\rho}_{\text{CCEPu}} - \rho\right) &= -\frac{1}{T}\left(\frac{1}{1-\rho} + \left(\frac{\theta}{1-\theta\rho} + \frac{\rho}{(1-\rho^2)}\right)\right)\left(1-\rho^2\right) + o_p\left(\frac{1}{T}\right), \\
&= -\frac{1}{T}\left(1 + 2\rho + \frac{\theta(1-\rho^2)}{1-\theta\rho}\right) + o_p\left(\frac{1}{T}\right). \qquad \text{(A-25)}
\end{aligned}
$$

Proof of Theorem 2. Letting $\left(\rho^0, F^0\right)$ denote the true parameter ρ and the true factor F respectively such that, after centering, the objective function in (20) is given by

$$
\begin{aligned}
S_{NT}(\rho, F) &= \frac{1}{NT}\sum_{i=1}^{N}\left(\widetilde{y}_i - \rho\widetilde{y}_{i,-1}\right)' M_{\widetilde{F}}\left(\widetilde{y}_i - \rho\widetilde{y}_{i,-1}\right) - \frac{1}{NT}\sum_{i=1}^{N}\widetilde{\varepsilon}_i' M_{\widetilde{F}^0}\widetilde{\varepsilon}_i, \\
&= \frac{1}{NT}\sum_{i=1}^{N}\left(\left(\rho^0-\rho\right)\widetilde{y}_{i,-1} + \lambda_i\widetilde{F}^0 + \widetilde{\varepsilon}_i\right)' M_{\widetilde{F}}\left(\left(\rho^0-\rho\right)\widetilde{y}_{i,-1} + \lambda_i\widetilde{F}^0 + \widetilde{\varepsilon}_i\right) - \frac{1}{NT}\sum_{i=1}^{N}\widetilde{\varepsilon}_i' M_{\widetilde{F}^0}\widetilde{\varepsilon}_i, \\
&= \breve{S}_{NT}(\rho, F) + 2\frac{\left(\rho^0-\rho\right)}{NT}\sum_{i=1}^{N}\widetilde{y}_{i,-1}' M_{\widetilde{F}}\widetilde{\varepsilon}_i + 2\frac{1}{NT}\sum_{i=1}^{N}\lambda_i\widetilde{F}^{0\prime} M_{\widetilde{F}}\widetilde{\varepsilon}_i + \frac{1}{NT}\sum_{i=1}^{N}\widetilde{\varepsilon}_i'\left(M_{\widetilde{F}} - M_{\widetilde{F}^0}\right)\widetilde{\varepsilon}_i,
\end{aligned}
$$

where

$$
\breve{S}_{NT}(\rho, F) = \frac{\left(\rho^0-\rho\right)^2}{NT}\sum_{i=1}^{N}\widetilde{y}_{i,-1}' M_{\widetilde{F}}\widetilde{y}_{i,-1} + \frac{1}{NT}\sum_{i=1}^{N}\lambda_i^2\widetilde{F}^{0\prime} M_{\widetilde{F}}\widetilde{F}^0 + 2\frac{\left(\rho^0-\rho\right)}{NT}\sum_{i=1}^{N}\widetilde{y}_{i,-1}' M_{\widetilde{F}}\widetilde{F}^0\lambda_i.
$$

Using that for $N, T \to \infty$

$$
\begin{aligned}
\frac{1}{NT}\sum_{i=1}^{N} y_{i,-1}' M_{\widetilde{F}}\widetilde{\varepsilon}_i &= \frac{1}{NT}\sum_{i=1}^{N} y_{i,-1}'\widetilde{\varepsilon}_i - \frac{1}{N}\sum_{i=1}^{N}\frac{y_{i,-1}'\widetilde{F}}{T}\left(\frac{\widetilde{F}'\widetilde{F}}{T}\right)^{-1}\frac{\widetilde{F}'\widetilde{\varepsilon}_i}{T} = o_p(1), \\
\frac{1}{NT}\sum_{i=1}^{N}\lambda_i F^{0\prime} M_{\widetilde{F}}\widetilde{\varepsilon}_i &= \frac{1}{NT}\sum_{i=1}^{N}\lambda_i F^{0\prime}\widetilde{\varepsilon}_i - \frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_i F^{0\prime}\widetilde{F}}{T}\left(\frac{\widetilde{F}'\widetilde{F}}{T}\right)^{-1}\frac{\widetilde{F}'\widetilde{\varepsilon}_i}{T} = o_p(1), \\
\frac{1}{NT}\sum_{i=1}^{N}\varepsilon_i'\left(M_{\widetilde{F}} - M_{\widetilde{F}^0}\right)\widetilde{\varepsilon}_i &= \frac{1}{N}\sum_{i=1}^{N}\frac{\varepsilon_i'\widetilde{F}^0}{T}\left(\frac{\widetilde{F}^{0\prime}\widetilde{F}^0}{T}\right)^{-1}\frac{\widetilde{F}^{0\prime}\widetilde{\varepsilon}_i}{T} - \frac{1}{N}\sum_{i=1}^{N}\frac{\varepsilon_i'\widetilde{F}}{T}\left(\frac{\widetilde{F}'\widetilde{F}}{T}\right)^{-1}\frac{\widetilde{F}'\widetilde{\varepsilon}_i}{T} = o_p(1).
\end{aligned}
$$

we have

$$
S_{NT}(\rho, F) = \breve{S}_{NT}(\rho, F) + o_p(1), \qquad \text{(A-26)}
$$

uniformly over ρ and F.

