Federal Reserve
Bank of Dallas

Pooled Bewley Estimator of Long-Run Relationships in Dynamic Heterogenous Panels

Alexander Chudik, M. Hashem Pesaran and Ron P. Smith

Pooled Bewley Estimator of Long-Run Relationships in Dynamic Heterogenous Panels[*]

Alexander Chudik[†], M. Hashem Pesaran[‡] and Ron P. Smith[§]

May 27, 2021
Revised: October 30, 2023

Abstract

Using a transformation of the autoregressive distributed lag model due to Bewley, a novel pooled Bewley (PB) estimator of long-run coefficients for dynamic panels with heterogeneous short-run dynamics is proposed. The PB estimator is directly comparable to the widely used Pooled Mean Group (PMG) estimator, and is shown to be consistent and asymptotically normal. Monte Carlo simulations show good small sample performance of PB compared to the existing estimators in the literature, namely PMG, panel dynamic OLS (PDOLS), and panel fully-modified OLS (FMOLS). Application of two bias-correction methods and a bootstrapping of critical values to conduct inference robust to cross-sectional dependence of errors are also considered. The utility of the PB estimator is illustrated in an empirical application to the aggregate consumption function.

Keywords: Heterogeneous dynamic panels, I(1) regressors, pooled mean group estimator (PMG), Autoregressive-Distributed Lag model (ARDL), Bewley transform, PDOLS, FMOLS, bias correction, robust inference, cross-sectional dependence.

JEL Classification: C12, C13, C23, C33

[†]Alexander Chudik, Federal Reserve Bank of Dallas, alexander.chudik@gmail.com.

[‡]M. Hashem Pesaran, University of Southern California, USA and Trinity College, Cambridge, UK, mhp1@cam.ac.uk.

[§]Ron P. Smith, Birkbeck, University of London, United Kingdom, r.smith@bbk.ac.uk.

1 Introduction

Estimation of cointegrating relationships in panels with heterogeneous short-run dynamics is important for empirical research in open economy macroeconomics as well as in other fields in economics. Existing single-equation panel estimators in the literature are panel Fully Modified OLS (FMOLS) by Pedroni (1996, 2001a, 2001b), panel Dynamic OLS (PDOLS) by Mark and Sul (2003), and the likelihood based Pooled Mean Group (PMG) estimator by Pesaran, Shin, and Smith (1999). Multi-equation (system) approach by Breitung (2005), and the related system PMG approach by Chudik, Pesaran, and Smith (2023) are another contributions in the literature on estimating cointegrating vectors in a panel context. In this paper, we propose a pooled Bewley (PB) estimator of long-run relationships, relying on the Bewley transform of an autoregressive distributed lag (ARDL) model (Bewley, 1979). See also Wickens and Breusch (1988) for a discussion of the Bewley transform.

Our setting is the same as that of Pesaran, Shin, and Smith (1999). Under this setting, any short-run feedbacks between the outcome variable (y) and regressors (x) are allowed, but the direction of the long run causality is assumed to go from x to y. Hence, as with the PMG, the PB estimator allows for heterogeneity in short-run feedbacks, but restricts the direction of long run causality. The PB estimator is computed analytically, and does not rely on numerical maximization of the likelihood function that underlies the PMG estimation. We derive the asymptotic distribution of the PB estimator when the cross-section dimension (n) and the time dimension (T) diverge to infinity jointly such that $n = \Theta\left(T^{\theta}\right)$, for $0 < \theta < 2$, where we use the notation $\Theta\left(.\right)$ to denote the same order of magnitude asymptotically, namely if $\{f_s\}_{s=1}^{\infty}$ and $\{g_s\}_{s=1}^{\infty}$ are both positive sequences of real numbers, then $f_s = \ominus\left(g_s\right)$ if there exists $S_0 \geq 1$ and positive finite constants C_0 and C_1, such that $\inf_{s \geq S_0}\left(f_s/g_s\right) \geq C_0$, and $\sup_{s \geq S_0}\left(f_s/g_s\right) \leq C_1$. Our asymptotic analysis is an advance over the theoretical results currently available for PMG, PDOLS and FMOLS estimators where it is assumed that n is small relative to T (which corresponds to the case where θ is close to zero).

How well individual estimators work in samples of interest in practice where n and/or T are often less than 50 is a different matter, which we shed light on using Monte Carlo experiments. Monte Carlo evidence shows PB estimator can be superior to PMG, PDOLS and FMOLS, in terms of its overall precision as measured by the Root Mean Square Error (RMSE), and in terms of accuracy of inference as measured by size distortions. These experiments reveal PB is a useful addition to the literature.

Monte Carlo evidence also shows that the time dimension is very important for the performance of these estimators, and that all of the four estimators under consideration suffer from the same two drawbacks: small sample bias and size distortion. Although the size distortions are found to be less serious for the PB estimator in our experiments, all four estimators exhibit notable over-rejections in sample sizes relevant in practice. In addition, all four estimators (perhaps unsurprisingly) suffer from bias in finite samples, albeit a rather small one. Both drawbacks diminish as T is increased relative to n.

To conduct reliable inference regardless of the cross-sectional dependence of errors, we make use of the sieve wild bootstrap procedure. To accommodate cross-sectional dependence, we resample the cross-section vectors of residuals, an idea that was originally proposed by Maddala and Wu (1999). We found the sieve wild bootstrap procedure to be remarkably effective for all four estimators, regardless of cross-sectional dependence of errors, and we therefore recommend using it in empirical research.

Regarding the small sample bias, we consider the application of two bias-correction methods taken from the literature, relying either on split-panel jackknife (Dhaene and Jochmans, 2015) or sieve wild bootstrap approaches. In contrast to split-panel approaches in panels without stochastic trends, as, for instance, considered by of Dhaene and Jochmans (2015) or Chudik, Pesaran, and Yang (2018), in this paper we need to combine the full sample and half-panel subsamples using different weighting due to the fact that the rate of convergence of the estimators of long run coefficients is faster, at the rate of $T\sqrt{n}$, as compared to the standard rate of $\sqrt{nT}$. We find that both of these approaches can be helpful in reducing the bias (for all four estimators). However, given that the bias is small to begin with, the value of bias correction methods is limited.

The relevance of choosing a particular estimation approach is illustrated in the context of a consumption function application for OECD economies taken from Pesaran, Shin, and Smith (1999). This application shows that quite a different conclusion would be reached when using PB estimator, which does not reject the zero long-run coefficient on inflation, in line with the long-run neutrality of monetary policy, whereas the PMG estimator results in a highly statistically significant negative long-run coefficient. Estimates of the long-run coefficient on real income are less diverse across estimators, but the inference on whether a unit long-run coefficient on real income (as suggested by balanced growth path models in the literature) can be rejected or not depends on

the choice of a particular estimator.

The remainder of this paper is organized as follows. Section 2 presents the model and assumptions, introduces the PB estimator, and provides asymptotic results. Application of bias correction methods and bootstrapping critical values are also discussed in Section 2. Section 3 presents Monte Carlo evidence. Section 4 revisits the aggregate consumption function empirical application in Pesaran, Shin, and Smith (1999). Section 5 concludes. Mathematical derivations and proofs are provided in Appendix A. Details on the implementation of individual estimators and bootstrapping, and additional Monte Carlo results are provided in Appendix B.

2 Pooled Bewley estimator of long-run relationships

We adopt the same setting as in Pesaran, Shin, and Smith (1999), and consider the following illustrative model

$$\Delta y_{it} = c_i - \alpha_i (y_{i,t-1} - \beta x_{i,t-1}) + u_{y,it}, \tag{1}$$

$$\Delta x_{it} = u_{x,it}, \tag{2}$$

for $i = 1, 2, ..., n$, and $t = 1, 2, ..., T$. For expositional clarity and notational simplicity, we focus on a single regressor and one lag, but it is understood that our analysis is applicable to multiple lags of $\Delta \mathbf{z}_{it} = (\Delta y_{it}, \Delta x_{it})'$ entering both equations (1)-(2), and the approach is also applicable to multiple x_{it}'s with a single long run relationship. We consider the following assumptions:

Assumption 1 *(Coefficients)* $\sup_i |1 - \alpha_i| < 1$.

Assumption 2 *(shocks)* $u_{x,it} \sim IID \left(0, \sigma_{xi}^2\right)$, *and* $u_{y,it}$ *is given by*

$$u_{y,it} = \delta_i u_{x,it} + v_{it}, \tag{3}$$

for all i and t, where $v_{it} \sim IID \left(0, \sigma_{vi}^2\right)$, and $u_{x,it}$ is independently distributed of $v_{i't'}$ for all i, i', t, and t'. In addition, $\sup_{i,t} E |v_{it}|^{16} < K$ and $\sup_{i,t} E |u_{x,it}|^8 < K$, and $\lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} \sigma_{xi}^2 = \sigma_x^2 > 0$ and $\lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} \sigma_{xi}^2 \sigma_{vi}^2 / \left(6\alpha_i^2\right) = \omega_v^2 > 0$ exist.

Assumption 3 *(Initial values and deterministic terms)* *The initial values,* $\mathbf{z}_{i,0} = (y_{i,0}, x_{i,0})'$,

3

follow the process

$$\mathbf{z}_{i0} = \boldsymbol{\mu}_i + \mathbf{C}_i^* (L) \, \mathbf{u}_0,$$

for all i and t, and $c_i = \alpha_i \mu_{i,1} - \alpha_i \beta \mu_{i,2}$ for all i, where $\mathbf{u}_0 = (u_{y,i,0}, u_{x,i,0})'$, $\boldsymbol{\mu}_i = (\mu_{i,1}, \mu_{i,2})'$, $\|\boldsymbol{\mu}_i\| < K$, and $\mathbf{C}_i^* (L)$ is defined in Section A.1 in Appendix A.

Remark 1 *Assumption 1 requires that $\alpha_i \neq 0$ for all i. In contrast PMG allows $\alpha_i = 0$ for some (but not all) units. Although the estimator works for $\alpha_i < 2$, empirically, $\alpha_i < 1$ is likely to be the relevant case.*

Remark 2 *Assumption 2 allows for $u_{x,it}$ to be correlated with $u_{y,it}$. Cross-section dependence of $\mathbf{u}_{it}$ is ruled out. Assumption 3 (together with the remaining assumptions) ensure that $\Delta \mathbf{z}_{it}$ and $(y_{it} - \beta x_{it})$ are covariance stationary.*

Remark 3 *In comparing our assumptions with the rest of the literature, it should be noted that the rest of the literature does not consider the case of joint convergence $n, T \rightarrow_j \infty$, but only the case where n is fixed as $T \rightarrow \infty$. Joint asymptotics typically requires more stringent assumptions on the errors and restrictions on the relative expansion rates of n and T.*

Substituting first (3) for $u_{y,it}$ in (1), and then substituting $u_{x,it} = \Delta x_{it}$, we obtain the following ARDL representation for y_{it}

$$\Delta y_{it} = c_i - \alpha_i \left(y_{i,t-1} - \beta x_{i,t-1} \right) + \delta_i \Delta x_{it} + v_{it}. \tag{4}$$

The pooled Bewley estimator takes advantage of the Bewley transform (Bewley, 1979). Subtracting $(1 - \alpha_i) \, y_{it}$ from both sides of (4) and re-arranging, we have

$$\alpha_i y_{it} = c_i - (1 - \alpha_i) \, \Delta y_{it} + \alpha_i \beta x_{it} + \delta_i \Delta x_{it} + v_{it},$$

or (noting that $\alpha_i > 0$ for all i and multiplying the equation above by α_i^{-1})

$$y_{it} = \alpha_i^{-1} c_i + \beta x_{it} + \boldsymbol{\psi}_i' \Delta \mathbf{z}_{it} + \alpha_i^{-1} v_{it}, \tag{5}$$

where $\Delta\mathbf{z}_{it} = (\Delta y_{it}, \Delta x_{it})'$, and $\boldsymbol{\psi}_i = \left(-\frac{1-\alpha_i}{\alpha_i}, \frac{\delta_i}{\alpha_i}\right)'$. Further, stacking (5) for $t = 1, 2, ..., T$, we have

$$\mathbf{y}_i = \alpha_i^{-1} c_i \boldsymbol{\tau}_T + \mathbf{x}_i \beta + \Delta\mathbf{Z}_i \boldsymbol{\psi}_i + \alpha_i^{-1} \mathbf{v}_i, \tag{6}$$

where $\mathbf{y}_i = (y_{i1}, y_{i2}, ..., y_{iT})'$, $\mathbf{x}_i = (x_{i1}, x_{i2}, ..., x_{iT})'$, $\Delta\mathbf{Z}_i = (\Delta\mathbf{z}_{i1}', \Delta\mathbf{z}_{i2}', ..., \Delta\mathbf{z}_{iT}')'$, $\mathbf{v}_i = (v_{i,1}, v_{i,2}, ..., v_{i,T})'$, and $\boldsymbol{\tau}_T$ is $T \times 1$ vector of ones. Define projection matrix $\mathbf{M}_\tau = \mathbf{I}_T - T^{-1}\boldsymbol{\tau}_T\boldsymbol{\tau}_T'$. This projection matrix subtracts the period average. Let $\tilde{\mathbf{y}}_i = (\tilde{y}_{i1}, \tilde{y}_{i2}, ..., \tilde{y}_{iT})' = \mathbf{M}_\tau\mathbf{y}_i$, and similarly $\tilde{\mathbf{x}}_i = (\tilde{x}_{i1}, \tilde{x}_{i2}, ..., \tilde{x}_{iT})' = \mathbf{M}_\tau\mathbf{x}_i$, $\Delta\tilde{\mathbf{Z}}_i = \mathbf{M}_\tau\Delta\mathbf{Z}_i$, and $\tilde{\mathbf{v}}_i = \mathbf{M}_\tau\mathbf{v}_i$. Multiplying (6) by $\mathbf{M}_\tau$, we have

$$\tilde{\mathbf{y}}_i = \tilde{\mathbf{x}}_i\beta + \Delta\tilde{\mathbf{Z}}_i\boldsymbol{\psi}_i + \alpha_i^{-1}\tilde{\mathbf{v}}_i.$$

Now consider the matrix of instruments

$$\tilde{\mathbf{H}}_i = (\tilde{\mathbf{y}}_{i,-1}, \tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_{i,-1}) = \mathbf{M}_\tau\mathbf{H}_i, \ \mathbf{H}_i = (\mathbf{y}_{i,-1}, \mathbf{x}_i, \mathbf{x}_{i,-1}), \tag{7}$$

where $\mathbf{y}_{i,-1} = (y_{i,1}, y_{i,1}, ..., y_{i,T-1})'$ is the data vector on the first lag of y_{it}, similarly $\mathbf{x}_{i,-1} = (x_{i,1}, x_{i,1}, ..., x_{i,T-1})'$. The PB estimator of β is given by

$$\hat{\beta} = \left(\sum_{i=1}^n \tilde{\mathbf{x}}_i'\mathbf{M}_i\tilde{\mathbf{x}}_i\right)^{-1} \left(\sum_{i=1}^n \tilde{\mathbf{x}}_i'\mathbf{M}_i\tilde{\mathbf{y}}_i\right), \tag{8}$$

where

$$\mathbf{M}_i = \mathbf{P}_i - \mathbf{P}_i\Delta\tilde{\mathbf{Z}}_i \left(\Delta\tilde{\mathbf{Z}}_i'\mathbf{P}_i\Delta\tilde{\mathbf{Z}}_i\right)^{-1} \Delta\tilde{\mathbf{Z}}_i'\mathbf{P}_i, \tag{9}$$

and

$$\mathbf{P}_i = \tilde{\mathbf{H}}_i \left(\tilde{\mathbf{H}}_i'\tilde{\mathbf{H}}_i\right)^{-1} \tilde{\mathbf{H}}_i', \tag{10}$$

is the projection matrix associated with $\tilde{\mathbf{H}}_i$.

In addition to Assumptions 1-3, we also require the following high-level conditions to hold in the derivations of the asymptotic distribution of the PB estimator under the joint asymptotics $n, T \to \infty$.

Assumption 4 *Let* $\mathbf{M}_\tau = \mathbf{I}_T - T^{-1}\boldsymbol{\tau}_T\boldsymbol{\tau}_T'$, $\tilde{\mathbf{y}}_i = \mathbf{M}_\tau\mathbf{y}_i$, $\tilde{\mathbf{x}}_i = \mathbf{M}_\tau\mathbf{x}_i$, $\Delta\tilde{\mathbf{Z}}_i = \mathbf{M}_\tau\Delta\mathbf{Z}_i = \mathbf{M}_\tau(\Delta\mathbf{z}_{i1}', \Delta\mathbf{z}_{i2}', ..., \Delta\mathbf{z}_{iT}')'$, *where* $\Delta\mathbf{z}_{it} = (\Delta y_{it}, \Delta x_{it})'$. *Then there exists* $T_0 \in \mathbb{N}$

5

such that the following conditions are satisfied:

(i) $\sup_{i \in \mathbb{N},\ T > T_0} E \left[\lambda_{\min}^{-2} (\mathbf{B}_{iT}) \right] < K$, *where* $\mathbf{B}_{iT} = \Delta \tilde{\mathbf{Z}}_i' \mathbf{P}_i \Delta \tilde{\mathbf{Z}}_i / T$, $\mathbf{P}_i$ *is given by (10).*

(ii) $\sup_{i \in \mathbb{N},\ T > T_0} E \left[\lambda_{\min}^{-2} \left(\mathbf{A}_T \tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{H}}_i^* \mathbf{A}_T \right) \right] < K$, *where* $\mathbf{H}_i^* = \left(\tilde{\mathbf{x}}_i, \Delta \tilde{\mathbf{x}}_i, \tilde{\boldsymbol{\xi}}_{i,-1} \right)$,

$$
\mathbf{A}_T = \begin{pmatrix} T^{-1} & 0 & 0 \\ 0 & T^{-1/2} & 0 \\ 0 & 0 & T^{-1/2} \end{pmatrix},
$$

$$
\tilde{\boldsymbol{\xi}}_{i,-1} = \left(\tilde{\xi}_{i,0}, \tilde{\xi}_{i,1}, ..., \tilde{\xi}_{i,T-1} \right)', \ \tilde{\xi}_{i,t-1} = \tilde{y}_{i,t-1} - \beta \tilde{x}_{i,t-1}.
$$

Remark 4 *Under Assumptions 1-3 (and without Assumption 4), we have* $\mathrm{plim}_{T \to \infty} \mathbf{B}_{i,T} = \mathbf{B}_i$, *where* $\mathbf{B}_i$ *is nonsingular (see Lemma A.7 in Appendix A). Similarly, it can be shown that Assumptions 1-3 are sufficient for* $\mathrm{plim}_{T \to \infty} \mathbf{A}_T \tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{H}}_i^* \mathbf{A}_T$ *to exist and to be nonsingular. However, these results are not sufficient for the moments of* $\left\| \mathbf{B}_i^{-1} \right\|$ *and* $\left\| (\mathbf{A}_T \mathbf{H}_i^{*\prime} \mathbf{H}_i^* \mathbf{A}_T)^{-1} \right\|$ *to exist, which we require for the derivations of the asymptotic distribution of the PB estimator. This is ensured by Assumption 4.*

