Estimation and Inference in Short Panel Vector Autoregressions with Unit Roots and Cointegration*

Michael Binder
Johann Wolfgang Goethe-University, Germany

Cheng Hsiao
University of Southern California

M. Hashem Pesaran
University of Southern California and Cambridge University

October 7, 2003

Abstract

This paper considers estimation and inference in panel vector autoregressions (PVARs) where (i) the individual effects are either random or fixed, (ii) the time-series properties of the model variables are unknown *a priori* and may feature unit roots and cointegrating relations, and (iii) the time dimension of the panel is short and its cross-sectional dimension is large. Generalized Method of Moments (GMM) and Quasi Maximum Likelihood (QML) estimators are obtained and then compared in terms of their asymptotic and finite sample properties. It is shown that GMM estimators based only on standard orthogonality conditions break down if the underlying time series contain unit roots. Extended GMM estimators making use of further moment conditions are not subject to this problem. However, their finite sample performance is shown to deteriorate as a ratio of cross-section to time-series variation is increased, while the performance of the fixed effects QML estimator is invariant to this ratio. The QML estimators also tend to outperform the various GMM estimators in finite sample. Overall, our findings favor the use of the fixed effects QML estimator, given that it does not impose any restrictions on the distribution generating the individual effects. The paper also shows how the fixed effects QML estimator can be used for unit root and cointegration tests in short panels.

Keywords: Panel Vector Autoregressions, Random/Fixed Effects, Unit Roots, Cointegration.
JEL-Classification: C12, C13, C33.

*We are grateful to Karim Abadir, Stephen Bond, Jinyong Hahn, Marc Nerlove, Ingmar Prucha, and, especially, Manuel Arellano, Peter Schmidt, Peter Phillips (the editor) and four anonymous referees for helpful and constructive comments. We have also benefited from useful suggestions by participants at various seminars and conferences. Correspondence addresses are: School of Business and Economics, Johann Wolfgang Goethe-University, Mertonstrasse 17, Box 79, 60054 Frankfurt , Germany; Department of Economics, University of Southern California, University Park, Los Angeles, CA 90089, USA; Trinity College, Cambridge, CB2 1TQ, England.

1 Introduction

Over the past decade important advances have been made in the study of dynamic panel data models where both the time dimension (T) and the cross-sectional dimension (N) are large. See, for example, the surveys by Baltagi and Kao (2000) and Phillips and Moon (2000) and the references cited therein. In this paper we are concerned with the more traditional panel literature where N is large and T is short (typically less than 10), which remains the prevalent setting for the majority of empirical microeconometric research.[1] However, this literature has primarily focussed on single equation dynamic panel data models whilst there are many applications that ideally require a simultaneous treatment of the decision problems faced by households, firms, and institutions. A natural starting point are vector autoregressive models (VARs) which have been extensively studied in the time-series literature. An early analysis of panel VARs (PVARs) with a short T was provided by Holtz-Eakin, Newey, and Rosen (1988). The fact that in such panels T is small does not mean that the underlying data could not have arisen from non-stationary and/or cointegarted processes. The slope homogeneity and the cross section independence assumptions of the traditional panel literature allows us to make inferences about the long- term properties of the model even though T is short. Moreover, the presence of cointegration provides a natural starting point for introducing cross–equation restrictions in panel VAR models.

As in single equation dynamic panel data models there are two main issues that need to be addressed in the study of PVARs. (i) The fact that T is fixed necessitates the modeling of the initial observations.[2] (ii) Presence of cross-sectional heterogeneity poses the important question of how to best model the unobserved individual-specific effects.[3] Here we shall consider both the random and fixed effects specifications. The fixed effects specification has the advantage of being robust to possible misspecification of the distribution of the individual effects. But it is still subject to the classical incidental parameters problem as in Neyman and Scott (1948), violating the regularity conditions needed for the consistency of the conventional Quasi Maximum Likelihood (QML) estimator.[4]

To overcome this problem Generalized Method of Moments (GMM) estimation has been suggested in the literature. It is useful to distinguish between the "standard" GMM estimators proposed by Holtz-Eakin, Newey, and Rosen (1988) and Arellano and Bond (1991), and their subsequent extensions by, for example, Ahn and Schmidt (1995), Arellano and Bover (1995), and

[1]For references to much of this empirical work see, for example, Baltagi (2001).

[2]For discussions of the initial observations in the single equation context see, for example, Anderson and Hsiao (1981, 1982), Bhargava and Sargan (1983), Blundell and Smith (1991), and Nerlove (1999).

[3]Dealing with possible slope coefficient heterogeneity poses further complications and might not be feasible in dynamic panels where T is very small. See, for example, Hsiao, Pesaran, and Tahmiscioglu (1999).

[4]See, for example, Anderson and Hsiao (1981) and Nickell (1981) for a discussion of this issue in the context of univariate models.

Blundell and Bond (1998). The "standard" GMM estimators are based on orthogonality conditions that interact the lagged values of the endogenous variables with first differences of the model's disturbances, while the "extended" GMM estimators augment these orthogonality conditions with additional moment conditions implied by homoskedasticity and initialization restrictions. The motivation behind the introduction of the extended GMM estimators has been twofold: (i) The standard GMM estimators, being based on a subset of the valid moment conditions, are asymptotically inefficient and are subject to the "weak instrument" problem if one or more roots of the characteristic equation of the model are close to unity (for example, Blundell and Bond, 1998). (ii) The standard GMM estimators break down in the presence of unit roots. A formal proof of this breakdown in the context of PVARs is provided in this paper. In contrast to the standard GMM estimators, the extended GMM estimators are consistent even if the unit root properties of the model are not known *a priori*. These estimators are discussed in Ahn and Schmidt (1995) and Arellano and Bover (1995) in the case of single equation models.

The paper develops random effects and fixed effects QML estimators (RE-QMLE and FE-QMLE, respectively) when it is not known *a priori* whether the underlying series are stationary, have unit roots or are cointegrated. It contributes to the discussion of the initialization of the dynamic models with a fixed T, by generalizing the stationarity restrictions proposed in the literature to settings involving unit roots and cointegration. New panel unit root and cointegration tests are proposed for panels with a short T. Under certain regularity conditions it is shown that the QML estimators is consistent and asymptotically normally distributed (as $N \to \infty$, with T fixed and small), irrespective of whether the underlying time series are (trend) stationary, integrated of order one, I(1), or I(1) and cointegrated. The paper also provides a generalization of the extended GMM estimators to PVAR models, and presents a comparative analysis of these estimation procedures in terms of their asymptotic properties as well as their finite sample performances using Monte Carlo experiments. The RE-QMLE is more efficient than the FE-QMLE, but it imposes moment homogeneity restrictions on the initial observations and requires the individual effects to be random draws from probability distributions with finite fourth-order moments. The standard and extended GMM estimators are also shown to impose restrictions on the distribution of the individual effects, not needed under the fixed effects specification. In the case of stationary PVAR models it is shown that the asymptotic variance of the standard GMM estimator (and by implication the extended ones) is an increasing function of the variance matrix of the individual effects. This is an important result and shows that in dynamic panels the quality of the GMM estimators can deteriorate often substantially in applications where the individual effects show considerable variations. By contrast, the FE-QMLE is invariant to the individual effects and hence is not subject to the same problem.

Finally, extensive Monte Carlo evidence on the finite sample properties of the QML and GMM estimators is provided, with the following two main conclusions:

(i) The FE-QMLE performs well under a variety of parameter configurations and is robust to the unit root properties of the underlying time-series processes. It also seems to be reasonably robust to non-normal errors.

(ii) As predicted by our theory, the finite sample properties of the GMM type estimators in general depend critically on τ, the ratio of the variance of the individual effects relative to the variance of the errors. When this ratio is well in excess of unity both the standard and extended GMM estimators tend to perform poorly, except for the pure unit root case. This is an important finding considering that the FE-QMLE is invariant to τ, and that it seems rather doubtful that τ in empirical work can generally be assumed to be small. Even under $\tau = 1$ and non-normal errors the GMM estimators are still outperformed (with the exception of the pure unit root case) by the FE-QMLE.

The remainder of this paper is organized as follows: Section 2 introduces the PVAR model. Sections 3 and 4 develop the QML estimators under random and fixed effects specifications, respectively. Section 5 proposes new tests for unit roots and cointegration in panels with short time dimension. Section 6 discusses GMM estimation of the PVAR model. Monte Carlo simulation results are presented in Section 7, and Section 8 concludes and provides some suggestions for future research. In Appendix A the restrictions on the distributions generating the individual and initialization effects needed for validity of the QML, and GMM estimators are contrasted. Appendix B provides a proof of the dependence of the asymptotic variance of the standard GMM estimator on the variance of individual effects. Appendix C describes the computational details and the related algorithms.

2 A Panel VAR Model

Let $\mathbf{w}_{it}$ be an $m \times 1$ vector of random variables for the i-th cross-sectional unit at time t, and suppose that the $\mathbf{w}_{it}$'s are generated by the following panel vector autoregressive model of order one, PVAR(1):

$$\mathbf{w}_{it} = (\mathbf{I}_m - \Phi)\,\boldsymbol{\mu}_i + \Phi\mathbf{w}_{i,t-1} + \boldsymbol{\varepsilon}_{it}, \tag{2.1}$$

for $i = 1, 2, \ldots, N$; and $t = 1, 2, \ldots, T$, where Φ denotes an $m \times m$ matrix of slope coefficients, $\boldsymbol{\mu}_i$ is an $m \times 1$ vector of individual-specific effects, $\boldsymbol{\varepsilon}_{it}$ is an $m \times 1$ vector of disturbances, and $\mathbf{I}_m$ denotes the identity matrix of dimension $m \times m$.

For simplicity we restrict our exposition to first-order PVAR models. However, the estimation and inference procedures discussed in the paper are extended to the p-th order case in an appendix available from the authors upon request. This appendix demonstrates that higher-order models can for most parts be treated in conceptually the same manner as first-order models.

We shall consider both random and fixed effects specifications of the individual-specific effects in the remainder of this paper, highlighting their differences, and the implications these differences have for estimation and inference. However, for both the random and fixed effects specifications we make the following *general* assumptions:

Assumption (G1) *The available observations are* $\mathbf{w}_{i0}$, $\mathbf{w}_{i1}$, $\ldots$, $\mathbf{w}_{iT}$, *with* $T \geq 2$ *but fixed as* $N \rightarrow \infty$.

Assumption (G2) *The disturbances* $\boldsymbol{\varepsilon}_{it}$, $t \leq T$, *are independently and identically distributed (i.i.d) for all* i *and* t *with* $E(\boldsymbol{\varepsilon}_{it}) = \mathbf{0}$, *and* $Var(\boldsymbol{\varepsilon}_{it}) = \Omega_{\varepsilon}$, Ω_{ε} *being a positive definite matrix.*

Under certain conditions it is possible to relax the cross-sectional independence assumption. Conley (1999), for example, presents in the context of a spatial model an economic distance metric to order the data over the cross section. In panels with N and T sufficiently large, Bai and Ng (2002), Moon and Perron (2003), Phillips and Sul (2002) and Pesaran (2002, 2003) consider cross-sectional dependence with a residual factor structure. Exploring the issue of cross-sectional dependence in the context of the PAVR model with T fixed is beyond the scope of the present paper. As for the assumption that the disturbances are identically distributed across t, we will discuss how this assumption can be weakened in Section 4 below.

Assumption (G3) The eigenvalues of Φ are either equal to unity or fall inside the unit circle.

Let

$$\boldsymbol{\xi}_{it} = \mathbf{w}_{it} - \boldsymbol{\mu}_i \tag{2.2}$$

Then (2.1) can be alternatively written as

$$(\mathbf{I}_m - \Phi L)\boldsymbol{\xi}_{it} = \boldsymbol{\varepsilon}_{it,}, \quad \text{for} \ \ t = 2, \ldots, T \tag{2.3}$$

and

$$\Delta \mathbf{w}_{i1} = -(\mathbf{I}_m - \Phi)(\mathbf{w}_{i0} - \boldsymbol{\mu}_i) + \boldsymbol{\varepsilon}_{i1} \tag{2.4}$$

When T is fixed, it is necessary to consider the initialization of the $\mathbf{w}_{it}$ process for estimation and inference. We assume that

Assumption (G4) The initial deviations, $\boldsymbol{\xi}_{i0}$, are identically and independently distributed across i, with zero means and the constant non-singular variance, $E(\boldsymbol{\xi}_{i0}\boldsymbol{\xi}_{i0}') = \Psi_{\boldsymbol{\xi}_0}$.

Under assumption G4, if all the eigenvalues of Φ are all inside the unit circle, the process 2.1 can either start from infinite past or finite past. If some of the eigenvalues of Φ are unity, then the nonstationary direction can only start from a finite past. For details, see Appendix A.

The PVAR(1) model (2.1) is the generalization of the univariate dynamic panel data model considered, for example, in Ahn and Schmidt (1995) to the multivariate context, except for the parameterization of the individual-specific effects. The multivariate counterpart of the Ahn and Schmidt formulation is given by

$$\mathbf{w}_{it} = \mathbf{a}_i + \Phi \mathbf{w}_{i,t-1} + \boldsymbol{\varepsilon}_{it}, \tag{2.5}$$

which would be equivalent to (2.1) when all eigenvalues of Φ fall inside the unit circle. However, in the presence of unit roots the two specifications (2.1) and (2.5) will have different trend properties, with the unrestricted intercepts specification (2.5) exhibiting linear trends whilst the restricted specification (2.1) does not. In what follows we adopt (2.1) as the data generating mechanism, although for estimation purposes it is often more convenient to work with (2.5).

3 Random Effects Specification

In this case the general assumptions, (**G1**) to (**G4**), need to be supplemented with additional assumptions on the individual-specific effects, $\boldsymbol{\mu}_i$. In particular, we shall make the following assumptions

Assumption (R1):

$$Var(\mathbf{a}_i) = \Omega_a \text{ and } Cov(\mathbf{a}_i, \boldsymbol{\varepsilon}_{it}) = \mathbf{0}, \text{ for all } i, \text{ and } t = 1, 2, ..., T. \tag{3.1}$$

This is a standard assumption in the random coefficient model and together with the general assumptions $(G1)$-$(G4)$ yields

$$\mathbf{r}_{it} = \begin{pmatrix} \mathbf{w}_{i0} \\ \mathbf{a}_i \\ \boldsymbol{\varepsilon}_{it} \end{pmatrix} \overset{i.i.d.}{\sim} (\mathbf{0}, \Omega_r), \text{ for all } i \text{ and } t = 1, 2, \ldots, T, \tag{3.2}$$

where

$$\Omega_r = \begin{pmatrix} \Omega_0 & \Omega_{0a} & \mathbf{0} \\ \Omega_{0a}' & \Omega_a & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Omega_\varepsilon \end{pmatrix}, \tag{3.3}$$

and Ω_0 and Ω_a are respectively positive definite and non-negative definite matrices.[5]

[5] Assumption (**R1**) could be relaxed for example to allow the individual effects $\mathbf{a}_i$ to have a common non-zero mean. Non-zero correlations between the disturbances $\boldsymbol{\varepsilon}_{it}$ and the initial observations $\mathbf{w}_{i0}$/the individual effects $\mathbf{a}_i$ could also be allowed for, but they will not be considered here since in general it is not possible to test whether these correlations are zero or non-zero. See Ahn and Schmidt (1995) for more detailed discussion of this in the univariate setting.

Assumption (R2) *All elements of the cross-product matrices* $\mathbf{r}_{it}\mathbf{r}_{it}'$, $t = 1, 2, \ldots, T$, *have finite second-order moments.*

Denote the $\left[2m^2 + 3m\left(m+1\right)/2\right] \times 1$ vector of unknown coefficients by $\boldsymbol{\theta}$,

$$\boldsymbol{\theta} = \left(\begin{array}{ccccc} \boldsymbol{\phi}', & \boldsymbol{\sigma}_{\varepsilon}', & \boldsymbol{\sigma}_{\mathbf{a}}', & \boldsymbol{\sigma}_{0}', & \boldsymbol{\sigma}_{0\mathbf{a}}' \end{array} \right)', \tag{3.4}$$

where $\boldsymbol{\phi} = vec\left(\Phi\right)$, $\boldsymbol{\sigma}_{\varepsilon} = vech\left(\Omega_{\varepsilon}\right)$, $\boldsymbol{\sigma}_{\mathbf{a}} = vech\left(\Omega_{\mathbf{a}}\right)$, $\boldsymbol{\sigma}_{0} = vech\left(\Omega_{0}\right)$, and $\boldsymbol{\sigma}_{0\mathbf{a}} = vec\left(\Omega_{0\mathbf{a}}\right)$.

Assumption (R3) $\boldsymbol{\theta} \in \Theta$, *where* Θ *is a compact subset of* $\Re^{2m^2+3m(m+1)/2}$, *and the true parameter vector,* $\boldsymbol{\theta}_0$, *falls in the interior of* Θ.

Assumption **R1** can be derived from more primitive assumptions concerning the initialization of the $\mathbf{w}_{it}$ process. For example, in the case when $\mathbf{w}_{it}$ is stationary and has started in the infinite past we have

$$\mathbf{w}_{i0} = \left(\mathbf{I}_m - \Phi\right)^{-1} \mathbf{a}_i + \sum_{j=0}^{\infty} \Phi^j \boldsymbol{\varepsilon}_{i,-j}, \tag{3.5}$$

and hence

$$\Omega_0 = Var\left(\mathbf{w}_{i0}\right) = \left(\mathbf{I}_m - \Phi\right)^{-1} \Omega_a \left(\mathbf{I}_m - \Phi\right)'^{-1} + \sum_{j=0}^{\infty} \Phi^j \Omega_{\varepsilon} \Phi'^j, \tag{3.6}$$

and

$$\Omega_{0a} = Cov\left(\mathbf{w}_{i0}, \mathbf{a}_i\right) = \Omega_a \left(\mathbf{I}_m - \Phi\right)'^{\,-1}.$$

See Appendix A for further details and other initialization examples.

To derive the RE-QMLE of $\boldsymbol{\theta}$, we let

$$\mathbf{w}_i = \begin{pmatrix} \mathbf{w}_{i0} \\ \mathbf{w}_{i1} \\ \vdots \\ \mathbf{w}_{iT} \end{pmatrix}, \qquad \text{and} \qquad \boldsymbol{\eta}_i = \begin{pmatrix} \mathbf{w}_{i0} \\ \mathbf{a}_i + \boldsymbol{\varepsilon}_{i1} \\ \mathbf{a}_i + \boldsymbol{\varepsilon}_{i2} \\ \vdots \\ \mathbf{a}_i + \boldsymbol{\varepsilon}_{iT} \end{pmatrix}, \tag{3.7}$$

and note that

$$\boldsymbol{\eta}_i = \mathbf{R}\mathbf{w}_i, \tag{3.8}$$

where $\mathbf{R}$ is a matrix of dimension $m\left(T+1\right) \times m\left(T+1\right)$ given by

$$\mathbf{R} = \begin{pmatrix} \mathbf{I}_m & & & \mathbf{0} \\ -\Phi & \mathbf{I}_m & & \\ & \ddots & \ddots & \\ \mathbf{0} & & -\Phi & \mathbf{I}_m \end{pmatrix}. \tag{3.9}$$

Clearly, $|\mathbf{R}| = 1$. From (3.2) we now have

$$E(\boldsymbol{\eta}_i) = \mathbf{0}, \qquad \text{and} \qquad Var(\boldsymbol{\eta}_i) = \Sigma_{\boldsymbol{\eta}}, \tag{3.10}$$

where

$$\Sigma_{\boldsymbol{\eta}} = \begin{pmatrix} \Omega_0 & \boldsymbol{\iota}_T' \otimes \Omega_{0a}' \\ \boldsymbol{\iota}_T \otimes \Omega_{0a} & \mathbf{I}_T \otimes \Omega_\varepsilon + \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \otimes \Omega_a \end{pmatrix}, \tag{3.11}$$

with $\boldsymbol{\iota}_T$ being a $T \times 1$ vector of ones. It follows that

$$E(\mathbf{w}_i) = \mathbf{0}, \qquad \text{and} \qquad Var(\mathbf{w}_i) = \Sigma_{\mathbf{w}} = \mathbf{R}^{-1} \Sigma_{\boldsymbol{\eta}} \mathbf{R}'^{-1}. \tag{3.12}$$

For RE-QML estimation of $\boldsymbol{\theta}$ we use the following log-likelihood function, which assumes normally distributed errors, as the criterion function:[6]

$$\mathfrak{L}(\boldsymbol{\theta}) = -\frac{mN(T+1)}{2} \log(2\pi) - \frac{N}{2} \log|\Sigma_{\boldsymbol{\eta}}| - \frac{N}{2} tr\left(\Sigma_{\mathbf{w}}^{-1} \mathbf{S}_{N,\mathbf{w}}\right), \tag{3.13}$$

where

$$\mathbf{S}_{N,\mathbf{w}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{w}_i \mathbf{w}_i'. \tag{3.14}$$

We then have the following proposition:[7]

Proposition 3.1 *Under assumptions (**G1**)-(**G4**), (**R1**), and (**R2**), and assuming that (2.1) holds, then as $N \to \infty$, $\mathbf{S}_{N,\mathbf{w}}$ converges almost surely to the non-stochastic matrix $\Sigma_{\mathbf{w}}$, and the random effects QML estimator (RE-QMLE) of $\boldsymbol{\theta}$, defined by*

$$\widehat{\boldsymbol{\theta}}_{QML} = \arg\max_{\boldsymbol{\theta}} \left[\mathfrak{L}(\boldsymbol{\theta})\right] \tag{3.15}$$

*is consistent. Furthermore, under the additional assumption (**R2**)*

$$\sqrt{N}\left(\widehat{\boldsymbol{\theta}}_{QML} - \boldsymbol{\theta}_0\right) \overset{a}{\backsim} N(\mathbf{0}, \mathcal{V}_{QML}), \tag{3.16}$$

where

$$\mathcal{V}_{QML} = \mathbf{H}_{\mathfrak{L}}^{-1} \mathbf{G}_{\mathfrak{L}} \mathbf{H}_{\mathfrak{L}}^{-1}, \tag{3.17}$$

$$\mathbf{H}_{\mathfrak{L}} = \lim_{N \to \infty} E\left[-\frac{1}{N} \frac{\partial^2 \mathfrak{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \, \partial \boldsymbol{\theta}'}\right], \quad \text{and} \quad \mathbf{G}_{\mathfrak{L}} = \lim_{N \to \infty} E\left[\frac{1}{N} \frac{\partial \mathfrak{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mathfrak{L}(\boldsymbol{\theta})'}{\partial \boldsymbol{\theta}}\right], \tag{3.18}$$

with $\mathbf{H}_{\mathfrak{L}}$ being a positive definite matrix.

[6] Notice that Σ_η is non-singular even if initialization restrictions such as $\Omega_{\mathbf{a}} = \mathbf{0}$ and $\Omega_{0\mathbf{a}} = \mathbf{0}$ that follow under $\Phi = \mathbf{I}_m$ are imposed. However, as noted above, in our estimation set-up $\Omega_{\mathbf{a}}$ and $\Omega_{0\mathbf{a}}$ will be treated as unrestricted coefficient matrices.

[7] A proof can be established using familiar techniques as reviewed, for example, in White (1994).