First note that as $M_{\widetilde{F}^0}\widetilde{F}^0 = 0$, $\breve{S}_{NT}\left(\rho^0, F^0\right) = 0$. Second, we show that for any $(\rho, F) \neq \left(\rho^0, F^0\right)$, $\breve{S}_{NT}(\rho, F) > 0$; thus $\breve{S}_{NT}\left(\rho^0, F^0\right)$ attains its unique minimum value at $(\rho, F) = \left(\rho^0, F^0\right)$. Define

$$
A = \frac{1}{NT}\sum_{i=1}^{N}\widetilde{y}_{i,-1}' M_{\widetilde{F}}\widetilde{y}_{i,-1}; \quad B = \frac{1}{NT}\sum_{i=1}^{N}\lambda_i^2; \quad C = \frac{1}{NT}\sum_{i=1}^{N}\lambda_i M_{\widetilde{F}}\widetilde{y}_{i,-1}.
$$

Then

$$\begin{aligned}\breve{S}_{NT}(\rho,F) &= \left(\rho^0-\rho\right)^2 A + \widetilde{F}^{0\prime} M_{\widetilde{F}} B M_{\widetilde{F}} \widetilde{F}^0 + 2\left(\rho^0-\rho\right) C' M_{\widetilde{F}} \widetilde{F}^0, \\ &= \left(\rho^0-\rho\right)^2 \left(A - C'B^{-1}C\right) + \left(\widetilde{F}^{0\prime} M_{\widetilde{F}} + \left(\rho^0-\rho\right) C'B^{-1}\right) B \left(M_{\widetilde{F}} \widetilde{F}^0 + B^{-1}C\left(\rho^0-\rho\right)\right), \\ &= \left(\rho^0-\rho\right)^2 D\left(\widetilde{F}\right) + \theta' B \theta, \\ &\geq 0,\end{aligned}$$

since $D(F) = A - C'B^{-1}C$ and B are both positive definite, where $\theta = M_{\widetilde{F}}\widetilde{F}^0 + B^{-1}C\left(\rho^0-\rho\right)$. Note that $\breve{S}_{NT}(\rho,F) > 0$ if either $\rho \neq \rho^0$ or $F \neq F^0$. This implies that $\widehat{\rho}_{\text{CCEPr}}$ is consistent for ρ.

Proof of Theorem 3. First note that using (12) and (23)

$$\begin{aligned}\widehat{F}_t &= \frac{1}{\overline{\gamma}}\left(\overline{y}_t - \widehat{\overline{\alpha}} - \widehat{\rho}_{\text{CCEPr}}\overline{y}_{t-1}\right) = \frac{1}{\overline{\gamma}}\left((\rho - \widehat{\rho}_{\text{CCEPr}})\overline{y}_{t-1} + (\overline{\alpha} - \widehat{\overline{\alpha}}) + \overline{\gamma}F_t + \overline{\varepsilon}_t\right), \\ F_t &= \widehat{F}_t + \frac{1}{\overline{\gamma}}\left((\widehat{\rho}_{\text{CCEPr}} - \rho)\overline{y}_{t-1} + (\widehat{\overline{\alpha}} - \overline{\alpha}) - \overline{\varepsilon}_t\right),\end{aligned}$$

which for $N \to \infty$ and using (14) and (15) reduces to

$$\begin{aligned}F_t &= \widehat{F}_t + \frac{1}{\overline{\gamma}}\left((\widehat{\rho}_{\text{CCEPr}} - \rho)\left((1-\rho)^{-1}\overline{\alpha} + \overline{\gamma}F^+_{t-1}\right) + (\widehat{\overline{\alpha}} - \overline{\alpha})\right), \\ &= \widehat{F}_t + (\widehat{\rho}_{\text{CCEPr}} - \rho)F^+_{t-1} + a^*,\end{aligned} \tag{A-27}$$

where $a^* = \frac{1}{\overline{\gamma}}\left((\widehat{\rho}_{\text{CCEPr}} - \rho)(1-\rho)^{-1}\overline{\alpha} + (\widehat{\overline{\alpha}} - \overline{\alpha})\right)$. Using (A-27) in (22)

$$\begin{aligned}\operatorname*{plim}_{N\to\infty}(\widehat{\rho}_{\text{CCEPr}} - \rho) &= \frac{\operatorname*{plim}_{N\to\infty}(1/NT)\sum_{i=1}^N y'_{i,-1} M_{\widehat{F}}\left(\gamma_i \operatorname*{plim}_{N\to\infty}(\widehat{\rho}_{\text{CCEPr}} - \rho)F^+_{-1} + \varepsilon_i\right)}{\operatorname*{plim}_{N\to\infty}(1/NT)\sum_{i=1}^N y'_{i,-1} M_{\widehat{F}} y_{i,-1}}, \\ &= \frac{\operatorname*{plim}_{N\to\infty}(1/NT)\sum_{i=1}^N y'_{i,-1} M_{\widehat{F}}\varepsilon_i}{\operatorname*{plim}_{N\to\infty}(1/NT)\sum_{i=1}^N y'_{i,-1} M_{\widehat{F}} y^+_{i,-1}}.\end{aligned} \tag{A-28}$$