2.1 Asymptotic results

Substituting $\tilde{\mathbf{y}}_i = \tilde{\mathbf{x}}_i \beta + \Delta \tilde{\mathbf{Z}}_i \boldsymbol{\psi}_i + \alpha_i^{-1} \tilde{\mathbf{v}}_i$ in (8), and using $\mathbf{M}_i \Delta \tilde{\mathbf{Z}}_i = \mathbf{0}$, we have

$$
T \sqrt{n} \left(\hat{\beta} - \beta \right) = \left(\frac{1}{n} \sum_{i=1}^{n} \frac{\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{x}}_i}{T^2} \right)^{-1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{v}}_i}{T \alpha_i} \right). \tag{11}
$$

Consider the first term on the right side of (11) first. Since $\mathbf{M}_i$ is an orthogonal projection matrix, $\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{x}}_i / T^2$ is bounded by $\tilde{\mathbf{x}}_i' \tilde{\mathbf{x}}_i / T^2$. The second moments of $\tilde{\mathbf{x}}_i' \tilde{\mathbf{x}}_i / T^2$ are bounded, and, in addition, $\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{x}}_i / T^2$ is cross-sectionally independent. It follows that $\frac{1}{n} \sum_{i=1}^{n} \tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{x}}_i / T^2$ converges to a constant, which we denote by ω_x^2, as $n, T \to \infty$. Lemma A.4 in Appendix A establishes the expression for $\omega_x^2 = \sigma_x^2 / 6$, where $\sigma_x^2 = \lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} \sigma_{xi}^2$, but the specific expression for ω_x^2 is not relevant for the inference approach that we adopt below. Consider the second term of (11)

6

next,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\tilde{\mathbf{x}}_i'\mathbf{M}_i\tilde{\mathbf{v}}_i}{\alpha_i T} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\frac{\tilde{\mathbf{x}}_i'\mathbf{M}_i\tilde{\mathbf{v}}_i}{\alpha_i T} - E\left(\frac{\tilde{\mathbf{x}}_i'\mathbf{M}_i\tilde{\mathbf{v}}_i}{\alpha_i T}\right)\right]$$
$$+ \frac{1}{\sqrt{n}}\sum_{i=1}^{n}E\left(\frac{\tilde{\mathbf{x}}_i'\mathbf{M}_i\tilde{\mathbf{v}}_i}{\alpha_i T}\right), \tag{12}$$

The term in the square brackets has zero mean and is independently distributed over i. For the asymptotic distribution to be correctly centered we need

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}E\left(\frac{\tilde{\mathbf{x}}_i'\mathbf{M}_i\tilde{\mathbf{v}}_i}{\alpha_i T}\right) \to 0, \tag{13}$$

as n and $T \to \infty$. This condition holds so long as $n = \Theta\left(T^\theta\right)$ for some $0 < \theta < 2$. See Lemma A.10 for a proof. The asymptotic distribution of the first term in (12) is in turn established by Lemma A.11, see (A.74). The following theorem now follows for the asymptotic distribution of $\hat{\beta}$.

Theorem 1 *Let (y_{it}, x_{it}) be generated by model (1)-(2), suppose Assumptions 1-4 hold, and $n, T \to \infty$ such that $n = \Theta\left(T^\theta\right)$, for some $0 < \theta < 2$. Consider the PB estimator $\hat{\beta}$ given by (8). Then,*

$$T\sqrt{n}\left(\hat{\beta} - \beta\right) \to_d N\left(0, \Omega\right), \ \Omega = \omega_x^{-4}\omega_v^2, \tag{14}$$

where $\omega_x^2 = \sigma_x^2/6$, $\sigma_x^2 = \lim_{n\to\infty} n^{-1}\sum_{i=1}^{n}\sigma_{xi}^2$ and $\omega_v^2 = \lim_{n\to\infty} n^{-1}\sum_{i=1}^{n}\sigma_{xi}^2\sigma_{vi}^2/\left(6\alpha_i^2\right)$.

Remark 5 *Like the PMG estimator in Pesaran, Shin, and Smith (1999), the PB estimator will also work when variables are integrated of order 0 (the I(0) case), which is not pursued in this paper. In the I(0) case, the PB estimator converges at the standard rate of $\sqrt{nT}$.*

To conduct inference, let

$$\hat{\omega}_x^2 = n^{-1}\sum_{i=1}^{n}\frac{\mathbf{x}_i'\mathbf{M}_i\mathbf{x}_i}{T^2}, \tag{15}$$

and $\hat{\omega}_v^2 = n^{-1}\sum_{i=1}^{n}\left(\frac{\mathbf{x}_i'\mathbf{M}_i\hat{\mathbf{v}}_i^*}{T}\right)^2$, where $\hat{\mathbf{v}}_i^* = \mathbf{M}_i\left(\mathbf{y}_i - \hat{\beta}\mathbf{x}_i\right)$, and $\mathbf{M}_i$ is defined by (9). Accordingly, we propose the following estimator of Ω:

$$\hat{\Omega} = \hat{\omega}_x^{-4}\hat{\omega}_v^2. \tag{16}$$

2.2 Bias mitigation and bootstrapping critical values for robust inference

When n is sufficiently large relative to T, specifically when $\sqrt{n}/T \to K > 0$, then $\sqrt{n}T\left(\hat{\beta} - \beta\right)$ is no longer asymptotically distributed with zero mean. The asymptotic bias is due to the nonzero mean of $T^{-1}\tilde{\mathbf{x}}_i'\mathbf{M}_i\tilde{\mathbf{v}}_i$, and it can be of some relevance for finite sample performance, as the Monte Carlo evidence in Section 3 illustrates. Monte Carlo evidence also reveals that the inference based on PB and other existing estimators in the literature can suffer from serious size distortions in finite samples. To deal with these problems, we consider bootstrapping critical values using sieve wild bootstrap for more accurate and more robust inference that allows for cross-sectional dependence of errors. In addition, we also consider two bias-correction techniques - a bootstrap one as well as the split-panel jackknife method. The same bias-correction methods are also applied to the three other estimators, namely PMG, PDOLS, and FMOLS, considered in the paper. In what follows we focus on the PB estimator. A description of bias-corrections applied to the other three estimators are given in Section B.2 of Appendix B.

2.2.1 Bootstrap bias reduction

Once an estimate of the bias of $\hat{\beta}$ is available, denoted as $\hat{b}$, then the bias-corrected PB estimator is given by

$$\tilde{\beta} = \hat{\beta} - \hat{b}. \tag{17}$$

One possibility of estimating the bias in the literature is by bootstrap. We adopt the following sieve wild bootstrap algorithm for generating simulated data.

1. Given $\hat{\beta}$, estimate the remaining unknown coefficients in (1)-(2) by least squares, and compute residuals denoted by $\hat{u}_{y,it}, \hat{u}_{x,it}$.

2. For each $r = 1, 2, ..., R$, generate new draws for $\hat{u}_{y,it}^{(r)} = a_t^{(r)}\hat{u}_{y,it}$, and $\hat{u}_{x,it}^{(r)} = a_t^{(r)}\hat{u}_{x,it}$, where $a_t^{(r)}$ is randomly drawn from Rademacher distribution (Liu, 1988),

$$a_t^{(r)} = \begin{cases} -1, & \text{with probability } 1/2 \\ 1, & \text{with probability } 1/2 \end{cases}.$$

Given the estimated parameters of (1)-(2) from Step 1, and the initial values $\{y_{i1}, x_{i1} \text{ for } i = 1, 2, ..., n\}$

generate the simulated series $y_{it}^{(r)}, x_{it}^{(r)}$ for $t = 2, 3, ..., T$, and $i = 1, 2, ..., n$, and the bootstrap estimates $\hat{\beta}^{(r)}$ for $r = 1, 2, ..., R$.

Using simulated data with $R = 10,000$, we compute an estimate of the bias $\hat{b}_R = \left[R^{-1} \sum_{r=1}^{R} \hat{\beta}^{(r)} - \hat{\beta} \right]$. We then compute the α percent critical values using the $1 - \alpha$ percent quantile of $\left\{ \left| t^{(r)} \right| \right\}_{r=1}^{R}$, where $t^{(r)} = \tilde{\beta}^{(r)} / se \left(\tilde{\beta}^{(r)} \right)$, $\tilde{\beta}^{(r)} = \hat{\beta}^{(r)} - \hat{b}$ is the bias-corrected estimate of β using the r-th draw of the simulated data, $se \left(\tilde{\beta}^{(r)} \right) = T^{-1} n^{-1/2} \hat{\Omega}^{(r)}$ is the corresponding standard error estimate, and $\hat{\Omega}^{(r)}$ is computed in the same way as $\hat{\Omega}$ in (16) but using the simulated data.

2.2.2 Jackknife bias reduction

The split-panel jackknife bias correction method is given by

$$\tilde{\beta}_{jk} = \tilde{\beta}_{jk} (\kappa) = \hat{\beta} - \kappa \left(\frac{\hat{\beta}_a + \hat{\beta}_b}{2} - \hat{\beta} \right), \tag{18}$$

where $\hat{\beta}$ is the full sample PB estimator, $\hat{\beta}_a$ and $\hat{\beta}_b$ are the first and the second half sub-sample PB estimators, and κ is a suitably chosen weighting parameter. In a stationary setting, where the bias is of order $O\left(T^{-1}\right)$, κ is chosen to be one, so that $\frac{K}{T} - \kappa \cdot \left(\frac{K}{T/2} - \frac{K}{T} \right) = 0$, for any arbitrary choice of K. See, for example, Dhaene and Jochmans (2015) and Chudik, Pesaran, and Yang (2018).

In general, when the bias is of order $O\left(T^{-\epsilon}\right)$ for some $\epsilon > 0$, then κ can be chosen to solve $\frac{K}{T^{\epsilon}} - \kappa \cdot \left(\frac{K}{(T/2)^{\epsilon}} - \frac{K}{T^{\epsilon}} \right) = 0$, which yields $\kappa = 1/(2^{\epsilon} - 1)$. Under our setup with I(1) variables, we need to correct $\hat{\beta}$ for its $O\left(T^{-2}\right)$ bias, namely $\epsilon = 2$, which yields $\kappa = 1/3$.

Inference using $\tilde{\beta}^{jk}$ can be conducted based on (16) but with $\hat{\omega}_v^2$ replaced by

$$\tilde{\omega}_v^2 = \hat{\omega}_v^2 = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\left[(1 + \kappa) \, \mathbf{x}_i' \mathbf{M}_i - 2\kappa \mathbf{x}_{ab,i}' \mathbf{M}_{ab,i} \right] \tilde{\mathbf{v}}_i^*}{T} \right)^2, \tag{19}$$

where $\tilde{\mathbf{v}}_i^* = \mathbf{M}_i \left(\mathbf{y}_i - \tilde{\beta}^{jk} \mathbf{x}_i \right)$,

$$\mathbf{x}_{ab,i}' = \begin{pmatrix} \mathbf{x}_{a,i}' \\ \mathbf{x}_{b,i}' \end{pmatrix}, \ \mathbf{M}_{ab,i} = \begin{pmatrix} \mathbf{M}_{a,i} \\ \mathbf{M}_{b,i} \end{pmatrix},$$

$\mathbf{x}_{a,i}' \left(\mathbf{x}_{b,i}' \right)$ and $\mathbf{M}_{a,i} \left(\mathbf{M}_{b,i} \right)$ are defined in the same way as $\mathbf{x}_i$, and $\mathbf{M}_i$ but using only the first

9

(second) half of the sample.

We compute bootstrapped critical values to conduct more accurate and robust small sample inference. Specifically, the α percent critical value is computed as the $1 - \alpha$ percent quantile of $\left\{ \left| t_{jk}^{(r)} \right| \right\}_{r=1}^{R}$, where $t_{jk}^{(r)} = \tilde{\beta}_{jk}^{(r)} / se\left(\tilde{\beta}_{jk}^{(r)}\right)$, $\tilde{\beta}_{jk}^{(r)}$ is the jackknife estimate of β using the r-th draw of the simulated data generated using the algorithm described in Subsection 2.2.1, $se\left(\tilde{\beta}_{jk}^{(r)}\right)$ is the corresponding standard error estimate, namely $se\left(\tilde{\beta}_{jk}^{(r)}\right) = T^{-1} n^{-1/2} \hat{\Omega}_{jk}^{(r)}$, $\hat{\Omega}_{jk}^{(r)} = \hat{\omega}_{x,(r)}^{-4} \tilde{\omega}_{v,(r)}^{2}$, in which $\tilde{\omega}_{v,(r)}$ and $\hat{\omega}_{x,(r)}^{2}$ are computed using the simulated data, based on expressions (19) and (15), respectively.

3 Monte Carlo Evidence

3.1 Design

The Data Generating Process (DGP) is given by (1)-(2), for $i = 1, 2, ..., n$, $T = 1, 2, ..., T$, with starting values satisfying Assumption 3 with $\boldsymbol{\mu}_i \sim IIDN\left(\boldsymbol{\tau}_2, \mathbf{I}_2\right)$, and $c_i = \alpha_i \mu_{i,1} - \alpha_i \beta \mu_{i,2}$. We generate $\alpha_i \sim IIDU\left[0.2, 0.3\right]$. We consider two DGPs based on the cross-sectional dependence of errors. In the cross-sectionally independent DGP, we generate $u_{y,it} = \sigma_{y,i} e_{y,it}$, $u_{x,it} = \sigma_{x,i} e_{x,it}$, $\sigma_{y,i}^2, \sigma_{x,i}^2 \sim IIDU\left[0.8, 1.2\right]$,

$$
\begin{pmatrix} e_{y,it} \\ e_{x,it} \end{pmatrix} \sim IIDN\left(\mathbf{0}_2, \boldsymbol{\Sigma}_e\right), \boldsymbol{\Sigma}_e \sim \begin{pmatrix} 1 & \rho_i \\ \rho_i & 1 \end{pmatrix}, \text{ and } \rho_i \sim IIDU\left[0.3, 0.7\right].
$$

In the DGP with cross-sectionally dependent errors, we generate $e_{y,it}$ to contain a factor structure including strong, semi-strong and weak factors:

$$
e_{y,it} = \varkappa_i \left(\varepsilon_{y,it} + \sum_{\ell=1}^{m} \gamma_{i\ell} f_{\ell t} \right),
$$

where $\varepsilon_{y,it} \sim IIDN\left(0, 1\right)$, $f_{\ell,t} \sim IIDN\left(0, 1\right)$, $\gamma_\ell \sim IIDU\left[0, \gamma_{\max,\ell}\right]$, for $\ell = 1, 2, ..., m$. We choose $m = 5$ factors and $\gamma_{\max,\ell} = 2n^{\alpha_\ell - 1}$ with $\alpha_\ell = 1, 0.9, 0.8, 0.7, 0.6$, for $\ell = 1, 2, ..., 5$, respectively. Scaling constant $\varkappa_i$ is set to ensure $E\left(e_{y,it}^2\right) = 1$, namely $\varkappa_i = \left(1 + \sum_{\ell=1}^{m} \gamma_{i\ell}^2\right)^{-1/2}$. We generate $e_{x,it}$ to ensure unit variance and $cov\left(e_{y,it}, e_{x,it}\right) = \rho_i$. Specifically, $e_{x,it} = \rho_i e_{y,it} + \sqrt{1 - \rho_i^2} \varepsilon_{x,it}$, $\varepsilon_{x,it} \sim IIDN\left(0, 1\right)$. Both designs features heteroskedastic (over i) and correlated (over y & x

equations) errors, namely $E\left(u_{y,it}^2\right) = \sigma_{y,i}^2$, $E\left(u_{x,it}^2\right) = \sigma_{x,i}^2$, and $corr\left(u_{y,it}, u_{x,it}\right) = \rho_i$. We consider $n, T = 20, 30, 40, 50$ and compute $R_{MC} = 2000$ Monte Carlo replications.

3.2 Bias, RMSE and inference

We report bias, root mean square error (RMSE), size ($H_0 : \beta = 1$, 5% nominal level) and power ($H_1 : \beta = 0.9$, 5% nominal level) findings for the PB estimator $\hat{\beta}$ given by (8), with variance estimated using (16). Moreover, we also report findings for the two bias corrected versions of PB estimator as described in Subsection 2.2 with bootstrapped critical values for inference robust to error cross-sectional dependence. We compare the performance of the PB estimator with the PMG estimator by Pesaran, Shin, and Smith (1999), panel dynamic OLS (PDOLS) estimator by Mark and Sul (2003), and the group-mean fully modified OLS (FMOLS) estimator by Pedroni (1996, 2001b). Similarly to the PB estimator, we also consider jackknife and bootstrap based bias-corrected versions of the PMG, PDOLS and FMOLS estimators with cross-sectionally robust bootstrapped critical values, described in Appendix B. We use $R_b = 10,000$ bootstrap replications (within each MC replication) for bootstrap bias correction and for computation of robust and more accurate bootstrapped critical values.

3.3 Findings

Table 1 report the results for the original (without bias-correction) estimators. PB estimator stands out as the most precise estimator in terms of having the lowest RMSE values among the four estimators. The second best is PMG estimator with RMSE values 1 to 21 percent larger compared with the PB estimator, the third is PDOLS with RMSE values 23 to 66 percent larger compared with PB, and the FMOLS comes last with RMSE values 95 to 180 percent larger compared with PB. In terms of the bias alone, the ordering of the estimators is slightly different with PMG and PB switching their places. For $T = 20$, the bias of PMG estimator is -0.016 to -0.020, the bias of PB estimator is in the range -0.034 to -0.037, the bias of the PDOLS estimator is in the range -0.052 to -0.056 and the bias of the FMOLS estimator is in the range -0.104 to -0.110. For such a small value of T, the bias is not very large, and, as expected, it declines with an increase in T.

All four estimators suffer from varying degrees of size distortions. The inference based on the PB estimator is the most accurate. Specifically, the size distortions for the PB estimator are lowest

among the four estimators - with reported size in the range between 9.9 and 25.2 percent, exceeding the chosen nominal value of 5 percent. Size distortions diminish with an increase in T.

Table 1: MC findings for the estimation of long-run coefficient β in experiments with cross-sectionally independent errors.

Estimators without bias correction and inference conducted using standard critical values.