Remark 3.1 *A consistent estimate of the variance-covariance matrix of* $\widehat{\boldsymbol{\theta}}_{QML}$ *robust to violations of the information matrix equality is given by*

$$\frac{1}{N}\widehat{\mathbf{H}}_N^{-1}\widehat{\mathbf{G}}_N\widehat{\mathbf{H}}_N^{-1}, \tag{3.19}$$

where $\widehat{\mathbf{H}}_N$ *and* $\widehat{\mathbf{G}}_N$ *are the matrices* $\mathbf{H}_N$ *and* $\mathbf{G}_N$ *defined below and evaluated at* $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}_{QML}$

$$\mathbf{H}_N = -\frac{1}{N}\frac{\partial^2 \mathfrak{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\, \partial \boldsymbol{\theta}'}, \quad \mathbf{G}_N = \frac{1}{N}\left(\frac{\partial \mathfrak{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)\left(\frac{\partial \mathfrak{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)'. \tag{3.20}$$

Remark 3.2 *If time-specific effects are present, and* $\mathbf{w}_{it}$ *is generated by*[8]

$$(\mathbf{I}_m - \Phi L)(\mathbf{w}_{it} - \boldsymbol{\mu}_i - \boldsymbol{\delta}_t) = \boldsymbol{\varepsilon}_{it}, \qquad i = 1, 2, \ldots, N;\ t = 1, 2, \ldots, T, \tag{3.21}$$

where $\boldsymbol{\delta}_t$ *is an* $m \times 1$ *vector of time-specific effects, then upon redefining for estimation purposes*

$$\mathbf{w}_i = \begin{pmatrix} \mathbf{w}_{i0} - \boldsymbol{\delta}_0 \\ \mathbf{w}_{i1} - \boldsymbol{\delta}_1 \\ \vdots \\ \mathbf{w}_{iT} - \boldsymbol{\delta}_T \end{pmatrix}, \quad \textit{and} \quad \boldsymbol{\eta}_i = \begin{pmatrix} \mathbf{w}_{i0} - \boldsymbol{\delta}_0 \\ \mathbf{a}_i + \boldsymbol{\varepsilon}_{i1} \\ \mathbf{a}_i + \boldsymbol{\varepsilon}_{i2} \\ \vdots \\ \mathbf{a}_i + \boldsymbol{\varepsilon}_{iT} \end{pmatrix}, \tag{3.22}$$

the log-likelihood function is again given by (**??**)*. It can be shown that the RE-QMLE of* $\boldsymbol{\delta}_t$ *is given by*

$$\widehat{\boldsymbol{\delta}}_t = \frac{1}{N}\sum_{i=1}^{N}\mathbf{w}_{it}, \qquad t = 0, 1, \ldots, T. \tag{3.23}$$

In the special case where $\boldsymbol{\delta}_t = \boldsymbol{\delta} t$, $t = 0, 1, \ldots, T$, *the RE-QMLE of* $\boldsymbol{\delta}$ *can be obtained using the following weighted average of the unrestricted estimates,* $\widehat{\boldsymbol{\delta}}_t$*:*

$$\widehat{\boldsymbol{\delta}} = \left(\sum_{t=0}^{T}\sum_{s=0}^{T}\Sigma_{\mathbf{w}}^{ts}\right)^{-1}\left\{\sum_{s=0}^{T}\sum_{t=0}^{T}\Sigma_{\mathbf{w}}^{ts}\widehat{\boldsymbol{\delta}}_s\right\}, \tag{3.24}$$

where we have partitioned $\Sigma_{\mathbf{w}}^{-1}$ *into* $(T+1)^2$ *blocks of dimension* $m \times m$*,*

$$\Sigma_{\mathbf{w}}^{-1} = \begin{pmatrix} \Sigma_{\mathbf{w}}^{00} & \Sigma_{\mathbf{w}}^{01} & \cdots & \Sigma_{\mathbf{w}}^{0T} \\ \Sigma_{\mathbf{w}}^{10} & \Sigma_{\mathbf{w}}^{11} & \cdots & \Sigma_{\mathbf{w}}^{1T} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{\mathbf{w}}^{T0} & \Sigma_{\mathbf{w}}^{T1} & \cdots & \Sigma_{\mathbf{w}}^{TT} \end{pmatrix}. \tag{3.25}$$

The RE-QMLE of the remaining parameters, $\boldsymbol{\theta}$*, can be computed using the concentrated log-likelihood function.*[9]

[8] The presence of N cross-sectional units allows us to consider a nonparametrically specified common trend for all cross-sectional units.

[9] Detailed derivations of (3.23) and (3.24) are contained in a note available from the authors upon request.

4 Fixed Effects Specification

Under the fixed effects specification no restrictions need to be placed on the probability distribution function generating the individual-specific effects $\boldsymbol{\mu}_i$ in (2.1) (or, in unrestricted form, $\mathbf{a}_i$ in (2.5)). In particular, assumptions (**R1**) and (**R2**) are no longer required. It can then be allowed, for example, that: (i) the individual effects are dependently distributed, (ii) the individual effects are heteroskedastic, (iii) the individual effects are (more generally) characterized by a joint probability distribution function with the number of unknown parameters increasing at the same rate as the number of cross-sectional observations in the panel, and (iv) the individual effects do not have moments.

Following standard practice the $\boldsymbol{\mu}_i$'s can be eliminated by first-differencing (2.1), namely[10]

$$\Delta\mathbf{w}_{it} = \Phi\Delta\mathbf{w}_{i,t-1} + \Delta\boldsymbol{\varepsilon}_{it}, \qquad t = 2, 3, \ldots, T. \tag{4.1}$$

The first-differenced model (4.1) allows us to obtain the probability distribution of $\Delta\mathbf{w}_{i2}$, $\Delta\mathbf{w}_{i3}$, $\ldots$, $\Delta\mathbf{w}_{iT}$, conditional on $\Delta\mathbf{w}_{i1}$. While it would be tempting to base the QML estimator of Φ on the associated conditional likelihood, the resultant estimator would be inconsistent as $N \longrightarrow \infty$ when T is finite, as discussed in the univariate setting by Hsiao, Pesaran, and Tahmiscioglu (2002). To obtain a consistent QML estimator one needs to work with the unconditional joint probability distribution of $(\Delta\mathbf{w}_{i1}, \Delta\mathbf{w}_{i2}, \ldots, \Delta\mathbf{w}_{iT})$, and ensure that this joint distribution is free of the incidental parameters problem. The latter condition is obviously satisfied if the unconditional distribution of $\Delta\mathbf{w}_{i1}$ does not depend on any incidental parameters. Therefore, for the fixed effects specification we shall supplement assumptions (**G1**)-(**G4**) with

Assumption (F1) The following moment restrictions are satisfied:

$$(\mathbf{I}_m - \Phi)\, E\left(\boldsymbol{\kappa}_{i0}\boldsymbol{\varepsilon}_{i1}'\right) = \mathbf{0}, \tag{4.2}$$

and

$$E\left(\boldsymbol{\kappa}_{i0}\Delta\boldsymbol{\varepsilon}_{it}'\right) = \mathbf{0}, \text{ for } t = 2, 3, ..., T, \tag{4.3}$$

where $\boldsymbol{\kappa}_{i0} = (\mathbf{I}_m - \Phi)(\mathbf{w}_{i0} - \boldsymbol{\mu}_i)$.

Combining this assumption with (**G1**)-(**G4**) and using (2.4) we now have[11]

$$\Delta\mathbf{w}_{i1} \overset{i.i.d.}{\sim} (\mathbf{0}, \Psi),$$

[10] Hsiao, Pesaran, and Tahmiscioglu (2002) in the univariate context show that the QML estimator is invariant to the choice of the $T \times (T+1)$ transformation matrix $\mathcal{P}$ that is of rank T and eliminates the individual-specific effects, namely has the property that $\mathcal{P}\mathfrak{c} = \mathbf{0}$, with $\mathfrak{c}$ being a vector of constants of dimension $(T+1) \times 1$. The argument in Hsiao, Pesaran, and Tahmiscioglu (2002) readily extends to the multivariate setting considered here.

[11] Assumptions (**G4**) and (**F1**) can be relaxed to allow for $\boldsymbol{\kappa}_{i0}$ to have a constant non-zero mean, and for $Cov(\boldsymbol{\kappa}_{i0}, \boldsymbol{\varepsilon}_{i1})$ and $Cov(\boldsymbol{\kappa}_{i0}, \Delta\boldsymbol{\varepsilon}_{it})$, for $t = 2, 3, \ldots, T$, to be non-zero and possibly time-varying (but still homogeneous across i).

and

$$Cov\left(\Delta\mathbf{w}_{i1}, \Delta\boldsymbol{\varepsilon}_{i2}\right) = -\Omega_{\varepsilon}, \text{and}, Cov\left(\Delta\mathbf{w}_{i1}, \Delta\boldsymbol{\varepsilon}_{it}\right) = \mathbf{0}, t = 3, 4, \ldots, T.$$

where

$$\Psi = \left(\mathbf{I}_m - \Phi\right)\Psi_{\xi_0}\left(\mathbf{I}_m - \Phi'\right) + \Omega_{\varepsilon}, \tag{4.4}$$

and Ψ_{ξ_0} is already defined in Assumption **G4**.

Remark 4.1 *It is clear that the individual effects $\boldsymbol{\mu}_i$ do not enter the initial first differences, $\Delta\mathbf{w}_{i1}$. The first-differencing operation simultaneously deals with the incidental parameters and unit root problems.*[12]

Remark 4.2 *It is important to note that assumption (**F1**) imposes homogeneity restrictions on a linear combination of the initial deviations, $\left(\mathbf{I}_m - \Phi\right)\boldsymbol{\xi}_{i0}$, and the initial error terms, $\boldsymbol{\varepsilon}_{i1}$, for all i, without imposing any such restrictions on the individual effects, $\boldsymbol{\mu}_i$, themselves.*

Finally, for the fixed effects specification we make the following moment and parameter space assumptions:

Assumption (F2) *The second moments of the cross-product matrix $\Delta\mathbf{r}_{it}\Delta\mathbf{r}_{it}'$, $t = 1, 2, \ldots, T$, with*

$$\Delta\mathbf{r}_{it} = \begin{pmatrix} \Delta\mathbf{w}_{i1} \\ \Delta\boldsymbol{\varepsilon}_{it} \end{pmatrix}, \tag{4.5}$$

exist.

Denote the $\left[m^2 + m\left(m+1\right)\right] \times 1$ vector of unknown coefficients by $\boldsymbol{\rho}$,

$$\boldsymbol{\rho} = \begin{pmatrix} \boldsymbol{\phi}', & \boldsymbol{\sigma}_{\varepsilon}', & \boldsymbol{\psi}' \end{pmatrix}', \tag{4.6}$$

where $\boldsymbol{\phi} = vec\left(\Phi\right)$, $\boldsymbol{\sigma}_{\varepsilon} = vech\left(\Omega_{\varepsilon}\right)$, and $\boldsymbol{\psi} = vech\left(\Psi\right)$.

Assumption (F3) *$\boldsymbol{\rho} \in \Xi$, where Ξ is a compact subset of $\Re^{m^2+m(m+1)}$, and the true parameter vector, $\boldsymbol{\rho}_0$, lies in the interior of Ξ.*

To derive the FE-QMLE of $\boldsymbol{\rho}$, we need to derive the second moment structure of $\left(\Delta\mathbf{w}_{i1}, \Delta\mathbf{w}_{i2}, \ldots, \Delta\mathbf{w}_{iT}\right)$. We let

$$\Delta\mathbf{w}_i = \begin{pmatrix} \Delta\mathbf{w}_{i1} \\ \Delta\mathbf{w}_{i2} \\ \vdots \\ \Delta\mathbf{w}_{iT} \end{pmatrix}, \quad \text{and} \quad \Delta\boldsymbol{\eta}_i = \begin{pmatrix} \Delta\mathbf{w}_{i1} \\ \Delta\boldsymbol{\varepsilon}_{i2} \\ \Delta\boldsymbol{\varepsilon}_{i3} \\ \vdots \\ \Delta\boldsymbol{\varepsilon}_{iT} \end{pmatrix}. \tag{4.7}$$

[12] As will be discussed in detail in Section 5, unlike in time-series models, first-differencing in panels with fixed time dimension still allows us to identify and estimate the long-run (level) relations that are of economic interest irrespective of the unit root and cointegrating properties of the $\mathbf{w}_{it}$ process.

From (4.7) it immediately follows that

$$\Delta\boldsymbol{\eta}_i = \mathbf{R}\,\Delta\mathbf{w}_i, \tag{4.8}$$

where $\mathbf{R}$ is given by (3.9), but its dimension now is $mT \times mT$. The mean and the variance-covariance matrix of $\Delta\boldsymbol{\eta}_i$ are now easily obtained. We have

$$E\left(\Delta\boldsymbol{\eta}_i\right) = \mathbf{0}, \qquad \text{and} \qquad Var\left(\Delta\boldsymbol{\eta}_i\right) = \Sigma_{\Delta\boldsymbol{\eta}}, \tag{4.9}$$

where

$$\Sigma_{\Delta\boldsymbol{\eta}} = \begin{pmatrix} \Psi & -\Omega_\varepsilon & & & & & \mathbf{0} \\ -\Omega_\varepsilon & 2\Omega_\varepsilon & -\Omega_\varepsilon & & & & \\ & -\Omega_\varepsilon & 2\Omega_\varepsilon & -\Omega_\varepsilon & & & \\ & & & \ddots & & & \\ & & & & -\Omega_\varepsilon & 2\Omega_\varepsilon & -\Omega_\varepsilon \\ \mathbf{0} & & & & & -\Omega_\varepsilon & 2\Omega_\varepsilon \end{pmatrix}. \tag{4.10}$$

It follows that

$$E\left(\Delta\mathbf{w}_i\right) = \mathbf{0}, \qquad \text{and} \qquad Var\left(\Delta\mathbf{w}_i\right) = \Sigma_{\Delta\mathbf{w}} = \mathbf{R}^{-1}\Sigma_{\Delta\boldsymbol{\eta}}\mathbf{R}'^{-1}. \tag{4.11}$$

We base the QML estimation of $\boldsymbol{\rho}$ on the following log-likelihood function derived from the joint probability distribution of $(\Delta\mathbf{w}_{i1}, \Delta\mathbf{w}_{i2}, \ldots, \Delta\mathbf{w}_{iT})$ under the normality assumption:[13]

$$\ell\left(\boldsymbol{\rho}\right) = -\frac{mNT}{2}\log\left(2\pi\right) - \frac{N}{2}\log\left|\Sigma_{\Delta\boldsymbol{\eta}}\right| - \frac{N}{2}tr\left(\Sigma_{\Delta\mathbf{w}}^{-1}\mathbf{S}_{N,\Delta\mathbf{w}}\right), \tag{4.12}$$

where

$$\mathbf{S}_{N,\Delta\mathbf{w}} = \frac{1}{N}\sum_{i=1}^{N}\Delta\mathbf{w}_i\,\Delta\mathbf{w}_i'. \tag{4.13}$$

The following proposition establishes the properties of the resultant QML estimator:

Proposition 4.1 *Under assumptions (***G1***)-(***G4***), (***F1***), and (***F2***), and assuming that (2.1) holds, then as $N \to \infty$, $\mathbf{S}_{N,\Delta\mathbf{w}}$ converges almost surely to the non-stochastic matrix $\Sigma_{\Delta\mathbf{w}}$, and the fixed effects QML estimator (FE-QMLE) of $\boldsymbol{\rho}$, defined by*

$$\widehat{\boldsymbol{\rho}}_{QML} = \arg\max_{\boldsymbol{\rho}}\ \left[\ell\left(\boldsymbol{\rho}\right)\right], \tag{4.14}$$

*is consistent. Furthermore, under the additional assumption (***F2)**

$$\sqrt{N}\left(\widehat{\boldsymbol{\rho}}_{QML} - \boldsymbol{\rho}_0\right) \overset{a}{\backsim} N(\mathbf{0}, \mathfrak{V}_{QML}), \tag{4.15}$$

[13] The likelihood function (4.11) holds whether T is finite or approaches to infinity. However, if $T \longrightarrow \infty$, one can estimate μ_i consistently, hence one may apply MLE to (2.1) instead of working with (4.11).

where

$$\mathfrak{V}_{QML} = \mathbf{H}_\ell^{-1}\mathbf{G}_\ell\mathbf{H}_\ell^{-1}, \tag{4.16}$$

$$\mathbf{H}_\ell = \lim_{N\to\infty} E\left[-\frac{1}{N}\frac{\partial^2\ell(\boldsymbol{\rho})}{\partial\boldsymbol{\rho}\,\partial\boldsymbol{\rho}'}\right], \quad \text{and} \quad \mathbf{G}_\ell = \lim_{N\to\infty} E\left[\frac{1}{N}\frac{\partial\ell(\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial\ell(\boldsymbol{\rho})'}{\partial\boldsymbol{\rho}}\right], \tag{4.17}$$

with $\mathbf{H}_\ell$ *being a positive definite matrix.*

A consistent estimate of the variance-covariance matrix of $\widehat{\boldsymbol{\rho}}_{QML}$ can be obtained using the counterpart of (3.19).

Remark 4.3 *If* $\mathbf{w}_{it}$ *is generated by the fixed effects counterpart of (3.21), so that time-specific effects are present, we have that*

$$(\mathbf{I}_m - \Phi L)(\Delta\mathbf{w}_{it} - \boldsymbol{\gamma}_t) = \Delta\boldsymbol{\varepsilon}_{it}, \qquad i = 1,2,\ldots,N;\ t = 2,3,\ldots,T, \tag{4.18}$$

with $\boldsymbol{\gamma}_t = \Delta\boldsymbol{\delta}_t$*. Upon redefining*

$$\Delta\mathbf{w}_i = \begin{pmatrix} \Delta\mathbf{w}_{i1} - \boldsymbol{\gamma}_1 \\ \Delta\mathbf{w}_{i2} - \boldsymbol{\gamma}_2 \\ \vdots \\ \Delta\mathbf{w}_{iT} - \boldsymbol{\gamma}_T \end{pmatrix}, \qquad \text{and} \qquad \Delta\boldsymbol{\eta}_i = \begin{pmatrix} \Delta\mathbf{w}_{i1} - \boldsymbol{\gamma}_1 \\ \Delta\boldsymbol{\varepsilon}_{i2} \\ \Delta\boldsymbol{\varepsilon}_{i3} \\ \vdots \\ \Delta\boldsymbol{\varepsilon}_{iT} \end{pmatrix}, \tag{4.19}$$

the log-likelihood function is again given by (4.12). Using similar derivations as in the random effects setting, it can be shown that the FE-QMLE of $\boldsymbol{\gamma}_t$ *is given by*

$$\widehat{\boldsymbol{\gamma}}_t = \frac{1}{N}\sum_{i=1}^{N}\Delta\mathbf{w}_{it}, \qquad t = 1,2,\ldots,T. \tag{4.20}$$

In the restricted case of $\boldsymbol{\gamma}_t = \boldsymbol{\gamma}$, $t = 1,2,\ldots,T$, *we have*

$$\widehat{\boldsymbol{\gamma}} = \left(\sum_{t=1}^{T}\sum_{s=1}^{T}\Sigma_{\Delta\mathbf{w}}^{ts}\right)^{-1}\left(\sum_{s=1}^{T}\sum_{t=1}^{T}\Sigma_{\Delta\mathbf{w}}^{ts}\widehat{\boldsymbol{\gamma}}_s\right), \tag{4.21}$$

where $\Sigma_{\Delta\mathbf{w}}^{-1}$ *is partitioned into* $m \times m$ *dimensional blocks* $\Sigma_{\Delta\mathbf{w}}^{ts}$, $t, s = 1,2,\ldots,T$, *analogous to the partition in (3.25).*

Remark 4.4 *Computation of the FE-QMLE is complicated by the fact that the matrix* $\Sigma_{\Delta\boldsymbol{\eta}}$ *will often be high-dimensional. However, to compute the determinant and inverse of* $\Sigma_{\Delta\boldsymbol{\eta}}$*, one may make use of the block-tridiagonal structure of* $\Sigma_{\Delta\boldsymbol{\eta}}$*. Applying the block* LDL' *factorization to* $\Sigma_{\Delta\boldsymbol{\eta}}$*, the latter may be factorized as* $\Sigma_{\Delta\boldsymbol{\eta}} = \mathbb{A}_L\mathbb{A}_D'\mathbb{A}_L'$*, where* $\mathbb{A}_D$ *is a block-diagonal matrix with* j*-th* *diagonal block given by* $\mathbb{A}_D^{(1)} = \Psi^{-1}$, $\mathbb{A}_D^{(j)} = \left[2\Omega_\varepsilon - \Omega_\varepsilon \mathbb{A}_D^{(j-1)}\Omega_\varepsilon\right]^{-1}$, $j = 2, 3, \ldots, T$, *and where* $\mathbb{A}_L = \mathbb{A}_D^{-1} - \mathbb{A}_U$, *with* $\mathbb{A}_U$ *being a block-subdiagonal matrix with all subdiagonal blocks equal to* Ω_ε.[14] *It then follows that*

$$\det(\Sigma_{\Delta\boldsymbol{\eta}}) = [\det(\mathbb{A}_L)]^2 \det(\mathbb{A}_D) = \frac{1}{\det(\mathbb{A}_D)} = \frac{1}{\det(\Psi)} \prod_{j=2}^{T} \det\left[\mathbb{A}_D^{(j)}\right]. \tag{4.22}$$

To compute the inverse of $\Sigma_{\Delta\boldsymbol{\eta}}$, *a computationally efficient scheme is to adapt the recursions based on Bowden's procedure in Binder and Pesaran (2000), which yields*

$$\Sigma^{-1}_{\Delta\boldsymbol{\eta},jl} = (T+1-l)\left[(j-1)\,\Omega_\varepsilon^{-1}\Psi - (j-2)\,\mathbf{I}_m\right]\left[T\Psi - (T-1)\,\Omega_\varepsilon\right]^{-1}, \qquad l \geq j, \tag{4.23}$$

and

$$\Sigma^{-1}_{\Delta\boldsymbol{\eta},jl} = (T+1-j)\left[T\Psi - (T-1)\,\Omega_\varepsilon\right]^{-1}\left[(l-1)\,\Psi\Omega_\varepsilon^{-1} - (l-2)\,\mathbf{I}_m\right], \qquad j > l, \tag{4.24}$$

where $\Sigma^{-1}_{\Delta\boldsymbol{\eta},jl}$ *denotes the* jl*-th block of* $\Sigma^{-1}_{\Delta\boldsymbol{\eta}}$, $j, l = 1, 2, \ldots, T$. *Further details of our numerical algorithm that renders computation of the FE-QMLE practically feasible even for high-dimensional systems are described in a note available upon request.*

As will be seen in Section 6 below, GMM estimators of Φ need not impose homoskedasticity of the error components $(\mathbf{I}_m - \Phi)\,\boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_{it}$, which, it may be argued, adds to their robustness. Since under the fixed effects specification no restrictions are placed on the distribution generating the individual effects $\boldsymbol{\mu}_i$, by default the FE-QMLE also allows for the possibility of cross-sectional heteroskedasticity in the combined error components, $(\mathbf{I}_m - \Phi)\,\boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_{it}$. Furthermore, the above analysis can readily accommodate intertemporal error variance heteroskedasticity. This can be done by relaxing (**G2**) so that the disturbances $\boldsymbol{\varepsilon}_{it}$ are distributed independently and identically for all i, and independently for all t with $Var(\boldsymbol{\varepsilon}_{it}) = \Omega_{\varepsilon_t}$, Ω_{ε_t} being positive definite matrices for all t. In this case $\Sigma_{\Delta\boldsymbol{\eta}}$ defined by (4.10) generalizes to

$$\Sigma_{\Delta\boldsymbol{\eta}} = \begin{pmatrix} \Psi & -\Omega_{\varepsilon_1} & & & & \\ -\Omega_{\varepsilon_1} & \Omega_{\varepsilon_1}+\Omega_{\varepsilon_2} & -\Omega_{\varepsilon_2} & & \mathbf{0} & \\ & & \ddots & & & \\ & \mathbf{0} & & -\Omega_{\varepsilon_{T-2}} & \Omega_{\varepsilon_{T-2}}+\Omega_{\varepsilon_{T-1}} & -\Omega_{\varepsilon_{T-1}} \\ & & & & -\Omega_{\varepsilon_{T-1}} & \Omega_{\varepsilon_{T-1}}+\Omega_{\varepsilon_T} \end{pmatrix}. \tag{4.25}$$

[14] For further details see, for example, Binder and Pesaran (2000), who in the context of the solution of multivariate linear rational expectations models discuss the block LDU factorization, of which the block LDL' factorization is a special case.

The FE-QMLE can now be derived under suitable parameterization of the error variance-covariance matrices Ω_{ε_t}, for $t = 1, 2, ..., T$.

Finally, it is also worth noting that under the random effects specification considered in Section 3 there are $m(T+1)(T+2)/2$ exploitable moment conditions, while under the fixed effects specification there are $mT(T+1)/2$ moment conditions, or $m(T+1)$ fewer moment restrictions. Therefore, in general one would expect the RE-QMLE to be asymptotically more efficient than the FE-QMLE. The finite sample importance of these additional moment conditions will be studied in Section 7, where the random and fixed effects QML estimators will be compared. Nevertheless, it should be clear that, in general, FE-QMLE is preferable to RE-QMLE; unless prior information is available that the individual effects are cross-sectionally homoskedastic and have finite moments of up to the fourth order.