Using (A-7)-(A-10), the numerator of (A-28) is given by

$$\begin{aligned}\operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^N y'_{i,-1} M_{\widehat{F}}\varepsilon_i &= \operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T y_{i,t-1}\left[(\varepsilon_{it} - \overline{\varepsilon}_i) - \sum_{s=1}^T \widehat{\overline{\tau}}_{st}(\varepsilon_{is} - \overline{\varepsilon}_i)\right], \\ &= \frac{1}{T}\sum_{t=1}^T \operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^N y_{i,t-1}\varepsilon_{it} - \frac{1}{T}\sum_{t=1}^T\left(1 - \sum_{s=1}^T \widehat{\overline{\tau}}_{st}\right)\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^N y_{i,t-1}\overline{\varepsilon}_i \\ &\quad - \frac{1}{T}\sum_{t=1}^T\sum_{s=1}^T \widehat{\overline{\tau}}_{st}\operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^N y_{i,t-1}\varepsilon_{is},\end{aligned}$$

$$= -\frac{1}{T}\sum_{t=1}^{T} \operatorname*{plim}_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} y_{i,t-1}\overline{\varepsilon}_i - \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\widehat{\widetilde{\tau}}_{st} \operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} y_{i,t-1}\varepsilon_{is},$$

$$= -\frac{\sigma_\varepsilon^2}{T}A(\rho,T) - \sigma_\varepsilon^2\frac{1}{T}\sum_{t=1}^{T-1}\rho^{t-1}\widetilde{g}_{\widehat{F},t}, \tag{A-29}$$

where $\widetilde{g}_{\widehat{F},t} = \sum_{s=t+1}^{T}\widehat{\widetilde{\tau}}_{s,s-t} = \sum_{s=t+1}^{T}\left(\widehat{F}_s - \overline{\widehat{F}}\right)\left(\widehat{F}_{s-t} - \overline{\widehat{F}}\right)\Big/\sum_{t=1}^{T}\left(\widehat{F}_t - \overline{\widehat{F}}\right)$.

Similarly to the derivation of (A-15), the denominator of equation (A-28) is given by

$$\begin{aligned}
&\operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N} y'_{i,-1}M_{\widehat{F}}y^+_{i,-1}\\
&\quad= \operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[y^+_{i,t-1} + \overline{\gamma}F^+_{t-1}\right]\left[\left(y^+_{i,t-1} - \overline{y}^+_{i,-1}\right) - \sum_{s=1}^{T}\widehat{\widetilde{\tau}}_{st}\left(y^+_{i,s-1} - \overline{y}^+_i\right)\right],\\
&\quad= \operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(y^+_{i,t-1} - \overline{y}^+_{i,-1}\right)^2 - \operatorname*{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\widehat{\widetilde{\tau}}_{st}y^+_{i,t-1}y^+_{i,s-1},\\
&\quad= \frac{\sigma_\epsilon^2}{1-\rho^2}\left(\frac{1-\rho^2}{T}B(\rho,T) - \frac{2}{T}\left(1 + \rho\sum_{t=1}^{T-1}\rho^{t-1}\widetilde{g}_{\widehat{F},t}\right)\right).
\end{aligned} \tag{A-30}$$

Deviding (A-29) by (A-30) yields the result in equation (25).

Next, letting $T\to\infty$, from (A-27) and theorem 2 follows

$$\widehat{F}_t = F_t + o_p(1). \tag{A-31}$$

Thus,

$$\widetilde{g}_{\widehat{F},t} = \frac{\sum_{s=1}^{T}\left(F_s - \overline{F}\right)\left(F_{s-t} - \overline{F}\right)}{\sum_{t=1}^{T}\left(F_t - \overline{F}\right)^2} + o_p(1) = \widetilde{g}_{F,t} + o_p(1). \tag{A-32}$$

The asymptotic representation of the CCEPr for $(N,T)_{\text{seq}}\to\infty$ then follows from substituting (A-31) in (25) and using similar derivations as to obtain (11).