$n\backslash T$	Bias ($\times$ 100)				RMSE ($\times$ 100)				Size (5% level)				Power (5% level)			
	20	30	40	50	20	30	40	50	20	30	40	50	20	30	40	50
PB																
20	-3.69	-1.75	-1.07	-0.73	6.43	4.12	3.04	2.45	18.40	13.35	11.80	11.40	34.00	68.30	89.95	97.70
30	-3.39	-1.79	-1.04	-0.74	5.55	3.54	2.58	2.03	19.40	14.50	11.95	10.10	43.70	81.50	96.50	99.60
40	-3.56	-1.87	-1.06	-0.74	5.18	3.25	2.33	1.81	21.25	15.55	12.30	10.45	45.90	87.35	99.05	99.95
50	-3.58	-1.90	-1.09	-0.74	4.96	3.05	2.18	1.66	25.20	15.55	13.40	9.95	54.05	93.25	99.65	100.00
PMG																
20	-1.97	-0.89	-0.51	-0.32	7.77	4.79	3.40	2.57	39.45	28.15	21.40	17.85	63.40	82.20	93.75	98.95
30	-1.56	-0.97	-0.41	-0.33	6.32	3.99	2.78	2.08	41.10	28.45	22.50	16.55	71.20	89.60	98.10	99.85
40	-1.64	-0.86	-0.44	-0.31	5.71	3.44	2.47	1.85	43.10	29.25	23.05	18.10	77.25	95.15	99.60	100.00
50	-1.70	-0.93	-0.48	-0.31	5.23	3.10	2.26	1.67	42.25	28.60	23.75	16.85	81.00	97.00	99.80	100.00
PDOLS																
20	-5.60	-3.64	-2.82	-2.32	7.93	5.28	4.02	3.31	21.75	19.10	17.30	19.30	17.35	43.15	72.95	90.15
30	-5.25	-3.60	-2.75	-2.26	7.05	4.81	3.64	2.96	24.10	23.10	23.10	24.20	21.40	55.90	84.85	97.30
40	-5.47	-3.77	-2.84	-2.31	6.78	4.69	3.53	2.86	29.90	30.60	28.30	30.10	21.70	62.60	91.35	99.10
50	-5.46	-3.78	-2.88	-2.30	6.57	4.52	3.45	2.75	35.10	34.30	34.40	36.45	25.25	72.40	95.85	99.85
FMOLS																
20	-11.01	-7.16	-5.45	-4.25	12.56	8.44	6.55	5.18	89.25	78.25	69.90	64.15	45.00	56.60	79.05	92.60
30	-10.44	-7.06	-5.30	-4.17	11.58	7.99	6.09	4.83	93.80	86.15	77.90	71.95	44.90	63.20	88.75	98.20
40	-10.78	-7.31	-5.50	-4.26	11.59	8.00	6.08	4.77	97.45	92.95	85.65	82.20	44.60	67.85	92.70	99.20
50	-10.76	-7.35	-5.50	-4.22	11.44	7.91	5.98	4.65	98.70	96.25	91.35	86.85	46.10	73.70	95.35	99.85

Notes: DGP is given by $\Delta y_{it} = c_i - \alpha_i (y_{i,t-1} - \beta x_{i,t-1}) + u_{y,it}$ and $\Delta x_{it} = u_{x,it}$, for $i = 1, 2, ..., n$, $T = 1, 2, ..., T$, with $\beta = 1$ and $\alpha_i \sim IIDU[0.2, 0.3]$. Errors $u_{y,it}$, $u_{x,it}$ are cross-sectionally independent, heteroskedastic over i, and correlated over y & x equations. See Section 3.1 for complete description of the DGP. The pooled Bewley estimator is given by (8), with variance estimated using (16). PMG is the Pooled Mean Group estimator proposed by Pesaran, Shin, and Smith (1999). PDOLS is panel dynamic OLS estimator by Mark and Sul (2003). FMOLS is the group-mean fully modified OLS estimator by Pedroni (1996, 2001b). The size and power findings are computed using 5% nominal level and the reported power is the rejection frequency for testing the hypothesis $\beta = 0.9$.

Table 2: MC findings for the estimation of long-run coefficient β in experiments with cross-sectionally independent errors.

Bias corrected estimators and inference conducted using bootstrapped critical values.

$n\backslash T$	Bias (× 100)				RMSE (× 100)				Size (5% level)				Power (5% level)			
	20	**30**	**40**	**50**	**20**	**30**	**40**	**50**	**20**	**30**	**40**	**50**	**20**	**30**	**40**	**50**
Jackknife bias-corrected estimators																
PB																
20	-1.60	-0.52	-0.25	-0.14	6.27	4.18	3.10	2.52	5.95	5.15	5.40	5.30	25.10	56.75	82.95	94.30
30	-1.34	-0.55	-0.25	-0.18	5.24	3.45	2.58	2.06	6.70	5.05	5.10	5.00	37.45	73.70	93.10	99.05
40	-1.45	-0.60	-0.24	-0.15	4.59	3.03	2.29	1.81	5.85	5.05	4.50	4.90	42.80	84.10	98.15	99.85
50	-1.53	-0.62	-0.26	-0.15	4.23	2.72	2.07	1.61	6.20	4.60	5.65	4.25	51.10	90.80	99.55	100.00
PMG																
20	-0.55	-0.20	-0.05	0.01	9.26	5.59	3.78	2.86	14.90	11.05	9.10	7.20	33.45	57.15	83.05	95.60
30	-0.10	-0.30	0.02	-0.04	7.35	4.48	3.09	2.28	14.00	10.95	8.50	7.35	45.60	72.05	93.10	99.30
40	-0.27	-0.13	-0.01	-0.01	6.73	3.83	2.72	2.06	15.40	9.30	8.90	8.05	48.90	82.95	97.10	99.75
50	-0.42	-0.21	-0.06	0.00	6.13	3.43	2.47	1.83	16.45	9.65	9.40	7.55	57.55	88.55	98.85	100.00
PDOLS																
20	-4.22	-2.56	-1.97	-1.59	8.00	5.05	3.77	3.04	8.40	5.80	5.40	4.75	8.70	27.50	55.25	77.90
30	-3.91	-2.55	-1.92	-1.56	6.89	4.42	3.30	2.64	8.70	5.75	5.95	4.70	10.15	37.30	69.75	91.00
40	-4.09	-2.69	-1.98	-1.60	6.35	4.17	3.08	2.47	7.70	5.80	4.90	4.05	10.05	41.10	75.20	94.95
50	-4.11	-2.70	-2.02	-1.58	6.01	3.91	2.93	2.30	7.90	5.90	4.55	3.55	11.30	45.75	83.75	97.60
FMOLS																
20	-8.70	-5.07	-3.70	-2.76	11.19	7.19	5.52	4.33	10.85	5.95	4.55	4.00	1.05	1.40	8.10	25.75
30	-8.17	-5.02	-3.60	-2.73	10.02	6.60	4.94	3.86	9.95	5.00	4.65	3.00	0.50	1.65	10.20	35.20
40	-8.49	-5.22	-3.78	-2.78	9.84	6.40	4.79	3.68	9.20	5.00	3.45	3.25	0.15	0.75	8.85	38.70
50	-8.50	-5.28	-3.78	-2.74	9.61	6.26	4.62	3.50	10.65	5.00	3.75	2.85	0.10	0.95	10.55	45.75
Bootstrap bias-corrected estimators																
PB																
20	-1.28	-0.33	-0.16	-0.11	5.87	3.93	2.93	2.39	7.25	6.20	6.05	5.75	36.20	67.60	88.30	96.95
30	-0.98	-0.39	-0.15	-0.13	4.90	3.24	2.43	1.93	7.75	6.35	6.35	5.80	52.20	84.00	96.90	99.55
40	-1.07	-0.43	-0.13	-0.10	4.24	2.83	2.15	1.70	7.00	6.55	5.95	5.10	59.75	91.35	99.45	99.95
50	-1.09	-0.44	-0.15	-0.10	3.89	2.56	1.96	1.52	8.95	6.25	6.45	5.55	69.40	95.50	99.90	100.00
PMG																
20	-1.28	-0.44	-0.21	-0.11	7.88	4.83	3.41	2.57	14.10	10.45	7.90	6.80	35.25	63.40	86.70	97.30
30	-0.88	-0.55	-0.12	-0.13	6.40	3.99	2.79	2.08	13.10	10.75	8.25	6.85	47.45	77.75	95.85	99.65
40	-0.96	-0.44	-0.14	-0.11	5.73	3.42	2.47	1.85	14.95	9.05	7.80	6.90	53.55	87.60	98.40	99.85
50	-1.02	-0.51	-0.19	-0.10	5.20	3.06	2.24	1.66	15.15	9.85	9.05	6.95	60.90	91.65	99.30	100.00
PDOLS																
20	-2.19	-0.90	-0.59	-0.42	6.75	4.31	3.12	2.55	10.35	7.80	7.25	7.25	28.70	64.45	86.65	95.85
30	-1.89	-0.96	-0.59	-0.43	5.70	3.62	2.63	2.08	10.75	8.90	7.70	7.15	39.90	77.10	95.35	99.30
40	-2.00	-1.04	-0.59	-0.41	4.98	3.24	2.34	1.84	10.05	8.65	7.05	7.35	46.40	85.75	98.65	99.95
50	-2.00	-1.05	-0.64	-0.40	4.61	2.93	2.14	1.65	10.70	9.40	8.65	6.80	52.85	91.55	99.50	100.00
FMOLS																
20	-4.59	-1.97	-1.30	-0.80	8.84	5.62	4.28	3.40	17.10	10.90	9.20	7.40	20.20	41.10	63.20	81.70
30	-4.19	-2.04	-1.28	-0.84	7.54	4.86	3.61	2.83	18.50	10.90	9.15	7.70	24.85	51.60	78.60	93.15
40	-4.36	-2.14	-1.34	-0.81	6.89	4.36	3.22	2.51	20.10	12.05	9.80	7.75	26.10	60.40	86.25	96.80
50	-4.41	-2.22	-1.35	-0.78	6.56	4.10	2.98	2.29	23.80	14.35	10.50	8.90	30.45	69.45	92.05	99.20

Notes: See the notes to Table 1. Bias-corrected versions of the PB estimator are described in Subsection 2.2. Bias-corrected versions of the PMG, PDOLS and FMOLS estimator are described in Appendix B. Inference is conducted using bootstrapped critical values.

We consider next the bias-corrected versions of the four estimators with inference carried out using robust bootstrap critical values. Upper panel of Table 2 reports findings for estimators corrected for bias using the jackknife procedure, and the bottom panel reports on bootstrap bias corrected estimators. Bias correction did not change the overall ranking of estimators – PB continues to be the most precise (lowest RMSE). Both bias correction approaches are quite effective in reducing the bias. The bias of PB and PMG estimators for any of the two bias corrections are very low. In addition to reducing the bias, in many cases the bias-correction also resulted in reduced RMSE values. In the case of the PB estimator, using bootstrap bias correction resulted in improved RMSE performance for all choices of n, T - by about 2 to 22 percent. Results in Table 2 also show notable improvement to inference comes from using bootstrapped critical values - with PB having virtually no size distortions and size distortions of the remaining estimators are relatively minor.

Last but not least, we consider the DGP with cross-sectionally correlated errors. The corresponding results, reported in Tables B1 and B2 in Appendix B, reveal the same ranking of the four estimators, and, importantly, the bootstrapped critical values continue to deliver correct size, despite the error cross-sectional dependence.

The Monte Carlo results show that PB estimator can perform better (in terms of overall precision as measured by RMSE, and in terms of accuracy of inference) than existing estimators (PMG, PDOLS, and FMOLS) in finite sample sizes of interest, whether or not bias correction is considered. Of' course, our results do not imply that PB estimator will always be better, but that it can be a useful addition to the existing literature as a complement to PMG, PDOLS, and FMOLS estimators. Bias corrections and bootstrapping critical values are helpful for all four estimators, resulting not only in reduced bias, but sometimes also in better RMSE. In all cases, they result in more accurate inference in our experiments.

4 Empirical Application

This section revisits consumption function empirical application undertaken by Pesaran, Shin, and Smith (1999), hereafter PSS. The long-run consumption function is assumed to be given by

$$c_{it} = d_i + \beta_1 y_{it}^d + \beta_2 \pi_{it} + \vartheta_{it},$$

for country $i = 1, 2, ..., n$, where c_{it} is the logarithm of real consumption per capita, y_{it}^d is the logarithm of real per capita disposable income, π_{it} is the rate of inflation, and ϑ_{it} is an $I(0)$ process. We take the dataset from PSS, which consists of $n = 24$ countries and a slightly unbalanced time period covering 1960-1993. PSS estimate β_1 and β_2 using an ARDL(1,1,1) specification, which can be written as error-correcting panel regressions

$$\Delta c_{it} = -\alpha_i \left(c_{i,t-1} - d_i - \beta_1 y_{i,t-1}^d - \beta_2 \pi_{i,t-1} \right) + \delta_{i1} \Delta y_{it}^d + \delta_{i2} \Delta \pi_{it} + v_{it}, \tag{20}$$

for $i = 1, 2, ..., n$, where all coefficients, except the long-run coefficients β_1 and β_2 are country-specific.

Table 3 presents alternative estimates of the long-run coefficients. The upper panel presents findings for estimators without bias correction and standard confidence intervals. The middle and lower panels present jackknife and bootstrap bias-corrected estimates with confidence intervals based on bootstrapped critical values. Results differ widely across different approaches to estimation and inference. Depending on which bias correction approach is conducted, the PB estimates of the long-run coefficient on real income (β_1) is estimated to be 0.921 or 0.926, and the long-run coefficient on the inflation variable (β_2) is estimated to be -0.120 or -0.125. The null hypothesis that the coefficient on y_{it}^d is unity cannot be rejected at the 5 percent nominal level, nor is the hypothesis that the long run coefficient on inflation is zero. From an economic perspective, unit long-run real income elasticity and no long-run effects of inflation on consumption seem both plausible - the former hypothesis is in line with balanced growth path models, and the latter in line with monetary policy neutrality in the long-run. A different conclusion would be reached according to PMG estimates - namely both the unit coefficients on the real income variable and zero coefficient on inflation would be rejected at the 5 percent nominal level. The results based on the PDOLS are in line with the PB estimates and do not reject unit real income and zero inflation long run coefficients. FMOLS estimates of β_1 are larger than the other estimates, but the unit coefficient on the income variable still cannot be rejected. The FMOLS estimates of β_2 are also quite large. The choice of estimation method clearly matters in this empirical illustration.

Table 3: Estimated consumption function coefficients for OECD countries

	β_1: Income	95% Conf. Int.	β_2: Inflation	95% Conf. Int.
	Estimator without bias correction			
PB	.912	[.845,.980]	-.134	[-.260,-.008]
PMG	.904	[.889,.919]	-.466	[-.566,-.365]
PDOLS	.923	[.798,1.047]	-.187	[-.407,.033]
FMOLS	.951	[.942,.959]	-.336	[-.408,-.265]
	Jackknife bias-corrected estimators			
PB	.926	[.835,1.017]	-.120	[-.345,.105]
PMG	.915	[.880,.949]	-.403	[-616.,-.190]
PDOLS	.940	[.737,1.143]	-.184	[-.530,.161]
FMOLS	.983	[.912,1.053]	-.397	[-1.370,.576]
	Bootstrap bias-corrected estimators			
PB	.921	[.830,1.012]	-.125	[-.314,.065]
PMG	.905	[.875,.936]	-.477	[-.657,-.297]
PDOLS	.932	[.746,1.118]	-.183	[-.499,.133]
FMOLS	.985	[.941,1.028]	-.438	[-1.047,.171]

Notes: This table revisits empirical application in Table 1 of Pesaran, Shin, and Smith (1999), reporting estimates of long-run income elasticity (β_1) and inflation effect (β_2) coefficients and their 95% confidence intervals in the ARDL(1,1,1) consumption functions (20) for OECD countries using the dataset from Pesaran, Shin, and Smith (1999). PB stands for pooled Bewley estimator developed in this paper. PMG is the Pooled Mean Group estimator proposed by Pesaran, Shin, and Smith (1999). PDOLS is panel dynamic OLS estimator by Mark and Sul (2003). FMOLS is the group-mean fully modified OLS estimator by Pedroni (1996, 2001b). Description of bias correction methods is provided in Subsection 2.2 for PB estimator and in Appendix B for PMG, PDOLS and FMOLS estimators. Inference in the case of original estimators uncorrected for bias is conducted using the standard asymptotic critical values, and it is valid only when errors are not cross-sectionally dependent. Inference in the case of bias-corrected estimators is conducted using bootstrapped critical values following Chudik, Pesaran, and Smith (2023), and it is robust to cross-section dependence of errors.

5 Conclusion

This paper proposes the pooled Bewley (PB) estimator of long-run relationships in heterogeneous dynamic panels. Relative to existing estimators in the literature – namely PMG, PDOLS and FMOLS – Monte Carlo evidence reveals that PB can perform well in small samples. While we developed the asymptotic theory of PB estimator under a similar setting to the PMG estimator, notably we assumed cross-sectionally independent errors, we have also shown the benefit of bootstrapping critical values for inference when errors are cross-sectionally correlated for all four estimators.

While the asymptotic distribution of the other estimators are derived for the case where n is fixed and $T \to \infty$, we derive the joint (n, T) asymptotic distribution of the PB estimator, when both n and T diverge to infinity jointly such that $n = \Theta\left(T^\theta\right)$, for $0 < \theta < 2$. This covers a broader range of empirical applications where both n and T are large. The small sample and asymptotic

results suggest that the PB estimator is a useful addition to estimators for long run effects in single equation dynamic heterogeneous panels, where the direction of long-run causality is known.

References

Bewley, R. A. (1979). The direct estimation of the equilibrium response in a linear dynamic model. *Economics Letters 3*, 357–361.

Breitung, J. (2005). A parametric approach to the estimation of cointegration vectors in panel data. *Econometric Reviews 24*, 151–173.

Chudik, A. and M. H. Pesaran (2013). Econometric analysis of high dimensional VARs featuring a dominant unit. *Econometric Reviews 32*, 592–649.

Chudik, A., M. H. Pesaran, and R. P. Smith (2023). Revisiting the Great Ratios Hypothesis. *Oxford Bulletin of Economics and Statistics 85*, 1023–1047.

Chudik, A., M. H. Pesaran, and J.-C. Yang (2018). Half-panel jackknife fixed-effects estimation of linear panels with weakly exogenous regressors. *Journal of Applied Econometrics 33*, 816–836.

Dhaene, G. and K. Jochmans (2015). Split-panel jackknife estimation of fixed-effect models. *Review of Economic Studies 82*, 991–1030.

Liu, R. (1988). Bootstrap procedure under some non-i.i.d. models. *Ann. Statist. 16.* 1696–1708.

Maddala, G. S. and S. Wu (1999). A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. *Oxford Bulletin of Economics and Statistics 61*, 603–767.

Mark, N. C. and D. Sul (2003). Cointegration vector estimation by panel DOLS and long-run money demand. *Oxford Bulletin of Economics and Statistics 65*, 655–680.