5 Unit Roots and Cointegration in PVARs

Since the issues of unit roots and cointegration can be of significant interest in economic modelling, it is desirable to have procedures available to test for unit roots and cointegration rank even though T is finite. The asymptotic properties of the QML estimators set out above hold irrespective of the location of the eigenvalues of Φ and the size of T. Therefore, one may use the results of sections 3 and 4 to test for the presence of unit roots and cointegration.

In order to be able to interpret the rank of the matrix Π as the number of linearly independent cointegrating relations, it is necessary to know whether each of the variables in $\mathbf{w}_{it}$ follows an I(1) process. Our framework can be easily adapted to test for unit roots in short panel univariate autoregressive models. For $m = 1$ the equation to be estimated is

$$w_{it} = a_i + \phi w_{i,t-1} + \varepsilon_{it}, \qquad \varepsilon_{it} \backsim iid\left(0, \sigma_\varepsilon^2\right), \tag{5.1}$$

where w_{it} is now a scalar variable.[15] The unit root hypothesis

$$H_0: \ \phi = 1 \qquad \text{vs.} \qquad H_1: \ \phi < 1, \tag{5.2}$$

can now be tested under both the random and fixed effects specifications. Denoting the QML estimator of the slope coefficient under either model specification as $\widehat{\phi}$, a Wald type statistic of testing H_0 versus H_1 will be

$$t_\phi = \frac{\widehat{\phi} - 1}{\text{se}(\widehat{\phi})}, \tag{5.3}$$

[15] As for unit root testing in the time-series context, the appropriate order of augmentation of w_{it} is important for the validity of the test. In practice one may therefore need to consider higher-order cases as well. Here we confine ourselves to $p = 1$ for simplicity of exposition.

where $\text{se}(\widehat{\phi})$ denotes the standard error of $\widehat{\phi}$. Under the null hypothesis t_{ϕ} is asymptotically distributed as a standard normal variate as $N \to \infty$, for a fixed $T \geq 3$. The alternative hypothesis considered is homogeneous and one-sided. This test can be extended to models with serially correlated errors (to models with $p > 1$), so long as the slope homogeneity assumption is maintained. Unit root tests for panels with slope heterogeneity and more complicated dynamics have been proposed in the literature but require large N and T panels and are not valid when the time dimension is short.[16] A short T panel unit root test has been proposed by Harris and Tzavalis (1999), but requires bias corrections and could be difficult to extend to models with serially correlated errors.

The natural next step after the unit root tests have been carried out is to test for cointegration. Consider again the PVAR(1) model in the m variables $\mathbf{w}_{it}$, now assumed to be I(1). The hypothesis that $\mathbf{w}_{it} - \boldsymbol{\mu}_i$ is cointegrated with rank r versus rank $r+1$, $r = 0, 1, \ldots, m-1$, can be formulated as

$$H_r : \Phi = \mathbf{I}_m + \boldsymbol{\alpha}_r \boldsymbol{\beta}_r' \qquad \text{vs.} \qquad H_{r+1} : \Phi = \mathbf{I}_m + \boldsymbol{\alpha}_{r+1} \boldsymbol{\beta}_{r+1}', \tag{5.4}$$

where $\boldsymbol{\alpha}_r$ and $\boldsymbol{\beta}_r$ are $m \times r$ matrices of full column rank r. Since $\boldsymbol{\alpha}_r \boldsymbol{\beta}_r' = \boldsymbol{\alpha}_r \mathbb{K}\mathbb{K}^{-1} \boldsymbol{\beta}_r'$ for any $r \times r$ nonsingular matrix $\mathbb{K}$, one needs, in the absence of short-run restrictions, r restrictions on each of the r columns of $\boldsymbol{\beta}_r$.[17] A convenient procedure for the identification of $\boldsymbol{\beta}_r$ is to let

$$\boldsymbol{\beta}_r = \mathbb{H}\boldsymbol{\delta}_r + \mathbf{b}_r, \tag{5.5}$$

where $\mathbb{H}$ and $\mathbf{b}_r$ are, respectively, $m \times q$ and $m \times r$ matrices, both with known coefficients, and $\boldsymbol{\delta}$ is a $q \times r$ matrix with unknown coefficients. For example, if one chooses (as we shall do in what follows) the Phillips (1991) exact identification restriction that

$$\boldsymbol{\beta}_r = \left(\begin{array}{cc} \mathbf{I}_r, & \widetilde{\boldsymbol{\beta}}_r' \end{array} \right)', \tag{5.6}$$

where $\widetilde{\boldsymbol{\beta}}_r'$ is an $r \times (m-r)$ matrix with unrestricted coefficients, then

$$\mathbb{H} = \left(\begin{array}{cc} \mathbf{0}, & \mathbf{I}_{m-r} \end{array} \right)', \quad \mathbf{b}_r = \left(\begin{array}{cc} \mathbf{I}_r, & \mathbf{0} \end{array} \right)', \quad \text{and} \quad \boldsymbol{\delta}_r = \widetilde{\boldsymbol{\beta}}_r. \tag{5.7}$$

The QML estimators restricting the rank of the matrix Π can now be set out as before, noting that in the random effects case the unknown coefficients are now given by

$$\boldsymbol{\theta}_{\Pi} = \left(\begin{array}{cccccc} vec\,(\boldsymbol{\alpha}_r)', & vec\left(\widetilde{\boldsymbol{\beta}}_r\right)', & \boldsymbol{\sigma}_{\varepsilon}', & \boldsymbol{\sigma}_{\mathbf{a}}', & \boldsymbol{\sigma}_0', & \boldsymbol{\sigma}_{0\mathbf{a}}' \end{array} \right)'$$

[16] See, for example, Levin, Lin and Chu (2002), Im, Pesaran, and Shin (2003), and Maddala and Wu (1999). Extensions of these tests to models with cross section dependence have also been considered by Bai and Ng (2002), Moon and Perron (2003), Phillips and Sul (2002) and Pesaran (2003).

[17] For a more detailed discussion see, for example, Pesaran and Shin (2002). Also note that the extrema of the QML and MD criterion functions under $rank\,(\Pi) = r$ are invariant to the choice of $\mathbb{K}$.

, and in the fixed effects case are defined by

$$\boldsymbol{\rho}_{\Pi} = \left(\; vec\,(\boldsymbol{\alpha}_r)', \quad vec\left(\widetilde{\boldsymbol{\beta}}_r\right)', \quad \boldsymbol{\sigma}'_{\varepsilon}, \quad \boldsymbol{\psi}' \; \right)'.$$

The likelihood ratio test statistic of H_r versus H_{r+1} is asymptotically chi-square distributed with $(m-r)^2 - (m-r-1)^2 = 2\,(m-r) - 1$ degrees of freedom. (Imposing Π to be of rank r leaves $m^2 - (m-r)^2$ unrestricted coefficients in Π.)

Additional parameter restrictions or overidentifying restrictions can be formulated in terms of

$$vec\,(\Phi) = \mathbb{G}\varkappa + \mathsf{f}, \tag{5.8}$$

where $\mathbb{G}$ is an $m^2 \times q$ matrix and f an $m^2 \times 1$ vector, both with known elements, and $\varkappa$ is a $q \times 1$ vector of free parameters. A likelihood ratio test of (5.8) will be asymptotically chi-square distributed with $m^2 - q$ degrees of freedom.

We will document the (perhaps surprisingly) good small sample properties of the unit root and cointegration tests proposed in this section when the tests are based on the QML estimator in Section 7 below.

6 GMM Estimation

There now exists an extensive literature on the GMM estimation of univariate dynamic panel data models (for example, Arellano and Bond, 1991, Ahn and Schmidt, 1995, 1997, Arellano and Bover, 1995, Blundell and Bond, 1998, Alonso-Borrego and Arellano, 1999). However, just like Three Stage Least Squares estimation of a system of equations can be more efficient than the single equation based Two Stage Least Squares, in this section we shall generalize GMM estimation to a systems context, and show that if the PVAR model (2.1) contains unit roots, then the standard GMM approach (for example, Arellano and Bond, 1991) of using lagged level variables as instruments that are orthogonal to the disturbances of the first-differenced form of the model breaks down. We then discuss how this problem may be overcome using additional moment conditions implied by homoskedasticity and initialization restrictions of the type suggested in the case of univariate models by Ahn and Schmidt (1995), Arellano and Bover (1995), and Blundell and Bond (1998).

Standard GMM and its Breakdown under Unit Roots

The standard GMM estimator of Arellano and Bond (1991) employs instruments that are orthogonal to the disturbances of the first-differenced form of the model. For the PVAR(1) model (2.1), such instruments are given by levels of the dependent variables, $\mathbf{w}_{it}$, lagged two or more periods. The resulting orthogonality conditions may be written as

$$E\left[\left(\Delta\mathbf{w}_{it} - \Phi\,\Delta\mathbf{w}_{i,t-1} \right) \mathbf{q}'_{it} \right] = \mathbf{0}, \qquad t = 2, 3, \ldots, T, \tag{6.1}$$

where $\mathbf{q}_{it}$ is the $m\,(t-1)\times 1$ vector defined by

$$\mathbf{q}_{it} = \left(\begin{array}{cccc} \mathbf{w}_{i0}', & \mathbf{w}_{i1}', & \ldots, & \mathbf{w}_{i,t-2}' \end{array} \right)'. \tag{6.2}$$

To derive the standard GMM estimator of Φ based on the moment conditions (6.1), it will be useful to rewrite these moment conditions in stacked form as:

$$E\left[\mathbf{Q}_i' \left(\Delta\mathbf{W}_i - \Delta\mathbf{W}_{i,-1}\Phi'\right)\right] = \mathbf{0}, \tag{6.3}$$

where $\mathbf{Q}_i'$ is a matrix of dimension $mT\,(T-1)/2 \times (T-1)$ given by

$$\mathbf{Q}_i' = \left(\begin{array}{ccccc} \mathbf{q}_{i2} & \mathbf{0} & \mathbf{0} & & \mathbf{0} \\ \mathbf{0} & \mathbf{q}_{i3} & \mathbf{0} & & \\ & & & \ddots & \\ & \mathbf{0} & & & \mathbf{q}_{iT} \end{array} \right), \tag{6.4}$$

and $\Delta\mathbf{W}_i$ and $\Delta\mathbf{W}_{i,-1}$ are $(T-1)\times m$ dimensional matrices,

$$\Delta\mathbf{W}_i = \left(\begin{array}{cccc} \Delta\mathbf{w}_{i2}, & \Delta\mathbf{w}_{i3}, & \ldots, & \Delta\mathbf{w}_{iT} \end{array} \right)', \tag{6.5}$$

and

$$\Delta\mathbf{W}_{i,-1} = \left(\begin{array}{cccc} \Delta\mathbf{w}_{i1}, & \Delta\mathbf{w}_{i2}, & \ldots, & \Delta\mathbf{w}_{i,T-1} \end{array} \right)'. \tag{6.6}$$

The standard GMM estimator of $\boldsymbol{\phi} = vec\,(\Phi)$ is now given by[18]

$$\widehat{\boldsymbol{\phi}}_{GMM} = \left(\mathbf{S}_{\mathbf{ZX}}'\mathbf{D}_{\widehat{\mathbf{e}}}^{-1}\mathbf{S}_{\mathbf{ZX}}\right)^{-1}\mathbf{S}_{\mathbf{ZX}}'\mathbf{D}_{\widehat{\mathbf{e}}}^{-1}\mathbf{S}_{\mathbf{Zy}}, \tag{6.7}$$

where

$$\mathbf{S}_{\mathbf{ZX}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{X}_i, \quad \mathbf{S}_{\mathbf{Zy}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{y}_i, \quad \mathbf{D}_{\widehat{\mathbf{e}}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i'\Upsilon_{\widehat{\mathbf{e}}}\mathbf{Z}_i, \quad \Upsilon_{\widehat{\mathbf{e}}} = \frac{1}{N}\sum_{i=1}^{N}\widehat{\mathbf{e}}_i\widehat{\mathbf{e}}_i', \tag{6.8}$$

$$\mathbf{Z}_i' = \mathbf{Q}_i'\otimes\mathbf{I}_m, \quad \mathbf{X}_i = \Delta\mathbf{W}_{i,-1}\otimes\mathbf{I}_m, \quad \mathbf{y}_i = vec\left(\Delta\mathbf{W}_i'\right), \quad \mathbf{e}_i = vec\left(\Delta\mathbf{E}_i'\right), \tag{6.9}$$

and $\widehat{\Delta\mathbf{E}}_i = \Delta\mathbf{W}_i - \Delta\mathbf{W}_{i,-1}\widehat{\Phi}_{IE}'$, where $\widehat{\Phi}_{IE}$ is an initial consistent estimate of Φ such as the generalized instrumental variables estimator obtained using the formula (6.7), but with $\mathbf{D}_{\widehat{\mathbf{e}}}$ replaced by $\Lambda_{\mathbf{Q}}\otimes\Omega_{\boldsymbol{\varepsilon}}$, where

$$\Lambda_{\mathbf{Q}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{Q}_i'\mathbb{V}\mathbf{Q}_i, \tag{6.10}$$

[18] An alternative estimator of $\mathbf{D}_{\widehat{\mathbf{e}}}$ also used in the literature is given by $N^{-1}\sum_{i=1}^{N}\mathbf{Z}_i'\widehat{\mathbf{e}}_i\widehat{\mathbf{e}}_i'\mathbf{Z}_i$. See, for example, Arellano and Honoré (2001) and Baltagi (2001). However, our Monte Carlo experiments suggest $\mathbf{D}_{\widehat{\mathbf{e}}}$ to be preferable in the settings we consider, particularly for purposes of hypothesis testing. Arellano and Honoré (2001) also discuss how auxiliary assumptions can be used to impose further restrictions on $\mathbf{D}_{\widehat{\mathbf{e}}}$.

and $\mathbb{V}$ is the $(T-1)\times(T-1)$ matrix,

$$\mathbb{V}=\begin{pmatrix} 2 & -1 & 0 & & & & \\ -1 & 2 & -1 & & & 0 & \\ & & & \ddots & & & \\ & 0 & & & -1 & 2 & -1 \\ & & & & 0 & -1 & 2 \end{pmatrix}. \tag{6.11}$$

Since the resultant instrumental variables estimator is invariant to the choice of Ω_{ε}, without loss of generality the estimator may be computed replacing $\mathbf{D}_{\widehat{\mathbf{e}}}$ by $\Lambda_{\mathbf{Q}}\otimes\mathbf{I}_m$. Using the standard formula, a consistent estimate of the variance-covariance matrix of $\widehat{\boldsymbol{\phi}}_{GMM}$ can be obtained as

$$\frac{1}{N}\left(\mathbf{S}'_{\mathbf{ZX}}\mathbf{D}_{\widehat{\mathbf{e}}}^{-1}\mathbf{S}_{\mathbf{ZX}}\right)^{-1}. \tag{6.12}$$

The standard GMM estimator is consistent if all eigenvalues of Φ fall inside the unit circle, but breaks down if some eigenvalues of Φ are equal to unity. Note that a necessary condition for the GMM estimator (6.7) to exist is that $rank(\mathbf{Q}'_i\,\Delta\mathbf{W}_{i,-1})=m$ as $N\to\infty$. In the case where $\Phi=\mathbf{I}_m$, $rank\,(\mathbf{Q}'_i\,\Delta\mathbf{W}_{i,-1})$ as $N\to\infty$ is less than m, however. This is because when $\Phi=\mathbf{I}_m$, for $t=2,3,\ldots,T$ we have $\Delta\mathbf{w}_{it}=\boldsymbol{\varepsilon}_{it}$, and $\mathbf{w}_{it}=\mathbf{w}_{i0}+\mathbf{s}_{it}$, with $\mathbf{s}_{it}=\sum_{q=1}^{t}\boldsymbol{\varepsilon}_{iq}$, and thus it follows that for $t=2,3,\ldots,T$, $l=0,1,\ldots,t-2$, as $N\to\infty$

$$\frac{1}{N}\sum_{i=1}^{N}\Delta\mathbf{w}_{i,t-1}\,\mathbf{w}'_{il}=\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_{i,t-1}\left(\mathbf{w}_{i0}+\mathbf{s}_{il}\right)'\xrightarrow{p}\mathbf{0}, \tag{6.13}$$

where $\xrightarrow{p}$ denotes convergence in probability. In other words, when $\Phi=\mathbf{I}_m$, the elements of $\mathbf{q}_{it}$ are not legitimate instruments.[19] As some of the eigenvalues of Φ approach unity, the $\mathbf{q}_{it}$'s become weak instruments in the terminology of Staiger and Stock (1997).

Extended GMM

Nevertheless, a consistent GMM type estimator may be obtained by making use of additional moment conditions. One possibility is the extended GMM estimator proposed by Ahn and Schmidt (1995, 1997) which augments the standard moment conditions with those implied by homoskedasticity assumptions as in (**G2**). These are legitimate instruments regardless of the unit root and cointegrating properties of $\mathbf{w}_{it}-\boldsymbol{\mu}_i$. In the context of the PVAR(1) model (2.1) invoking homoskedasticity (over time) of the $\boldsymbol{\varepsilon}_{it}$'s yields the following two sets of moment conditions:

$$E\left[\left(\overline{\mathbf{w}}_i-\Phi\overline{\mathbf{w}}_{i,-1}\right)\left(\Delta\mathbf{w}_{it}-\Phi\Delta\mathbf{w}_{i,t-1}\right)'\right]=\mathbf{0},\quad t=2,3,\ldots,T, \tag{6.14}$$

[19] The same conclusion holds for PVAR(p) with more complicated derivation. A note containing a detailed argument is available from the authors upon request.

and

$$E\left[(\Delta \mathbf{w}_{i,t-1} - \Phi\Delta\mathbf{w}_{i,t-2})\,\mathbf{w}'_{i,t-2} - (\Delta\mathbf{w}_{it} - \Phi\Delta\mathbf{w}_{i,t-1})\,\mathbf{w}'_{i,t-1}\right] = \mathbf{0}, \quad t = 3,4,\ldots,T, \tag{6.15}$$

where

$$\overline{\mathbf{w}}_i = \frac{1}{T}\sum_{t=1}^{T}\mathbf{w}_{it}, \qquad \text{and} \qquad \overline{\mathbf{w}}_{i,-1} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{w}_{i,t-1}. \tag{6.16}$$

Note that the moment conditions (6.14) are nonlinear in Φ. Stacking the moment conditions (6.1), (6.14), and (6.15) as a $\left[m^2T\,(T-1)\,/2 + m^2\,(2T-3)\right] \times 1$ dimensional column vector, $\mathbf{m}(\mathbf{w}_i, \boldsymbol{\phi})$, the moment conditions can be rewritten as

$$E\left[\mathbf{m}(\mathbf{w}_i, \boldsymbol{\phi})\right] = \mathbf{0}. \tag{6.17}$$

Ahn and Schmidt's (1995, 1997) extended GMM estimator applied to the PVAR model (2.1) is then given by

$$\widehat{\boldsymbol{\phi}}_{GMM} = \arg\min_{\boldsymbol{\phi}} \left\{\mathbf{M}'_N(\boldsymbol{\phi})\left[\mathbf{W}_N(\boldsymbol{\phi})\right]^{-1}\mathbf{M}_N(\boldsymbol{\phi})\right\}, \tag{6.18}$$

where

$$\mathbf{M}_N(\boldsymbol{\phi}) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{m}(\mathbf{w}_i, \boldsymbol{\phi}), \quad \text{and} \quad \mathbf{W}_N(\boldsymbol{\phi}) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{m}(\mathbf{w}_i, \widehat{\boldsymbol{\phi}}_{IE})\,\mathbf{m}(\mathbf{w}_i, \widehat{\boldsymbol{\phi}}_{IE})', \tag{6.19}$$

with $\widehat{\boldsymbol{\phi}}_{IE}$ being an initial consistent estimate of $\boldsymbol{\phi}$. One possibility would be to use for this purpose the generalized instrumental variables estimator applied to the linear moment conditions (6.1) and (6.15) only.

Arellano and Bover (1995) and Blundell and Bond (1998) proposed an additional set of moment conditions which when applied to the PVAR model (6.20) can be written as

$$E\left[(\mathbf{w}_{it} - \Phi\mathbf{w}_{i,t-1})\,\Delta\mathbf{w}'_{i,t-1}\right] = \mathbf{0}, \qquad t = 2,3,\ldots,T. \tag{6.20}$$

It is readily seen that these conditions require that

$$(\mathbf{I}_m - \Phi)\,E\left[\boldsymbol{\mu}_i\,(\mathbf{w}_{i0} - \boldsymbol{\mu}_i)'\right](\Phi - \mathbf{I}_m)' = \mathbf{0}. \tag{6.21}$$

Thus the moment conditions (6.20) involve restrictions on the distribution of the initial observations, $\mathbf{w}_{i0}$, unless of course $\Phi = \mathbf{I}_m$. The Ahn and Schmidt (1995, 1997) homoskedasticity implied moment restrictions, (6.14) and (6.15), and the Arellano and Bover (1995) and Blundell and Bond (1998) initialization restrictions implied moment conditions, (6.20), can now be combined, after eliminating redundant conditions, to yield the *linear* moment conditions

$$E\left[\left(\mathbf{w}_{it}-\Phi\mathbf{w}_{i,t-1}\right)\mathbf{w}_{it}'-\left(\mathbf{w}_{i,t-1}-\Phi\mathbf{w}_{i,t-2}\right)\mathbf{w}_{i,t-1}'\right]=\mathbf{0},\qquad t=2,3,\ldots,T, \tag{6.22}$$

and (6.20).