Appendix B Tables Monte Carlo simulations

Table 1: $\rho = 0.6$, $\beta = 0.4$, $\theta = 0.4$, $\psi_1 = 1$, $\psi_2 = 1$, $\psi_3 = 0.8$ and $\xi = 3$

T	N		bias ρ	stde ρ	stdv ρ	rmse ρ	bias β	stde β	stdv β	rmse β
5	20	WGn	-0.118	0.072	0.104	0.157	0.004	0.038	0.045	0.045
		WGi	-0.121	0.072	0.089	0.150	-0.019	0.038	0.041	0.045
		CCEPu	-0.355	0.170	0.306	0.469	-0.074	0.069	0.095	0.121
		CCEPr	-0.206	0.077	0.151	0.255	-0.034	0.036	0.051	0.062
10	20	WGn	-0.034	0.041	0.068	0.076	0.022	0.025	0.030	0.037
		WGi	-0.054	0.037	0.040	0.067	-0.005	0.023	0.023	0.024
		CCEPu	-0.118	0.055	0.073	0.139	-0.015	0.027	0.029	0.033
		CCEPr	-0.083	0.042	0.057	0.101	-0.009	0.022	0.025	0.026
20	20	WGn	0.009	0.026	0.048	0.048	0.026	0.018	0.021	0.033
		WGi	-0.025	0.023	0.023	0.035	-0.001	0.015	0.015	0.015
		CCEPu	-0.050	0.030	0.033	0.060	-0.004	0.016	0.017	0.017
		CCEPr	-0.038	0.026	0.030	0.048	-0.003	0.015	0.016	0.016
30	20	WGn	0.022	0.021	0.039	0.045	0.027	0.014	0.017	0.032
		WGi	-0.017	0.018	0.018	0.024	-0.001	0.012	0.012	0.012
		CCEPu	-0.032	0.022	0.023	0.040	-0.002	0.013	0.013	0.013
		CCEPr	-0.025	0.021	0.022	0.033	-0.002	0.012	0.012	0.012
40	20	WGn	0.030	0.018	0.035	0.046	0.027	0.012	0.015	0.031
		WGi	-0.012	0.015	0.015	0.019	-0.000	0.010	0.010	0.010
		CCEPu	-0.023	0.018	0.019	0.030	-0.001	0.011	0.011	0.011
		CCEPr	-0.018	0.017	0.019	0.026	-0.001	0.010	0.011	0.011
50	20	WGn	0.035	0.016	0.031	0.046	0.028	0.011	0.013	0.031
		WGi	-0.010	0.013	0.013	0.016	-0.000	0.009	0.009	0.009
		CCEPu	-0.018	0.016	0.017	0.025	-0.001	0.009	0.010	0.010
		CCEPr	-0.014	0.015	0.016	0.021	-0.001	0.009	0.009	0.009
5	50	WGn	-0.114	0.045	0.090	0.146	0.006	0.024	0.032	0.033
		WGi	-0.120	0.045	0.069	0.139	-0.018	0.024	0.027	0.032
		CCEPu	-0.358	0.105	0.267	0.446	-0.074	0.043	0.075	0.105
		CCEPr	-0.203	0.049	0.115	0.234	-0.033	0.023	0.034	0.047
10	50	WGn	-0.032	0.026	0.059	0.067	0.022	0.016	0.022	0.031
		WGi	-0.052	0.023	0.026	0.059	-0.005	0.014	0.015	0.016
		CCEPu	-0.119	0.035	0.055	0.131	-0.016	0.017	0.019	0.025
		CCEPr	-0.082	0.027	0.039	0.091	-0.009	0.014	0.016	0.018
20	50	WGn	0.011	0.017	0.042	0.043	0.026	0.011	0.015	0.030
		WGi	-0.025	0.014	0.015	0.029	-0.001	0.010	0.010	0.010
		CCEPu	-0.051	0.019	0.022	0.055	-0.004	0.010	0.010	0.011
		CCEPr	-0.037	0.017	0.019	0.042	-0.002	0.010	0.010	0.010
30	50	WGn	0.023	0.013	0.036	0.043	0.027	0.009	0.012	0.030
		WGi	-0.016	0.011	0.011	0.019	-0.001	0.008	0.008	0.008
		CCEPu	-0.031	0.014	0.015	0.035	-0.002	0.008	0.008	0.008
		CCEPr	-0.023	0.013	0.014	0.027	-0.001	0.008	0.008	0.008
40	50	WGn	0.030	0.011	0.031	0.044	0.027	0.008	0.011	0.029
		WGi	-0.012	0.009	0.009	0.015	-0.001	0.007	0.007	0.007
		CCEPu	-0.023	0.012	0.012	0.026	-0.001	0.007	0.007	0.007
		CCEPr	-0.017	0.011	0.012	0.021	-0.001	0.007	0.007	0.007
50	50	WGn	0.036	0.010	0.029	0.046	0.028	0.007	0.010	0.029
		WGi	-0.009	0.008	0.008	0.012	-0.000	0.006	0.006	0.006
		CCEPu	-0.018	0.010	0.011	0.021	-0.000	0.006	0.006	0.006
		CCEPr	-0.014	0.010	0.010	0.017	-0.001	0.006	0.006	0.006

Table 2: $N = 50$, $\beta = 1 - \rho$, $\theta = 0.4$, $\psi_1 = 1$, $\psi_2 = 1$, $\psi_3 = 0.8$ and $\xi = 3$