Pedroni, P. (1996). Fully Modified OLS for Heterogeneous Cointegrated Panels and the Case of Purchasing Power Parity. Indiana University working papers in economics no. 96-020 (June 1996).

Pedroni, P. (2001a). Fully modified OLS for heterogeneous cointegrated panels. In B. Baltagi, T. Fomby, and R. C. Hill (Eds.), *Nonstationary Panels, Panel Cointegration, and Dynamic Panels, (Advances in Econometrics, Vol. 15)*, pp. 93–130. Emerald Group Publishing Limited, Bingley.

Pedroni, P. (2001b). Purchasing Power Parity Tests in Cointegrated Panels. *The Review of Economics and Statistics 83*, 727–731.

Pesaran, M. H., Y. Shin, and R. P. Smith (1999). Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American Statistical Association 94*, 621–634.

Phillips, P. C. B. and H. R. Moon (1999). Linear regression limit theory for nonstationary panel data. *Econometrica 67*, 1057–1011.

Wickens, M. R. and T. S. Breusch (1988). Dynamic specification, the long-run and the estimation of transformed regression models. *The Economic Journal 98*, 189–205.

Appendices

A   Mathematical derivations

This appendix is organized in four sections. Section A.1 introduces some notations and definitions. Section A.2 presents lemmas and proofs needed for the proof of Theorem 1 presented in the body of the paper.

A.1   Notations and definitions

Let $\mathbf{z}_{it} = (y_{it}, x_{it})'$, and define $\mathbf{C}_i(L) = \sum_{\ell=0}^{\infty} \mathbf{C}_{i\ell} L^{\ell}$ and $\mathbf{C}_i^*(L) = \sum_{\ell=0}^{\infty} \mathbf{C}_{i\ell}^* L^{\ell}$, where

$$\mathbf{C}_{i0} = \mathbf{I}_2,$$
$$\mathbf{C}_{i\ell} = (\boldsymbol{\Phi}_i - \mathbf{I}_2)\boldsymbol{\Phi}_i^{\ell-1}, \ \ell = 1, 2, ...,$$

$$\boldsymbol{\Phi}_i = \begin{pmatrix} 1 - \alpha_i & \alpha_i \beta \\ 0 & 1 \end{pmatrix}, \tag{A.1}$$

$$\mathbf{C}_i(1) = \mathbf{C}_{i0} + \mathbf{C}_{i1} + ... = \lim_{\ell \to \infty} \boldsymbol{\Phi}_i^{\ell} = \begin{pmatrix} 0 & \beta \\ 0 & 1 \end{pmatrix},$$

and

$$\mathbf{C}_{i0}^* = \mathbf{C}_{i0} - \mathbf{C}_i(1) = \begin{pmatrix} 1 & -\beta \\ 0 & 0 \end{pmatrix},$$
$$\mathbf{C}_{i\ell}^* = \mathbf{C}_{i,\ell-1}^* + \mathbf{C}_{i\ell} = \begin{pmatrix} (1-\alpha_i)^{\ell} & -(1-\alpha_i)^{\ell}\beta \\ 0 & 0 \end{pmatrix}, \ \text{for } \ell = 1, 2, ...$$

Model (1)-(2) can be equivalently written as

$$\boldsymbol{\Phi}_i(L)\mathbf{z}_{it} = \mathbf{c}_i + \mathbf{u}_{it},$$

for $i = 1, 2, ..., n$ and $t = 1, 2, ..., T$, where $\mathbf{c}_i = (c_i, 0)'$,

$$\boldsymbol{\Phi}_i(L) = \mathbf{I}_2 - \boldsymbol{\Phi}_i L, \tag{A.2}$$

and $\mathbf{I}_2$ is a 2×2 identity matrix. The lag polynomial $\boldsymbol{\Phi}_i(L)$ can be re-written in the following (error correcting) form

$$\boldsymbol{\Phi}_i(L) = -\Pi_i L + (1 - L)\mathbf{I}_2, \tag{A.3}$$

where

$$\Pi_i = -(\mathbf{I}_2 - \boldsymbol{\Phi}_i) = \begin{pmatrix} -\alpha_i & \alpha_i \beta \\ 0 & 0 \end{pmatrix}. \tag{A.4}$$

The VAR model (A.5) can be also rewritten in the following form

$$\boldsymbol{\Phi}_i(L)(\mathbf{z}_{it} - \boldsymbol{\mu}_i) = \mathbf{u}_{it}, \tag{A.5}$$

where $\mathbf{c}_i = -\Pi_i \boldsymbol{\mu}_i = (c_i, 0)'$, namely $c_i = \alpha_i \mu_{i,1} - \alpha_i \beta \mu_{i,2}$.

Using Granger representation theorem, the process $\mathbf{z}_{it}$ under the assumptions 1-3 has representation

$$
\begin{aligned}
y_{it} &= \mu_{yi} + \beta s_{it} + \sum_{\ell=0}^{\infty} (1 - \alpha_i)^{\ell} \left(u_{y,i,t-\ell} - \beta u_{x,i,t-\ell} \right), & \text{(A.6)} \\
x_{it} &= \mu_{xi} + s_{it}, & \text{(A.7)}
\end{aligned}
$$

where

$$
s_{it} = \sum_{\ell=1}^{t} u_{x,it}, \tag{A.8}
$$

is the stochastic trend.

A.2 Lemmas: Statements and proofs

Lemma A.1 *Suppose Assumptions 2 and 3 hold, and consider $\tilde{\mathbf{x}}_i = (\tilde{x}_{i,1}, \tilde{x}_{i,2}, ..., \tilde{x}_{i,T})'$, where $\tilde{x}_{it} = x_{it} - \bar{x}_i$, $x_{it} = \sum_{s=1}^{t} u_{x,it}$, and $\bar{x}_i = T^{-1} \sum_{t=1}^{T} x_{it}$. Then*

$$
n^{-1} \sum_{i=1}^{n} \frac{\tilde{\mathbf{x}}_i' \tilde{\mathbf{x}}_i}{T^2} \to_p \omega_x^2 = \frac{\sigma_x^2}{6}, \text{ as } n, T \to \infty, \tag{A.9}
$$

where $\sigma_x^2 = \lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} \sigma_{xi}^2$.

Proof. Recall that $\mathbf{M}_\tau = \mathbf{I}_T - T^{-1} \boldsymbol{\tau}_T \boldsymbol{\tau}_T'$, where $\mathbf{I}_T$ is $T \times T$ identity matrix and $\boldsymbol{\tau}_T$ is $T \times 1$ vector of ones. Since $\tilde{\mathbf{x}}_i = \mathbf{M}_\tau \mathbf{x}_i$, and $\mathbf{M}_\tau$ is symmetric and idempotent $(\mathbf{M}_\tau' \mathbf{M}_\tau = \mathbf{M}_\tau = \mathbf{M}_\tau')$ we can write $\tilde{\mathbf{x}}_i' \tilde{\mathbf{x}}_i$ as $\tilde{\mathbf{x}}_i' \tilde{\mathbf{x}}_i = \mathbf{x}_i' \mathbf{M}_\tau' \mathbf{M}_\tau \mathbf{x}_i = \mathbf{x}_i' \mathbf{M}_\tau' \mathbf{x}_i = \tilde{\mathbf{x}}_i' \mathbf{x}_i$. Denote $S_{i,T} = \tilde{\mathbf{x}}_i' \mathbf{x}_i / T^2$. We have

$$
n^{-1} \sum_{i=1}^{n} \frac{\tilde{\mathbf{x}}_i' \tilde{\mathbf{x}}_i}{T^2} = n^{-1} \sum_{i=1}^{n} S_{i,T} = n^{-1} \sum_{i=1}^{n} E(S_{i,T}) + n^{-1} \sum_{i=1}^{n} [S_{i,T} - E(S_{i,T})]. \tag{A.10}
$$

Consider $E(S_{i,T})$ first. Noting that $\tilde{x}_{it} = \sum_{s=1}^{t} u_{x,it} - \bar{x}_i$, $\bar{x}_i = T^{-1} \sum_{s=1}^{T} (T - s + 1) u_{st}$, and $x_{it} = \sum_{s=1}^{t} u_{x,it}$, $S_{i,T}$ can be written as

$$
\begin{aligned}
S_{i,T} &= \frac{1}{T^2} \sum_{t=1}^{T} \tilde{x}_{it} x_{it}, \\
&= \frac{1}{T^2} \sum_{t=1}^{T} \left[\left(\sum_{s=1}^{t} u_{x,is} \right)^2 - \bar{x}_i \sum_{s=1}^{t} u_{x,is} \right], \\
&= \frac{1}{T^2} \sum_{t=1}^{T} \left[\left(\sum_{s=1}^{t} u_{x,is} \right)^2 - \sum_{s=1}^{t} \frac{T - s + 1}{T} u_{x,is} \cdot \sum_{s=1}^{t} u_{x,is} \right].
\end{aligned}
$$

Taking expectations, we obtain

$$
E(S_{i,T}) = \frac{\sigma_{xi}^2}{T^2} \sum_{t=1}^{T} \left[t - \sum_{s=1}^{t} \frac{T - s + 1}{T} \right].
$$

Using $\sum_{s=1}^{t} \frac{T - s + 1}{T} = \sum_{s=1}^{t} (1 - s/T + 1/T) = t - (t + 1) t / (2T) + t/T$, we have

$$
E(S_{i,T}) = \frac{\sigma_{xi}^2}{T^2} \sum_{t=1}^{T} \left[t - t + \frac{(t + 1) t}{2T} - \frac{t}{T} \right] = \frac{\sigma_{xi}^2}{T^2} \sum_{t=1}^{T} \frac{(t + 1) t}{2T} - \frac{t}{T}.
$$

20

Finally, noting that $\sum_{t=1}^{T} (t+1) t = (T+2) (T+1) T/3$, and $\sum_{t=1}^{T} t = (T+1) T/2$, we obtain

$$E(S_{i,T}) = \sigma_{xi}^2 \varkappa_T < K < \infty, \tag{A.11}$$

for all $T > 0$, where

$$\varkappa_T = \left[\frac{(T+2)(T+1)T}{6T^3} - \frac{(T+1)T}{2T^3} \right]. \tag{A.12}$$

In addition, $\varkappa_T \to 1/6$, as $T \to \infty$, and

$$\frac{1}{n} \sum_{i=1}^{n} E(S_{i,T}) = \varkappa_T \frac{1}{n} \sum_{i=1}^{n} \sigma_{xi}^2 \to \frac{\sigma_x^2}{6},$$

as $n, T \to \infty$. This establishes the limit of the first term on the right side of (A.10). Consider the second term next. Since $E[S_{i,T} - E(S_{i,T})] = 0$, and $S_{i,T}$ is independent over i, we have

$$E\left\{ n^{-1} \sum_{i=1}^{n} [S_{i,T} - E(S_{i,T})] \right\}^2 = \frac{1}{n^2} \sum_{i=1}^{n} E(S_{i,T}^2) - \frac{1}{n^2} \sum_{i=1}^{n} [E(S_{i,T})]^2.$$

But it follows from (A.11) that there exist finite positive constant $K_1 < \infty$ (which does not depend on n, T) such that $[E(S_{i,T})]^2 < K_1$. In addition, due to existence of uniformly bounded fourth moments of $u_{x,it}$, it also can be shown that $E(S_{i,T}^2) < K_2 < \infty$. Hence, $E\left\{ n^{-1} \sum_{i=1}^{n} [S_{i,T} - E(S_{i,T})] \right\}^2 = O(n^{-1})$, which implies $n^{-1} \sum_{i=1}^{n} [S_{i,T} - E(S_{i,T})] \to_p 0$, as $n, T \to \infty$. This completes the proof. ∎

Lemma A.2 *Suppose Assumptions 1-2 hold. Then there exists finite positive constant K that does not depend on i and/or T such that*

$$E\left(\frac{1}{T} \sum_{t=1}^{T} u_{x,it} \tilde{x}_{it} \right)^\varrho < K, \tag{A.13}$$

and

$$E\left(\frac{1}{T} \sum_{t=1}^{T} \Delta y_{it} \tilde{x}_{it} \right)^\varrho < K, \tag{A.14}$$

for $\varrho = 4$, where $\tilde{x}_{it} = x_{it} - \bar{x}_i$, $x_{it} = \sum_{s=1}^{t} u_{x,it}$, $\bar{x}_i = T^{-1} \sum_{t=1}^{T} x_{it}$, and $\Delta y_{it} = \delta_i u_{x,it} + v_{it} - \alpha_i \sum_{\ell=1}^{\infty} (1-\alpha_i)^{\ell-1} [v_{i,t-\ell} + (\delta_i - \beta) u_{x,i,t-\ell}].$

Proof. Consider $\sum_{t=1}^{T} u_{it} \tilde{x}_{it}/T$ and $\varrho = 2$ first, and note that $\tilde{x}_{it} = \sum_{s=1}^{t} u_{x,is} - \bar{x}_i$, where $\bar{x}_i = T^{-1} \sum_{s=1}^{T} (T - s + 1) u_{x,is}$. We have

$$\left(\frac{1}{T} \sum_{t=1}^{T} u_{x,it} \tilde{x}_{it} \right)^2 = \frac{1}{T^2} \sum_{t=1}^{T} \sum_{t'=1}^{T} u_{x,it} u_{x,it'} \tilde{x}_{it} \tilde{x}_{it'},$$

$$= \frac{1}{T^2} \sum_{t=1}^{T} \sum_{t'=1}^{T} u_{x,it} u_{x,it'} \left(\sum_{s=1}^{t} u_{x,is} - \bar{x}_i \right) \left(\sum_{s=1}^{t'} u_{x,is} - \bar{x}_i \right),$$

$$= A_{i,T,1} + A_{i,T,2} - A_{i,T,3} - A_{i,T,4},$$

where

$$A_{i,T,1} = \frac{1}{T^2} \sum_{t=1}^{T} \sum_{t'=1}^{T} u_{x,it} u_{x,it'} \left(\sum_{s=1}^{t} u_{x,is} \right) \left(\sum_{s=1}^{t'} u_{x,is} \right),$$

$$A_{i,T,2} = \frac{1}{T^2} \sum_{t=1}^{T} \sum_{t'=1}^{T} u_{x,it} u_{x,it'} \bar{x}_i^2,$$

$$A_{i,T,3} = \frac{1}{T^2} \sum_{t=1}^{T} \sum_{t'=1}^{T} u_{x,it} u_{x,it'} \bar{x}_i \sum_{s=1}^{t} u_{x,is},$$

$$A_{i,T,4} = \frac{1}{T^2} \sum_{t=1}^{T} \sum_{t'=1}^{T} u_{x,it} u_{x,it'} \bar{x}_i \sum_{s=1}^{t'} u_{x,is}.$$

Taking expectations and noting that $u_{x,it}$ is independent of $u_{x,it'}$ for any $t \neq t'$, we have

$$E(A_{i,T,1}) = \frac{1}{T^2} \left(\sum_{t=1}^{T} \sum_{t'=1}^{t-1} \sigma_{ix}^2 + \sum_{t=1}^{T} E\left(u_{x,it}^4 \right) + \sum_{t=1}^{T} \sum_{t'=1}^{t-1} \sigma_{ix}^2 \right).$$

Under Assumption 2, there exists a finite constant K that does not depend on i and/or t, such that $\sigma_{ix}^2 < K$ and $E\left(u_{x,it}^4 \right) < K$. Hence $|E(A_{i,T,1})| < K$. Similarly, we can bound the remaining elements, $|E(A_{i,T,j})| < K$, for $j = 2, 3, 4$. It now follows that $E\left(\frac{1}{T} \sum_{t=1}^{T} u_{x,it} \tilde{x}_{it} \right)^2 < K$, where the upper bound K does not depend on i or T. This establishes (A.13) hold for $\varrho = 2$. Sufficient conditions for (A.13) to hold when $\varrho = 4$ are: $E\left(A_{i,T,j}^2 \right) < K$ for $j = 1, 2, 3, 4$. These conditions follow from uniformly bounded eights moments of $u_{x,it}$. This completes the proof of (A.13). Result (A.14) can be established in the same way by using the first difference of representation (A.6). ∎

Lemma A.3 *Suppose Assumptions 1-4 hold, and consider s_{iT} given by*

$$s_{iT} = \tilde{\mathbf{x}}_i' \Delta \tilde{\mathbf{Z}}_i \left(\Delta \tilde{\mathbf{Z}}_i' \mathbf{P}_i \Delta \tilde{\mathbf{Z}}_i \right)^{-1} \Delta \tilde{\mathbf{Z}}_i' \tilde{\mathbf{x}}_i, \tag{A.15}$$

where $\mathbf{P}_i$ is given by (10), and $\tilde{\mathbf{x}}_i$ and $\Delta \tilde{\mathbf{Z}}_i$ are defined below (6). Then,

$$n^{-1} \sum_{i=1}^{n} \frac{s_{iT}}{T^2} \to_p 0, \text{ as } n, T \to \infty. \tag{A.16}$$

Proof. Consider $s_{i,T}/T$, which can be written as

$$\frac{s_{iT}}{T} \leq \mathbf{a}_{iT}' \mathbf{B}_{iT}^{-1} \mathbf{a}_{iT}, \tag{A.17}$$

where

$$\mathbf{a}_{iT} = \frac{\Delta \tilde{\mathbf{Z}}_i' \tilde{\mathbf{x}}_i}{T} = \frac{\Delta \mathbf{Z}_i' \tilde{\mathbf{x}}_i}{T}, \tag{A.18}$$

and

$$\mathbf{B}_{iT} = \frac{\Delta \tilde{\mathbf{Z}}_i' \mathbf{P}_i \Delta \tilde{\mathbf{Z}}_i}{T}. \tag{A.19}$$

Using these notations, we have

$$E \left| \frac{1}{n} \sum_{i=1}^{n} \frac{s_{iT}}{T^2} \right| \le \frac{1}{nT} \sum_{i=1}^{n} E \left| \mathbf{a}_{iT}' \mathbf{B}_{iT}^{-1} \mathbf{a}_{iT} \right|.$$

Using $\left| \mathbf{a}_{iT}' \mathbf{B}_{iT}^{-1} \mathbf{a}_{iT} \right| \le \lambda_{\min}^{-1} (\mathbf{B}_{iT}) \mathbf{a}_{iT}' \mathbf{a}_{iT}$, and Cauchy-Schwarz inequality, we obtain

$$E \left| \frac{1}{n} \sum_{i=1}^{n} \frac{s_{iT}}{T^2} \right| \le \frac{1}{nT} \sum_{i=1}^{n} \sqrt{E \left[(\mathbf{a}_{iT}' \mathbf{a}_{iT})^2 \right]} \sqrt{E \left[\lambda_{\min}^{-2} (\mathbf{B}_{iT}) \right]}.$$

Lemma A.2 implies the fourth moments of the individual elements of $\mathbf{a}_{i,T}$ are uniformly bounded in i and T, which is sufficient for $E \left[(\mathbf{a}_{iT}' \mathbf{a}_{iT})^2 \right] < K$. In addition, $E \left[\lambda_{\min}^{-2} (\mathbf{B}_{iT}) \right] < K$ by Assumption 4. Hence, there exists $K < \infty$, which does not depend on (n, T) such that

$$E \left| n^{-1} \sum_{i=1}^{n} \frac{s_{iT}}{T^2} \right| < \frac{K}{T}, \tag{A.20}$$

and result (A.16) follows. ∎

Lemma A.4 *Suppose Assumptions 1-4 hold. Then*

$$n^{-1} \sum_{i=1}^{n} \frac{\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{x}}_i}{T^2} \to_p \omega_x^2 = \frac{\sigma_x^2}{6}, \text{ as } n, T \to \infty, \tag{A.21}$$

where $\sigma_x^2 = \lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} \sigma_{xi}^2$, $\mathbf{M}_i$ is defined in (9) and $\tilde{\mathbf{x}}_i$ is defined below (6).