To derive the extended GMM estimator of Φ based on the linear moment conditions (6.1), (6.20), and (6.22), it will be useful to rewrite the three sets of moment conditions in stacked form as:

$$E\left[\mathbf{P}_i'\left(\mathbf{W}_i-\mathbf{W}_{i,-1}\Phi'\right)\right]=\mathbf{0}, \tag{6.23}$$

where $\mathbf{P}_i'$ is a matrix of dimension $\left[mT\left(T-1\right)/2+2m\left(T-1\right)\right]\times T$,

$$\mathbf{P}_i'=\begin{pmatrix}\mathbf{P}_{1i}', & \mathbf{P}_{2i}', & \mathbf{P}_{3i}'\end{pmatrix}', \tag{6.24}$$

with $\mathbf{P}_{1i}$ a matrix of dimension $mT\left(T-1\right)/2\times T$ given by

$$\mathbf{P}_{1i}=\begin{pmatrix}-\mathbf{q}_{i2} & \mathbf{q}_{i2} & \mathbf{0} & \mathbf{0} & & \mathbf{0} & \\ \mathbf{0} & -\mathbf{q}_{i3} & \mathbf{q}_{i3} & \mathbf{0} & & & \\ & & & & \ddots & & \\ & \mathbf{0} & & & & \mathbf{0} & -\mathbf{q}_{iT} & \mathbf{q}_{iT}\end{pmatrix}, \tag{6.25}$$

$\mathbf{P}_{2i}$ a matrix of dimension $m\left(T-1\right)\times T$ given by

$$\mathbf{P}_{2i}=\begin{pmatrix}\mathbf{0} & \Delta\mathbf{w}_{i1} & \mathbf{0} & \mathbf{0} & & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \Delta\mathbf{w}_{i2} & \mathbf{0} & & \\ & & & & \ddots & \\ & \mathbf{0} & & & \mathbf{0} & \Delta\mathbf{w}_{i,T-1}\end{pmatrix}, \tag{6.26}$$

and $\mathbf{P}_{3i}$ a matrix of dimension $m\left(T-1\right)\times T$ given by

$$\mathbf{P}_{3i}=\begin{pmatrix}-\mathbf{w}_{i1} & \mathbf{w}_{i2} & \mathbf{0} & \mathbf{0} & & \mathbf{0} & \\ \mathbf{0} & -\mathbf{w}_{i2} & \mathbf{w}_{i3} & \mathbf{0} & & & \\ & & & & \ddots & & \\ & \mathbf{0} & & & & \mathbf{0} & -\mathbf{w}_{i,T-1} & \mathbf{w}_{iT}\end{pmatrix}, \tag{6.27}$$

and $\mathbf{W}_i$ and $\mathbf{W}_{i,-1}$ are $T\times m$ dimensional matrices,

$$\mathbf{W}_i=\begin{pmatrix}\mathbf{w}_{i1}, & \mathbf{w}_{i2}, & \ldots, & \mathbf{w}_{iT}\end{pmatrix}', \tag{6.28}$$

and

$$\mathbf{W}_{i,-1} = \begin{pmatrix} \mathbf{w}_{i0}, & \mathbf{w}_{i1}, & \ldots, & \mathbf{w}_{i,T-1} \end{pmatrix}'. \tag{6.29}$$

The extended GMM estimator of $\boldsymbol{\phi}$ based on the orthogonality, homoskedasticity, and initialization restrictions implied moment conditions (6.1), (6.20), and (6.22) is now given by

$$\widehat{\phi}_{GMM} = \left(\mathbf{S}'_{\mathring{\mathbf{Z}}\mathring{\mathbf{X}}}\mathbf{D}_{\widehat{\mathbf{u}}}^{-1}\mathbf{S}_{\mathring{\mathbf{Z}}\mathring{\mathbf{X}}}\right)^{-1}\mathbf{S}'_{\mathring{\mathbf{Z}}\mathring{\mathbf{X}}}\mathbf{D}_{\widehat{\mathbf{u}}}^{-1}\mathbf{S}_{\mathring{\mathbf{Z}}\mathring{\mathbf{y}}}, \tag{6.30}$$

where

$$\mathbf{S}_{\mathring{\mathbf{Z}}\mathring{\mathbf{X}}} = \frac{1}{N}\sum_{i=1}^{N}\mathring{\mathbf{Z}}'_i\mathring{\mathbf{X}}_i, \quad \mathbf{S}_{\mathring{\mathbf{Z}}\mathring{\mathbf{y}}} = \frac{1}{N}\sum_{i=1}^{N}\mathring{\mathbf{Z}}'_i\mathring{\mathbf{y}}_i, \quad \mathbf{D}_{\widehat{\mathbf{u}}} = \frac{1}{N}\sum_{i=1}^{N}\mathring{\mathbf{Z}}'_i\Upsilon_{\widehat{\mathbf{u}}}\mathring{\mathbf{Z}}_i, \quad \Upsilon_{\widehat{\mathbf{u}}} = \frac{1}{N}\sum_{i=1}^{N}\widehat{\mathbf{u}}_i\widehat{\mathbf{u}}'_i, \tag{6.31}$$

$$\mathring{\mathbf{Z}}'_i = \mathbf{P}'_i \otimes \mathbf{I}_m, \qquad \mathring{\mathbf{X}}_i = \mathbf{W}_{i,-1} \otimes \mathbf{I}_m, \qquad \mathring{\mathbf{y}}_i = vec\left(\mathbf{W}'_i\right), \qquad \widehat{\mathbf{u}}_i = vec\left(\widehat{\mathbf{U}}'_i\right), \tag{6.32}$$

and $\widehat{\mathbf{U}}_i = \mathbf{W}_i - \mathbf{W}_{i,-1}\widehat{\Phi}'_{IE}$, where $\widehat{\Phi}_{IE}$ is an initial consistent estimator of Φ, for example the generalized instrumental variables estimator based on (6.30), but with $\mathbf{D}_{\widehat{\mathbf{u}}}$ replaced by $\Lambda_{\mathbf{P}} \otimes \mathbf{I}_m$, where

$$\Lambda_{\mathbf{P}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{P}'_i\mathbf{P}_i. \tag{6.33}$$

Using the standard formula, a consistent estimate of the variance-covariance matrix of the extended GMM estimator (6.30) can be obtained as

$$\frac{1}{N}\left(\mathbf{S}'_{\mathring{\mathbf{Z}}\mathring{\mathbf{X}}}\mathbf{D}_{\widehat{\mathbf{u}}}^{-1}\mathbf{S}_{\mathring{\mathbf{Z}}\mathring{\mathbf{X}}}\right)^{-1}. \tag{6.34}$$

Remark 6.1 *The GMM estimators (6.7), (6.18), and (6.30) require that the second moments of* $\boldsymbol{\mu}_i$ *exist. For asymptotic normality of these estimators it will also be required that the fourth moments of* $\boldsymbol{\mu}_i$ *exist. The existence of these moments is not implied by any of the assumptions we had invoked for QML estimation under the fixed effects specification. Moreover, the number of moment conditions for GMM increases at the order of* T^2*, while the orthogonality conditions for QMLE remain the same as* T *increases, which can have implications for the finite sample performance of the two types of estimators.*

Remark 6.2 *Due to the use of levels variables (*$\mathbf{w}_{it}$*,*$t = 0, 1, ..., T-2$*) as instruments, the variance of the GMM estimators will depend on the distribution of the unobserved individual effects,,* $\boldsymbol{\mu}_i$*, a dependence that does not disappear with* $N \to \infty$*. As shown in Appendix B the asymptotic variance of the standard GMM estimator is in fact an increasing function of* Ω_μ*, the variance matrix of individual effects,* $\boldsymbol{\mu}_i$*, in the sense that for any two variance matrices,* $\Omega_\mu^{(1)}$ *and* $\Omega_\mu^{(2)}$*,*

$$AsyV\left(\hat{\phi}_{GMM}\left[\Omega_\mu^{(1)}\right]\right) - AsyV\left(\hat{\phi}_{GMM}\left[\Omega_\mu^{(2)}\right]\right) \geq \mathbf{0},$$

if $\Omega_\mu^{(1)} - \Omega_\mu^{(2)} \geq \mathbf{0}$*, and vice versa, where "* $\geq$ *" stands for a positive semi-definite matrix. In particular, in the special case where* $\Omega_\mu = \tau\Omega_\varepsilon$ *the precision of the GMM estimators deteriorates with* τ*. In contrast, the asymptotic variance of the FE-QMLE discussed in Section 4 does not depend on* Ω_μ*.*

Remark 6.3 *The pure unit root case presents an exception. Under (2.1), when* $\Phi = \mathbf{I}_m$*, no individual effects,* $\boldsymbol{\mu}_i$*, are present. However, if one were to assume that the initialization of series differ across the individuals, then from* $\mathbf{w}_{it} = \mathbf{w}_{i0} + \sum_{s=1}^{t} \boldsymbol{\varepsilon}_{st}$*, one can again deduce that the efficiency of (extended) GMM estimators in the pure unit root case will depend on the magnitude of* $V(\mathbf{w}_{i0})$*, the cross-section variation of the initial values relative to the time series dimension variations as given by* Ω_ε*.*

Remark 6.4 *The asymptotic efficiency arguments in Ahn and Schmidt (1995) carry over to the extended GMM estimator (6.30) set out above, provided that the fourth moments of* $\boldsymbol{\mu}_i$ *exist. However, as is well known from the instrumental variables literature, such asymptotic results need not carry over to small or even moderate sized samples, particularly when the number of moment conditions is large relative to the number of observations. The extended GMM estimators seem to be subject to such a shortcoming. This is because in the absence of prior information on the unit root properties of* $\mathbf{w}_{it}$ *all moment conditions could be informative,*[20] *and as a result the extended GMM estimators tend to use moment conditions well in excess of the number of unknown parameters.*[21] *Therefore, the extended GMM estimators are likely to be subject to important small sample bias. This issue is taken up in the next section.*

7 Finite Sample Evidence

In this section we provide evidence on the finite sample properties of the QML estimators, and standard and extended GMM estimators by means of Monte Carlo experiments.[22] While we consider a fairly broad range of model specifications, our Monte Carlo analysis is, given the scope of the paper, necessarily limited in nature. Nevertheless, we conjecture that our conclusions are likely to be of general validity.

[20] In the univariate context, Wansbeek and Bekker (1996) argue the importance of using all applicable moment conditions; Hahn (1999) argues that the information content of the homoskedasticity implied moment conditions is significantly augmented if initialization restrictions are imposed.

[21] The use of more moment conditions can lead to an increase in the bias of the GMM estimators in finite sample, for example, see Ziliak (1997). Also note that for both the standard and extended GMM estimators the number of orthogonality implied moment conditions increases quadratically with the time dimension of the panel.

[22] In the univariate context Monte Carlo studies of the finite sample properties of various GMM estimators include Kiviet (1995), Blundell and Bond (1998), and Alonso-Borrego and Arellano (1999).

7.1 Monte Carlo Design

We consider three types of designs for the matrix of slope coefficients, Φ. These designs distinguish between stationary , pure unit root, and cointegrated PVAR models.[23] In the case of stationary designs we consider three sub-cases with Φ having maximum eigenvalues equal; to 0.60, 0.80, and 0.95. For all designs we set $m = 2.$, and to make the Monte Carlo results from the various designs comparable, we specify (where appropriate) different error variance matrices for different designs so as to obtain similar population R^2 values for both equations of the PVAR model and across all designs:

Design 1a: Stationary PVAR with maximum eigenvalue of Φ *equal to* 0.6

$$\Phi = \begin{pmatrix} 0.4 & 0.2 \\ 0.2 & 0.4 \end{pmatrix}, \qquad \Omega_\varepsilon = \begin{pmatrix} 0.07 & 0.05 \\ 0.05 & 0.07 \end{pmatrix}.$$

The other eigenvalue of Φ is 0.2, and the population R^2 values are given by $R^2_{\triangle w_{lit}} = 0.2364$, $l = 1, 2$, $i = 1, 2, ..., N$, and $t = 2, 3, \ldots, T$, where

$$R^2_{\triangle w_{lit}} = 1 - \frac{[\Omega_\varepsilon]_{ll}}{\left[\sum_{j=0}^{\infty} \mathbf{C}_j \Omega_\varepsilon \mathbf{C}_j'\right]_{ll}}, \tag{7.1}$$

with $\mathbf{C}_0 = \mathbf{I}_m$, $\mathbf{C}_1 = -(\mathbf{I}_m - \Phi)$, and $\mathbf{C}_j = \mathbf{C}_{j-1}\Phi$, $j = 2, 3, \ldots$, and $[\mathbb{S}]_{ll}$ denoting the element in the l-th row and l-th column of the matrix $\mathbb{S}$.[24]

Design 1b: Stationary PVAR with maximum eigenvalue of Φ *equal to* 0.8

$$\Phi = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.6 \end{pmatrix}, \qquad \Omega_\varepsilon = \begin{pmatrix} 0.07 & -0.02 \\ -0.02 & 0.07 \end{pmatrix}.$$

The other eigenvalue of Φ is 0.4, and the population R^2 values are given by $R^2_{\triangle w_{lit}} = 0.2396$, $l = 1, 2$, $t = 2, 3, \ldots, T$, where $R^2_{\triangle w_{lit}}$ are computed as in (7.1).

Design 1c: Stationary PVAR with maximum eigenvalue of Φ *equal to* 0.95

$$\Phi = \begin{pmatrix} 0.7 & 0.25 \\ 0.25 & 0.7 \end{pmatrix}, \qquad \Omega_\varepsilon = \begin{pmatrix} 0.08 & -0.05 \\ -0.05 & 0.08 \end{pmatrix}.$$

The other eigenvalue of Φ is 0.45, and the population R^2 values are given by $R^2_{\triangle w_{lit}} = 0.2383$, $l = 1, 2$, $t = 2, 3, \ldots, T$, where $R^2_{\triangle w_{lit}}$ are computed as in (7.1).

Design 2: PVAR with unit roots (but non-cointegrated)

[23] An earlier version of this paper also included two additional designs which are dropped to save space.

[24] See Pesaran, Shin, and Smith (2000) for a discussion of the computation of R^2 values for (possibly cointegrated) VARs.

$$\Phi = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \Omega_{\varepsilon} = \begin{pmatrix} 0.08 & -0.05 \\ -0.05 & 0.08 \end{pmatrix}.$$

Design 3: Cointegrated PVAR

$$\Phi = \begin{pmatrix} 0.5 & 0.1 \\ -0.5 & 1.1 \end{pmatrix}, \qquad \Omega_{\varepsilon} = \begin{pmatrix} 0.05 & 0.03 \\ 0.03 & 0.05 \end{pmatrix}.$$

The eigenvalues of Φ in this case are given by 1 and 0.6, and the implied vectors/matrices $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and Π are given by

$$\boldsymbol{\alpha} = \begin{pmatrix} -0.5 \\ -0.5 \end{pmatrix}, \qquad \boldsymbol{\beta} = \begin{pmatrix} 1 \\ -0.2 \end{pmatrix}, \qquad \Pi = \begin{pmatrix} -0.5 & 0.1 \\ -0.5 & 0.1 \end{pmatrix}.$$

The population R^2 values are given by $R^2_{\triangle w_{lit}} = 0.2381$, $l = 1, 2$, $t = 2, 3, \ldots, T$, where $R^2_{\triangle w_{lit}}$ are computed as in (7.1).

The baseline settings across all five designs for the remaining model parameters are as follows: We take the $\boldsymbol{\varepsilon}_{it}$'s to be normally distributed, and generate the individual-specific effects as

$$\boldsymbol{\mu}_i = \sqrt{\tau}\left(\frac{\mathfrak{q}_i - 1}{\sqrt{2}}\right)\mathfrak{n}_i, \qquad \mathfrak{q}_i \overset{iid}{\sim} \chi^2(1), \qquad \text{and} \qquad \mathfrak{n}_i \overset{iid}{\sim} N(\mathbf{0}, \Omega_{\varepsilon}), \tag{7.2}$$

with $\mathfrak{q}_i$ and $\mathfrak{n}_i$ being distributed independently of $\boldsymbol{\varepsilon}_{it}$ for all i and t. In this way the individual effects will not be normally distributed. Clearly, the particular way that the individual effects are generated has no consequence for the FE-QMLE but could be important for the GMM type estimators. For τ we consider two value, $\tau = 1$ and 5. It should be recalled that τ measures the degree of cross-section to the time-series variations, which tends to be quite large for most economic data sets. The Monte Carlo studies of GMM estimators in the univariate context typically set $\tau = 1$, and to our knowledge we are the first to consider implications of changes in τ for the GMM estimators. The FE-QMLE does not depend on τ.

The $\mathbf{w}'_{it}s$ were generated using (2.1) and the initialization $(I3)$ as set out in Appendix A, with $M = 25$ and $\Omega_{\mathfrak{z}} = \Omega_{\varepsilon}$. Note that under this initialization we have that $R^2_{\triangle w_{li1}} = R^2_{\triangle w_{lit}}$, for $t = 2, 3, \ldots, T$, and $l = 1, 2$. We set $N = (50,\ 250)$, $T = (3,\ 10)$, and carry out $1,000$ replications for all baseline experiments, computing the FE-QML as well as the standard and extended GMM estimators.

In further experiments we consider a couple of deviations from the baseline scenario. As a partial analysis of the performance of the QML estimator under non-normal disturbances we also consider the cases of t- and chi square distributed disturbances: We generate t distributed disturbances $\boldsymbol{\varepsilon}_{it}$ as

$$\boldsymbol{\varepsilon}_{it} = \sqrt{\frac{3}{5}}\mathbf{P}'_{\varepsilon}\begin{pmatrix} \varsigma_{1it} \\ \varsigma_{2it} \end{pmatrix}, \tag{7.3}$$

where $\mathbf{P}_{\varepsilon}$ is the (upper triangular) Cholesky factor of Ω_{ε}, and ς_{lit}, $l = 1, 2$, are (for all l, i, and t) independently distributed standard t variates with five degrees of freedom. Chi square distributed disturbances are generated as

$$\boldsymbol{\varepsilon}_{it} = \sqrt{\frac{1}{2}}\mathbf{P}'_{\varepsilon}\begin{pmatrix} \chi^2_{1it} - 1 \\ \chi^2_{2it} - 1 \end{pmatrix}, \tag{7.4}$$

where χ^2_{1it} and χ^2_{2it} are independently distributed chi square variates with one degree of freedom. Also, as a partial analysis of the information content of the moment conditions available under the random but not the fixed effects specification, we compare the fixed and random effects QML estimators under cross-sectionally homoskedastic individual-specific effects, generating the latter as

$$\boldsymbol{\mu}_i = \sqrt{\tau}\mathfrak{n}_i, \qquad \mathfrak{n}_i \overset{iid}{\sim} N\left(\mathbf{0}, \Omega_{\varepsilon}\right), \tag{7.5}$$

with $\mathfrak{n}_i$ again being distributed independently of $\boldsymbol{\varepsilon}_{it}$ for all i and t, and $\tau = (1,\ 5)$.

In what follows we compare the various estimators in terms of their biases and root mean square errors (RMSEs). We also investigate the finite sample performance of a number of tests based on these estimators. For Designs 1 and 2 we compute the various estimators with Π unrestricted, and for Design 3 we compute the QMLE both with and without imposing rank restrictions on Π. In what follows we refer to the GMM estimator that uses only the orthogonality and initialization restrictions implied moment conditions as the "Extended GMM Estimator I", and the GMM estimator that uses only the orthogonality and homoskedasticity implied moment conditions as the "Extended GMM Estimator II". Finally, the GMM estimator that uses the orthogonality, initialization restrictions, and homoskedasticity implied moment conditions will be referred to as the "Extended GMM Estimator III". A summary of the computational details is provided in Appendix C.

7.2 The Results

The evidence on the finite sample properties of the various estimators in the case of normally distributed disturbances and when no rank restrictions are imposed on the matrix Π are summarized in Tables 1 and 2. Performance of each estimator is evaluated according to the familiar four criteria, namely bias, RMSE, size, and power. Table 1 reports the bias and RMSEs of the various estimators. To economize on space we focus on the results for the elements in the first column of Φ, namely ϕ_{11} and ϕ_{21}. The results for ϕ_{12} and ϕ_{22} are qualitatively similar and are available upon request. Size and power of the tests are reported in Tables 2a-2d. The nominal size is set to 5%. Once again to save space these tables only report the results for design 1a, and 3. For the GMM type estimators we report the results for $\tau = 1$ and 5.

The results in Tables 1 and 2 clearly show that the performance of the GMM type estimators tends to deteriorate with increases in τ, except in the pure unit root case. These simulation results are in line with our theoretical derivations discussed in Remark 6.2 and 6.3, and will not disappear if larger sample sizes are considered. The FE-QMLE is invariant to changes in τ and its finite sample performance is therefore unaffected by the choice of τ. Except for the pure unit root case where $\Phi = \mathbf{I}_2$, in which our data generating process no longer depends on individual effects. The dependence of the GMM type estimators on τ and/or on whether $\Phi = \mathbf{I}_2$, complicates the comparison of the various estimators. However, our Monte Carlo results suggest that even when τ is relatively small, $\tau = 1$, the FE-QMLE tends to perform significantly better than the GMM type estimators, possibly with the exception of the extended GMM estimators in the pure unit root case. While for a small number of scenarios with $\tau = 1$ one or more of the extended GMM estimators on a subset of our four evaluation criteria perform slightly better than the FE-QMLE, the differences in performance in those cases tend to be small, and are outweighed or at least offset by reverse ranking on one or more of the other evaluation criteria.

The results in Tables 1 and 2 also confirm the breakdown of the standard GMM estimator in the presence of unit roots, and document its deterioration as the eigenvalues of Φ approach unity, even for the relatively large sample size of $N = 250$, $T = 10$. None of the extended GMM estimators suffers from this problem. In fact, *ceteris paribus* the extended GMM estimators perform best in the pure unit root case. Of the various extended GMM estimators, the one using the homoskedasticity but not the initialization restrictions (Extended GMM Estimator II) is least sensitive to changes in τ. The tests based on Extended GMM Estimators I and III suffer from a considerable degree of over-rejection when $T = 10$, particularly as τ is increased. This finding should not be too surprising given that the standard orthogonality conditions as well as the initialization restrictions implied moment conditions involve interaction terms involving both levels and first differences, whereas the homoskedasticity implied moment conditions only involve levels terms. The finding that the Extended GMM Estimator II is relatively robust to changes in τ is unfortunately of limited use for empirical analysis, however, as the Extended GMM Estimator II performs worse than the other two extended GMM estimators when smaller values of τ are considered. Also, it is worth noting that while for $\tau = 1$ the extended GMM estimators outperform the standard GMM estimator, as τ increases this ranking is reversed in some instances. Finally, the results show that the performance of the GMM type estimators need not improve as T is increased. This is due to the rapid increase in the number of legitimate instruments with T, and stands in contrast to the FE-QMLE whose performance invariably improve with T. In summary, the Monte Carlo results suggest that the

GMM estimators are likely to perform well if prior knowledge were available regarding the location of the eigenvalues of Φ and/or if it were known that τ is small (so that the most suitable moment conditions could be picked). However, even in cases where the GMM estimators perform reasonably well in terms of bias and RMSE, they tend to be outperformed by the FE-QMLE in terms of size and power of the tests, except in the pure unit root case. The extended GMM estimators, however, provide useful consistent initial estimates for the QMLE iterations.

Overall, the results show that the FE-QMLE performs well, and is remarkably robust to the time-series properties of the underlying variables. In particular, the performance of the FE-QMLE is generally unaffected by whether the maximal eigenvalue of Φ is moderately sized, close, or equal to unity.

Table 3 presents evidence on the finite sample properties of the FE-QMLE in the case of a cointegrated PVAR model (Design 3). We did not compute any of the GMM estimators for this design: The main virtue of the GMM estimators, their computational simplicity, is lost in the presence of rank restrictions on Π, as in such cases the GMM estimators would have to be computed using iterative optimization techniques. The results in Table 3 show that the FE-QMLE continues to perform reasonably well under rank restrictions on Π. Nevertheless, it should be noted that in the smallest sample ($N = 50$, $T = 3$) the RMSEs for the FE-QMLE tend to be larger than for the other designs, and the test of cointegration rank is undersized. For larger sample sizes featuring a larger N and/or T, bias and RMSE diminish rather rapidly, however, and size and power properties of the tests improve considerably.

Table 4 reports on the performance of the FE-QMLE under two types of departures from normally distributed disturbances, namely when the disturbances are t (5) or $\chi^2(1)$ distributed.[25] For the case of t (5) distributed disturbances the FE-QMLE on all four evaluation criteria performs just about the same as under normally distributed disturbances. For $\chi^2(1)$ distributed disturbances the same tends to be true, specifically for bias and RMSE, except that there is now significant evidence of over-rejection when $T = 3$.[26] The size does quickly tend towards its nominal value as T is increased, though. Consider now the size of the tests when normality is (erroneously) assumed in the computation of the standard errors. With one exception the Monte Carlo results do not favor the use of the robust estimator of the variance-covariance matrix. The exception is that under $\chi^2(1)$ distributed disturbances, when T is small ($T = 3$) and N relatively large ($N = 250$), the use of the sandwich formula helps in correcting the over-rejection problem.

All of the above arguments carry over to the random effects setting. To economize on space,

[25] Since the two types departures from normality considered here cover both the possibility of fat tails and the possibility of an asymmetric/skewed shock distribution, it is very likely that the results reported here could be of greater generality.

[26] We have also obtained broadly similar results for the various GMM estimators. These are available from the authors upon request.

however, Table 5, which provides our findings under the random effects specification of the individual effects, focuses on the comparison of the random effects QML (RE-QML) and the FE-QML estimators. The table reveals that under the random effects model the FE-QMLE performs on par with the RE-QMLE even for the smallest sample size ($N = 50$, $T = 3$). The differences between the two estimators are very small across all the four evaluation criteria, often even favoring the FE-QMLE.[27] Thus, the argument often advanced concerning the efficiency loss involved in the first-differencing operation that underlies the FE-QMLE as compared to the RE-QMLE does not appear to be important for the estimation of PVARs using finite samples. The RE-QMLE, however, remains the estimator of choice if the primary purpose of the analysis is the identification and estimation of the effects of time-invariant variables in short panels. In that case great caution must be exercised since the random effects model imposes strong assumptions on the distribution of the individual-specific effects. For the identification and estimation of the effects of time-varying variables our findings favor the use of FE-QMLE on grounds of its robustness to any form of specification of the distribution generating the individual-specific effects.

8 Conclusion

In this paper, we have extended the analysis of linear dynamic panel data models with short time dimension in a number of respects. We have generalized the extended GMM estimators, hitherto studied in the literature in a single equation context, to a multivariate PVAR setting. We have derived random and fixed effects QML estimators, and have shown that the QML estimators are consistent and asymptotically normally distributed when the cross-sectional dimension of the panel approaches infinity, irrespective of whether the underlying time series are (trend) stationary, pure I(1), or I(1) and cointegrated. Furthermore, we have proposed new QML based procedures for conducting tests for unit roots and cointegration rank in panels with short time dimension, and shown that the limiting distributions of the relevant test statistics follow standard chi square and normal distributions.