ρ	T		bias ρ	stde ρ	stdv ρ	rmse ρ	bias β	stde β	stdv β	rmse β
0.4	5	WGn	-0.058	0.039	0.071	0.091	0.018	0.028	0.037	0.042
		WGi	-0.073	0.038	0.050	0.089	-0.015	0.028	0.030	0.034
		CCEPu	-0.205	0.086	0.187	0.278	-0.067	0.052	0.083	0.106
		CCEPr	-0.106	0.040	0.072	0.128	-0.023	0.027	0.034	0.041
	10	WGn	-0.004	0.024	0.051	0.051	0.033	0.019	0.026	0.042
		WGi	-0.032	0.021	0.023	0.040	-0.004	0.017	0.017	0.017
		CCEPu	-0.062	0.029	0.038	0.073	-0.010	0.019	0.021	0.023
		CCEPr	-0.045	0.023	0.029	0.054	-0.006	0.017	0.017	0.018
	20	WGn	0.024	0.016	0.040	0.047	0.036	0.013	0.019	0.041
		WGi	-0.017	0.013	0.014	0.021	-0.001	0.011	0.011	0.011
		CCEPu	-0.029	0.017	0.018	0.034	-0.002	0.012	0.012	0.012
		CCEPr	-0.022	0.015	0.016	0.028	-0.002	0.011	0.011	0.012
	30	WGn	0.033	0.013	0.033	0.047	0.036	0.010	0.015	0.039
		WGi	-0.011	0.011	0.011	0.015	-0.000	0.009	0.009	0.009
		CCEPu	-0.018	0.013	0.013	0.022	-0.001	0.009	0.009	0.009
		CCEPr	-0.014	0.012	0.013	0.019	-0.001	0.009	0.009	0.009
	40	WGn	0.038	0.011	0.029	0.048	0.037	0.009	0.013	0.039
		WGi	-0.008	0.009	0.009	0.012	-0.000	0.008	0.008	0.008
		CCEPu	-0.013	0.011	0.011	0.017	-0.000	0.008	0.008	0.008
		CCEPr	-0.011	0.010	0.011	0.015	-0.001	0.008	0.008	0.008
	50	WGn	0.041	0.010	0.027	0.049	0.037	0.008	0.012	0.039
		WGi	-0.006	0.008	0.008	0.010	-0.000	0.007	0.007	0.007
		CCEPu	-0.011	0.009	0.010	0.014	-0.000	0.007	0.007	0.007
		CCEPr	-0.009	0.009	0.009	0.013	-0.001	0.007	0.007	0.007
0.8	5	WGn	-0.309	0.059	0.150	0.343	0.024	0.038	0.057	0.062
		WGi	-0.285	0.061	0.132	0.315	-0.025	0.039	0.042	0.049
		CCEPu	-0.794	0.132	0.380	0.880	-0.080	0.057	0.072	0.108
		CCEPr	-0.630	0.071	0.237	0.673	-0.054	0.035	0.045	0.070
	10	WGn	-0.131	0.033	0.092	0.160	0.057	0.026	0.044	0.072
		WGi	-0.125	0.030	0.051	0.135	-0.008	0.025	0.025	0.026
		CCEPu	-0.353	0.050	0.121	0.373	-0.028	0.027	0.030	0.041
		CCEPr	-0.281	0.039	0.101	0.299	-0.020	0.024	0.026	0.032
	20	WGn	-0.038	0.019	0.057	0.068	0.071	0.019	0.031	0.077
		WGi	-0.055	0.017	0.022	0.059	-0.002	0.017	0.017	0.017
		CCEPu	-0.158	0.026	0.047	0.165	-0.009	0.017	0.018	0.020
		CCEPr	-0.121	0.023	0.039	0.127	-0.006	0.016	0.017	0.018
	30	WGn	-0.010	0.015	0.043	0.044	0.073	0.015	0.026	0.077
		WGi	-0.035	0.012	0.015	0.038	-0.001	0.013	0.013	0.013
		CCEPu	-0.098	0.019	0.028	0.102	-0.004	0.014	0.014	0.015
		CCEPr	-0.073	0.017	0.024	0.077	-0.003	0.013	0.013	0.014
	40	WGn	0.007	0.012	0.037	0.037	0.075	0.013	0.023	0.079
		WGi	-0.025	0.010	0.012	0.027	-0.001	0.011	0.012	0.012
		CCEPu	-0.071	0.015	0.021	0.074	-0.002	0.012	0.012	0.012
		CCEPr	-0.053	0.014	0.018	0.056	-0.002	0.011	0.012	0.012
	50	WGn	0.014	0.010	0.032	0.035	0.075	0.012	0.021	0.078
		WGi	-0.019	0.009	0.010	0.022	-0.001	0.010	0.010	0.010
		CCEPu	-0.055	0.013	0.016	0.057	-0.002	0.010	0.010	0.011
		CCEPr	-0.040	0.012	0.015	0.043	-0.002	0.010	0.010	0.010

Table 3: $N = 50$, $\rho = 0.6$, $\beta = 0.4$, $\psi_1 = 1$, $\psi_2 = 1$, $\psi_3 = 0.8$ and $\xi = 3$