Proof. Noting that $\tilde{\mathbf{x}}_i$ is one of the column vectors of $\mathbf{H}_i$, we have $\mathbf{P}_i \tilde{\mathbf{x}}_i = \tilde{\mathbf{x}}_i$, and $\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{x}}_i$ can be written as

$$\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{x}}_i = \tilde{\mathbf{x}}_i' \tilde{\mathbf{x}}_i - s_{i,T}, \tag{A.22}$$

where $s_{i,T}$ is given by (A.15). Sufficient conditions for result (A.21) are:

$$n^{-1} \sum_{i=1}^{n} \frac{\tilde{\mathbf{x}}_i' \tilde{\mathbf{x}}_i}{T^2} \to_p \omega_x^2 = \frac{\sigma_x^2}{6}, \text{ as } n, T \to \infty, \tag{A.23}$$

and

$$n^{-1} \sum_{i=1}^{n} \frac{s_{i,T}}{T^2} \to_p 0, \text{ as } n, T \to \infty. \tag{A.24}$$

Condition (A.23) is established by Lemma A.1, and condition (A.24) is established by Lemma A.3. ∎

Lemma A.5 *Let Assumptions 1-3 hold. Then*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\tilde{\mathbf{x}}_i' \tilde{\mathbf{v}}_i}{T \alpha_i} \to_d N \left(0, \omega_v^2 \right), \text{ as } n, T \to \infty, \tag{A.25}$$

where $\omega_v^2 = \lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} \sigma_{xi}^2 \sigma_{vi}^2 / \left(6\alpha_i^2 \right)$, and $\tilde{\mathbf{x}}_i$ and $\tilde{\mathbf{v}}_i$ are defined below (6).

Proof. Recall $\mathbf{M}_\tau = \mathbf{I}_T - T^{-1} \boldsymbol{\tau}_T \boldsymbol{\tau}_T'$, where $\mathbf{I}_T$ is $T \times T$ identity matrix and $\boldsymbol{\tau}_T$ is $T \times 1$ vector of ones. Since $\mathbf{M}_\tau' \mathbf{M}_\tau = \mathbf{M}_\tau'$, we have

$$\tilde{\mathbf{x}}_i' \tilde{\mathbf{v}}_i = \mathbf{x}_i' \mathbf{M}_\tau' \mathbf{M}_\tau \mathbf{v}_i = \mathbf{x}_i' \mathbf{M}_\tau' \mathbf{v}_i = \tilde{\mathbf{x}}_i' \mathbf{v}_i.$$

Let $C_i = \frac{\sigma_{xi}\sigma_{vi}}{\alpha_i}$ and $Q_{i,T} = C_i^{-1}\frac{\tilde{\mathbf{x}}_i'\mathbf{v}_i}{T\alpha_i}$. We have $E(Q_{i,T}) = 0$, and (under independence of v_{it} over t and independence of v_{it} and $u_{x,it'}$ for any t,t')

$$E\left[\left(\frac{\tilde{\mathbf{x}}_i'\mathbf{v}_i}{T\alpha_i}\right)^2\right] = \frac{1}{T^2\alpha_i^2}\sum_{t=1}^{T}E\left(\tilde{x}_{it}^2\right)E\left(v_{it}^2\right),$$

where $E\left(v_{it}^2\right) = \sigma_{vi}^2$. In addition, (A.11) established that $\frac{1}{T^2}\sum_{t=1}^{T}E\left(\tilde{x}_{it}^2\right) = \sigma_{xi}^2\varkappa_T$, where $\varkappa_T$ is given by (A.12). Hence,

$$E\left[\left(\frac{\tilde{\mathbf{x}}_i'\mathbf{v}_i}{T\alpha_i}\right)^2\right] = \frac{\sigma_{vi}^2\sigma_{xi}^2}{\alpha_i^2}\varkappa_T = C_i^2\varkappa_T.$$

It follows that

$$E\left(Q_{i,T}^2\right) = \varkappa_T,$$

where $\varkappa_T \to 1/6 < \infty$. Finite fourth moments of $u_{x,it}$ and v_{it} imply $Q_{i,T}^4$ is uniformly bounded in T, and therefore $Q_{i,T}^2$ is uniformly integrable in T. We can apply Theorem 3 of Phillips and Moon (1999) to obtain

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}C_iQ_{i,T} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\tilde{\mathbf{x}}_i'\mathbf{v}_i}{T\alpha_i} \to_d N\left(0,\omega_v^2\right), \text{ as } n,T\to\infty,$$

where $\omega_v^2 = \lim_{n\to\infty}C_i^2\varkappa_T = \lim_{n\to\infty}n^{-1}\sum_{i=1}^{n}\sigma_{xi}^2\sigma_{vi}^2/\left(6\alpha_i^2\right)$. $\blacksquare$

Lemma A.6 *Suppose Assumptions 1-4 hold, and consider* $\mathbf{q}_{iT} = \alpha_i^{-1}\Delta\tilde{\mathbf{Z}}_i'\mathbf{P}_i\tilde{\mathbf{v}}_i/\sqrt{T}$. *Then,*

$$E\|\mathbf{q}_{iT}\|_2^4 < K, \tag{A.26}$$

and

$$|E(\mathbf{q}_{iT})| < \frac{K}{\sqrt{T}}. \tag{A.27}$$

Proof. Denote the individual elements of 2×1 vector $\mathbf{q}_{iT}$ as $q_{iT,j}$, $j = 1,2$. Sufficient conditions for (A.26) to hold are

$$E\left(q_{iT,j}\right)^4 < K, \text{ for } j = 1,2. \tag{A.28}$$

We establish (A.28) for $j = 1$ first. We have

$$q_{iT,1} = \frac{\Delta\tilde{\mathbf{y}}_i'\mathbf{P}_i\tilde{\mathbf{v}}_i}{\alpha_i\sqrt{T}},$$

where $\Delta\mathbf{y}_i$ can be written as

$$\Delta\tilde{\mathbf{y}}_i = -\alpha_i\tilde{\boldsymbol{\xi}}_{i,-1} + \delta_i\Delta\tilde{\mathbf{x}}_i + \mathbf{v}_i, \tag{A.29}$$

where $\tilde{\boldsymbol{\xi}}_{i,-1} = \tilde{\mathbf{y}}_{i,-1} - \tilde{\mathbf{x}}_{i,-1}$. Note that $\mathbf{P}_i = \tilde{\mathbf{H}}_i\left(\tilde{\mathbf{H}}_i'\tilde{\mathbf{H}}_i\right)^{-1}\tilde{\mathbf{H}}_i'$ and $\tilde{\mathbf{H}}_i = (\tilde{\mathbf{y}}_{i,-1},\tilde{\mathbf{x}}_i,\tilde{\mathbf{x}}_{i,-1})$. Hence $\Delta\tilde{\mathbf{x}}_i'\mathbf{P}_i = \Delta\tilde{\mathbf{x}}_i'$ and $\tilde{\boldsymbol{\xi}}_{i,-1}'\mathbf{P}_i = \tilde{\boldsymbol{\xi}}_{i,-1}'$, since $\Delta\tilde{\mathbf{x}}_i$ and $\tilde{\boldsymbol{\xi}}_{i,-1}$ can be both obtained as a linear combinations of the column vectors of $\tilde{\mathbf{H}}_i$. Hence

$$q_{iT,1} = -\frac{\tilde{\boldsymbol{\xi}}_{i,-1}'\tilde{\mathbf{v}}_i}{\sqrt{T}} + \frac{\Delta\tilde{\mathbf{x}}_i'\tilde{\mathbf{v}}_i}{\alpha_i\sqrt{T}} + \frac{\tilde{\mathbf{v}}_i'\mathbf{P}_i\tilde{\mathbf{v}}_i}{\alpha_i\sqrt{T}} \equiv \varsigma_{a,iT} + \varsigma_{b,iT} + \varsigma_{c,iT}, \tag{A.30}$$

where we simplified notations by introducing $\varsigma_{a,iT} = -\tilde{\boldsymbol{\xi}}_{i,-1}'\tilde{\mathbf{v}}_i/\sqrt{T}$, $\varsigma_{b,iT} = \alpha_i^{-1}\Delta\tilde{\mathbf{x}}_i'\tilde{\mathbf{v}}_i/\sqrt{T}$ and $\varsigma_{c,iT} = $

24

$\alpha_i^{-1}\Delta\tilde{\mathbf{v}}_i'\mathbf{P}_i\tilde{\mathbf{v}}_i/\sqrt{T}$ to denote the individual terms in the expression (A.30) for $q_{iT,1}$. Sufficient conditions for $E\left(q_{iT,1}^4\right) < K$ are $E\left(\varsigma_{s,iT}^4\right) < K$ for $s \in \{a,b,c\}$.

For $s = a$, we have

$$\varsigma_{a,iT} = -\frac{\tilde{\boldsymbol{\xi}}_{i,-1}'\tilde{\mathbf{v}}_i}{\sqrt{T}} = -\frac{1}{\sqrt{T}}\sum_{i=1}^{T}\left(\xi_{i,t-1} - \bar{\xi}_{i,-1}\right)\left(v_{it} - \bar{v}_i\right) = -\frac{1}{\sqrt{T}}\sum_{i=1}^{T}\xi_{i,t-1}v_{it} + \sqrt{T}\bar{\xi}_{i,-1}\bar{v}_i,$$

where $\bar{\xi}_{i,-1} = T^{-1}\sum_{t=1}^{T}\xi_{i,t-1}$, and

$$
\begin{aligned}
\xi_{it} &= \sum_{\ell=0}^{\infty}(1-\alpha_i)^\ell\left(u_{y,i,t-\ell} - \beta u_{x,i,t-\ell}\right), \\
&= \sum_{\ell=0}^{\infty}(1-\alpha_i)^\ell\left(\delta_i - \beta\right)u_{x,i,t-\ell} + \sum_{\ell=0}^{\infty}(1-\alpha_i)^\ell v_{it}.
\end{aligned}
$$

Noting that $\sup_i|1-\alpha_i| < 1$ under Assumption 1, and fourth moments of $u_{x,i,t}$ and eights moments of v_{it} are bounded, we obtain

$$E\left[\left(\frac{1}{\sqrt{T}}\sum_{i=1}^{T}\xi_{i,t-1}v_{it}\right)^4\right] \leq K,$$

and

$$T^2 \cdot E\left(\bar{\xi}_{i,-1}^4\bar{v}_i^4\right) \leq K,$$

which are sufficient conditions for $E\left(\varsigma_{a,iT}^4\right) \leq K$.

For $s = b$, we have

$$\varsigma_{b,iT} = \frac{\Delta\tilde{\mathbf{x}}_i'\tilde{\mathbf{v}}_i}{\alpha_i\sqrt{T}} = \frac{1}{\alpha_i\sqrt{T}}\sum_{t=1}^{T}\left(u_{x,it} - \bar{u}_{x,i}\right)\left(v_{it} - \bar{v}_i\right) = \frac{1}{\alpha_i\sqrt{T}}\sum_{t=1}^{T}u_{x,it}v_{it} - \frac{\sqrt{T}}{\alpha_i}\bar{u}_{x,i}\bar{v}_i.$$

Using Assumption 2, we obtain the following upper bound

$$E\left(\varsigma_{b,iT}^4\right) \leq \left|\alpha_i^{-4}\right|\frac{1}{T}\sum_{t=1}^{T}E\left(u_{x,it}^4\right)E\left(v_{it}^4\right) + \left|\alpha_i^{-4}\right|TE\left(\bar{u}_{x,i}^4\right)E\left(\bar{v}_i^4\right) \leq K, \tag{A.31}$$

where $\left|\alpha_i^{-4}\right| < K$, $E\left(u_{x,it}^4\right) < K$, $E\left(v_{it}^4\right) < K$, $E\left(\bar{u}_{x,i}^4\right) < K/T^2$, and $E\left(\bar{v}_i^4\right) < K/T^2$.

For $s = c$, we have

$$\varsigma_{c,iT} = \frac{\tilde{\mathbf{v}}_i'\mathbf{P}_i\tilde{\mathbf{v}}_i}{\alpha_i\sqrt{T}} = \frac{\Delta\tilde{\mathbf{v}}_i'\tilde{\mathbf{H}}_i\left(\tilde{\mathbf{H}}_i'\tilde{\mathbf{H}}_i\right)^{-1}\tilde{\mathbf{H}}_i'\tilde{\mathbf{v}}_i}{\alpha_i\sqrt{T}}.$$

Consider $\tilde{\mathbf{H}}_i^* = \left(\tilde{\mathbf{x}}_i, \Delta\tilde{\mathbf{x}}_i, \tilde{\boldsymbol{\xi}}_{i,-1}\right)$ and note that

$$\tilde{\mathbf{H}}_i = \left(\tilde{\mathbf{y}}_{i,-1}, \tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_{i,-1}\right) = \mathbf{B}^*\tilde{\mathbf{H}}_i^*,$$

where

$$\mathbf{B}^* = \begin{pmatrix} -\beta & \beta & 1 \\ 1 & 0 & 0 \\ -1 & 1 & 0 \end{pmatrix},$$

25

is nonsingular (for any β). Hence $\mathbf{P}_i = \tilde{\mathbf{H}}_i \left(\tilde{\mathbf{H}}_i' \tilde{\mathbf{H}}_i \right)^{-1} \tilde{\mathbf{H}}_i' = \tilde{\mathbf{H}}_i^* \left(\tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{H}}_i^* \right)^{-1} \tilde{\mathbf{H}}_i^{*\prime}$, and we can write $\varsigma_{c,iT}$ as

$$\varsigma_{c,iT} = \frac{\tilde{\mathbf{v}}_i' \tilde{\mathbf{H}}_i^* \left(\tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{H}}_i^* \right)^{-1} \tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{v}}_i}{\alpha_i \sqrt{T}}.$$

Consider the scaling matrix

$$\mathbf{A}_T = \begin{pmatrix} T^{-1} & 0 & 0 \\ 0 & T^{-1/2} & 0 \\ 0 & 0 & T^{-1/2} \end{pmatrix}. \tag{A.32}$$

We have

$$\varsigma_{c,iT} = \frac{1}{\alpha_i \sqrt{T}} \tilde{\mathbf{v}}_i' \tilde{\mathbf{H}}_i^* \mathbf{A}_T \left(\mathbf{A}_T \tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{H}}_i^* \mathbf{A}_T \right)^{-1} \mathbf{A}_T \tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{v}}_i \geq 0.$$

Using the inequality $\mathbf{x}' \mathbf{A}^{-1} \mathbf{x} \leq \lambda_{\min} (\mathbf{A}) \|\mathbf{x}\|^2$, we have

$$0 \leq \varsigma_{c,iT} \leq \frac{1}{\alpha_i \sqrt{T}} \lambda_{\min}^{-1} \left(\mathbf{A}_T \tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{H}}_i^* \mathbf{A}_T \right) \left\| \mathbf{A}_T \tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{v}}_i \right\|_2^2.$$

Using Cauchy-Schwarz inequality, we obtain

$$E \left(\varsigma_{c,iT}^4 \right) \leq \frac{1}{\alpha_i^4 T^2} \sqrt{E \left[\lambda_{\min}^{-4} \left(\mathbf{A}_T \tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{H}}_i^* \mathbf{A}_T \right) \right]} \sqrt{E \left\| \mathbf{A}_T \tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{v}}_i \right\|_2^8}.$$

But $\alpha_i^{-4} < K$ under Assumption 1, and $E \left[\lambda_{\min}^{-4} \left(\mathbf{A}_T \tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{H}}_i^* \mathbf{A}_T \right) \right] < K$ under Assumption 4. It follows

$$E \left(\varsigma_{c,iT}^4 \right) \leq \frac{K}{T^2} \sqrt{E \left\| \mathbf{A}_T \tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{v}}_i \right\|_2^8}.$$

Let $\mathbf{A}_T \tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{v}}_i \equiv \mathbf{h}_{viT}$ and consider the individual elements of $\mathbf{h}_{viT}$, denoted as $h_{viT,j}$ for $j = 1, 2, 3$,

$$\mathbf{h}_{viT} = \mathbf{A}_T \tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{v}}_i = \begin{pmatrix} h_{viT,1} \\ h_{viT,2} \\ h_{viT,3} \end{pmatrix} = \begin{pmatrix} \frac{1}{T} \sum_{t=1}^T \tilde{x}_{it} \tilde{v}_{it} \\ \frac{1}{\sqrt{T}} \sum_{t=1}^T \tilde{u}_{it} \tilde{v}_{it} \\ \frac{1}{\sqrt{T}} \sum_{t=1}^T \bar{\xi}_{i,t-1} \tilde{v}_{it} \end{pmatrix}.$$

Under Assumption 2, it can be shown that

$$E \left(h_{viT,j}^8 \right) < K, \text{ for } j = 1, 2, 3,$$

which is sufficient for $E \left\| \mathbf{A}_T \tilde{\mathbf{H}}_i^{*\prime} \tilde{\mathbf{v}}_i \right\|_2^8 < K$. It follows that

$$E \left(\varsigma_{c,iT}^4 \right) < \frac{K}{T^2}. \tag{A.33}$$

This completes the proof of (A.26) for $j = 1$. Consider next (A.26) for $j = 2$, and note $q_{iT,2}$ is the same as $\varsigma_{b,iT}$, namely

$$q_{iT,2} = \frac{\Delta \tilde{\mathbf{x}}_i' \mathbf{P}_i \tilde{\mathbf{v}}_i}{\alpha_i \sqrt{T}} = \frac{\Delta \tilde{\mathbf{x}}_i' \tilde{\mathbf{v}}_i}{\alpha_i \sqrt{T}} = \varsigma_{b,iT}.$$

But $E \left(\varsigma_{b,iT}^4 \right) < K$, see (A.31). This completes the proof of (A.26).