Asymptotic considerations would suggest that the extended GMM estimator making use of the full set of moment conditions, when applicable, would in general be superior to the other estimators. However, the validity of this argument requires that certain assumptions on the unobserved individual effects are satisfied. From the perspective of empirical analysis, these assumptions could be restrictive in the case of the extended GMM estimator. In addition, the Monte Carlo evidence presented in Section 7 suggests that such asymptotic efficiency considerations do not generally carry over to finite samples. Our results favor the fixed effects QML estimator over the various GMM estimators, even under important departures from normally distributed disturbances. The

[27]While for space reasons Table 5 reports the RE-QMLE results only for $\tau = 5$, we have found similar results as reported in that table for $\tau = 1$ also.

finite sample performance of the various GMM estimators depends critically on a ratio reflecting unobserved cross-section variation in the data relative to unobserved time-series variation, but even if this ratio is relatively small the GMM estimators are outperformed in finite sample by the QML estimators. A theoretical rationale is also provided for this result whereby it is shown that asymptotic variance of the Standard GMM estimator is an increasing function of the variance of the individual effects, while the distribution of the FE-QMLE is invariant to the size of this variance.

The use of likelihood based procedures for estimation and inference in VAR models is standard in the time-series literature. This paper has provided theoretical as well as operational arguments for the application of likelihood based methods to Panel VAR models. The ultimate test of our approach lies in the application of the proposed techniques to substantive economic problems. This is the next stage of our research, and hopefully that of others. The likelihood approach can also be used to address other theoretical issues of interest, such as model selection and conditional estimation and inference in PVARs. It would also be of interest to compare the finite sample performance of the QML estimators with other types of GMM estimators, such as continuously updated GMM and iterated GMM estimators discussed in the literature. In this way a fairer comparison with GMM type estimators could be provided.

Appendix A: Technical Issues Regarding the Initializations

The moment restrictions on $\mathbf{w}_{i0}$ and $\Delta\mathbf{w}_{i1}$ assumed in the paper could be motivated directly without necessarily relating them to the assumed data generating process for $\mathbf{w}_{it}$. However, further insight into the moment homogeneity restrictions can be obtained by solving for $\mathbf{w}_{i0}$ and $\Delta\mathbf{w}_{i1}$ in terms of individual-specific initializations of the data generating process. We first note from (2.1) that

$$\Delta\mathbf{w}_{i1} = -\left(\mathbf{I}_m - \Phi\right)\left(\mathbf{w}_{i0} - \boldsymbol{\mu}_i\right) + \boldsymbol{\varepsilon}_{i1}. \tag{A.1}$$

Hence the properties of both $\mathbf{w}_{i0}$ and $\Delta\mathbf{w}_{i1}$ can be examined by considering the deviations $\boldsymbol{\xi}_{it}$,

$$\boldsymbol{\xi}_{it} = \mathbf{w}_{it} - \boldsymbol{\mu}_i. \tag{A.2}$$

Suppose that the $\boldsymbol{\xi}_{it}$ process for cross-sectional unit i started at time $t = -M_i$, $M_i \geq 0$, with given $\boldsymbol{\xi}_{i,-M_i}$. Then from (2.1) and (A.2) we obtain

$$\boldsymbol{\xi}_{i0} = \Phi^{M_i}\boldsymbol{\xi}_{i,-M_i} + \sum_{j=0}^{M_i-1}\Phi^j\boldsymbol{\varepsilon}_{i,-j}. \tag{A.3}$$

To ensure that $\boldsymbol{\xi}_{i0}$ exists irrespective of the unit root properties of (2.1) it will be useful to distinguish between three main cases, where $\{\boldsymbol{\xi}_{it}\}$ is stationary, pure I(1), or I(1) and cointegrated.

In the case where all eigenvalues of Φ fall inside the unit circle, it then follows from (A.3) and assumption (**G2**) that $\boldsymbol{\xi}_{i0}$ will exist for all M_i, including the case where the $\boldsymbol{\xi}_{it}$ process has been in operation for a long period of time, namely when $M_i \to \infty$. In this latter case (A.3) becomes

$$\boldsymbol{\xi}_{i0} = \sum_{j=0}^{\infty}\Phi^j\boldsymbol{\varepsilon}_{i,-j}, \tag{A.4}$$

which is independent of the initialization deviations, $\boldsymbol{\xi}_{i,-M_i}$. However, when M_i is finite for all i and all eigenvalues of Φ fall inside the unit circle, then the distribution of $\boldsymbol{\xi}_{i0}$ will depend on $\boldsymbol{\xi}_{i,-M_i}$ and homogeneity assumptions regarding the cross-sectional distribution of $\boldsymbol{\xi}_{i,-M_i}$ will be required both under the random and fixed effects specifications. See assumption (**R1**) in Section 3, and assumption (**F1**) in Section 4. Intermediate cases where $M_i \to \infty$ only for some i could also be entertained.

At the other extreme where all eigenvalues of Φ are equal to unity, it follows from (A.3) that

$$\boldsymbol{\xi}_{i0} = \boldsymbol{\xi}_{i,-M_i} + \sum_{j=0}^{M_i-1}\boldsymbol{\varepsilon}_{i,-j}, \tag{A.5}$$

and to ensure that $\boldsymbol{\xi}_{i0}$ exists, the $\boldsymbol{\xi}_{it}$ process *must* have started in a finite period in the past for all i. While homogeneity assumptions regarding the cross-sectional distribution of $\boldsymbol{\xi}_{i,-M_i}$ will then be required for the random effects specification which is based on $\mathbf{w}_{i0}$, under the fixed effects specification, which is based on $\Delta \mathbf{w}_{i1}$, it is clear from (A.1) that no restrictions will be required regarding $\boldsymbol{\xi}_{i,-M_i}$.

It remains to consider the case where some eigenvalues of Φ are equal to unity, and the remaining ones fall inside the unit circle. For this case it is helpful to note that in terms of the deviations $\boldsymbol{\xi}_{it}$ the model (2.1) can also be written as

$$\Delta \boldsymbol{\xi}_{it} = \Pi \boldsymbol{\xi}_{i,t-1} + \boldsymbol{\varepsilon}_{it}, \qquad t = -M_i + 1, -M_i + 2, \ldots, T, \tag{A.6}$$

where

$$\Pi = -\left(\mathbf{I}_m - \Phi\right). \tag{A.7}$$

As is well known from the time-series literature on cointegrated systems the key to the analysis of these systems lies in the rank of the long-run multiplier matrix, Π, which we denote by r. When $r = 0$, $\boldsymbol{\xi}_{it}$ is a pure random walk process. When $r = m$, Π is of full rank and $\boldsymbol{\xi}_{it}$ is a stationary process. In the intermediate case where $r = 1, 2, \ldots, m-1$ there exist $m \times r$ matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ such that

$$\Pi = \boldsymbol{\alpha}\boldsymbol{\beta}^{'}, \tag{A.8}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ have full column rank.[28] The above discussion is formalized in the following assumptions:

Assumption (I1) *The eigenvalues of Φ are either equal to unity or fall inside the unit circle.*

Assumption (I2) *Assume rank* $(\Pi) = r$, *and rank* $(\alpha'_{\perp}\beta_{\perp}) = m - r$ for some $r = 1, 2, ..., m-1$, *where $\boldsymbol{\alpha}_{\perp}$ and $\boldsymbol{\beta}_{\perp}$ are $m \times (m-r)$ matrices of full column rank such that $\boldsymbol{\alpha}'\boldsymbol{\alpha}_{\perp} = \mathbf{0}$ and $\boldsymbol{\beta}'\boldsymbol{\beta}_{\perp} = \mathbf{0}$.*

Under assumptions (**I1**) and (**I2**) the elements of $\boldsymbol{\xi}_{it}$ are either I(0) or I(1). To separate the stochastic trend components in $\boldsymbol{\xi}_{it}$ from the cointegrating relations, we follow Johansen (1995, Ch. 4) and define

$$\mathbf{C} = \boldsymbol{\beta}_{\perp}\left(\boldsymbol{\alpha}'_{\perp}\boldsymbol{\beta}_{\perp}\right)^{-1}\boldsymbol{\alpha}'_{\perp}. \tag{A.9}$$

If $rank\,(\Pi) = r$, $r = 1, 2, \ldots, m-1$, then the matrix $\mathbf{C}$ has rank $m - r$, and there are $m - r$ common stochastic trend components in $\boldsymbol{\xi}_{it}$, given by $\boldsymbol{\beta}'_{\perp}\boldsymbol{\xi}_{it} \sim I(1)$, $t = -M_i, -M_i + 1, \ldots, T$. To ensure that $\boldsymbol{\xi}_{i0}$ exists, the $m - r$ common stochastic trend components, $\boldsymbol{\beta}'_{\perp}\boldsymbol{\xi}_{it} \backsim I(1)$, must have started in a finite period in the past. For consistency with (A.4), the r cointegrating relations, $\boldsymbol{\beta}'\boldsymbol{\xi}_{i0}$, must be stationary. The following assumption ensures this, irrespective of the number of common stochastic trend components in $\boldsymbol{\xi}_{it}$.

[28] See, for example, Johansen (1995, Ch. 4).

Assumption (I3) *The initial deviations* $\boldsymbol{\xi}_{i,-M_i}$ *for* $i = 1, 2, \ldots, N$ *are given by*

$$\boldsymbol{\xi}_{i,-M_i} = \sum_{j=0}^{\infty} \left(\Phi^j - \mathbf{C}\right) \boldsymbol{\varepsilon}_{i,-M_i-j} + \mathbf{C}\mathfrak{z}_i, \tag{A.10}$$

where $\mathfrak{z}_i$ *is an* $m \times 1$ *vector of individual-specific initialization effects.*[29]

Substituting (A.10) back into (A.3), and noting from the definition of $\mathbf{C}$ that $\Pi\mathbf{C} = \mathbf{0}$, and thus $\mathbf{C} = \Phi\mathbf{C}$, we have that

$$\boldsymbol{\xi}_{i0} = \sum_{j=0}^{\infty} \left(\Phi^j - \mathbf{C}\right) \boldsymbol{\varepsilon}_{i,-j} + \mathbf{C}\left(\sum_{j=0}^{M_i-1} \boldsymbol{\varepsilon}_{i,-j}\right) + \mathbf{C}\mathfrak{z}_i, \tag{A.11}$$

and therefore also

$$\boldsymbol{\beta}'\boldsymbol{\xi}_{i0} = \boldsymbol{\beta}' \sum_{j=0}^{\infty} \Phi^j \boldsymbol{\varepsilon}_{i,-j}. \tag{A.12}$$

It is thus seen that assumption (**I3**) indeed ensures that the cointegrating relations $\boldsymbol{\beta}'\boldsymbol{\xi}_{i0}$ are stationary, irrespective of the number of common trend components in the $\boldsymbol{\xi}_{it}$ process.[30] Furthermore, in the case where all eigenvalues of Φ fall inside the unit circle, $\mathbf{C} = \mathbf{0}$, and (A.11) reduces to (A.4). In the case where all eigenvalues of Φ are equal to unity, $\mathbf{C} = \Phi = \mathbf{I}_m$, and (A.11) reduces to (A.5), with $\boldsymbol{\xi}_{i,-M_i} = \mathfrak{z}_i$. Finally, since $\left\{\Phi^j - \mathbf{C}\right\}_{j=0}^{\infty}$ is absolutely summable irrespective of the number of eigenvalues of Φ that are equal to unity,[31] it follows from (A.11) that $\boldsymbol{\xi}_{i0}$ exists. We shall discuss the implications of the initialization (**I3**) for random and fixed effects QML and MD estimation in the next two sections.

[29] *According to assumption (***I3***), there are* $m - r$ *linearly independent initialization effects in* $\boldsymbol{\xi}_{i,-M_i}$*, corresponding to the number of common stochastic trend components in* $\boldsymbol{\xi}_{it}$*. Also, since assumption (***I3***) effectively characterizes the* r *stationary components of* $\boldsymbol{\xi}_{it}$ *as having started a long time ago, initialization effects in the latter components will, irrespective of the effects' properties, have vanished at time* $t = -M_i$*,* M_i *finite, and thus do not feature in (A.10).*

[30] The following is in fact also true: The same processes generating the stationary components of $\{\boldsymbol{\xi}_{it}\}_{t=1}^{T}$ also generate those of $\boldsymbol{\xi}_{i0}$, and the same processes generating the common stochastic trend components of $\{\boldsymbol{\xi}_{it}\}_{t=1}^{T}$ also generate those of $\boldsymbol{\xi}_{i0}$.

[31] For a proof of the absolute summability property of $\left\{\Phi^j - \mathbf{C}\right\}_{j=0}^{\infty}$ in the context of a p-th order VAR model, see Johansen (1995, Ch. 4).

Appendix B: The Asymptotic Variance Matrix of the Standard GMM Estimator

In this appendix we derive the asymptotic variance of $\hat{\boldsymbol{\phi}}_{GMM}$, the standard GMM estimator defined by (6.7), for $T = 3$, and as $N \to \infty$, and show that it is an increasing function of Ω_μ, the variance matrix of the individual effects, $\boldsymbol{\mu}_i$, in the sense that if $\Omega_\mu^{(1)} - \Omega_\mu^{(2)}$ is a positive semi-definite (p.s.d.) matrix, so will be

$$AsyV\left(\hat{\boldsymbol{\phi}}_{GMM}\left[\Omega_\mu^{(1)}\right]\right) - AsyV\left(\hat{\boldsymbol{\phi}}_{GMM}\left[\Omega_\mu^{(2)}\right]\right),$$

where $\Omega_\mu^{(i)}$, $i = 1, 2$ are two different variance matrices for the individual effects.[32]

The asymptotic variance of $\hat{\boldsymbol{\phi}}_{GMM}$ is given by (also see (6.12))

$$AsyV\left(\hat{\boldsymbol{\phi}}_{GMM}\right) = \left(\mathbf{S}'\mathbf{D}^{-1}\mathbf{S}\right)^{-1} \otimes \Omega_\varepsilon, \tag{B.1}$$

where $\Omega_\varepsilon = E\left(\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{it}'\right)$,

$$\mathbf{D} = p\lim_{N\to\infty}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{Q}_i'\mathbb{V}\mathbf{Q}_i\right), \tag{B.2}$$

$$\mathbf{S} = p\lim_{N\to\infty}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{Q}_i'\Delta\mathbf{W}_{i,-1}\right), \tag{B.3}$$

$\mathbf{Q}_i$, $\Delta\mathbf{W}_{i,-1}$,and $\mathbb{V}$ are defined by (6.4), (6.6), and (6.11), respectively.

To simplify the derivations we suppose that assumptions **G1** and **G2** hold, all eigenvalues of Φ lie inside the unit circle and the $\mathbf{w}_{it}$ process has started in the infinite past. Under these assumptions[33]

$$\mathbf{w}_{it} = \boldsymbol{\mu}_i + \sum_{j=0}^{\infty}\Phi^j\boldsymbol{\varepsilon}_{i,t-j} = \boldsymbol{\mu}_i + \boldsymbol{\xi}_{it}, \text{ for all } t, \tag{B.4}$$

and it is easily seen that

$$E\left(\mathbf{w}_{it}\mathbf{w}_{i,t+s}'\right) = \Omega_\mu + \Omega_{\boldsymbol{\xi}}\Phi'^s,\ E\left(\mathbf{w}_{it}\Delta\mathbf{w}_{i,t+s}'\right) = -\Omega_{\boldsymbol{\xi}}\left(\mathbf{I}_m - \Phi'\right)\Phi'^{s-1}, \tag{B.5}$$

where $\Omega_{\boldsymbol{\xi}} = \sum_{j=0}^{\infty}\Phi^j\Omega_\varepsilon\Phi'^j$. Also under Assumption **G2**

$$\mathbf{S} = E\left(\mathbf{Q}_i'\Delta\mathbf{W}_{i,-1}\right), \text{ and } \mathbf{D} = E\left(\mathbf{Q}_i'\mathbb{V}\mathbf{Q}_i\right),$$

[32] The more complicated case of the dependence of extended GMM on Ω_μ can be derived similarly noting that $\mathbf{w}_{it} - \Phi\mathbf{w}_{i,t-1} = (\mathbf{I} - \Phi)\boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_{it}$ and making use of the relations (6.20) and (6.22).

[33] The results will be qualitatively unaffected if we consider other initializations of the $\mathbf{w}_{it}$ process discussed in Appendix A.

and using (B.5) we have (for $T = 3$)

$$\mathbf{S} = \begin{pmatrix} -\Omega_{\xi}\left(\mathbf{I}_m - \Phi'\right) \\ -\Omega_{\xi}\left(\mathbf{I}_m - \Phi'\right)\Phi' \\ -\Omega_{\xi}\left(\mathbf{I}_m - \Phi'\right) \end{pmatrix}, \ \mathbf{D} = \Omega_{\mu} \otimes \mathbb{V} + \mathbf{H}, \tag{B.6}$$

where

$$\mathbf{H} = \begin{pmatrix} 2\Omega_{\xi} & -\Omega_{\xi} & -\Omega_{\xi}\Phi' \\ -\Omega_{\xi} & 2\Omega_{\xi} & 2\Omega_{\xi}\Phi' \\ -\Phi\Omega_{\xi} & 2\Phi\Omega_{\xi} & 2\Omega_{\xi} \end{pmatrix}.$$

It is clear that $AsyV\left(\hat{\boldsymbol{\phi}}_{GMM}\right)$ will depend on Ω_{μ} only through matrix $\mathbf{D}$, and for any two variance matrices, $\Omega_{\mu}^{(i)}$, $i = 1, 2$,

$$AsyV\left(\hat{\boldsymbol{\phi}}_{GMM}\left[\Omega_{\mu}^{(1)}\right]\right) \geq AsyV\left(\hat{\boldsymbol{\phi}}_{GMM}\left[\Omega_{\mu}^{(1)}\right]\right), \tag{B.7}$$

if

$$\left(\mathbf{S}'\mathbf{D}_1^{-1}\mathbf{S}\right)^{-1} \geq \left(\mathbf{S}'\mathbf{D}_2^{-1}\mathbf{S}\right)^{-1}, \tag{B.8}$$

where $\mathbf{D}_i = \Omega_{\mu}^{(i)} \otimes \mathbb{V} + \mathbf{H}$, and $\mathbf{A} \geq \mathbf{B}$ denotes that $\mathbf{A} - \mathbf{B}$ is a p.s.d. matrix. This condition implies

$$\mathbf{S}'\mathbf{D}_1^{-1}\mathbf{S} \leq \mathbf{S}'\mathbf{D}_2^{-1}\mathbf{S},$$

and since $\mathbf{S}$ does not depend on $\Omega_{\mu}^{(i)}$, using (B.6) the condition (B.7) will be satisfied if

$$\Omega_{\mu}^{(1)} \otimes \mathbb{V} + \mathbf{H} \geq \Omega_{\mu}^{(2)} \otimes \mathbb{V} + \mathbf{H},$$

or if $\Omega_{\mu}^{(1)} \geq \Omega_{\mu}^{(2)}$. The above sequence can be reversed to show that if $\Omega_{\mu}^{(1)} \geq \Omega_{\mu}^{(2)}$ then (B.7) will follow.

In the simple case where $m = 1$, we have

$$\mathbf{D} = \boldsymbol{\sigma}_{\mu}^2 \begin{pmatrix} 1 & -1 & -1 \\ -1 & 2 & 2 \\ -1 & 2 & 2 \end{pmatrix} + \frac{\sigma_{\varepsilon}^2}{1-\phi^2} \begin{pmatrix} 1 & -1 & -\phi \\ -1 & 2 & 2\phi \\ -\phi & 2\phi & 2 \end{pmatrix}, \mathbf{S} = \frac{-\sigma_{\varepsilon}^2}{1+\phi} \begin{pmatrix} 1 \\ \phi \\ 1 \end{pmatrix},$$

and after some algebra it follows that

$$AsyV\left(\hat{\phi}_{GMM}\right) = 2\left(\frac{1+\phi}{1-\phi}\right)\left(\frac{\left[2\tau(1-\phi)+1\right]\left[\tau(1-\phi^2)+1\right]}{\tau^2(1-\phi^2)(1-\phi)^2 + 2\tau(1-\phi)\left[\phi^2+4\phi+5\right] + \left[\phi^2+4\phi+5\right]}\right),$$

where $\tau = \sigma_{\mu}^2/\sigma_{\varepsilon}^2$. It is interesting to note that $AsyV\left(\hat{\phi}_{GMM}\right)$ depends on the ratio τ and not the error variances, $\boldsymbol{\sigma}_{\mu}^2$ and σ_{ε}^2 separately. It is also easily established that $AsyV\left(\hat{\phi}_{GMM}\right)$ is an increasing function of τ, for all values of $|\phi| < 1$ and $\tau > 0$.

Appendix C: Computation of the Various Estimators for the Monte Carlo Analysis

In this appendix we provide details on how the various GMM andQML estimators were computed in our Monte Carlo experiments.

For the computation of the standard GMM estimator we use (6.7), with the initial estimate of Φ computed using (6.7), but with $\mathbf{D}_{\hat{\mathbf{e}}}$ replaced by $\Lambda_{\mathbf{Q}} \otimes \mathbf{I}_m$, where $\Lambda_{\mathbf{Q}}$ is given by (6.10). The variance-covariance matrix of the standard GMM estimator is computed using (6.12). The Extended GMM Estimators I, II, and III are obtained using (6.30), but with $\mathbf{P}_i'$ defined as $\mathbf{P}_i' = \left(\begin{array}{cc} \mathbf{P}_{1i}', & \mathbf{P}_{2i}' \end{array} \right)'$, $\mathbf{P}_i' = \left(\begin{array}{cc} \mathbf{P}_{1i}', & \mathbf{P}_{3i}' \end{array} \right)'$, and $\mathbf{P}_i' = \left(\begin{array}{ccc} \mathbf{P}_{1i}', & \mathbf{P}_{2i}', & \mathbf{P}_{3i}' \end{array} \right)'$, respectively. See (6.24)-(6.27). We compute initial estimates of the extended GMM estimators using (with the appropriate definitions of $\mathbf{P}_i'$) (6.30), but with $\mathbf{D}_{\hat{\mathbf{u}}}$ replaced by $\Lambda_{\mathbf{P}} \otimes \mathbf{I}_m$, where $\Lambda_{\mathbf{P}}$ is defined by (6.33). The variance-covariance matrices of these estimators are computed using (6.34), again with the appropriate definition of $\mathbf{P}_i'$.

The FE-QMLE for Designs 1-3 are computed using (4.14) with Ψ given by

$$\Psi = \Omega_\varepsilon + \sum_{j=0}^{\infty} \mathbf{G}_j \Omega_\varepsilon \mathbf{G}_j', \qquad \text{with} \qquad \mathbf{G}_j = (\mathbf{I}_m - \Phi)\, \Phi^j. \tag{C.1}$$

When all eigenvalues of Φ fall inside the unit circle, observing that

$$\begin{aligned} \Phi \Psi \Phi' &= \Phi \Omega_\varepsilon \Phi' + \sum_{j=0}^{\infty} \Phi^{j+1} (\mathbf{I}_m - \Phi)\, \Omega_\varepsilon (\mathbf{I}_m - \Phi)' \, \Phi'^{\,j+1} \\ &= \Phi \Omega_\varepsilon \Phi' + \Psi - \Omega_\varepsilon - (\mathbf{I}_m - \Phi)\, \Omega_\varepsilon (\mathbf{I}_m - \Phi)', \end{aligned} \tag{C.2}$$

Ψ can then be computed directly by

$$\boldsymbol{\psi} = \mathbb{D}_m^+ \left(\mathbf{I}_{m^2} - \Phi \otimes \Phi \right)^{-1} \mathbb{D}_m vech \left(2\Omega_\varepsilon - \Phi \Omega_\varepsilon - \Omega_\varepsilon \Phi' \right), \tag{C.3}$$

where $\mathbb{D}_m^+$ denotes the generalized inverse of the $m^2 \times m(m+1)/2$ dimensional duplication matrix $\mathbb{D}_m$ defined such that $vec(\mathbb{M}) = \mathbb{D}_m vech(\mathbb{M})$ for any symmetric $m \times m$ dimensional matrix $\mathbb{M}$. When some eigenvalues of Φ are equal to unity, then $(\mathbf{I}_{m^2} - \Phi \otimes \Phi)$ is singular; Ψ may then be computed using recursions that invoke an appropriate stopping rule to truncate $\sum_{j=0}^{\infty} \mathbf{G}_j \Omega_\varepsilon \mathbf{G}_j'$. The variance-covariance matrix of the FE-QMLE is based on the fixed effects counterpart of (3.19). As initial estimates of Φ we use the Extended GMM I estimates, which we denote by $\widehat{\Phi}^{(0)}$. Furthermore, we obtain initial estimates of Ω_ε as

$$\widehat{\Omega}_\varepsilon^{(0)} = \frac{1}{2N(T-1)} \sum_{t=2}^{T} \sum_{i=1}^{N} \left[\Delta \mathbf{w}_{it} - \widehat{\Phi}^{(0)} \Delta \mathbf{w}_{i,t-1} \right] \left[\Delta \mathbf{w}_{it} - \widehat{\Phi}^{(0)} \Delta \mathbf{w}_{i,t-1} \right]', \tag{C.4}$$

and of Ψ from (C.1), replacing Φ by $\widehat{\Phi}^{(0)}$, and Ω_ε by $\widehat{\Omega}_\varepsilon^{(0)}$.