θ	T		bias ρ	stde ρ	stdv ρ	rmse ρ	bias β	stde β	stdv β	rmse β
0	5	WGn	-0.144	0.044	0.095	0.173	0.001	0.022	0.028	0.028
		WGi	-0.080	0.039	0.050	0.095	-0.012	0.020	0.021	0.024
		CCEPu	-0.241	0.091	0.217	0.324	-0.049	0.036	0.060	0.077
		CCEPr	-0.122	0.042	0.080	0.146	-0.018	0.019	0.025	0.031
	10	WGn	-0.071	0.026	0.058	0.091	0.016	0.015	0.019	0.025
		WGi	-0.034	0.020	0.022	0.041	-0.003	0.012	0.012	0.012
		CCEPu	-0.071	0.029	0.040	0.081	-0.008	0.014	0.015	0.017
		CCEPr	-0.048	0.023	0.028	0.056	-0.004	0.012	0.012	0.013
	20	WGn	-0.034	0.017	0.040	0.052	0.021	0.010	0.013	0.025
		WGi	-0.017	0.013	0.013	0.021	-0.001	0.008	0.008	0.008
		CCEPu	-0.032	0.016	0.018	0.036	-0.002	0.008	0.009	0.009
		CCEPr	-0.023	0.014	0.015	0.027	-0.001	0.008	0.008	0.008
	30	WGn	-0.023	0.013	0.033	0.040	0.022	0.008	0.010	0.024
		WGi	-0.011	0.010	0.010	0.014	-0.000	0.006	0.006	0.006
		CCEPu	-0.020	0.012	0.013	0.024	-0.001	0.006	0.007	0.007
		CCEPr	-0.015	0.011	0.012	0.019	-0.001	0.006	0.006	0.006
	40	WGn	-0.017	0.011	0.028	0.033	0.022	0.007	0.009	0.024
		WGi	-0.008	0.008	0.008	0.011	-0.000	0.005	0.005	0.005
		CCEPu	-0.015	0.010	0.010	0.018	-0.000	0.005	0.006	0.006
		CCEPr	-0.011	0.010	0.010	0.014	-0.000	0.005	0.005	0.005
	50	WGn	-0.013	0.010	0.025	0.028	0.023	0.006	0.008	0.024
		WGi	-0.006	0.007	0.007	0.010	-0.000	0.005	0.005	0.005
		CCEPu	-0.012	0.009	0.009	0.015	-0.000	0.005	0.005	0.005
		CCEPr	-0.008	0.008	0.009	0.012	-0.000	0.005	0.005	0.005
0.8	5	WGn	-0.140	0.048	0.075	0.159	-0.003	0.028	0.033	0.033
		WGi	-0.214	0.057	0.102	0.237	-0.035	0.032	0.036	0.051
		CCEPu	-0.498	0.117	0.309	0.586	-0.104	0.053	0.091	0.138
		CCEPr	-0.335	0.057	0.144	0.364	-0.057	0.029	0.046	0.073
	10	WGn	-0.036	0.028	0.053	0.064	0.019	0.018	0.024	0.031
		WGi	-0.097	0.030	0.040	0.105	-0.011	0.019	0.020	0.022
		CCEPu	-0.195	0.042	0.077	0.210	-0.027	0.022	0.026	0.038
		CCEPr	-0.145	0.031	0.058	0.156	-0.017	0.019	0.022	0.028
	20	WGn	0.018	0.017	0.041	0.045	0.027	0.013	0.018	0.032
		WGi	-0.047	0.018	0.020	0.051	-0.003	0.013	0.013	0.013
		CCEPu	-0.088	0.023	0.030	0.093	-0.008	0.014	0.014	0.016
		CCEPr	-0.067	0.020	0.027	0.072	-0.006	0.013	0.013	0.014
	30	WGn	0.038	0.013	0.035	0.051	0.030	0.010	0.014	0.033
		WGi	-0.031	0.014	0.014	0.034	-0.001	0.010	0.010	0.010
		CCEPu	-0.056	0.017	0.020	0.060	-0.003	0.011	0.011	0.011
		CCEPr	-0.044	0.015	0.019	0.047	-0.003	0.010	0.010	0.011
	40	WGn	0.049	0.011	0.032	0.058	0.030	0.009	0.013	0.033
		WGi	-0.023	0.011	0.012	0.026	-0.001	0.009	0.009	0.009
		CCEPu	-0.041	0.014	0.015	0.044	-0.002	0.009	0.009	0.009
		CCEPr	-0.033	0.013	0.015	0.036	-0.003	0.009	0.009	0.009
	50	WGn	0.054	0.010	0.029	0.061	0.031	0.008	0.011	0.033
		WGi	-0.018	0.010	0.010	0.021	-0.000	0.008	0.008	0.008
		CCEPu	-0.032	0.012	0.013	0.035	-0.001	0.008	0.008	0.008
		CCEPr	-0.025	0.011	0.013	0.029	-0.002	0.008	0.008	0.008

Table 4: $N = 50$, $\rho = 0.6$, $\beta = 0.4$, $\theta = 0.4$, $\psi_1 = 1$, $\psi_3 = 0.8$ and $\xi = 3$