Next we establish (A.27). As before we consider the individual elements of 2×1 vector $\mathbf{q}_{iT}$, denoted as

$q_{iT,s}$ for $s = 1, 2$, separately. For $s = 1$ we have (using the individual terms in expression (A.30))

$$|E(q_{iT,1})| = \left| E\left(-\frac{\tilde{\boldsymbol{\xi}}'_{i,-1}\tilde{\mathbf{v}}_i}{\sqrt{T}} + \frac{\Delta\tilde{\mathbf{x}}'_i\tilde{\mathbf{v}}_i}{\alpha_i\sqrt{T}} + \frac{\Delta\tilde{\mathbf{v}}'_i\mathbf{P}_i\tilde{\mathbf{v}}_i}{\alpha_i\sqrt{T}} \right) \right| \le |E(\varsigma_{a,iT})| + |E(\varsigma_{b,iT})| + |E(\varsigma_{c,iT})|. \tag{A.34}$$

For the first term in (A.34), we obtain

$$\begin{aligned}
|E(\varsigma_{a,iT})| &= \left| E\left[-\frac{\tilde{\boldsymbol{\xi}}'_{i,-1}\tilde{\mathbf{v}}_i}{\sqrt{T}} = -\frac{1}{\sqrt{T}}\sum_{i=1}^{T}\left(\xi_{i,t-1} - \bar{\xi}_{i,-1}\right)\left(v_{it} - \bar{v}_i\right) \right] \right|, \\
&\le \frac{1}{\sqrt{T}}\sum_{i=1}^{T}\left|E\left(\xi_{i,t-1}v_{it}\right)\right| + \sqrt{T}E\left|\bar{\xi}_{i,-1}\bar{v}_i\right|.
\end{aligned}$$

But $E\left(\xi_{i,t-1}v_{it}\right) = 0$ and $E\left|\bar{\xi}_{i,-1}\bar{v}_i\right| < K/T$ under Assumptions 1-2. Hence,

$$|E(\varsigma_{a,iT})| \le \frac{K}{\sqrt{T}}.$$

For the second term in (A.34), we obtain

$$\begin{aligned}
|E(\varsigma_{b,iT})| &= \left| E\left(\frac{1}{\alpha_i\sqrt{T}}\sum_{t=1}^{T}u_{x,it}v_{it} - \frac{\sqrt{T}}{\alpha_i}\bar{u}_{x,i}\bar{v}_i \right) \right|, \\
&\le K\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left|E\left(u_{x,it}v_{it}\right)\right| + K\sqrt{T}\left|E\left(\bar{u}_{x,i}\bar{v}_i\right)\right|.
\end{aligned}$$

But $E\left(u_{x,it}v_{it}\right) = 0$ and $E\left(\bar{u}_{x,i}\bar{v}_i\right) = 0$ under Assumption 2. Hence

$$|E(\varsigma_{b,iT})| = 0.$$

Finally, for the last term we note that

$$|E(\varsigma_{c,iT})| \le E\left|\varsigma_{c,iT}\right| \le \sqrt{E\left(\varsigma_{c,iT}^2\right)},$$

and using result (A.33), we obtain

$$|E(\varsigma_{c,iT})| < \frac{K}{\sqrt{T}}.$$

It now follows that $|E(q_{iT,1})| < K/\sqrt{T}$, as desired.

Consider $|E(q_{iT,s})|$ for $s = 2$ next. We have

$$|E(q_{iT,2})| = |E(\varsigma_{b,iT})| = 0.$$

This completes the proof of result (A.27). ∎

Lemma A.7 *Let Assumptions 1-4 hold, and consider* $\mathbf{B}_{iT}$ *defined by (A.19). Then we have*

$$T^{\varphi}\left\|\mathbf{B}_{iT} - \mathbf{B}_i\right\| \to_p 0 \text{ as } T \to \infty, \text{ for any } \varphi < 1/2, \tag{A.35}$$

where

$$\mathbf{B}_i = \plim_{T \to \infty} \mathbf{B}_{iT} = \begin{pmatrix} \alpha_i^2 E\left(\xi_{it}^2\right) + \delta_i^2 \sigma_{xi}^2 & \delta_i \sigma_{xi}^2 \\ \delta_i \sigma_{xi}^2 & \sigma_{xi}^2 \end{pmatrix}, \tag{A.36}$$

and $\xi_{it} = \sum_{\ell=0}^{\infty} (1-\alpha_i)^{\ell} (u_{y,i,t-\ell} - \beta u_{x,i,t-\ell})$.

Proof. We have

$$\mathbf{B}_{iT} = \frac{\Delta \tilde{\mathbf{Z}}_i' \mathbf{P}_i \Delta \tilde{\mathbf{Z}}_i}{T} = \frac{1}{T} \begin{pmatrix} \Delta \tilde{\mathbf{y}}_i' \mathbf{P}_i \Delta \tilde{\mathbf{y}}_i & \Delta \tilde{\mathbf{y}}_i' \mathbf{P}_i \Delta \tilde{\mathbf{x}}_i \\ \Delta \tilde{\mathbf{x}}_i' \mathbf{P}_i \Delta \tilde{\mathbf{y}}_i & \Delta \tilde{\mathbf{x}}_i' \mathbf{P}_i \Delta \tilde{\mathbf{x}}_i \end{pmatrix} = \begin{pmatrix} b_{iT,11} & b_{iT,12} \\ b_{iT,21} & b_{iT,22} \end{pmatrix}.$$

Consider the element $b_{iT,22}$ first. Since $\Delta \tilde{\mathbf{x}}_i' \mathbf{P}_i = \Delta \tilde{\mathbf{x}}_i'$, and $\Delta \tilde{x}_{it} = u_{x,it} - \bar{u}_{x,i}$, we have

$$b_{iT,12} = \left(\frac{1}{T} \sum_{t=1}^{T} u_{x,it}^2\right) - \bar{u}_{x,i}^2.$$

Under Assumption 2, $u_{x,it} \sim IID\left(0, \sigma_{xi}^2\right)$ with finite fourth order moments, and therefore

$$T^{\varphi} \left(\frac{1}{T} \sum_{t=1}^{T} u_{x,it}^2 - \sigma_{xi}^2\right) \overset{p}{\to} 0, \text{ for any } \varphi < 1/2.$$

In addition, $E\left(\bar{u}_{x,i}^2\right) < K/T$, which implies $T^{\varphi} \bar{u}_{x,i}^2 \overset{p}{\to} 0$, for any $\varphi < 1/2$. It follows

$$T^{\varphi} \left(b_{iT,22} - \sigma_{xi}^2\right) \overset{p}{\to} 0, \text{ for any } \varphi < 1/2. \tag{A.37}$$

Consider the element $b_{iT,11}$ next. We will use similar arguments as in the proof of Lemma A.6. In particular, $\Delta \tilde{\mathbf{y}}_i$ can be written as in (A.29), and, since $\mathbf{P}_i \tilde{\boldsymbol{\xi}}_{i,-1} = \tilde{\boldsymbol{\xi}}_{i,-1}$ and $\mathbf{P}_i \Delta \tilde{\mathbf{x}}_i = \Delta \tilde{\mathbf{x}}_i$, we have

$$b_{i,T,11} = \frac{\Delta \tilde{\mathbf{y}}_i' \mathbf{P}_i \Delta \tilde{\mathbf{y}}_i}{T} = \zeta_{aa,iT} + \zeta_{bb,iT} + \zeta_{cc,iT} + 2\zeta_{ab,iT} + 2\zeta_{ac,iT} + 2\zeta_{bc,iT}, \tag{A.38}$$

where

$$\begin{aligned}
\zeta_{aa,iT} &= \alpha_i^2 \frac{\tilde{\boldsymbol{\xi}}_{i,-1}' \tilde{\boldsymbol{\xi}}_{i,-1}}{T}, \; \zeta_{bb,iT} = \delta_i^2 \frac{\Delta \tilde{\mathbf{x}}_i' \Delta \tilde{\mathbf{x}}_i}{T}, \\
\zeta_{cc,iT} &= \frac{\tilde{\mathbf{v}}_i' \mathbf{P}_i \tilde{\mathbf{v}}_i}{T},
\end{aligned}$$

and the cross-product terms are

$$\zeta_{ab,iT} = \alpha_i \delta_i \frac{\tilde{\boldsymbol{\xi}}_{i,-1}' \Delta \tilde{\mathbf{x}}_i}{T}, \; \zeta_{ac,iT} = \alpha_i \frac{\tilde{\boldsymbol{\xi}}_{i,-1}' \tilde{\mathbf{v}}_i}{T}, \; \zeta_{bc,iT} = \delta_i \frac{\Delta \tilde{\mathbf{x}}_i' \tilde{\mathbf{v}}_i}{T}.$$

We consider these individual terms ζ next. Note that

$$\begin{aligned}
\xi_{it} &= \sum_{\ell=0}^{\infty} (1-\alpha_i)^{\ell} (u_{y,i,t-\ell} - \beta u_{x,i,t-\ell}), \\
&= \sum_{\ell=0}^{\infty} (1-\alpha_i)^{\ell} v_{i,t-\ell} + (\delta_i - \beta) \sum_{\ell=0}^{\infty} (1-\alpha_i)^{\ell} u_{x,i,t-\ell},
\end{aligned}$$

where $\sup_i |1-\alpha_i| < 1$ under Assumption 1, and innovations v_{it} and u_{xit} have finite fourth order moments

28

under Assumption 2. Hence, $T^\varphi \left[T^{-1} \sum_{t=1}^{T} \xi_{i,t-1}^2 - E\left(\xi_{i,t-1}^2 \right) \right] \rightarrow_p 0$, $E\left(\bar{\xi}_{i,-1}^2 \right) < K/T$, and we obtain

$$T^\varphi \left[\zeta_{aa,iT} - \alpha_i^2 E\left(\bar{\xi}_{i,-1}^2 \right) \right] \rightarrow_p 0, \text{ for any } \varphi < 1/2. \tag{A.39}$$

Noting that $\zeta_{bb,iT} = \delta_i^2 b_{i,T,12}$, and using result (A.37), we have

$$T^\varphi \left[\zeta_{bb,iT} - \delta_i^2 \sigma_{xi}^2 \right] \rightarrow_p 0, \text{ for any } \varphi < 1/2. \tag{A.40}$$

Consider $\zeta_{cc,iT}$ and note that $\zeta_{cc,iT} = \frac{\alpha_i}{\sqrt{T}} \varsigma_{c,iT}$, where $\varsigma_{c,iT} = \alpha_i^{-1} \Delta \tilde{\mathbf{v}}_i' \mathbf{P}_i \tilde{\mathbf{v}}_i / \sqrt{T}$ was introduced in (A.30) in proof of Lemma A.6. But $E\left(\varsigma_{c,iT}^2 \right) < \frac{K}{T}$ by (A.33), and it follows

$$T^\varphi \zeta_{cc,iT} \rightarrow_p 0, \text{ for any } \varphi < 1/2. \tag{A.41}$$

Using similar arguments, we obtain for the cross-product terms,

$$T^\varphi \zeta_{ab,iT} \rightarrow_p 0, \ T^\varphi \zeta_{ac,iT} \rightarrow_p 0, \text{ and } T^\varphi \zeta_{bc,iT} \rightarrow_p 0, \text{ for any } \varphi < 1/2, \text{ as } T \rightarrow \infty. \tag{A.42}$$

Using (A.39)-(A.42) in (A.38), we obtain

$$T^\varphi \left(b_{i,T,11} - \alpha_i^2 E\left(\xi_{it}^2 \right) - \delta_i^2 \sigma_{xi}^2 \right) \xrightarrow{p} 0, \text{ for any } \varphi < 1/2. \tag{A.43}$$

Using the same arguments for the last term $b_{i,T,12} = b_{i,T,21}$, we obtain

$$T^\varphi \left(b_{i,T,12} - \delta_i \sigma_{xi}^2 \right) \xrightarrow{p} 0, \text{ for any } \varphi < 1/2.$$

This completes the proof of (A.35). ∎

Lemma A.8 *Let Assumptions 1-4 hold, and consider* $\mathbf{B}_{iT}$ *defined by (A.19) and* $\mathbf{B}_i = \text{plim}_{T\rightarrow\infty} \mathbf{B}_{iT}$ *defined by (A.36). Then we have*

$$T^\varphi \left\| \mathbf{B}_{iT}^{-1} - \mathbf{B}_i^{-1} \right\| \rightarrow_p 0, \ \text{as } T \rightarrow \infty, \text{ for any } \varphi < 1/2. \tag{A.44}$$

Proof. This proof closely follows proof of Lemma A.8 in Chudik and Pesaran (2013). Let $p = \left\| \mathbf{B}_i^{-1} \right\|$, $q = \left\| \mathbf{B}_{iT}^{-1} - \mathbf{B}_i^{-1} \right\|$, and $r = \left\| \mathbf{B}_{iT} - \mathbf{B}_i \right\|$. We suppressed subscripts i, T to simplify the notations, but it is understood that the terms p, q, r depend on (i, T). Using the triangle inequality and the submultiplicative property of matrix norm $\|.\|$, we have

$$\begin{aligned} q &= \left\| \mathbf{B}_{iT}^{-1} \left(\mathbf{B}_i - \mathbf{B}_{iT} \right) \mathbf{B}_i^{-1} \right\|, \\ &\leq \left\| \mathbf{B}_{iT}^{-1} \right\| rp, \\ &\leq \left\| \left(\mathbf{B}_{iT}^{-1} - \mathbf{B}_i^{-1} \right) + \mathbf{B}_i^{-1} \right\| rp, \\ &\leq (p + q) \, rp. \end{aligned}$$

Subtracting rpq from both sides and multiplying by T^φ, we have, for any $\varphi < 1/2$,

$$(1 - rp)\left(T^\varphi q \right) \leq p^2 \left(T^\varphi r \right). \tag{A.45}$$

Note that $T^\varphi r \xrightarrow{p} 0$ by Lemma A.7, and $|p| < K$ since $\mathbf{B}_i$ is invertible and $\lambda_{\min}(\mathbf{B}_i)$ is bounded away from zero (this follows from observing that both σ_{xi}^2 and $E(\xi_{it}^2)$ as well as α_i^2 in (A.36) are bounded away from zero). Hence,

$$(1 - rp) \xrightarrow{p} 1, \tag{A.46}$$

and

$$p^2 (T^\varphi r) \xrightarrow{p} 0. \tag{A.47}$$

(A.45)-(A.47) imply $T^\varphi q \xrightarrow{p} 0$. This establishes result (A.44). ∎

Lemma A.9 *Let Assumptions 1-4 hold, and consider ξ_{iT} defined by*

$$\xi_{iT} = \frac{\tilde{\mathbf{x}}_i' \Delta \tilde{\mathbf{Z}}_i}{T} \left(\frac{\Delta \tilde{\mathbf{Z}}_i' \mathbf{P}_i \Delta \tilde{\mathbf{Z}}_i}{T} \right)^{-1} \frac{\Delta \tilde{\mathbf{Z}}_i' \mathbf{P}_i \tilde{\mathbf{v}}_i}{\alpha_i \sqrt{T}}. \tag{A.48}$$

where $\mathbf{P}_i$ is given by (10), and $\tilde{\mathbf{x}}_i$ and $\Delta \tilde{\mathbf{Z}}_i$ are defined below (6). Then

$$\frac{1}{\sqrt{nT}} \sum_{i=1}^n \xi_{iT} \to_p 0, \tag{A.49}$$

as $n, T \to \infty$ such that $n = \Theta(T^\theta)$ for some $0 < \theta < 2$.

Proof. Term ξ_{iT} can be written as

$$\xi_{iT} = \mathbf{a}_{iT}' \mathbf{B}_{iT}^{-1} \mathbf{q}_{iT}, \tag{A.50}$$

where $\mathbf{a}_{iT}$ is given by (A.18), $\mathbf{B}_{iT}$ is given by (A.19), and

$$\mathbf{q}_{iT} = \frac{\Delta \tilde{\mathbf{Z}}_i' \mathbf{P}_i \tilde{\mathbf{v}}_i}{\alpha_i \sqrt{T}}. \tag{A.51}$$

We have

$$\frac{1}{\sqrt{nT}} \sum_{i=1}^n \xi_{i,T} = \frac{1}{\sqrt{nT}} \sum_{i=1}^n \mathbf{a}_{iT}' \left(\mathbf{B}_{iT}^{-1} - \mathbf{B}_i^{-1} \right) \mathbf{q}_{iT} + \frac{1}{\sqrt{nT}} \sum_{i=1}^n \mathbf{a}_{iT}' \mathbf{B}_i^{-1} \mathbf{q}_{iT}. \tag{A.52}$$

Consider the two terms on the right side of (A.52) in turn. Lemma A.2 established fourth moments of $\mathbf{a}_{iT}$ are bounded, which is sufficient for $\|\mathbf{a}_{iT}\| = O_p(1)$. Result (A.26) of Lemma A.6 established second moments of individual elements of $\mathbf{q}_{iT}$ are bounded, which is sufficient for $\|\mathbf{q}_{iT}\| = O_p(1)$. In addition, Lemma A.8 established

$$T^\varphi \left\| \mathbf{B}_{iT}^{-1} - \mathbf{B}_i^{-1} \right\| \to_p 0 \text{ as } T \to \infty, \text{ for any } \varphi < 1/2.$$

Set $\varphi = (\theta - 1)/2 < 1/2$. Then we obtain

$$\frac{1}{\sqrt{nT}} \sum_{i=1}^n \mathbf{a}_{iT}' \left(\mathbf{B}_{iT}^{-1} - \mathbf{B}_i^{-1} \right) \mathbf{q}_{iT} = \frac{\sqrt{n}}{\sqrt{T}} \frac{1}{T^\varphi} \frac{1}{n} \sum_{i=1}^n \mathbf{a}_{iT}' \left[T^\varphi \left(\mathbf{B}_{iT}^{-1} - \mathbf{B}_i^{-1} \right) \right] \mathbf{q}_{iT},$$

$$\le \frac{\sqrt{n}}{T^{\theta/2}} \left(\frac{1}{n} \sum_{i=1}^n \|\mathbf{a}_{iT}\| \left(T^\varphi \left\| \mathbf{B}_{iT}^{-1} - \mathbf{B}_i^{-1} \right\| \right) \|\mathbf{q}_{iT}\| \right) \to_p 0, \quad (A.53)$$

as $n, T \to \infty$ such that $\theta < 2$, where we used $\sqrt{T} T^\varphi = T^{1/2 + \varphi} = T^{\theta/2}$, and $\frac{\sqrt{n}}{T^{\theta/2}} < K$ since $n = \Theta(T^\theta)$.