For Design 3 we compute the FE-QMLE under different rank restrictions on Π. This is achieved by (i) leaving the rank of Π unrestricted; (ii) setting $\Phi = \boldsymbol{\alpha\beta}' + \mathbf{I}_m$, where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $m \times r$ vectors, $r = 1, 2, \ldots, m-1$; and, finally (iii) setting $\Phi = \mathbf{I}_m$. Under (ii), to obtain initial estimates of $\boldsymbol{\beta}$, which we denote by $\widehat{\boldsymbol{\beta}}^{(0)}$, we run cross-section regressions in the elements of $\mathbf{w}_{it}$ under the normalization restriction (5.6). Also from (A.1) and noting that

$$\Delta^2 \mathbf{w}_{it} = \boldsymbol{\alpha\beta}' \Delta \mathbf{w}_{i,t-1} + \Delta \boldsymbol{\varepsilon}_{it}, \qquad t = 2, 3, \ldots, T, \tag{C.5}$$

we then obtain the following initial estimate of $\boldsymbol{\alpha}$:

$$vec\left[\widehat{\boldsymbol{\alpha}}^{(0)}\right] = \left\{ \sum_{i=1}^{N} \mathbf{H}_i' \left[\widehat{\Sigma}_{\triangle \boldsymbol{\eta}}^{(0)}\right]^{-1} \mathbf{H}_i \right\}^{-1} \left\{ \sum_{i=1}^{N} \mathbf{H}_i' \left[\widehat{\Sigma}_{\triangle \boldsymbol{\eta}}^{(0)}\right]^{-1} \mathbf{g}_i \right\}, \tag{C.6}$$

where

$$\mathbf{H}_i = \mathbf{L}_i' \widehat{\boldsymbol{\beta}}^{(0)} \otimes \mathbf{I}_m, \qquad \mathbf{L}_i = \begin{pmatrix} \mathbf{0}, & \Delta \mathbf{w}_{i1}, & \Delta \mathbf{w}_{i2}, & \ldots, & \Delta \mathbf{w}_{i,T-1} \end{pmatrix},$$

$$\mathbf{g}_i = vec \begin{pmatrix} \Delta \mathbf{w}_{i1}, & \Delta^2 \mathbf{w}_{i2}, & \Delta^2 \mathbf{w}_{i3}, & \ldots, & \Delta^2 \mathbf{w}_{iT} \end{pmatrix},$$

and $\widehat{\Sigma}_{\triangle \boldsymbol{\eta}}^{(0)}$ is obtained from (4.10), with Ψ replaced by $\widehat{\Psi}^{(0)}$, and with $\Omega_{\boldsymbol{\varepsilon}}$ replaced by $\widehat{\Omega}_{\boldsymbol{\varepsilon}}^{(0)}$, where we compute $\widehat{\Psi}^{(0)}$ and $\widehat{\Omega}_{\boldsymbol{\varepsilon}}^{(0)}$ as described above.

For the RE-QMLE we concentrate on Designs 1 and 2 and compute it. We compute the estimators using (3.15), with the variance-covariance given by (3.19). We restrict Ω_0 and $\Omega_{0\mathbf{a}}$ as in (3.6), use the Extended GMM I estimates as the initial estimates of Φ, and compute initial estimates for $\Omega_{\boldsymbol{\varepsilon}}$ using (C.4). For the initial estimate of Ω_0, we use

$$\widehat{\Omega}_0^{(0)} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{w}_{i0} \mathbf{w}_{i0}'. \tag{C.7}$$

Observing that under (3.6)

$$\boldsymbol{\sigma}_{\mathbf{a}} = \mathbb{D}_m^+ \left[(\mathbf{I}_m - \Phi) \otimes (\mathbf{I}_m - \Phi) \right] \mathbb{D}_m \left[\boldsymbol{\sigma}_0 + \left(\mathbf{I}_{m^2} - \Phi \otimes \Phi \right)^{-1} \boldsymbol{\sigma}_{\boldsymbol{\varepsilon}} \right], \tag{C.8}$$

we obtain initial estimates of $\Omega_{\mathbf{a}}$, $\widehat{\Omega}_{\mathbf{a}}^{(0)}$, from (C.8), replacing Φ by $\widehat{\Phi}^{(0)}$, and $\Omega_{\boldsymbol{\varepsilon}}$ by $\widehat{\Omega}_{\boldsymbol{\varepsilon}}^{(0)}$. Finally, for the initial estimate of $\Omega_{0\mathbf{a}}$ we use

$$\widehat{\Omega}_{0\mathbf{a}}^{(0)} = \widehat{\Omega}_{\mathbf{a}}^{(0)} \left(\mathbf{I}_m - \widehat{\Phi}^{(0)} \right)^{\prime\, -1}.$$

The numerical optimization routine we employ for computation of the QML estimators is based on a trust region method type algorithm, and is described in some detail in a note available upon request.

References

Ahn, S.C., and P. Schmidt (1995): Efficient Estimation of Models for Dynamic Panel Data, *Journal of Econometrics*, 68, 5-27.

Ahn, S.C., and P. Schmidt (1997): Efficient Estimation of Dynamic Panel Data Models: Alternative Assumptions and Simplified Estimation, *Journal of Econometrics*, 76, 309-321.

Alonso-Borrego, C., and M. Arellano (1999): Symmetrically Normalized Instrumental Variable Estimation Using Panel Data, *Journal of Business & Economic Statistics*, 17, 36-49.

Anderson, T.W., and C. Hsiao (1981): Estimation of Dynamic Models with Error Components, *Journal of the American Statistical Association*, 76, 598-606.

Anderson, T.W., and C. Hsiao (1982): Formulation and Estimation of Dynamic Models Using Panel Data, *Journal of Econometrics*, 18, 47-82.

Arellano, M., and S.R. Bond (1991): Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies*, 58, 277-297.

Arellano, M., and O. Bover (1995): Another Look at the Instrumental Variable Estimation of Error-Components Models, *Journal of Econometrics*, 68, 29-51.

Arellano, M., and B. Honoré (2001): Panel Data Models: Some Recent Developments, in: J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Vol. 5, Amsterdam: North-Holland (Forthcoming).

Bai, J. and S. Ng, (2002): A Panic Attack on Unit Roots and Cointegration, *Department of Economics, Boston College*, Unpublished Manuscript.

Baltagi, B.H. (2001): *Econometric Analysis of Panel Data*, 2^{nd} Edition, New York: John Wiley.

Baltagi, B.H., and C. Kao (2000): Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey, *Advances in Econometrics*, 15, 7-51.

Bhargava, A., and J.D. Sargan (1983): Estimating Dynamic Random Effects Models from Panel Data Covering Short Time Periods, *Econometrica*, 51, 1635-1659.

Binder, M., and M.H. Pesaran (2000): Solution of Finite-Horizon Multivariate Linear Rational Expectations Models and Sparse Linear Systems, *Journal of Economic Dynamics and Control*, 24, 325-346.

Blundell, R., and S.R. Bond (1998): Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, *Journal of Econometrics*, 87, 115-143.

Blundell, R., and R.J. Smith (1991): Initial Conditions and Efficient Estimation in Dynamic Panel Data Models, *Annales d'Economie et de Statistique*, 20/21, 109-123.

Conley, T.G. (1999): GMM Estimation with Cross Sectional Dependence, *Journal of Econometrics*, 92, 1-45.

Donald, S.G., and W.K. Newey (2001): Choosing the Number of Instruments, *Econometrica*, 69, 1161-1191.

Hahn, J. (1999): How Informative is the Initial Condition in the Dynamic Panel Data Model with Fixed Effects?, *Journal of Econometrics*, 93, 309-326.

Harris, R.D.F., and E. Tzavalis (1999): Inference for Unit Roots in Dynamic Panels where the Time Dimension is Fixed, *Journal of Econometrics*, 91, 201-226.

Holtz-Eakin, D., W.K. Newey, and H.S. Rosen (1988): Estimating Vector Autoregressions with Panel Data, *Econometrica*, 56, 1371-1395.

Hsiao, C., M.H. Pesaran, and A.K. Tahmiscioglu (1999): Bayes Estimation of Short-run Coefficients in Dynamic Panel Data Models, in C. Hsiao, K. Lahiri, L-F Lee and M.H. Pesaran (eds), *Analysis of Panels and Limited Dependent Variables*, A Volume in Honour of G S Maddala Cambridge University Press, Cambridge, 1999, chapter 11

Hsiao, C., M.H. Pesaran, and A.K. Tahmiscioglu (2002): Maximum Likelihood Estimation of Fixed Effects Dynamic Panel Data Models Covering Short Time Periods, *Journal of Econometrics, 109, 107-150.*

Im, K., Pesaran, H., and Y. Shin, (2003): Testing for Unit Roots in Heterogeneous Panels, *Journal of Econometrics* 115, 53-74.

Johansen, S. (1995): *Likelihood Based Inference on Cointegration in the Vector Autoregressive Model*, Oxford: Oxford University Press.

Kiviet, J.F. (1995): On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models, *Journal of Econometrics*, 68, 53-78.

Levin, A., Lin, F. and C. Chu, (2002): Unit root tests in panel data: asymptotic and finite-sample properties, *Journal of Econometrics*, 108, 1-24.

Maddala, G.S., Wu, S., (1999): A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test, *Oxford Bulletin of Economics and Statistics,* 61, 631-652.

Moon, H.R., and B. Perron, (2003): Testing for a Unit Root in Panels with Dynamic Factors, *Journal of Econometrics* (forthcoming).

Nerlove, M. (1999): Likelihood Inference for Dynamic Panel Data Models, *Annales d'Economie et de Statistique*, 55/56, 370-410.

Neyman, J., and E. Scott (1948): Consistent Estimates Based on Partially Consistent Observations, *Econometrica*, 16, 1-32.

Nickell, S. (1981): Biases in Dynamic Models with Fixed Effects, *Econometrica*, 49, 1399-1416.

Pesaran, M.H., (2002) "Estimation and Inference in Large Heterogeneous Panels with Cross Section Dependence", *University of Cambridge DAE Working Paper No.0305* and CESifo Working Paper Series No 869, October 2002.

Pesaran, M.H., (2003) "A Simple Panel Unit Root Test in the Presence of Cross Section Dependence", *University of Southern California, unpublished manuscript, September 2003.*

Pesaran, M.H., and Y. Shin (2002): Long-Run Structural Modelling, *Econometric Reviews*, 21, 49-87.

Pesaran, M.H., Y. Shin, and R.J. Smith (2000): Structural Analysis of Vector Error Correction Models with Exogenous I(1) Variables, *Journal of Econometrics*, 97, 293-343.

Phillips, P.C.B. (1991): Optimal Inference in Cointegrated Systems, *Econometrica*, 59, 283-306.

Phillips, P.C.B., and H.R. Moon (2000): Nonstationary Panel Data Analysis: An Overview of Some Recent Developments, *Econometric Reviews*, 19, 263-286.

Phillips, P.C.B. and D. Sul, (2002): "Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence, *Cowles Foundation Discussion Paper No. 1362*, Yale University.

Staiger, D., and J.H. Stock (1997): Instrumental Variables Regression with Weak Instruments, *Econometrica*, 65, 557-586.

Wansbeek, T., and P. Bekker (1996): On IV, GMM and ML in a Dynamic Panel Data Model, *Economics Letters*, 51, 145-152.

White, H. (1994): *Estimation, Inference and Specification Analysis,* Cambridge: Cambridge University Press.

Table 1: Bias and RMSE of Alternative Estimators of Panel VAR[1]

$\lambda_{max} = 0.6$

	Estimator	$\phi_{11} = 0.4$								$\phi_{21} = 0.2$							
	True Value	*Bias*				*RMSE*				*Bias*				*RMSE*			
		N = 50		*N = 250*		*N = 50*		*N = 250*		*N = 50*		*N = 250*		*N = 50*		*N = 250*	
		$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$
T = 3	*SGMM*	0.0884	0.1643	0.0211	0.0443	0.3349	0.4459	0.1461	0.2059	0.0261	0.0604	0.0145	0.0271	0.3265	0.4323	0.1432	0.1975
	EGMM I	-0.0109	-0.0947	-0.0044	-0.0245	0.2066	0.2628	0.0964	0.1258	0.0092	0.0259	0.0052	0.0084	0.2064	0.2399	0.0943	0.1171
	EGMM II	0.0222	-0.0415	0.0024	-0.0216	0.2729	0.3150	0.1231	0.1468	0.0150	0.0311	0.0063	0.0055	0.2729	0.3147	0.1203	0.1411
	EGMM III	-0.0234	-0.1346	-0.0085	-0.0391	0.2130	0.2824	0.0958	0.1256	0.0159	0.0415	0.0023	0.0003	0.2130	0.2400	0.0936	0.1127
	FE-QML	0.0027		0.0003		0.1969		0.0898		0.0027		0.0008		0.1969		0.0809	
T = 10	*SGMM*	0.0894	0.1133	0.0199	0.0266	0.1334	0.1544	0.0491	0.0561	0.0222	0.0354	0.0069	0.0113	0.1051	0.1159	0.0458	0.0515
	EGMM I	-0.0656	-0.2004	-0.0072	-0.0291	0.1193	0.2441	0.0382	0.0550	0.0205	0.0588	0.0024	0.0014	0.0965	0.1315	0.0367	0.0446
	EGMM I	0.0502	-0.0036	0.0107	0.0038	0.1064	0.1127	0.0431	0.0449	0.0173	0.0332	0.0048	0.0057	0.0994	0.1164	0.0423	0.0461
	EGMM II	-0.0702	-0.2181	-0.0068	-0.0323	0.1265	0.2615	0.0383	0.0577	0.0238	0.0685	0.0024	0.0021	0.1020	0.1395	0.0367	0.0452
	FE-QML	0.0023		0.0027		0.0737		0.0327		0.0005		0.0019		0.0706		0.0303	

$\lambda_{max} = 0.8$

	Estimator	$\phi_{11} = 0.6$								$\phi_{21} = 0.2$							
	True Value	*Bias*				*RMSE*				*Bias*				*RMSE*			
		N = 50		*N = 250*		*N = 50*		*N = 250*		*N = 50*		*N = 250*		*N = 50*		*N = 250*	
		$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$
T = 3	*SGMM*	0.1704	0.2842	0.0395	0.0798	0.4057	0.5450	0.1614	0.2327	0.0781	0.1198	0.0179	0.0340	0.3777	0.5090	0.1561	0.2200
	EGMM I	0.0161	-0.0397	0.0083	-0.0038	0.1710	0.2041	0.0828	0.1071	0.0121	0.0393	0.0035	0.0075	0.1750	0.2041	0.0800	0.1018
	EGMM II	0.0594	0.0199	0.0135	-0.0023	0.2405	0.2634	0.1131	0.1306	0.0126	0.0306	0.0070	0.0124	0.2238	0.2558	0.1087	0.1288
	EGMM III	0.0023	-0.0735	0.0029	-0.0216	0.1649	0.2041	0.0799	0.1033	0.0137	0.0515	0.0043	0.0094	0.1646	0.1933	0.0784	0.0974
	FE-QML	0.0105		0.0040		0.1694		0.0736		0.0037		0.0014		0.1385		0.0607	
T = 10	*SGMM*	0.1228	0.1591	0.0292	0.0427	0.1474	0.1852	0.0466	0.0620	0.0390	0.0582	0.0103	0.0198	0.0960	0.1203	0.0363	0.0481
	EGMM I	-0.0394	-0.1328	-0.0047	-0.0242	0.0779	0.1601	0.0282	0.0446	0.0288	0.0747	0.0024	0.0048	0.0688	0.1063	0.0264	0.0341
	EGMM II	0.0691	0.0314	0.0139	0.0086	0.0983	0.0928	0.0341	0.0373	0.0185	0.0367	0.0030	0.0053	0.0726	0.0907	0.0303	0.0372
	EGMM III	-0.0415	-0.1425	-0.0046	-0.0266	0.0818	0.1700	0.0283	0.0464	0.0304	0.0817	0.0022	0.0057	0.0724	0.1135	0.0267	0.0349
	FE-QML	0.0038		0.0022		0.0513		0.0231		0.0023		-0.0003		0.0445		0.0198	

For details of the Monte Carlo design, see Section 7.1. The data generating process is given by $(\boldsymbol{I}_2 - \boldsymbol{\Phi}L)(\boldsymbol{w}_{it} - \boldsymbol{\mu}_i) = \boldsymbol{\varepsilon}_{it}$. λ_{max} denotes the maximum eigenvalue of $\boldsymbol{\Phi}$, and ϕ_{jk} the element in the j-th row and k-th column of $\boldsymbol{\Phi}$, $j, k = 1, 2$. 'RMSE' denotes the root mean square error, 'SGMM' the Standard GMM Estimator, 'EGMM I' the Extended GMM Estimator I, 'EGMM II' the Extended GMM Estimator II, 'EGMM III' the Extended GMM Estimator III and 'FE-QML' the Fixed Effects Quasi Maximum Likelihood Estimator. See Section 7.2 for further details.

Table 1 (Continued)
Bias and RMSE of Alternative Estimators of Panel VAR

$\lambda_{max} = 0.95$

	Estimator	$\phi_{11} = 0.7$								$\phi_{21} = 0.25$							
	True Value	*Bias*				*RMSE*				*Bias*				*RMSE*			
		N = 50		*N = 250*		*N = 50*		*N = 250*		*N = 50*		*N = 250*		*N = 50*		*N = 250*	
		$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$
T = 3	*SGMM*	0.3585	0.4193	0.1150	0.1596	0.8923	0.9174	0.4020	0.4966	0.2331	0.2297	0.0521	0.0787	0.8527	0.8501	0.3792	0.4688
	EGMM I	0.0260	-0.0149	0.0153	0.0091	0.1982	0.2396	0.1104	0.1396	0.0196	0.0524	0.0006	0.0056	0.1987	0.2369	0.1042	0.1339
	EGMM II	0.0711	0.0494	0.0246	0.0151	0.2979	0.3223	0.1519	0.1743	0.0186	0.0416	0.0033	0.0121	0.2836	0.3187	0.1474	0.1723
	EGMM III	0.0124	-0.0426	0.0090	-0.0069	0.1743	0.2101	0.0994	0.1241	0.0171	0.0660	0.0027	0.0138	0.1775	0.2140	0.0957	0.1206
	FE-QML	0.0175		0.0064		0.2026		0.0873		0.0134		0.0006		0.1674		0.0727	
T = 10	*SGMM*	0.1892	0.2151	0.0587	0.0754	0.2336	0.2605	0.0879	0.1081	0.1057	0.1113	0.0379	0.0488	0.1778	0.1897	0.0740	0.0905
	EGMM I	-0.0248	-0.0945	-0.0017	-0.0145	0.0707	0.1268	0.0330	0.0492	0.0328	0.0946	0.0040	0.0125	0.0707	0.1229	0.0325	0.0468
	EGMM I	0.0754	0.0464	0.0163	0.0129	0.1157	0.1147	0.0429	0.0508	0.0179	0.0395	0.0035	0.0066	0.0862	0.1064	0.0388	0.0493
	EGMM II	-0.0264	-0.1022	-0.0014	-0.0161	0.0727	0.1336	0.0332	0.0497	0.0350	0.1024	0.0038	0.0139	0.0732	0.1298	0.0328	0.0475
	FE-QML	0.0053		0.0024		0.0558		0.0259		0.0025		-0.0004		0.0486		0.0217	

$\lambda_{max} = 1$

	Estimator	$\phi_{11} = 1$								$\phi_{21} = 0$							
	True Value	*Bias*				*RMSE*				*Bias*				*RMSE*			
		N = 50		*N = 250*		*N = 50*		*N = 250*		*N = 50*		*N = 250*		*N = 50*		*N = 250*	
		$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$	$\tau = 1$	$\tau = 5$
T = 3	*SGMM*	0.9731	0.9576	0.8874	0.8656	1.3102	1.2884	1.2042	1.1845	-0.0295	-0.0225	0.0499	0.0681	0.9415	0.9780	0.9263	0.9374
	EGMM I	0.0354	0.0358	0.0034	0.0036	0.1598	0.1635	0.0571	0.0571	0.0001	-0.0001	-0.0016	-0.0016	0.1387	0.1421	0.0527	0.0539
	EGMM II	0.1019	0.1029	0.0187	0.0184	0.2837	0.2834	0.0962	0.0968	-0.0155	-0.0138	-0.0039	-0.0039	0.2268	0.2319	0.0897	0.0915
	EGMM III	0.0172	0.0180	0.0033	0.0035	0.1055	0.1050	0.0514	0.0513	-0.0039	-0.0034	-0.0013	-0.0015	0.1027	0.1018	0.0484	0.0488
	FE-QML	0.0234		0.0069		0.2031		0.1012		-0.0015		-0.0031		0.1562		0.0693	
T = 10	*SGMM*	0.4874	0.4879	0.5283	0.5288	0.5257	0.5267	0.5727	0.5728	0.0075	0.0077	0.0125	0.0130	0.2017	0.2022	0.2176	0.2184
	EGMM I	0.0023	0.0024	0.0003	0.0003	0.0244	0.0235	0.0117	0.0116	0.0013	0.0011	-0.0000	-0.0000	0.0234	0.0225	0.0117	0.0117
	EGMM II	0.0354	0.0358	0.0034	0.0036	0.0712	0.0713	0.0188	0.0189	0.0034	0.0028	-0.0002	-0.0002	0.0549	0.0543	0.0174	0.0175
	EGMM III	0.0024	0.0025	0.0004	0.0005	0.0239	0.0231	0.0116	0.0115	0.0016	0.0015	-0.0001	-0.0001	0.0234	0.0227	0.0115	0.0115
	FE-QML	0.0091		0.0006		0.0623		0.0274		0.0021		0.0003		0.0443		0.0182	

Table 2a : Size and Power Properties of Tests for ϕ_{11} Under Alternative Estimators of Panel VAR[2,3], $\lambda_{max} = 0.6$

	Estimator		$\phi_{11} = 0.1$	$\phi_{11} = 0.2$	$\phi_{11} = 0.3$	$\boldsymbol{\phi_{11} = 0.4}$	$\phi_{11} = 0.5$	$\phi_{11} = 0.6$	$\phi_{11} = 0.7$
N = 50, T = 3	*SGMM*	$\tau = 1$	0.1220	0.0790	0.0710	**0.0990**	0.1730	0.2610	0.3790
		$\tau = 5$	0.0600	0.0530	0.0750	**0.1260**	0.1860	0.2580	0.3450
	EGMM I	$\tau = 1$	0.4120	0.2560	0.1440	**0.0880**	0.1100	0.2250	0.3670
		$\tau = 5$	0.4920	0.3870	0.2750	**0.1790**	0.1180	0.1410	0.2210
	EGMM II	$\tau = 1$	0.2130	0.1310	0.0810	**0.0930**	0.1400	0.2360	0.3690
		$\tau = 5$	0.2900	0.1960	0.1350	**0.1010**	0.1130	0.1610	0.2500
	EGMM III	$\tau = 1$	0.4700	0.3000	0.1760	**0.1250**	0.1430	0.2280	0.0420
		$\tau = 5$	0.6240	0.5090	0.3640	**0.2750**	0.1930	0.1820	0.2640
	FE-QML (Normal)		0.2960	0.1440	0.0700	**0.0640**	0.0950	0.1880	0.3530
	FE-QML (Robust)		0.3080	0.1670	0.0930	**0.0730**	0.1120	0.2050	0.3480
N = 50, T = 10	*SGMM*	$\tau = 1$	0.7630	0.3800	0.1400	**0.2910**	0.6710	0.9400	0.9920
		$\tau = 5$	0.6290	0.2760	0.1310	**0.3490**	0.6970	0.9410	0.9910
	EGMM I	$\tau = 1$	0.9990	0.9410	0.6810	**0.3030**	0.2660	0.6360	0.9020
		$\tau = 5$	1.0000	0.9890	0.9210	**0.7110**	0.4860	0.4250	0.6190
	EGMM II	$\tau = 1$	0.8950	0.5760	0.2190	**0.2180**	0.5890	0.8930	0.9890
		$\tau = 5$	0.9410	0.7120	0.3740	**0.2280**	0.4120	0.7330	0.9280
	EGMM III	$\tau = 1$	0.9950	0.9470	0.7030	**0.3410**	0.2960	0.6430	0.8990
		$\tau = 5$	1.0000	0.9930	0.9400	**0.7800**	0.5480	0.4680	0.6320
	FE-QML (Normal)		0.9810	0.7860	0.2870	**0.0660**	0.2770	0.7930	0.9860
	FE-QML (Robust)		0.9720	0.7520	0.3040	**0.0760**	0.2800	0.7820	0.9790
N = 250, T = 3	*SGMM*	$\tau = 1$	0.5220	0.2600	0.0810	**0.0750**	0.1920	0.4200	0.6640
		$\tau = 5$	0.2570	0.0980	0.0510	**0.0790**	0.1670	0.3310	0.4980
	EGMM I	$\tau = 1$	0.8890	0.6010	0.2340	**0.0600**	0.1810	0.5630	0.8950
		$\tau = 5$	0.7750	0.5280	0.2520	**0.1020**	0.1160	0.3370	0.6970
	EGMM II	$\tau = 1$	0.7180	0.3830	0.1270	**0.0630**	0.1880	0.4540	0.7150
		$\tau = 5$	0.6270	0.3600	0.1570	**0.0720**	0.1180	0.2970	0.5630
	EGMM III	$\tau = 1$	0.9230	0.6400	0.2480	**0.0700**	0.2010	0.5850	0.8910
		$\tau = 5$	0.8550	0.6110	0.3010	**0.1320**	0.1340	0.3750	0.7280
	FE-QML (Normal)		0.8400	0.5350	0.1720	**0.0440**	0.1870	0.5570	0.8250
	FE-QML (Robust)		0.7410	0.4560	0.1610	**0.0590**	0.1730	0.4990	0.7380
N = 250, T = 10	*SGMM*	$\tau = 1$	1.0000	0.9900	0.5140	**0.0980**	0.8120	1.0000	1.0000
		$\tau = 5$	1.0000	0.9690	0.4000	**0.1220**	0.7790	0.9990	1.0000
	EGMM I	$\tau = 1$	1.0000	1.0000	0.8560	**0.0680**	0.7640	0.9990	1.0000
		$\tau = 5$	1.0000	1.0000	0.8900	**0.1760**	0.5010	0.9770	1.0000
	EGMM II	$\tau = 1$	1.0000	0.9990	0.6430	**0.0890**	0.8110	1.0000	1.0000
		$\tau = 5$	1.0000	0.9960	0.6420	**0.0840**	0.7010	0.9990	1.0000
	EGMM III	$\tau = 1$	1.0000	1.0000	0.8540	**0.0700**	0.7670	1.0000	1.0000
		$\tau = 5$	1.0000	1.0000	0.9040	**0.2040**	0.4920	0.9720	1.0000
	FE-QML (Normal)		0.9960	0.9960	0.8020	**0.0460**	0.8510	0.9960	0.9960
	FE-QML (Robust)		0.9900	0.9750	0.7310	**0.0480**	0.7570	0.9810	0.9930

[2] See the footnote to Table 1 for a description of the data generating process and the notation used in this table.
[3] The table reports the fraction of rejections for tests of H_0: $\phi_{11} = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7\}$, versus two-sided alternatives. The true value of ϕ_{11} is equal to *0.4*.