ψ_2	T		bias ρ	stde ρ	stdv ρ	rmse ρ	bias β	stde β	stdv β	rmse β
0.5	5	WGn	-0.128	0.045	0.069	0.145	-0.002	0.024	0.029	0.029
		WGi	-0.147	0.050	0.077	0.166	-0.023	0.026	0.029	0.037
		CCEPu	-0.379	0.108	0.277	0.470	-0.079	0.045	0.077	0.110
		CCEPr	-0.228	0.050	0.127	0.261	-0.037	0.025	0.036	0.052
	10	WGn	-0.045	0.026	0.044	0.063	0.015	0.016	0.020	0.025
		WGi	-0.064	0.026	0.030	0.071	-0.006	0.016	0.016	0.017
		CCEPu	-0.127	0.036	0.058	0.139	-0.016	0.018	0.021	0.026
		CCEPr	-0.090	0.028	0.043	0.100	-0.009	0.015	0.017	0.019
	20	WGn	-0.006	0.017	0.032	0.032	0.020	0.011	0.014	0.024
		WGi	-0.031	0.016	0.016	0.035	-0.002	0.010	0.011	0.011
		CCEPu	-0.055	0.020	0.023	0.060	-0.004	0.011	0.011	0.012
		CCEPr	-0.041	0.017	0.021	0.046	-0.003	0.010	0.011	0.011
	30	WGn	0.007	0.013	0.026	0.027	0.021	0.009	0.011	0.023
		WGi	-0.020	0.012	0.012	0.024	-0.001	0.008	0.008	0.008
		CCEPu	-0.035	0.015	0.016	0.038	-0.002	0.009	0.009	0.009
		CCEPr	-0.026	0.014	0.015	0.030	-0.001	0.008	0.008	0.009
	40	WGn	0.013	0.011	0.022	0.026	0.021	0.008	0.010	0.023
		WGi	-0.015	0.010	0.011	0.018	-0.000	0.007	0.007	0.007
		CCEPu	-0.025	0.012	0.013	0.028	-0.001	0.007	0.007	0.007
		CCEPr	-0.019	0.012	0.012	0.023	-0.001	0.007	0.007	0.007
	50	WGn	0.017	0.010	0.020	0.026	0.021	0.007	0.008	0.023
		WGi	-0.012	0.009	0.009	0.015	-0.000	0.006	0.006	0.006
		CCEPu	-0.020	0.011	0.011	0.023	-0.001	0.006	0.007	0.007
		CCEPr	-0.015	0.010	0.011	0.019	-0.001	0.006	0.006	0.006
2	5	WGn	-0.101	0.044	0.117	0.154	0.010	0.024	0.035	0.036
		WGi	-0.091	0.040	0.058	0.108	-0.014	0.021	0.023	0.027
		CCEPu	-0.316	0.100	0.246	0.400	-0.066	0.039	0.069	0.095
		CCEPr	-0.166	0.046	0.098	0.192	-0.027	0.021	0.029	0.040
	10	WGn	-0.018	0.026	0.077	0.079	0.026	0.016	0.024	0.035
		WGi	-0.039	0.020	0.023	0.045	-0.004	0.013	0.013	0.013
		CCEPu	-0.101	0.032	0.049	0.112	-0.013	0.015	0.017	0.022
		CCEPr	-0.068	0.025	0.033	0.076	-0.007	0.013	0.014	0.015
	20	WGn	0.027	0.016	0.054	0.061	0.030	0.011	0.016	0.034
		WGi	-0.018	0.012	0.013	0.022	-0.001	0.008	0.008	0.009
		CCEPu	-0.043	0.017	0.020	0.047	-0.003	0.009	0.009	0.010
		CCEPr	-0.031	0.015	0.017	0.036	-0.002	0.008	0.009	0.009
	30	WGn	0.041	0.013	0.047	0.062	0.031	0.009	0.013	0.034
		WGi	-0.011	0.009	0.010	0.015	-0.000	0.007	0.007	0.007
		CCEPu	-0.026	0.013	0.014	0.030	-0.001	0.007	0.007	0.007
		CCEPr	-0.020	0.012	0.013	0.023	-0.001	0.007	0.007	0.007
	40	WGn	0.049	0.011	0.041	0.064	0.031	0.008	0.011	0.033
		WGi	-0.009	0.008	0.008	0.012	-0.000	0.006	0.006	0.006
		CCEPu	-0.019	0.011	0.011	0.022	-0.001	0.006	0.006	0.006
		CCEPr	-0.014	0.010	0.011	0.018	-0.001	0.006	0.006	0.006
	50	WGn	0.055	0.010	0.037	0.066	0.031	0.007	0.010	0.033
		WGi	-0.007	0.007	0.007	0.010	-0.000	0.005	0.005	0.005
		CCEPu	-0.015	0.009	0.010	0.018	-0.000	0.005	0.005	0.005
		CCEPr	-0.012	0.009	0.009	0.015	-0.000	0.005	0.005	0.005

Table 5: $N = 50$, $\rho = 0.6$, $\beta = 0.4$, $\theta = 0.4$, $\psi_1 = 1$, $\psi_2 = 1$ and $\xi = 3$