Consider next the second term on the right side of (A.52). Let $\mu_{iT}^* = E\left(\mathbf{a}_{iT}' \mathbf{B}_i^{-1} \mathbf{q}_{iT} \right)$, and consider the

variance of $(nT)^{-1/2} \sum_{i=1}^{n} \mathbf{a}'_{iT} \mathbf{B}_i^{-1} \mathbf{q}_{iT}$. By independence of $\mathbf{a}'_{iT} \mathbf{B}_i^{-1} \mathbf{q}_{iT}$ across i,

$$
\begin{aligned}
Var \left(\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \mathbf{a}'_{iT} \mathbf{B}_i^{-1} \mathbf{q}_{iT} \right) &= \frac{1}{nT} \sum_{i=1}^{n} Var \left(\mathbf{a}'_{iT} \mathbf{B}_i^{-1} \mathbf{q}_{iT} \right), \\
&\leq \frac{1}{nT} \sum_{i=1}^{n} E \left(\mathbf{a}'_{iT} \mathbf{B}_i^{-1} \mathbf{q}_{iT} \right)^2.
\end{aligned}
\tag{A.54}
$$

Denoting individual elements of $\mathbf{B}_i^{-1}$ as $b_{i,sj}^-$, individual elements of $\mathbf{a}_{iT}$ as $a_{iT,j}$, and individual elements of $\mathbf{q}_{iT}$ as $q_{iT,s}$, for $s, j = 1, 2$, we have

$$
\begin{aligned}
\mathbf{a}'_{iT} \mathbf{B}_i^{-1} \mathbf{q}_{iT} &= \sum_{s=1}^{2} \sum_{j=1}^{2} b_{i,sj}^- a_{iT,s} q_{iT,j}, \\
&= b_{i,11}^- a_{iT,1} q_{iT,1} + b_{i,21}^- a_{iT,2} q_{iT,1} + b_{i,12}^- a_{iT,1} q_{iT,2} + b_{i,22}^- a_{iT,2} q_{iT,2},
\end{aligned}
\tag{A.55}
$$

where

$$
a_{iT,1} = \frac{1}{T} \sum_{t=1}^{T} \tilde{x}_{it} \Delta \tilde{y}_{it} = \frac{1}{T} \sum_{t=1}^{T} (x_{it} - \bar{x}_i) \Delta y_{it},
\tag{A.56}
$$

$$
a_{iT,2} = \frac{1}{T} \sum_{t=1}^{T} \tilde{x}_{it} \Delta \tilde{x}_{it} = \frac{1}{T} \sum_{t=1}^{T} (x_{it} - \bar{x}_i) u_{x,it},
\tag{A.57}
$$

$$
q_{iT,1} = \frac{\Delta \tilde{\mathbf{y}}'_i \mathbf{P}_i \tilde{\mathbf{v}}_i}{\alpha_i \sqrt{T}},
\tag{A.58}
$$

and

$$
q_{iT,2} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{\tilde{u}_{x,it} \tilde{v}_{it}}{\alpha_i}.
\tag{A.59}
$$

Note that $\mathbf{B}_i$ is invertible and $\inf_i \lambda_{\min}(\mathbf{B}_i)$ is bounded away from zero (this follows from observing that both σ_{xi}^2 and $E(\xi_{it}^2)$ as well as α_i^2 in (A.36) are bounded away from zero). It follows $\sup_i \|\mathbf{B}_i^{-1}\| < K$, and therefore $\left|(b_{i,sj}^-)^2\right| < K$. Using this result and Cauchy-Schwarz inequality for the individual summands on the right side of (A.55), we obtain

$$
E \left(\mathbf{a}'_{iT} \mathbf{B}_i^{-1} \mathbf{q}_{iT} \right)^2 \leq K \sum_{s=1}^{2} \sum_{j=1}^{2} \sqrt{E \left(a_{iT,s}^4 \right)} \sqrt{E \left(q_{iT,j}^4 \right)} < K,
\tag{A.60}
$$

where $E \left(a_{iT,s}^4 \right) < K$ by Lemma A.2, and $E \left(q_{iT,j}^4 \right) < K$ by result (A.26) of Lemma A.6. Using (A.60) in (A.54), it follows that

$$
Var \left(\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \mathbf{a}'_{iT} \mathbf{B}_i^{-1} \mathbf{q}_{iT} \right) < \frac{K}{T},
$$

and therefore

$$
\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \left(\mathbf{a}'_{iT} \mathbf{B}_i^{-1} \mathbf{q}_{iT} - \mu_{iT}^* \right) \to_{q.m.} 0 \text{ as } n, T \to \infty.
\tag{A.61}
$$

We establish an upper bound for $|\mu_{iT}^*|$ next. We have (using (A.55) and noting that $\left|b_{i,sj}^-\right| < K$)

$$
|\mu_{iT}^*| < K \cdot \sum_{s=1}^{2} \sum_{j=1}^{2} |E(a_{iT,s} q_{iT,j})|.
$$

It follows that if we can show that

$$|E\left(a_{iT,s}q_{iT,j}\right)| < \frac{K}{\sqrt{T}}, \tag{A.62}$$

holds for all $s, j = 1, 2$, then

$$|\mu_{iT}^*| < \frac{K}{\sqrt{T}}, \tag{A.63}$$

hold. We establish (A.62) for $s = j = 2$, first, which is the most convenient case to consider. We have

$$E\left(a_{iT,2}q_{iT,2}\right) = E\left(\frac{1}{T}\sum_{t=1}^{T}\left(x_{it} - \bar{x}_i\right)u_{x,it} \cdot \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{u_{x,it}v_{it}}{\alpha_i}\right) = 0, \tag{A.64}$$

since v_{it} is independently distributed of $u_{x,it'}$ for any t, t'. Consider next $s = 1$, $j = 2$. We have

$$E\left(a_{iT,1}q_{iT,2}\right) = E\left(\frac{1}{T}\sum_{t=1}^{T}\left(x_{it} - \bar{x}_i\right)\Delta y_{it} \cdot \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{u_{x,it}v_{it}}{\alpha_i}\right), \tag{A.65}$$

where (first-differencing (A.6) and substituting (3))

$$\begin{aligned}
\Delta y_{it} &= \delta_i u_{x,it} + v_{it} - \alpha_i \sum_{\ell=1}^{\infty}\left(1 - \alpha_i\right)^{\ell-1}\left[v_{i,t-\ell} + \left(\delta_i - \beta\right)u_{x,i,t-\ell}\right], \\
&= \eta_{u,it} + \eta_{v,it}, \tag{A.66}
\end{aligned}$$

in which

$$\eta_{u,it} = \delta_i u_{x,it} - \alpha_i \sum_{\ell=1}^{\infty}\left(1 - \alpha_i\right)^{\ell-1}\left(\delta_i - \beta\right)u_{x,i,t-\ell}, \tag{A.67}$$

and

$$\eta_{v,it} = v_{it} - \alpha_i \sum_{\ell=1}^{\infty}\left(1 - \alpha_i\right)^{\ell-1}v_{i,t-\ell}. \tag{A.68}$$

Hence, $E\left(a_{iT,1}q_{iT,2}\right)$ can be written as

$$\begin{aligned}
E\left(a_{iT,1}q_{iT,2}\right) &= E\left(\frac{1}{T}\sum_{t=1}^{T}\left(x_{it} - \bar{x}_i\right)\eta_{u,it} \cdot \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{u_{x,it}v_{it}}{\alpha_i}\right) \\
&\quad + E\left(\frac{1}{T}\sum_{t=1}^{T}\left(x_{it} - \bar{x}_i\right)\eta_{v,it} \cdot \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{u_{x,it}v_{it}}{\alpha_i}\right).
\end{aligned}$$

The first term is equal to 0, since v_{it} is independently distributed of $u_{x,it'}$ for any t, t'. Consider the second term. Noting that $E\left[\left(x_{it} - \bar{x}_i\right)u_{x,is}\right] < K$ and $\left|\alpha_i^{-1}\right| < K$ for any i, t, s, we obtain

$$\begin{aligned}
E\left(\frac{1}{T}\sum_{t=1}^{T}\left(x_{it} - \bar{x}_i\right)\eta_{v,it} \cdot \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{u_{x,it}v_{it}}{\alpha_i}\right) &= \frac{1}{T^{3/2}}\sum_{t=1}^{T}\sum_{s=1}^{T}\alpha_i^{-1}E\left[\left(x_{it} - \bar{x}_i\right)u_{x,is}\right]E\left(\eta_{v,is}v_{it}\right), \\
&\leq \frac{K}{T^{3/2}}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left(\eta_{v,is}v_{it}\right).
\end{aligned}$$

But

$$E\left(\eta_{v,is}v_{it}\right) = \begin{cases} 0, & \text{for } s < t, \\ \sigma_{vi}^2 < K, & \text{for } s = t, \\ \leq K\rho^{s-t}, & \text{for } s > t, \end{cases}$$

32

where $\rho \equiv \sup_i |1 - \alpha_i| < 1$ by Assumption 1. Hence $\left| \sum_{s=1}^T E\left(\eta_{v,is} v_{it} \right) \right| < K$ for any $t = 1, 2, ...T$, and

$$\left| E\left(\frac{1}{T} \sum_{t=1}^T (x_{it} - \bar{x}_i)\, \eta_{v,it} \cdot \frac{1}{\sqrt{T}} \sum_{t=1}^T \frac{u_{x,it} v_{it}}{\alpha_i} \right) \right| \leq \frac{K}{\sqrt{T}}, \tag{A.69}$$

as desired. This establish (A.62) hold for $s = 1$, $j = 2$.

Consider next (A.62) for $s \in \{1, 2\}$ and $j = 1$. Using expression (A.30), we can write $a_{iT,s} q_{iT,1}$, for $s = 1, 2$, as

$$a_{iT,s} q_{iT,1} = a_{iT,s}\varsigma_{a,iT} + a_{iT,s}\varsigma_{b,iT} + a_{iT,s}\varsigma_{c,iT}, \tag{A.70}$$

where as in the proof of Lemma A.6 $\varsigma_{a,iT} = -\tilde{\boldsymbol{\xi}}'_{i,-1}\tilde{\mathbf{v}}_i/\sqrt{T}$, $\varsigma_{b,iT} = \alpha_i^{-1}\Delta\tilde{\mathbf{x}}'_i\tilde{\mathbf{v}}_i/\sqrt{T}$ and $\varsigma_{c,iT} = \alpha_i^{-1}\Delta\tilde{\mathbf{v}}'_i\mathbf{P}_i\tilde{\mathbf{v}}_i/\sqrt{T}$. Using similar arguments as in establishing (A.69), we obtain

$$|E\left(a_{iT,s}\varsigma_{a,iT}\right)| < \frac{K}{\sqrt{T}}, \text{ for } s = 1, 2.$$

Noting next that $\varsigma_{b,iT} = \alpha_i^{-1} q_{i,T,2}$, it directly follows from results (A.64) and (A.69) that

$$|E\left(a_{iT,s}\varsigma_{b,iT}\right)| < \frac{K}{\sqrt{T}}, \text{ for } s = 1, 2.$$

Consider the last term, $a_{i,T,s}\varsigma_{c,iT}$, for $s = 1, 2$. Using Cauchy-Schwarz inequality we have

$$|E\left(a_{iT,s}\varsigma_{c,iT}\right)| \leq \sqrt{E\left(a_{iT,s}^2\right)}\sqrt{E\left(\varsigma_{c,iT}^2\right)}, \text{ for } s = 1, 2.$$

But $E\left(a_{iT,s}^2\right) < K$, for $s = 1, 2$ by Lemma A.2, and $E\left(\varsigma_{c,iT}^2\right) < K/T$ is implied by (A.33). Hence

$$|E\left(a_{iT,s}\varsigma_{c,iT}\right)| \leq \frac{K}{\sqrt{T}}, \text{ for } s = 1, 2.$$

This completes the proof of (A.62) for all $s, j = 1, 2$, and therefore (A.63) holds. Using (A.63), we

$$\left| \frac{1}{\sqrt{nT}} \sum_{i=1}^n \mu_{iT}^* \right| \leq \frac{1}{\sqrt{nT}} \sum_{i=1}^n |\mu_{iT}^*| < \frac{1}{\sqrt{nT}} \sum_{i=1}^n \frac{K}{\sqrt{T}} = K\frac{\sqrt{n}}{T} \to 0, \tag{A.71}$$

as $n, T \to \infty$ such that $\sqrt{n}/T \to 0$. Results (A.61) and (A.71) imply

$$\frac{1}{\sqrt{nT}} \sum_{i=1}^n \mathbf{a}'_{iT}\mathbf{B}_i^{-1}\mathbf{q}_{iT} \to_p 0, \tag{A.72}$$

as $n, T \to \infty$ such that $\sqrt{n}/T \to 0$. Finally, using (A.53) and (A.72) in (A.52), we obtain (A.61), as desired.
■

Lemma A.10 *Let Assumptions 1-4 hold. Then*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n E\left(\frac{\tilde{\mathbf{x}}'_i\mathbf{M}_i\tilde{\mathbf{v}}_i}{\alpha_i T} \right) \to 0,$$

as $n, T \to \infty$ such that $n = \Theta\left(T^\theta\right)$ for some $0 < \theta < 2$, where $\mathbf{M}_i$ is defined in (9), and $\tilde{\mathbf{x}}_i$ and $\tilde{\mathbf{v}}_i$ are defined below (6).

Proof. We have

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{v}}_i}{\alpha_i T} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\tilde{\mathbf{x}}_i' \tilde{\mathbf{v}}_i}{\alpha_i T} + \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \xi_{iT}, \tag{A.73}$$

where ξ_{iT} is defined by (A.48). For the first term, $E(\tilde{\mathbf{x}}_i' \tilde{\mathbf{v}}_i) = 0$. For the second term, Lemma A.9 established $\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \xi_{iT} \rightarrow_p 0$, as $n, T \rightarrow \infty$ such that $n = \Theta(T^\theta)$, for some $0 < \theta < 2$. Hence it follows that $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} E\left(\frac{\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{v}}_i}{\alpha_i T}\right) \rightarrow 0$, as $n, T \rightarrow \infty$ such that $n = \Theta(T^\theta)$, for some $0 < \theta < 2$. ∎

Lemma A.11 *Suppose conditions of Theorem 1 hold. Then*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\frac{\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{v}}_i}{\alpha_i T} - E\left(\frac{\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{v}}_i}{\alpha_i T} \right) \right] \rightarrow_d N\left(0, \omega_v^2\right), \tag{A.74}$$

where $\omega_v^2 = \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^{n} \sigma_{xi}^2 \sigma_{vi}^2 / (6\alpha_i^2)$, $\mathbf{M}_i$ is defined by (9), and $\tilde{\mathbf{x}}_i$ and $\tilde{\mathbf{v}}_i$ are defined below (6).

Proof. It is convenient to use (A.73) in (A.74) to obtain

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\frac{\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{v}}_i}{\alpha_i T} - E\left(\frac{\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{v}}_i}{\alpha_i T} \right) \right] = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\tilde{\mathbf{x}}_i' \tilde{\mathbf{v}}_i}{\alpha_i T} + \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \left[\xi_{iT} - E\left(\xi_{iT}\right) \right],$$

where $E(\tilde{\mathbf{x}}_i' \tilde{\mathbf{v}}_i) = 0$. It follows from Lemmas A.9 and A.10 that $\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} [\xi_{iT} - E(\xi_{iT})] \rightarrow_p 0$, as $n, T \rightarrow \infty$ such that $n = \Theta(T^\theta)$, for some $0 < \theta < 2$. Hence, under the conditions of Theorem 1, the asymptotic distribution of $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\frac{\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{v}}_i}{\alpha_i T} - E\left(\frac{\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{v}}_i}{\alpha_i T} \right) \right]$ is given by the first term, $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\tilde{\mathbf{x}}_i' \tilde{\mathbf{v}}_i}{\alpha_i T}$, alone. Lemma A.5 establishes the asymptotic normality of this term.

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\tilde{\mathbf{x}}_i' \mathbf{M}_i \tilde{\mathbf{v}}_i}{\alpha_i T} \rightarrow_d N\left(0, \omega_v^2\right), \tag{A.75}$$

as $n, T \rightarrow \infty$. This completes the proof. ∎

B   Estimation algorithms

This appendix describes implementation of the Pooled Mean Group (PMG) estimator, which we compute iteratively, in Section B.1. Section B.2 discusses implementation of bias-correction methods and bootstrapping of critical values for the PMG, PDOLS and FMOLS estimators. Section B.3 provides tables with Monte Carlo findings for experiments with cross-sectionally dependent errors.

B.1   Computation of PMG estimator

Consider the same illustrative panel ARDL model as in the paper, namely the model given by equations (1)-(2). PMG estimator of the long-run coefficient β, as originally proposed by Pesaran, Shin, and Smith (1999), is computed by solving the following equations iteratively:

$$\hat{\beta}_{PMG} = -\left(\sum_{i=1}^{n} \frac{\hat{\phi}_i^2}{\hat{\sigma}_i^2} \mathbf{x}_i' \mathbf{H}_{x,i} \mathbf{x}_i \right)^{-1} \sum_{i=1}^{n} \frac{\hat{\phi}_i^2}{\hat{\sigma}_i^2} \mathbf{x}_i' \mathbf{H}_{x,i} \left(\Delta \mathbf{y}_i - \hat{\phi}_i \mathbf{y}_{i,-1} \right), \tag{B.1}$$

$$\hat{\phi}_i = \left(\hat{\boldsymbol{\xi}}_i' \mathbf{H}_{x,i} \hat{\boldsymbol{\xi}}_i \right)^{-1} \hat{\boldsymbol{\xi}}_i' \mathbf{H}_{x,i} \Delta \mathbf{y}_i, \ i = 1, 2, ..., n, \tag{B.2}$$

and

$$\hat{\sigma}_i^2 = T^{-1} \left(\Delta\mathbf{y}_i - \hat{\phi}_i \hat{\boldsymbol{\xi}}_i \right)' \mathbf{H}_{x,i} \left(\Delta\mathbf{y}_i - \hat{\phi}_i \hat{\boldsymbol{\xi}}_i \right), \ i = 1, 2, ..., n, \tag{B.3}$$

where $\hat{\boldsymbol{\xi}}_i = \mathbf{y}_{i,-1} - \mathbf{x}_i \hat{\beta}_{PMG}$, $\mathbf{x}_i = (x_{i,1}, x_{i,2}, ..., x_{i,T})'$, $\Delta\mathbf{y}_i = \mathbf{y}_i - \mathbf{y}_{i,-1}$, $\mathbf{y}_i = (y_{i,1}, y_{i,2}, ..., y_{i,T})'$, $\mathbf{y}_{i,-1} = (y_{i,0}, y_{i,1}, ..., y_{i,T-1})'$, $\mathbf{H}_{x,i} = \mathbf{I}_T - \Delta\mathbf{x}_i (\Delta\mathbf{x}_i' \Delta\mathbf{x}_i)^{-1} \Delta\mathbf{x}_i'$, $\Delta\mathbf{x}_i = \mathbf{x}_i - \mathbf{x}_{i,-1}$, and $\mathbf{x}_{i,-1} = (x_{i,0}, x_{i,1}, ..., x_{i,T-1})'$.