Table 2b: Size and Power Properties of Tests for ϕ_{21} Under Alternative Estimators of Panel VAR[4,5], $\lambda_{max} = 0.6$

	Estimator		$\phi_{21} = -0.1$	$\phi_{21} = 0$	$\phi_{21} = 0.1$	$\boldsymbol{\phi_{21} = 0.2}$	$\phi_{21} = 0.3$	$\phi_{21} = 0.4$	$\phi_{21} = 0.5$
N = 50, T = 3	*SGMM*	$\tau = 1$	0.1680	0.1180	0.0740	**0.0800**	0.1120	0.1640	0.2580
		$\tau = 5$	0.1110	0.0800	0.0560	**0.0670**	0.1070	0.1500	0.2060
	EGMM I	$\tau = 1$	0.3650	0.2160	0.1100	**0.0730**	0.1160	0.2300	0.4100
		$\tau = 5$	0.3190	0.1880	0.1050	**0.0820**	0.1360	0.2640	0.4140
	EGMM II	$\tau = 1$	0.2590	0.1620	0.0990	**0.0820**	0.1190	0.1950	0.3180
		$\tau = 5$	0.2220	0.1340	0.0880	**0.0860**	0.1330	0.2030	0.3100
	EGMM III	$\tau = 1$	0.3960	0.2410	0.1250	**0.1010**	0.1670	0.2890	0.4570
		$\tau = 5$	0.3830	0.2240	0.1400	**0.1460**	0.2350	0.3520	0.5090
	FE-QML (Normal)		0.3870	0.2040	0.0980	**0.0600**	0.0990	0.2150	0.3800
	FE-QML (Robust)		0.3790	0.2280	0.1020	**0.0670**	0.1030	0.2160	0.3840
N = 50, T = 10	*SGMM*	$\tau = 1$	0.9000	0.6410	0.2780	**0.1570**	0.4180	0.7690	0.9620
		$\tau = 5$	0.8370	0.5380	0.2260	**0.1720**	0.4170	0.7590	0.9480
	EGMM I	$\tau = 1$	0.9550	0.7640	0.3770	**0.2230**	0.5070	0.8870	0.9860
		$\tau = 5$	0.8850	0.6600	0.4130	**0.3980**	0.6590	0.9020	0.9870
	EGMM II	$\tau = 1$	0.9330	0.7110	0.3430	**0.1730**	0.4470	0.8070	0.9770
		$\tau = 5$	0.8840	0.6330	0.3150	**0.2330**	0.4880	0.8080	0.9680
	EGMM III	$\tau = 1$	0.9540	0.7640	0.3970	**0.2670**	0.5420	0.8870	0.9880
		$\tau = 5$	0.8820	0.6580	0.4460	**0.4770**	0.7070	0.9210	0.9880
	FE-QML (Normal)		0.9960	0.8010	0.2940	**0.0650**	0.3050	0.8240	0.9850
	FE-QML (Robust)		0.9850	0.7950	0.3120	**0.0690**	0.3060	0.8080	0.9800
N = 250, T = 3	*SGMM*	$\tau = 1$	0.5580	0.2950	0.0940	**0.0620**	0.1690	0.3750	0.6290
		$\tau = 5$	0.3220	0.1460	0.0520	**0.0640**	0.1260	0.2670	0.4530
	EGMM I	$\tau = 1$	0.8880	0.5710	0.1830	**0.0630**	0.2110	0.6130	0.9030
		$\tau = 5$	0.7310	0.4110	0.1330	**0.0680**	0.1680	0.4600	0.7830
	EGMM II	$\tau = 1$	0.7080	0.3850	0.1330	**0.0560**	0.1720	0.4370	0.7560
		$\tau = 5$	0.5830	0.3030	0.1100	**0.0600**	0.1540	0.3540	0.6120
	EGMM III	$\tau = 1$	0.9090	0.6050	0.2120	**0.0600**	0.2270	0.6310	0.9090
		$\tau = 5$	0.8100	0.4960	0.1990	**0.0750**	0.2030	0.5100	0.8130
	FE-QML (Normal)		0.8880	0.6130	0.2190	**0.0530**	0.2170	0.6290	0.9020
	FE-QML (Robust)		0.8030	0.5310	0.1950	**0.0600**	0.2070	0.5360	0.8020
N = 250, T = 10	*SGMM*	$\tau = 1$	1.0000	0.9940	0.6190	**0.0890**	0.7360	0.9980	1.0000
		$\tau = 5$	1.0000	0.9790	0.5250	**0.0920**	0.6970	0.9910	1.0000
	EGMM I	$\tau = 1$	1.0000	1.0000	0.7960	**0.0650**	0.8360	1.0000	1.0000
		$\tau = 5$	1.0000	0.9980	0.7150	**0.1110**	0.7480	0.9950	1.0000
	EGMM II	$\tau = 1$	1.0000	0.9980	0.6860	**0.0790**	0.7600	0.9980	1.0000
		$\tau = 5$	1.0000	0.9930	0.6330	**0.0880**	0.7010	0.9950	1.0000
	EGMM III	$\tau = 1$	1.0000	1.0000	0.8000	**0.0680**	0.8430	1.0000	1.0000
		$\tau = 5$	1.0000	0.9980	0.7420	**0.1260**	0.7620	0.9980	1.0000
	FE-QML (Normal)		0.9970	0.9960	0.8410	**0.0530**	0.8880	0.9960	0.9960
	FE-QML (Robust)		0.9930	0.9780	0.7720	**0.0510**	0.8120	0.9820	0.9940

[4] See the footnote to Table 1 for a description of the data generating process and the notation used in this table.
[5] The table reports the fraction of rejections for tests of H_0: $\phi_{21} = \{-0.1, 0, 0.1, 0.2, 0.3, 0.4, 0.5\}$, versus two-sided alternatives. The true value of ϕ_{21} is equal to *0.2*.

Table 2c: Size and Power Properties of Tests for ϕ_{11} Under Alternative Estimators of Panel VAR[6,7], $\lambda_{max} = 1$

	Estimator		$\phi_{11} = 0.7$	$\phi_{11} = 0.8$	$\phi_{11} = 0.9$	$\phi_{11} = 1$	$\phi_{11} = 1.1$	$\phi_{11} = 1.2$	$\phi_{11} = 1.3$
N = 50, T = 3	*SGMM*	$\tau = 1$	0.1990	0.2290	0.2730	**0.3150**	0.3760	0.4310	0.4880
		$\tau = 5$	0.1860	0.2260	0.2720	**0.3080**	0.3610	0.4050	0.4650
	EGMM I	$\tau = 1$	0.6980	0.4900	0.1740	**0.0340**	0.2200	0.5980	0.8200
		$\tau = 5$	0.7160	0.4920	0.1690	**0.0320**	0.2250	0.6170	0.8240
	EGMM II	$\tau = 1$	0.3560	0.2120	0.0860	**0.0780**	0.1880	0.3970	0.5990
		$\tau = 5$	0.3690	0.2170	0.0870	**0.0770**	0.1910	0.3940	0.5980
	EGMM III	$\tau = 1$	0.8460	0.6510	0.2680	**0.0390**	0.2980	0.7270	0.9210
		$\tau = 5$	0.8500	0.6560	0.2770	**0.0380**	0.3110	0.7490	0.9210
	FE-QML (Normal)		0.3600	0.2110	0.1160	**0.0840**	0.1390	0.2540	0.4260
	FE-QML (Robust)		0.3720	0.2430	0.1270	**0.1090**	0.1590	0.2740	0.4250
N = 50, T = 10	*SGMM*	$\tau = 1$	0.4700	0.6760	0.8490	**0.9480**	0.9860	0.9970	1.0000
		$\tau = 5$	0.4710	0.6670	0.8510	**0.9430**	0.9850	0.9970	1.0000
	EGMM I	$\tau = 1$	1.0000	0.9990	0.9770	**0.1210**	0.9960	1.0000	1.0000
		$\tau = 5$	1.0000	0.9990	0.9800	**0.1280**	0.9970	1.0000	1.0000
	EGMM II	$\tau = 1$	0.9820	0.9200	0.6010	**0.3020**	0.9350	0.9950	1.0000
		$\tau = 5$	0.9830	0.9140	0.5890	**0.3150**	0.9390	0.9960	1.0000
	EGMM III	$\tau = 1$	1.0000	1.0000	0.9810	**0.1570**	0.9970	1.0000	1.0000
		$\tau = 5$	1.0000	1.0000	0.9810	**0.1610**	0.9980	1.0000	1.0000
	FE-QML (Normal)		0.9700	0.8860	0.4900	**0.0870**	0.5860	0.9460	0.9750
	FE-QML (Robust)		0.9700	0.8760	0.5130	**0.1040**	0.6030	0.9330	0.9760
N = 250, T = 3	*SGMM*	$\tau = 1$	0.1650	0.1980	0.2370	**0.2840**	0.3220	0.3790	0.4310
		$\tau = 5$	0.1550	0.1950	0.2440	**0.2860**	0.3210	0.3680	0.4150
	EGMM I	$\tau = 1$	0.9830	0.9220	0.5630	**0.0310**	0.5560	0.9570	0.9930
		$\tau = 5$	0.9820	0.9140	0.5600	**0.0270**	0.5580	0.9560	0.9890
	EGMM II	$\tau = 1$	0.8570	0.6330	0.2170	**0.0720**	0.3280	0.7470	0.9250
		$\tau = 5$	0.8530	0.6120	0.2110	**0.0690**	0.3320	0.7450	0.9130
	EGMM III	$\tau = 1$	0.9900	0.9410	0.6010	**0.0400**	0.6200	0.9710	0.9980
		$\tau = 5$	0.9910	0.9400	0.6010	**0.0390**	0.6320	0.9760	0.9970
	FE-QML (Normal)		0.7210	0.5120	0.2010	**0.0660**	0.2240	0.5520	0.7490
	FE-QML (Robust)		0.6740	0.4800	0.2020	**0.0820**	0.2220	0.4890	0.6940
N = 250, T = 10	*SGMM*	$\tau = 1$	0.4830	0.6630	0.8320	**0.9350**	0.9790	0.9940	0.9990
		$\tau = 5$	0.4850	0.6660	0.8260	**0.9360**	0.9810	0.9970	0.9990
	EGMM I	$\tau = 1$	1.0000	1.0000	1.0000	**0.0580**	1.0000	1.0000	1.0000
		$\tau = 5$	1.0000	1.0000	1.0000	**0.0550**	1.0000	1.0000	1.0000
	EGMM II	$\tau = 1$	1.0000	1.0000	1.0000	**0.0930**	1.0000	1.0000	1.0000
		$\tau = 5$	1.0000	1.0000	0.9990	**0.0940**	1.0000	1.0000	1.0000
	EGMM III	$\tau = 1$	1.0000	1.0000	1.0000	**0.0670**	1.0000	1.0000	1.0000
		$\tau = 5$	1.0000	1.0000	1.0000	**0.0710**	1.0000	1.0000	1.0000
	FE-QML (Normal)		0.9630	0.9560	0.9140	**0.0500**	0.9340	0.9570	0.9640
	FE-QML (Robust)		0.9730	0.9620	0.8900	**0.0600**	0.9100	0.9620	0.9730

[6] See the footnote to Table 1 for a description of the data generating process and the notation used in this table.
[7] The table reports the fraction of rejections for tests of H_0: $\phi_{11} = \{0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3\}$, versus two-sided alternatives. The true value of ϕ_{11} is equal to *1*.

Table 2d: Size and Power Properties of Tests for ϕ_{21} Under Alternative Estimators of Panel VAR[8,9], $\lambda_{max} = 1$

	Estimator		ϕ_{21} = -0.3	ϕ_{21} = -0.2	ϕ_{21} = -0.1	ϕ_{21} = 0	ϕ_{21} = 0.1	ϕ_{21} = 0.2	ϕ_{21} = 0.3
N = 50, T = 3	*SGMM*	$\tau = 1$	0.0920	0.0740	0.0630	**0.0500**	0.0500	0.0620	0.0790
		$\tau = 5$	0.0850	0.0690	0.0650	**0.0580**	0.0560	0.0670	0.0780
	EGMM I	$\tau = 1$	0.7530	0.5380	0.1990	**0.0300**	0.1620	0.5420	0.7540
		$\tau = 5$	0.7510	0.5380	0.1980	**0.0240**	0.1690	0.5390	0.7570
	EGMM II	$\tau = 1$	0.4640	0.2690	0.1040	**0.0410**	0.0900	0.2540	0.4430
		$\tau = 5$	0.4750	0.2750	0.1120	**0.0420**	0.1010	0.2470	0.4450
	EGMM III	$\tau = 1$	0.8870	0.6750	0.2810	**0.0410**	0.2630	0.6880	0.8730
		$\tau = 5$	0.8930	0.6900	0.2940	**0.0400**	0.2730	0.6940	0.8750
	FE-QML (Normal)		0.4200	0.2440	0.1050	**0.0620**	0.1050	0.2230	0.4100
	FE-QML (Robust)		0.4380	0.2610	0.1290	**0.0780**	0.1230	0.2350	0.4300
	FE-MD		0.3700	0.2110	0.1170	**0.0790**	0.1350	0.2500	0.3740
N = 50, T = 10	*SGMM*	$\tau = 1$	0.6680	0.4950	0.3790	**0.3240**	0.3890	0.5290	0.6910
		$\tau = 5$	0.6750	0.5060	0.3520	**0.3160**	0.3840	0.5230	0.7070
	EGMM I	$\tau = 1$	1.0000	1.0000	0.9830	**0.1210**	0.9870	1.0000	1.0000
		$\tau = 5$	1.0000	1.0000	0.9850	**0.1160**	0.9890	1.0000	1.0000
	EGMM II	$\tau = 1$	0.9990	0.9730	0.7710	**0.2050**	0.8090	0.9820	0.9990
		$\tau = 5$	0.9990	0.9750	0.7710	**0.2120**	0.8040	0.9810	0.9990
	EGMM III	$\tau = 1$	1.0000	1.0000	0.9880	**0.1560**	0.9900	1.0000	1.0000
		$\tau = 5$	1.0000	1.0000	0.9900	**0.1480**	0.9910	1.0000	1.0000
	FE-QML (Normal)		0.9730	0.9520	0.5810	**0.0600**	0.6220	0.9470	0.9750
	FE-QML (Robust)		0.9720	0.9410	0.5830	**0.0780**	0.6050	0.9450	0.9710
N = 250, T = 3	*SGMM*	$\tau = 1$	0.0760	0.0600	0.0460	**0.0470**	0.0520	0.0580	0.0730
		$\tau = 5$	0.0770	0.0550	0.0430	**0.0460**	0.0470	0.0570	0.0820
	EGMM I	$\tau = 1$	0.9940	0.9490	0.5480	**0.0300**	0.5430	0.9370	0.9870
		$\tau = 5$	0.9910	0.9400	0.5500	**0.0260**	0.5390	0.9360	0.9840
	EGMM II	$\tau = 1$	0.9200	0.7020	0.2490	**0.0340**	0.2530	0.6820	0.8960
		$\tau = 5$	0.9090	0.7030	0.2540	**0.0310**	0.2520	0.6700	0.8890
	EGMM III	$\tau = 1$	0.9970	0.9750	0.6160	**0.0350**	0.6000	0.9590	0.9980
		$\tau = 5$	0.9970	0.9740	0.6180	**0.0360**	0.6150	0.9550	0.9970
	FE-QML (Normal)		0.8170	0.6240	0.2610	**0.0480**	0.2320	0.6050	0.8030
	FE-QML (Robust)		0.7580	0.5490	0.2500	**0.0620**	0.2250	0.5530	0.7460
N = 250, T = 10	*SGMM*	$\tau = 1$	0.6250	0.4350	0.2980	**0.2650**	0.3300	0.4900	0.6640
		$\tau = 5$	0.6140	0.4350	0.3040	**0.2690**	0.3380	0.4920	0.6550
	EGMM I	$\tau = 1$	1.0000	1.0000	1.0000	**0.0550**	1.0000	1.0000	1.0000
		$\tau = 5$	1.0000	1.0000	1.0000	**0.0540**	1.0000	1.0000	1.0000
	EGMM II	$\tau = 1$	1.0000	1.0000	1.0000	**0.0630**	0.9990	1.0000	1.0000
		$\tau = 5$	1.0000	1.0000	1.0000	**0.0640**	0.9990	1.0000	1.0000
	EGMM III	$\tau = 1$	1.0000	1.0000	1.0000	**0.0570**	1.0000	1.0000	1.0000
		$\tau = 5$	1.0000	1.0000	1.0000	**0.0550**	1.0000	1.0000	1.0000
	FE-QML (Normal)		0.9640	0.9600	0.9400	**0.0520**	0.9410	0.9610	0.9650
	FE-QML (Robust)		0.9780	0.9680	0.9260	**0.0580**	0.9350	0.9710	0.9790

[8] See the footnote to Table 1 for a description of the data generating process and the notation used in this table.
[9] The table reports the fraction of rejections for tests of H_0: $\phi_{21} = \{-0.3, -0.2, -0.1, 0, 0.1, 0.2, 0.3\}$, versus two-sided alternatives. The true value of ϕ_{21} is equal to *0*.

Table 3a: Bias and RMSE of Fixed Effects QML Estimator of Cointegrated Panel VAR[10]

	α_1 = -0.5		α_2 = -0.5		β_2 = -0.2	
	Bias	*RMSE*	*Bias*	*RMSE*	*Bias*	*RMSE*
N = 50, T = 3	0.0438	0.4422	0.0068	0.3883	-0.0186	0.4982
N = 50, T = 10	0.0058	0.2572	0.0024	0.2263	-0.0021	0.2483
N = 250, T = 3	0.0045	0.3044	0.0009	0.2579	-0.0075	0.3188
N = 250, T = 10	0.0019	0.1661	0.0015	0.1495	-0.0007	0.1604

Table 3b: Size and Power Properties of Cointegration Rank Tests Based on Fixed Effects QML Estimator[11]

	Size: H_1 vs. H_2	*Power: H_0 vs. H_1*
N = 50, T = 3	0.0230	0.8560
N = 50, T = 10	0.0500	1.0000
N = 250, T = 3	0.0460	1.0000
N = 250, T = 10	0.0460	1.0000

[10] See the footnote to Table 1 for a description of the data generating process, where now $\boldsymbol{\Phi} = \boldsymbol{I_2} + \boldsymbol{\alpha\beta}'$, with $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)'$ and $\boldsymbol{\beta} = (\beta_1, \beta_2)'$. The remaining notation is as described in the footnote to Table 1.
[11] The table reports the fraction of rejections for tests of H_r: $rank(\Pi) = r$ versus H_{r+1}: $rank(\Pi) = r+1$, $r = 0, 1$, where the true rank of Π is equal to 1.