ψ_3	T		bias ρ	stde ρ	stdv ρ	rmse ρ	bias β	stde β	stdv β	rmse β
0.4	5	WGn	-0.119	0.045	0.089	0.148	-0.002	0.024	0.029	0.030
		WGi	-0.124	0.046	0.069	0.142	-0.019	0.024	0.027	0.033
		CCEPu	-0.356	0.105	0.266	0.444	-0.074	0.043	0.073	0.104
		CCEPr	-0.203	0.048	0.118	0.235	-0.035	0.023	0.033	0.048
	10	WGn	-0.035	0.027	0.058	0.067	0.014	0.016	0.020	0.025
		WGi	-0.053	0.024	0.027	0.059	-0.005	0.014	0.015	0.015
		CCEPu	-0.116	0.035	0.055	0.128	-0.015	0.017	0.019	0.025
		CCEPr	-0.081	0.026	0.040	0.090	-0.010	0.014	0.016	0.019
	20	WGn	0.008	0.017	0.043	0.044	0.019	0.011	0.014	0.023
		WGi	-0.026	0.014	0.015	0.030	-0.001	0.010	0.010	0.010
		CCEPu	-0.050	0.019	0.022	0.055	-0.004	0.010	0.011	0.011
		CCEPr	-0.037	0.016	0.020	0.042	-0.004	0.009	0.010	0.011
	30	WGn	0.021	0.013	0.036	0.042	0.019	0.009	0.011	0.022
		WGi	-0.016	0.011	0.011	0.020	-0.001	0.008	0.008	0.008
		CCEPu	-0.032	0.014	0.015	0.035	-0.002	0.008	0.008	0.008
		CCEPr	-0.024	0.013	0.015	0.028	-0.003	0.008	0.008	0.009
	40	WGn	0.029	0.011	0.031	0.042	0.019	0.008	0.010	0.022
		WGi	-0.012	0.009	0.010	0.016	-0.000	0.007	0.007	0.007
		CCEPu	-0.023	0.012	0.012	0.026	-0.001	0.007	0.007	0.007
		CCEPr	-0.018	0.011	0.012	0.022	-0.003	0.007	0.007	0.008
	50	WGn	0.033	0.010	0.028	0.044	0.019	0.007	0.009	0.021
		WGi	-0.010	0.008	0.008	0.013	-0.000	0.006	0.006	0.006
		CCEPu	-0.018	0.010	0.011	0.021	-0.001	0.006	0.006	0.006
		CCEPr	-0.015	0.010	0.010	0.018	-0.003	0.006	0.006	0.007

Table 6: $N = 50$, $\rho = 0.6$, $\beta = 0.4$, $\theta = 0.4$, $\psi_1 = 1$, $\psi_2 = 1$ and $\psi_3 = 0.8$

ξ	T		bias ρ	stde ρ	stdv ρ	rmse ρ	bias β	stde β	stdv β	rmse β
10	5	WGn	-0.045	0.028	0.034	0.056	-0.003	0.012	0.013	0.013
		WGi	-0.074	0.035	0.047	0.088	-0.012	0.014	0.015	0.020
		CCEPu	-0.182	0.078	0.170	0.249	-0.039	0.027	0.046	0.060
		CCEPr	-0.087	0.031	0.061	0.106	-0.014	0.013	0.019	0.023
	10	WGn	-0.010	0.016	0.024	0.026	0.003	0.008	0.009	0.009
		WGi	-0.033	0.017	0.020	0.039	-0.004	0.008	0.008	0.009
		CCEPu	-0.055	0.023	0.031	0.063	-0.008	0.010	0.010	0.013
		CCEPr	-0.037	0.017	0.023	0.044	-0.004	0.008	0.009	0.010
	20	WGn	0.010	0.010	0.021	0.023	0.006	0.005	0.006	0.008
		WGi	-0.016	0.010	0.011	0.020	-0.001	0.005	0.005	0.005
		CCEPu	-0.024	0.012	0.014	0.028	-0.002	0.006	0.006	0.006
		CCEPr	-0.018	0.011	0.012	0.021	-0.001	0.005	0.005	0.006
	30	WGn	0.018	0.008	0.019	0.026	0.006	0.004	0.005	0.008
		WGi	-0.011	0.008	0.008	0.014	-0.000	0.004	0.004	0.004
		CCEPu	-0.015	0.009	0.010	0.018	-0.001	0.004	0.004	0.005
		CCEPr	-0.012	0.008	0.009	0.015	-0.001	0.004	0.004	0.004
	40	WGn	0.022	0.007	0.018	0.028	0.007	0.004	0.005	0.008
		WGi	-0.008	0.007	0.007	0.011	-0.000	0.004	0.004	0.004
		CCEPu	-0.011	0.007	0.008	0.014	-0.001	0.004	0.004	0.004
		CCEPr	-0.009	0.007	0.007	0.012	-0.001	0.004	0.004	0.004
	50	WGn	0.024	0.006	0.016	0.029	0.007	0.003	0.004	0.008
		WGi	-0.006	0.006	0.006	0.009	-0.000	0.003	0.003	0.003
		CCEPu	-0.009	0.006	0.007	0.011	-0.000	0.003	0.003	0.003
		CCEPr	-0.007	0.006	0.006	0.010	-0.000	0.003	0.003	0.003