To solve (B.1)-(B.3) iteratively, we set $\hat{\beta}_{PMG,(0)}$ to the pooled Engle-Granger estimator, and given the initial estimate $\hat{\beta}_{PMG,(0)}$, we compute $\hat{\boldsymbol{\xi}}_{i,(0)} = \mathbf{y}_{i,-1} - \mathbf{x}_i \hat{\beta}_{PMG,(0)}$, $\hat{\phi}_{i,(0)}$ and $\hat{\sigma}_{i,(0)}^2$, for $i = 1, 2, ..., n$ using (B.2)-(B.3). Next we compute $\hat{\beta}_{PMG,(1)}$ using (B.1) and given values $\hat{\phi}_{i,(0)}$ and $\hat{\sigma}_{i,(0)}^2$. Then we iterate - for a given value of $\hat{\beta}_{PMG,(\ell)}$ we compute $\hat{\boldsymbol{\xi}}_{i,(\ell)}$, $\hat{\phi}_{i,(\ell)}$ and $\hat{\sigma}_{i,(\ell)}^2$; and for given values of $\hat{\phi}_{i,(\ell)}$ and $\hat{\sigma}_{i,(\ell)}^2$ we compute $\hat{\beta}_{PMG,(\ell+1)}$. If convergence is not achieved, we increase ℓ by one and repeat. We define convergence by $\left| \hat{\beta}_{PMG,(\ell+1)} - \hat{\beta}_{PMG,(\ell)} \right| < 10^{-4}$.

Inference is conducted using equation (17) of Pesaran, Shin, and Smith (1999). In particular,

$$T\sqrt{n} \left(\hat{\beta}_{PMG} - \beta_0 \right) \sim N \left(0, \Omega_{PMG} \right),$$

where

$$\Omega_{PMG} = \left(\frac{1}{n} \sum_{i=1}^n \frac{\phi_{i,0}}{\sigma_{i,0}^2} r_{x_i, x_i} \right)^{-1}, \text{ and } r_{x_i, x_i} = plim_{T \to \infty} T^{-2} \mathbf{x}_i' \mathbf{H}_{x,i} \mathbf{x}_i \ .$$

Standard error of $\hat{\beta}_{PMG}$, denoted as $se\left(\hat{\beta}_{PMG}\right)$, is estimated as

$$\widehat{se} \left(\hat{\beta}_{PMG} \right) = T^{-1} n^{-1/2} \hat{\Omega}_{PMG},$$

where

$$\hat{\Omega}_{PMG} = \left(\frac{1}{n} \sum_{i=1}^n \frac{\hat{\phi}_{i,0}}{\hat{\sigma}_{i,0}^2} \hat{r}_{x_i, x_i} \right)^{-1} \text{ and } \hat{r}_{x_i, x_i} = T^{-2} \mathbf{x}_i' \mathbf{H}_{x,i} \mathbf{x}_i. \tag{B.4}$$

B.2 Bias-corrected PMG, PDOLS and FMOLS estimators, and bootstrapped critical values

Similarly to the bootstrap bias-corrected PB estimator, we consider the following bootstrap bias-corrected PMG, PDOLS and FMOLS estimators. Let the original (uncorrected) estimators be denoted as $\hat{\beta}_e$ for $e = PMG, PDOLS$, and $FMOLS$, respectively. Bootstrap bias corrected version of these estimators is given by

$$\tilde{\beta}_e = \hat{\beta}_e - \hat{b}_e, \tag{B.5}$$

for $e = PMG, PDOLS$, and $FMOLS$, where $\hat{b}_e$ an estimate of the bias obtained by the following sieve wild bootstrap algorithm, which resembles the algorithm in Subsection 2.2.1. For e= $PMG, PDOLS$, and $FMOLS$:

1. Compute $\hat{\beta}_e$. Given $\hat{\beta}_e$, estimate the remaining unknown coefficients of (1)-(2) by least squares, and compute residuals $\hat{u}_{y,it}^e, \hat{u}_{x,it}^e$.

2. For each $r = 1, 2, ..., R$, generate new draws for $\hat{u}_{y,it}^{e,(r)} = a_t^{(r)} \hat{u}_{y,it}^e$, and $\hat{u}_{x,it}^{e,(r)} = a_t^{(r)} \hat{u}_{x,it}^e$, where $a_t^{(r)}$ are randomly drawn from Rademacher distribution (Liu, 1988) namely

$$a_t^{(r)} = \begin{cases} -1, & \text{with probability } 1/2 \\ 1, & \text{with probability } 1/2 \end{cases} .$$

35

Given the estimated parameters of (1)-(2) from Step 1, and initial values y_{i1}, x_{i1} generate simulated data $y_{it}^{e,(r)}, x_{it}^{e,(r)}$ for $t = 2, 3, ..., T$ and $i = 1, 2, ..., n$. Using the generated data compute $\hat{\beta}_e^{(r)}$.

3. Compute $\hat{b}_e = \left[R^{-1} \sum_{r=1}^{R} \hat{\beta}_e^{(r)} - \hat{\beta}_e \right]$.

The possibility of iterating the algorithm above by using the bias-corrected estimate $\tilde{\beta}_e$ in Step 1 is not considered in this paper.

We conduct inference by using the $1 - \alpha$ confidence interval $C_{1-\alpha}\left(\tilde{\beta}_e\right) = \tilde{\beta}_e \pm \hat{k}_e \widehat{se}\left(\hat{\beta}_e\right) = \tilde{\beta}_e \pm T^{-1}n^{-1/2}\hat{k}_e \hat{\Omega}_e$, where $\hat{k}_e$ is the $1 - \alpha$ percent quantile of $\left\{ \left| t_e^{(r)} \right| \right\}_{r=1}^{R}$, in which $t_e^{(r)} = \tilde{\beta}_e^{(r)}/\widehat{se}\left(\hat{\beta}_e^{(r)}\right) = T^{-1}n^{-1/2}\tilde{\beta}_e^{(r)}/\hat{\Omega}_e^{(r)}$, $\tilde{\beta}_e^{(r)} = \hat{\beta}_e^{(r)} - \hat{b}_e$ is the bias-corrected PMG estimate of β in the r-th draw of the bootstrap data in the algorithm above, and $\hat{\Omega}_e^{(r)}$ is estimated standard error using the bootstrap data.

B.2.1   Jackknife bias-corrections

We consider similar jackknife bias correction for PMG, PDOLS and FMOLS estimator as for the PB estimator in Section 2.2. In particular,

$$\tilde{\beta}_{jk,e} = \tilde{\beta}_{jk,e}\left(\kappa\right) = \hat{\beta}_e - \kappa\left(\frac{\hat{\beta}_{e,a} + \hat{\beta}_{e,b}}{2} - \hat{\beta}_e\right),$$

for $e = PMG, PDOLS$, and $FMOLS$, where $\hat{\beta}_e$ is the full sample estimator, $\hat{\beta}_{e,a}$ and $\hat{\beta}_{e,b}$ are the first and the second half sub-sample estimators, and $\kappa = 1/3$ is the same weighting parameter as in Section 2.2.

We conduct inference by using the $1 - \alpha$ confidence interval $C_{1-\alpha}\left(\tilde{\beta}_{jk,e}\right) = \tilde{\beta}_{jk,e} \pm \hat{k}_{jk,e} \widehat{se}\left(\hat{\beta}_e\right) = \tilde{\beta}_{jk,e} \pm \hat{k}_{jk,e}T^{-1}n^{-1/2}\hat{\Omega}_e$, where $\hat{k}_{jk}$ is the $1 - \alpha$ percent quantile of $\left\{ \left| t_{jk,e}^{(r)} \right| \right\}_{r=1}^{R}$, in which $t_{jk,e}^{(r)} = \tilde{\beta}_{jk,e}^{(r)}/\widehat{se}\left(\hat{\beta}_e^{(r)}\right) = T^{-1}n^{-1/2}\tilde{\beta}_{jk,e}^{(r)}/\hat{\Omega}_e^{(r)}$, $\tilde{\beta}_{jk,e}^{(r)}$ is the jackknife bias-corrected estimate of β using the r-th draw of the bootstrap data generated using the same algorithm as in Subsection B.2, and $\hat{\Omega}_e^{(r)}$ is estimated standard error using the bootstrap data.

B.3   Monte Carlo results for experiments with cross-sectionally dependent errors

Table B1: MC findings for the estimation of long-run coefficient β in experiments with cross-sectionally dependent errors.

Estimators without bias correction and inference conducted using standard critical values.

	Bias (× 100)				RMSE (× 100)				Size (5% level)				Power (5% level)			
$n \backslash T$	20	30	40	50	20	30	40	50	20	30	40	50	20	30	40	50
PB																
20	-3.83	-1.93	-1.01	-0.66	7.58	4.92	3.57	2.78	28.00	21.95	17.55	16.80	37.30	64.50	86.90	96.85
30	-3.57	-1.98	-1.17	-0.69	6.67	4.43	3.32	2.46	28.95	23.55	22.10	18.80	45.10	75.25	92.70	99.00
40	-3.94	-2.04	-1.10	-0.74	6.58	4.20	3.08	2.37	33.60	27.25	23.75	22.90	46.60	82.00	96.70	99.65
50	-3.88	-1.99	-1.10	-0.76	6.35	4.01	2.89	2.23	38.30	29.55	25.30	23.30	50.70	85.60	97.75	99.95
PMG																
20	-1.96	-0.90	-0.40	-0.26	9.05	5.49	3.86	3.00	45.75	34.00	26.30	23.90	62.45	78.50	92.45	98.15
30	-1.60	-1.05	-0.56	-0.31	7.76	4.83	3.54	2.58	46.35	37.25	32.75	27.35	70.50	85.25	95.60	99.65
40	-1.79	-1.02	-0.50	-0.32	7.07	4.51	3.27	2.44	49.40	40.80	35.00	29.90	73.05	90.25	97.65	99.85
50	-1.76	-0.89	-0.46	-0.32	6.70	4.21	3.00	2.28	54.30	42.95	34.20	31.55	76.40	93.90	98.90	100.00
PDOLS																
20	-6.20	-4.10	-3.08	-2.40	9.50	6.31	4.67	3.67	31.40	27.30	25.70	26.30	22.35	39.95	65.20	85.80
30	-5.76	-4.15	-3.08	-2.37	8.65	5.89	4.39	3.41	34.85	33.45	32.15	33.35	27.95	48.60	77.05	93.55
40	-6.29	-4.24	-3.12	-2.44	8.67	5.82	4.27	3.35	42.95	41.85	38.60	38.80	26.95	52.35	82.15	97.05
50	-6.12	-4.19	-3.09	-2.47	8.46	5.62	4.16	3.27	47.90	45.45	45.15	45.05	31.50	59.55	88.25	98.65
FMOLS																
20	-10.89	-7.25	-5.42	-4.20	13.27	9.09	6.79	5.37	85.10	76.30	68.35	63.05	54.40	54.40	73.30	88.35
30	-10.32	-7.26	-5.29	-4.06	12.28	8.70	6.45	5.04	88.60	82.70	76.00	71.35	57.00	60.70	82.20	95.25
40	-10.94	-7.60	-5.45	-4.29	12.57	8.76	6.41	5.10	91.90	88.20	82.85	77.50	60.10	61.75	84.95	96.65
50	-10.72	-7.42	-5.38	-4.30	12.26	8.52	6.28	5.00	93.50	89.15	85.40	82.35	63.55	68.15	89.30	98.80

Notes: DGP is given by $\Delta y_{it} = c_i - \alpha_i (y_{i,t-1} - \beta x_{i,t-1}) + u_{y,it}$ and $\Delta x_{it} = u_{x,it}$, for $i = 1, 2, ..., n$, $T = 1, 2, ..., T$, with $\beta = 1$ and $\alpha_i \sim IIDU[0.2, 0.3]$. Errors $u_{y,it}$, $u_{x,it}$ are cross-sectionally dependent, heteroskedastic over i, and also correlated over y & x equations. See Section 3.1 for complete description of the DGP. The pooled Bewley estimator is given by (8), with variance estimated using (16). PMG is the Pooled Mean Group estimator proposed by Pesaran, Shin, and Smith (1999). PDOLS is panel dynamic OLS estimator by Mark and Sul (2003). FMOLS is the group-mean fully modified OLS estimator by Pedroni (1996, 2001b). The size and power findings are computed using 5% nominal level and the reported power is the rejection frequency for testing the hypothesis $\beta = 0.9$.

Table B2: MC findings for the estimation of long-run coefficient β in experiments with cross-sectionally dependent errors.

Bias corrected estimators and inference conducted using bootstrapped critical values.

$n\backslash T$	Bias ($\times$ 100)				RMSE ($\times$ 100)				Size (5% level)				Power (5% level)			
	20	**30**	**40**	**50**	**20**	**30**	**40**	**50**	**20**	**30**	**40**	**50**	**20**	**30**	**40**	**50**
	Jackknife bias-corrected estimators															
	PB															
20	-1.57	-0.54	-0.11	-0.01	7.83	5.24	3.92	3.04	7.65	6.55	5.55	4.65	20.55	41.35	65.35	83.50
30	-1.29	-0.67	-0.34	-0.05	6.76	4.63	3.59	2.72	6.30	6.60	6.40	5.25	25.80	51.35	73.90	91.45
40	-1.67	-0.62	-0.20	-0.06	6.48	4.35	3.36	2.58	6.75	6.15	6.20	5.40	25.55	54.30	80.25	94.45
50	-1.54	-0.58	-0.21	-0.09	6.16	4.12	3.15	2.43	6.80	6.20	5.85	4.65	27.65	59.40	84.95	96.70
	PMG															
20	-0.54	-0.15	0.02	0.06	10.66	6.26	4.45	3.44	16.05	10.80	8.55	7.90	29.10	45.05	69.15	84.75
30	-0.25	-0.39	-0.19	0.03	9.10	5.50	4.02	2.94	14.80	10.40	9.05	6.90	36.30	53.85	76.40	93.40
40	-0.57	-0.27	-0.08	0.03	8.25	5.13	3.78	2.78	14.10	10.50	9.15	7.80	35.10	59.50	81.25	95.55
50	-0.41	-0.10	-0.06	0.03	7.81	4.81	3.44	2.63	14.85	11.05	8.65	7.65	39.55	64.65	85.90	97.15
	PDOLS															
20	-4.61	-2.84	-2.06	-1.54	9.89	6.23	4.51	3.51	8.75	7.50	6.45	4.95	10.05	21.55	39.00	63.80
30	-4.09	-2.96	-2.15	-1.55	8.86	5.72	4.21	3.20	8.70	7.30	6.55	5.40	12.20	24.05	46.10	73.10
40	-4.66	-2.94	-2.13	-1.57	8.62	5.53	4.00	3.06	8.80	6.55	6.45	4.55	10.00	23.90	47.70	77.80
50	-4.41	-2.91	-2.11	-1.63	8.30	5.29	3.84	2.95	9.85	6.50	6.25	5.10	11.95	24.40	52.10	80.80
	FMOLS															
20	-8.65	-5.12	-3.61	-2.69	12.18	7.99	5.79	4.56	11.25	5.65	4.20	3.85	2.00	4.55	9.50	25.50
30	-8.06	-5.27	-3.57	-2.55	11.00	7.43	5.42	4.14	9.65	6.85	5.20	3.35	2.10	3.95	12.65	31.60
40	-8.65	-5.53	-3.70	-2.74	11.08	7.35	5.25	4.08	11.50	6.50	4.55	3.85	1.45	3.35	12.25	32.05
50	-8.44	-5.36	-3.63	-2.78	10.72	7.07	5.07	3.94	12.45	6.70	4.55	3.40	1.75	4.40	13.70	34.30
	Bootstrap bias-corrected estimators															
	PB															
20	-1.26	-0.43	-0.05	-0.01	7.34	4.80	3.56	2.78	10.10	7.95	6.45	5.85	29.35	52.15	76.10	90.15
30	-1.04	-0.51	-0.23	-0.05	6.29	4.21	3.24	2.42	10.20	7.35	7.20	5.40	37.20	64.30	85.00	95.90
40	-1.31	-0.51	-0.12	-0.08	5.98	3.93	3.01	2.32	10.15	7.50	6.75	5.80	39.00	69.80	89.80	98.15
50	-1.26	-0.45	-0.12	-0.09	5.78	3.72	2.78	2.17	10.20	7.85	7.05	6.00	42.90	75.25	93.50	99.20
	PMG															
20	-1.23	-0.44	-0.09	-0.03	9.24	5.54	3.91	3.03	16.00	10.50	7.60	6.70	31.35	51.80	76.10	90.80
30	-0.92	-0.60	-0.27	-0.10	7.89	4.84	3.57	2.59	14.70	9.95	8.45	5.85	38.20	60.85	83.80	97.10
40	-1.10	-0.56	-0.19	-0.11	7.14	4.51	3.30	2.45	14.80	10.00	8.90	7.00	39.35	67.05	88.90	98.10
50	-1.06	-0.45	-0.16	-0.10	6.77	4.23	3.02	2.29	14.90	10.45	8.25	6.00	44.90	71.40	91.75	99.15
	PDOLS															
20	-2.34	-1.13	-0.66	-0.37	8.73	5.46	3.92	3.05	12.05	10.05	7.75	6.90	22.40	45.25	71.30	87.75
30	-2.03	-1.29	-0.74	-0.41	7.73	4.85	3.53	2.68	12.15	8.60	8.45	7.65	26.95	51.50	79.30	94.15
40	-2.46	-1.28	-0.71	-0.41	7.40	4.68	3.31	2.54	12.65	10.20	9.05	7.75	26.15	55.55	84.05	96.90
50	-2.32	-1.22	-0.68	-0.44	7.28	4.41	3.15	2.39	15.00	9.60	8.40	7.40	30.15	59.85	88.40	98.55
	FMOLS															
20	-4.39	-2.01	-1.24	-0.77	10.37	6.73	4.85	3.83	18.45	10.90	7.80	6.85	17.80	31.75	50.55	70.75
30	-4.00	-2.25	-1.25	-0.71	9.02	6.02	4.42	3.43	17.25	12.00	9.60	7.70	20.95	37.25	61.40	82.55
40	-4.46	-2.44	-1.27	-0.84	8.82	5.67	4.11	3.21	19.35	12.10	9.60	9.05	20.10	38.50	67.65	85.95
50	-4.30	-2.27	-1.22	-0.86	8.48	5.41	3.92	3.02	21.20	13.90	10.10	7.80	22.85	44.85	72.20	90.40

Notes: See the notes to Table B1. Bias-corrected versions of the PB estimator are described in Subsection 2.2. Bias-corrected versions of the PMG, PDOLS and FMOLS estimator are described in Appendix B. Inference is conducted using bootstrapped critical values.