Table 3c: Size and Power Properties of Tests for α_1 Under Fixed Effects QML Estimator of Cointegrated Panel VAR[12]

	Estimator	$\alpha_1 = -0.8$	$\alpha_1 = -0.7$	$\alpha_1 = -0.6$	$\alpha_1 = -0.5$	$\alpha_1 = -0.4$	$\alpha_1 = -0.3$	$\alpha_1 = -0.2$
N = 50, T = 3	*FE-QML (Normal)*	0.2240	0.1330	0.0870	**0.0750**	0.1250	0.2230	0.3730
	FE-QML (Robust)	0.2690	0.1640	0.1010	**0.0920**	0.1450	0.2480	0.3840
N = 50, T = 10	*FE-QML (Normal)*	0.9740	0.8240	0.3240	**0.0650**	0.3600	0.8670	0.9810
	FE-QML (Robust)	0.9490	0.7690	0.3360	**0.0920**	0.3670	0.8320	0.9600
N = 250, T = 3	*FE-QML (Normal)*	0.6530	0.4270	0.1710	**0.0750**	0.1870	0.4840	0.7190
	FE-QML (Robust)	0.6440	0.4430	0.2020	**0.1060**	0.2100	0.4620	0.6720
N = 250, T = 10	*FE-QML (Normal)*	0.9580	0.9520	0.8330	**0.0440**	0.8560	0.9500	0.9590
	FE-QML (Robust)	0.9490	0.9120	0.7320	**0.0600**	0.7680	0.9150	0.9510

Table 3d: Size and Power Properties of Tests for α_2 Under Fixed Effects QML Estimator of Cointegrated Panel VAR[13]

	Estimator	$\alpha_2 = -0.8$	$\alpha_2 = -0.7$	$\alpha_2 = -0.6$	$\alpha_2 = -0.5$	$\alpha_2 = -0.4$	$\alpha_2 = -0.3$	$\alpha_2 = -0.2$
N = 50, T = 3	*FE-QML (Normal)*	0.4330	0.2150	0.0910	**0.0590**	0.1310	0.2760	0.4910
	FE-QML (Robust)	0.4170	0.2250	0.1050	**0.0640**	0.1280	0.2840	0.4850
N = 50, T = 10	*FE-QML (Normal)*	0.9960	0.9460	0.4710	**0.0580**	0.4940	0.9690	0.9950
	FE-QML (Robust)	0.9850	0.9160	0.4770	**0.0620**	0.4850	0.9430	0.9840
N = 250, T = 3	*FE-QML (Normal)*	0.8510	0.7130	0.2890	**0.0660**	0.3010	0.7210	0.8590
	FE-QML (Robust)	0.8210	0.6530	0.2690	**0.0680**	0.2660	0.6510	0.8260
N = 250, T = 10	*FE-QML (Normal)*	0.9650	0.9630	0.9250	**0.0460**	0.9450	0.9630	0.9650
	FE-QML (Robust)	0.9710	0.9530	0.8620	**0.0510**	0.8760	0.9540	0.9710

Table 3e: Size and Power Properties of Tests for β_2 Under Fixed Effects QML Estimator of Cointegrated Panel VAR[14]

	Estimator	$\beta_2 = -0.5$	$\beta_2 = -0.4$	$\beta_2 = -0.3$	$\beta_2 = -0.2$	$\beta_2 = -0.1$	$\beta_2 = 0$	$\beta_2 = 0.1$
N = 50, T = 3	*FE-QML (Normal)*	0.2410	0.1130	0.0560	**0.0740**	0.1610	0.2530	0.3950
	FE-QML (Robust)	0.3030	0.1760	0.1080	**0.1070**	0.1760	0.2670	0.4080
N = 50, T = 10	*FE-QML (Normal)*	0.9880	0.9100	0.3560	**0.0450**	0.4010	0.8300	0.9710
	FE-QML (Robust)	0.9670	0.8460	0.3680	**0.0690**	0.4040	0.8050	0.9450
N = 250, T = 3	*FE-QML (Normal)*	0.6970	0.4170	0.1440	**0.0730**	0.2020	0.4640	0.6410
	FE-QML (Robust)	0.6670	0.4170	0.1820	**0.1050**	0.2210	0.4500	0.6250
N = 250, T = 10	*FE-QML (Normal)*	0.9620	0.9250	0.8800	**0.0470**	0.8690	0.9500	0.9580
	FE-QML (Robust)	0.9560	0.9310	0.7890	**0.0630**	0.7840	0.9300	0.9540

[12] The table reports the fraction of rejections for tests of H_0: $\alpha_1 = \{-0.8, -0.7, -0.6, -0.5, -0.4, -0.3, -0.2\}$, versus two-sided alternatives. The true value of α_1 is equal to *-0.5*.
[13] The table reports the fraction of rejections for tests of H_0: $\alpha_2 = \{-0.8, -0.7, -0.6, -0.5, -0.4, -0.3, -0.2\}$, versus two-sided alternatives. The true value of α_2 is equal to *-0.5*.
[14] The table reports the fraction of rejections for tests of H_0: $\beta_2 = \{-0.5, -0.4, -0.3, -0.2, -0.1, 0, 0.1\}$, versus two-sided alternatives. The true value of β_2 is equal to *–0.2*.

Table 4a: Bias and RMSE of Fixed Effects QML Estimator of Panel VAR With Non-Normal Disturbances[15], $\lambda_{max} = 0.6$

	Estimator								
		t Distributed Disturbances							
	True Value	$\phi_{11} = 0.4$				$\phi_{21} = 0.2$			
		Bias		*RMSE*		*Bias*		*RMSE*	
		N = 50	*N = 250*	*N = 50*	*N = 250*	*N = 50*	*N = 250*	*N = 50*	*N = 250*
T = 3	*FE-QML*	-0.0026	-0.0050	0.2100	0.0961	0.0005	-0.0049	0.1862	0.0888
T = 10	*FE-QML*	0.0044	0.0001	0.0728	0.0319	0.0025	0.0011	0.0688	0.0312
		Chi Square Distributed Disturbances							
	True Value	$\phi_{11} = 0.4$				$\phi_{21} = 0.2$			
		Bias		*RMSE*		*Bias*		*RMSE*	
		N = 50	*N = 250*	*N = 50*	*N = 250*	*N = 50*	*N = 250*	*N = 50*	*N = 250*
T = 3	*FE-QML*	-0.0085	-0.0001	0.2380	0.1070	0.0120	0.0010	0.2178	0.0969
T = 10	*FE-QML*	0.0041	-0.0005	0.0758	0.0325	0.0026	0.0003	0.0710	0.0304

[15] For details of the Monte Carlo design, see Section 7.1. The data generating process is given by $(\boldsymbol{I}_2 - \boldsymbol{\Phi}L)(\boldsymbol{w}_{it} - \boldsymbol{\mu}_i) = \boldsymbol{\varepsilon}_{it}$. See the footnote to Table 1 for a description of the notation used in this table.

Table 4b: Size and Power Properties of Tests for ϕ_{11} Based on Fixed Effects QML Estimator of Panel VAR in the Case of Non-Normal Disturbances[16,17], $\lambda_{max} = 0.6$

	Estimator	$\phi_{11} = 0.1$	$\phi_{11} = 0.2$	$\phi_{11} = 0.3$	$\boldsymbol{\phi_{11} = 0.4}$	$\phi_{11} = 0.5$	$\phi_{11} = 0.6$	$\phi_{11} = 0.7$
					t Distributed Disturbances			
N = 50, T = 3	*FE-QML (Normal)*	0.3200	0.1550	0.0780	**0.0660**	0.1120	0.2090	0.3390
	FE-QML (Robust)	0.3100	0.1620	0.0960	**0.0750**	0.1170	0.2290	0.3420
N = 50, T = 10	*FE-QML (Normal)*	0.9780	0.7510	0.2550	**0.0530**	0.3150	0.8000	0.9720
	FE-QML (Robust)	0.9590	0.7340	0.2540	**0.0630**	0.3240	0.7650	0.9510
N = 250, T = 3	*FE-QML (Normal)*	0.8390	0.5370	0.2000	**0.0660**	0.1720	0.5230	0.8040
	FE-QML (Robust)	0.6860	0.4330	0.1610	**0.0570**	0.1550	0.4160	0.6730
N = 250, T = 10	*FE-QML (Normal)*	0.9950	0.9940	0.8290	**0.0490**	0.8330	0.9940	0.9950
	FE-QML (Robust)	0.9910	0.9750	0.7570	**0.0540**	0.7550	0.9760	0.9870
					Chi Square Distributed Disturbances			
N = 50, T = 3	*FE-QML (Normal)*	0.3220	0.1990	0.1320	**0.0950**	0.1410	0.2130	0.3590
	FE-QML (Robust)	0.2940	0.1840	0.1070	**0.0960**	0.1450	0.2250	0.3420
N = 50, T = 10	*FE-QML (Normal)*	0.9730	0.7530	0.2690	**0.0620**	0.3120	0.7690	0.9770
	FE-QML (Robust)	0.9350	0.7230	0.2660	**0.0690**	0.3310	0.7440	0.9420
N = 250, T = 3	*FE-QML (Normal)*	0.8000	0.5100	0.2090	**0.0770**	0.2400	0.5440	0.7850
	FE-QML (Robust)	0.6160	0.3580	0.1390	**0.0510**	0.1510	0.3850	0.6340
N = 250, T = 10	*FE-QML (Normal)*	0.9940	0.9930	0.8330	**0.0540**	0.8170	0.9930	0.9940
	FE-QML (Robust)	0.9890	0.9740	0.7300	**0.0510**	0.7160	0.9710	0.9880

[16] See the footnote to Table 4a for a description of the data generating process used in this table, and the footnote to Table 1 for a description of the notation.

[17] The table reports the fraction of rejections for tests of H_0: $\phi_{11} = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7\}$, versus two-sided alternatives. The true value of ϕ_{11} is equal to *0.4*.

Table 4c: Size and Power Properties of Tests for ϕ_{21} Based on Fixed Effects QML Estimator of Panel VAR in the Case of Non-Normal Disturbances[18,19], $\lambda_{max} = 0.6$

	Estimator	ϕ_{21} = -0.1	ϕ_{21} = 0	ϕ_{21} = 0.1	**ϕ_{21} = 0.2**	ϕ_{21} = 0.3	ϕ_{21} = 0.4	ϕ_{21} = 0.5
					t Distributed Disturbances			
N = 50, T = 3	*FE-QML (Normal)*	0.3840	0.2100	0.1020	**0.0730**	0.1000	0.2190	0.3890
	FE-QML (Robust)	0.3510	0.2190	0.1090	**0.0680**	0.1050	0.2070	0.3570
N = 50, T = 10	*FE-QML (Normal)*	0.9920	0.8040	0.2960	**0.0460**	0.3280	0.8220	0.9880
	FE-QML (Robust)	0.9740	0.7840	0.2990	**0.0620**	0.3360	0.7980	0.9690
N = 250, T = 3	*FE-QML (Normal)*	0.8800	0.6140	0.2510	**0.0760**	0.2180	0.5900	0.8640
	FE-QML (Robust)	0.7530	0.4990	0.1890	**0.0510**	0.1760	0.4610	0.7340
N = 250, T = 10	*FE-QML (Normal)*	0.9950	0.9930	0.8530	**0.0530**	0.8650	0.9950	0.9950
	FE-QML (Robust)	0.9900	0.9780	0.7770	**0.0490**	0.7890	0.9800	0.9900
					Chi Square Distributed Disturbances			
N = 50, T = 3	*FE-QML (Normal)*	0.3620	0.2320	0.1490	**0.0980**	0.1450	0.2370	0.3860
	FE-QML (Robust)	0.3140	0.1930	0.1210	**0.0900**	0.1200	0.2120	0.3400
N = 50, T = 10	*FE-QML (Normal)*	0.9790	0.7990	0.2920	**0.0590**	0.3040	0.7960	0.9860
	FE-QML (Robust)	0.9530	0.7640	0.2980	**0.0730**	0.3160	0.7810	0.9590
N = 250, T = 3	*FE-QML (Normal)*	0.8530	0.5880	0.2460	**0.0910**	0.2420	0.6050	0.8590
	FE-QML (Robust)	0.6570	0.4170	0.1660	**0.0590**	0.1490	0.4320	0.6870
N = 250, T = 10	*FE-QML (Normal)*	0.9940	0.9930	0.8680	**0.0470**	0.8550	0.9930	0.9940
	FE-QML (Robust)	0.9890	0.9740	0.7760	**0.0490**	0.7610	0.9760	0.9900
	FE-MD	0.8790	0.6030	0.2260	**0.0580**	0.2040	0.5690	0.8670

[18] See the footnote to Table 4a for a description of the data generating process used in this table, and the footnote to Table 1 for a description of the notation.
[19] The table reports the fraction of rejections for tests of H_0: $\phi_{21} = \{-0.1, 0, 0.1, 0.2, 0.3, 0.4, 0.5\}$, versus two-sided alternatives. The true value of ϕ_{21} is equal to *0.2*.

Table 5a: Bias and RMSE of Random and Fixed Effects QML Estimators of Panel VAR in the Case of Random Individual Effects[20], $\lambda_{max} = 0.6$

	Estimator										
	True Value		$\phi_{11} = 0.4$		$\phi_{21} = 0.2$		$\phi_{12} = 0.2$		$\phi_{22} = 0.4$	
			Bias	*RMSE*	*Bias*	*RMSE*	*Bias*	*RMSE*	*Bias*	*RMSE*
N = 50, T = 3	*RE-QML*	$\tau = 5$	0.0036	0.1971	0.0032	0.1815	0.0059	0.1697	0.0049	0.1987
	FE-QML		0.0033	0.1971	0.0027	0.1829	0.0044	0.1698	0.0023	0.1986
N = 50, T = 10	*RE-QML*	$\tau = 5$	0.0023	0.0737	0.0005	0.0707	0.0006	0.0669	0.0040	0.0716
	FE-QML		0.0023	0.0737	0.0005	0.0706	0.0006	0.0669	0.0040	0.0717
N = 250, T = 3	*RE-QML*	$\tau = 5$	0.0007	0.0884	0.0013	0.0800	0.0024	0.0794	0.0016	0.0895
	FE-QML		0.0003	0.0898	0.0008	0.0806	0.0022	0.0809	0.0015	0.0898
N = 250, T = 10	*RE-QML*	$\tau = 5$	0.0028	0.0327	0.0019	0.0303	-0.0012	0.0314	-0.0009	0.0304
	FE-QML		0.0027	0.0327	0.0019	0.0303	-0.0012	0.0314	-0.0010	0.0304

[20] For details of the Monte Carlo design, see Section 7.1. The data generating process is given by $(\boldsymbol{I}_2 - \boldsymbol{\Phi}L)(\boldsymbol{w}_{it} - \boldsymbol{\mu}_i) = \boldsymbol{\varepsilon}_{it}$. See the footnote to Table 1 for a description of the notation used in this table. See Section 7.2 for an explanation of any discrepancies between results for the FE-QML estimator reported in this table and those reported in Table 1.

Table 5b: Size and Power Properties of Tests for ϕ_{11} Under Random and Fixed Effects QML Estimators of Panel VAR in the Case of Random Individual Effects[21,22], λ_{max} = 0.6

	Estimator		$\phi_{11} = 0.1$	$\phi_{11} = 0.2$	$\phi_{11} = 0.3$	$\boldsymbol{\phi_{11} = 0.4}$	$\phi_{11} = 0.5$	$\phi_{11} = 0.6$	$\phi_{11} = 0.7$
N = 50, T = 3	*RE-QML (Normal)*	$\tau = 5$	0.3060	0.1530	0.0740	**0.0630**	0.1070	0.1910	0.3330
	FE-QML (Normal)		0.3020	0.1430	0.0640	**0.0600**	0.0980	0.1950	0.3350
	RE-QML (Robust)	$\tau = 5$	0.3010	0.1700	0.0850	**0.0750**	0.1140	0.1900	0.3090
	FE-QML (Robust)		0.3000	0.1700	0.0860	**0.0730**	0.1170	0.2080	0.3450
N = 50, T = 10	*RE-QML (Normal)*	$\tau = 5$	0.9780	0.7880	0.2810	**0.0670**	0.2690	0.8010	0.9820
	FE-QML (Normal)		0.9760	0.7940	0.2750	**0.0680**	0.2850	0.7850	0.9840
	RE-QML (Robust)	$\tau = 5$	0.9540	0.7520	0.2830	**0.0700**	0.2740	0.7850	0.9740
	FE-QML (Robust)		0.9640	0.7790	0.2770	**0.0870**	0.2860	0.7690	0.9680
N = 250, T = 3	*RE-QML (Normal)*	$\tau = 5$	0.8590	0.5410	0.1860	**0.0540**	0.2070	0.5640	0.8330
	FE-QML (Normal)		0.8410	0.5290	0.1670	**0.0410**	0.1900	0.5340	0.8290
	RE-QML (Robust)	$\tau = 5$	0.7540	0.4610	0.1680	**0.0520**	0.1850	0.4630	0.7480
	FE-QML (Robust)		0.7320	0.4510	0.1640	**0.0540**	0.1690	0.4710	0.7290
N = 250, T = 10	*RE-QML (Normal)*	$\tau = 5$	1.0000	1.0000	0.8460	**0.0480**	0.8500	1.0000	1.0000
	FE-QML (Normal)		0.9980	0.9970	0.8080	**0.0520**	0.8480	0.9950	0.9980
	RE-QML (Robust)	$\tau = 5$	1.0000	1.0000	0.8070	**0.0480**	0.8500	1.0000	1.0000
	FE-QML (Robust)		0.9890	0.9800	0.7390	**0.0610**	0.7770	0.9750	0.9920

[21] See the footnote to Table 5a for a description of the data generating process and the notation used in this table. Also, see Section 7.2 for an explanation of any discrepancies between results for the FE-QML estimator reported in this table and those reported in Table 2a.

[22] The table reports the fraction of rejections for tests of H_0: $\phi_{11} = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7\}$, versus two-sided alternatives. The true value of ϕ_{11} is equal to *0.4*.

Table 5c : Size and Power Properties of Tests for ϕ_{21} Under Random and Fixed Effects QML Estimators of Panel VAR in the Case of Random Individual Effects[23,24] , $\lambda_{max} = 0.6$

	Estimator		$\phi_{21} = -0.1$	$\phi_{21} = 0$	$\phi_{21} = 0.1$	$\boldsymbol{\phi_{21} = 0.2}$	$\phi_{21} = 0.3$	$\phi_{21} = 0.4$	$\phi_{21} = 0.5$
N = 50, T = 3	*RE-QML (Normal)*	$\tau = 5$	0.3760	0.2100	0.1010	**0.0670**	0.1000	0.2080	0.3740
	FE-QML (Normal)		0.3960	0.2130	0.0900	**0.0670**	0.0990	0.2080	0.3740
	RE-QML (Robust)	$\tau = 5$	0.3600	0.2150	0.0960	**0.0730**	0.0970	0.2040	0.3550
	FE-QML (Robust)		0.4010	0.2180	0.0900	**0.0700**	0.1110	0.2090	0.3830
N = 50, T = 10	*RE-QML (Normal)*	$\tau = 5$	0.9890	0.8040	0.2890	**0.0580**	0.3080	0.8180	0.9820
	FE-QML (Normal)		0.9700	0.7940	0.2970	**0.0670**	0.3040	0.7930	0.9720
	RE-QML (Robust)	$\tau = 5$	0.9920	0.8150	0.3010	**0.0610**	0.3160	0.8220	0.9850
	FE-QML (Robust)		0.9780	0.7970	0.3130	**0.0770**	0.3130	0.7880	0.9750
N = 250, T = 3	*RE-QML (Normal)*	$\tau = 5$	0.8880	0.6300	0.2300	**0.0540**	0.2150	0.6420	0.9050
	FE-QML (Normal)		0.8890	0.6010	0.2220	**0.0460**	0.2060	0.6370	0.8990
	RE-QML (Robust)	$\tau = 5$	0.7910	0.5370	0.2160	**0.0510**	0.1950	0.5490	0.7940
	FE-QML (Robust)		0.7840	0.5290	0.2020	**0.0530**	0.1740	0.5630	0.8040
N = 250, T = 10	*RE-QML (Normal)*	$\tau = 5$	1.0000	1.0000	0.8900	**0.0520**	0.9140	1.0000	1.0000
	FE-QML (Normal)		0.9980	0.9970	0.8500	**0.0440**	0.8770	0.9950	0.9990
	RE-QML (Robust)	$\tau = 5$	1.0000	1.0000	0.8660	**0.0460**	0.8770	1.0000	1.0000
	FE-QML (Robust)		0.9910	0.9810	0.7760	**0.0490**	0.8050	0.9820	0.9910

[23] See the footnote to Table 5a for a description of the data generating process and the notation used in this table. Also, see Section 7.2 for an explanation of any discrepancies between results for the FE-QML estimator reported in this table and those reported in Table 2b.

[24] The table reports the fraction of rejections for tests of H_0: $\phi_{21} = \{-0.1, 0, 0.1, 0.2, 0.3, 0.4, 0.5\}$, versus two-sided alternatives. The true value of ϕ_{21} is equal to *0.2*.

Table 5d: Size and Power Properties of Tests for ϕ_{12} Under Random and Fixed Effects QML Estimators of Panel VAR in the Case of Random Individual Effects[25,26], $\lambda_{max} = 0.6$

	Estimator		$\phi_{12} = -0.1$	$\phi_{12} = 0$	$\phi_{12} = 0.1$	$\phi_{12} = 0.2$	$\phi_{12} = 0.3$	$\phi_{12} = 0.4$	$\phi_{12} = 0.5$
$N = 50, T = 3$	RE-QML (Normal)	$\tau = 5$	0.3480	0.1960	0.0850	**0.0520**	0.0800	0.1990	0.3920
	FE-QML (Normal)		0.3500	0.2020	0.0850	**0.0470**	0.0710	0.1930	0.3930
	RE-QML (Robust)	$\tau = 5$	0.3300	0.1900	0.0910	**0.0580**	0.0950	0.2130	0.3660
	FE-QML (Robust)		0.3400	0.1960	0.0940	**0.0620**	0.0950	0.2110	0.3670
$N = 50, T = 10$	RE-QML (Normal)	$\tau = 5$	0.9930	0.8130	0.2780	**0.0420**	0.2970	0.8260	0.9900
	FE-QML (Normal)		0.9940	0.8150	0.2810	**0.0450**	0.3030	0.8300	0.9860
	RE-QML (Robust)	$\tau = 5$	0.9790	0.7880	0.2690	**0.0500**	0.3030	0.7930	0.9730
	FE-QML (Robust)		0.9780	0.8030	0.2790	**0.0510**	0.3080	0.8080	0.9720
$N = 250, T = 3$	RE-QML (Normal)	$\tau = 5$	0.8880	0.6260	0.2250	**0.0500**	0.2110	0.6460	0.9010
	FE-QML (Normal)		0.8850	0.6040	0.2080	**0.0550**	0.2190	0.6250	0.9000
	RE-QML (Robust)	$\tau = 5$	0.7910	0.5600	0.2100	**0.0530**	0.1910	0.5580	0.8070
	FE-QML (Robust)		0.7840	0.5390	0.1970	**0.0590**	0.1820	0.5460	0.7970
$N = 250, T = 10$	RE-QML (Normal)	$\tau = 5$	1.0000	1.0000	0.8890	**0.0530**	0.8740	1.0000	1.0000
	FE-QML (Normal)		0.9980	0.9960	0.8630	**0.0470**	0.8500	0.9970	0.9990
	RE-QML (Robust)	$\tau = 5$	1.0000	1.0000	0.8640	**0.0510**	0.8430	1.0000	1.0000
	FE-QML (Robust)		0.9890	0.9790	0.7920	**0.0520**	0.7770	0.9790	0.9900

Table 5e: Size and Power Properties of Tests for ϕ_{22} Under Random and Fixed Effects QML Estimators of Panel VAR in the Case of Random Individual Effects[27,28], $\lambda_{max} = 0.6$

	Estimator		$\phi_{22} = 0.1$	$\phi_{22} = 0.2$	$\phi_{22} = 0.3$	$\phi_{22} = 0.4$	$\phi_{22} = 0.5$	$\phi_{22} = 0.6$	$\phi_{22} = 0.7$
$N = 50, T = 3$	RE-QML (Normal)	$\tau = 5$	0.2920	0.1460	0.0800	**0.0540**	0.1020	0.1970	0.3570
	FE-QML (Normal)		0.2870	0.1440	0.0750	**0.0530**	0.0990	0.1980	0.3560
	RE-QML (Robust)	$\tau = 5$	0.2950	0.1690	0.0950	**0.0780**	0.1090	0.1950	0.3430
	FE-QML (Robust)		0.3030	0.1700	0.0900	**0.0760**	0.1140	0.2210	0.3590
$N = 50, T = 10$	RE-QML (Normal)	$\tau = 5$	0.9790	0.7760	0.2500	**0.0460**	0.3040	0.7940	0.9870
	FE-QML (Normal)		0.9820	0.7680	0.2610	**0.0440**	0.2940	0.7970	0.9850
	RE-QML (Robust)	$\tau = 5$	0.9640	0.7530	0.2530	**0.0560**	0.3030	0.7720	0.9680
	FE-QML (Robust)		0.9640	0.7530	0.2750	**0.0590**	0.3080	0.7730	0.9750
$N = 250, T = 3$	RE-QML (Normal)	$\tau = 5$	0.8360	0.5360	0.1910	**0.0460**	0.1890	0.5600	0.8320
	FE-QML (Normal)		0.8410	0.5240	0.1760	**0.0480**	0.1960	0.5420	0.8350
	RE-QML (Robust)	$\tau = 5$	0.7260	0.4790	0.1880	**0.0480**	0.1710	0.4680	0.7290
	FE-QML (Robust)		0.7240	0.4600	0.1590	**0.0570**	0.2060	0.4860	0.7250
$N = 250, T = 10$	RE-QML (Normal)	$\tau = 5$	1.0000	1.0000	0.8860	**0.0360**	0.8740	1.0000	1.0000
	FE-QML (Normal)		0.9980	0.9950	0.8420	**0.0330**	0.8390	0.9950	0.9990
	RE-QML (Robust)	$\tau = 5$	1.0000	1.0000	0.8530	**0.0400**	0.8410	1.0000	1.0000
	FE-QML (Robust)		0.9900	0.9770	0.7770	**0.0400**	0.7690	0.9760	0.9870

[25] See the footnote to Table 5a for a description of the data generating process and the notation used in this table.
[26] The table reports the fraction of rejections for tests of H_0: $\phi_{12} = \{-0.1, 0, 0.1, 0.2, 0.3, 0.4, 0.5\}$, versus two-sided alternatives. The true value of ϕ_{12} is equal to *0.2*.
[27] See the footnote to Table 5a for a description of the data generating process and the notation used in this table.
[28] The table reports the fraction of rejections for tests of H_0: $\phi_{22} = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7\}$, versus two-sided alternatives. The true value of ϕ_{22} is equal to *0.4*.