Econometric Theory, **22**, 2006, 499–512. Printed in the United States of America.
DOI: 10.1017/S0266466606060245

REDUCING BIAS OF MLE IN A DYNAMIC PANEL MODEL

JINYONG HAHN
UCLA

HYUNGSIK ROGER MOON
University of Southern California

This paper investigates a simple dynamic linear panel regression model with both fixed effects and time effects. Using "large n and large T" asymptotics, we approximate the distribution of the fixed effect estimator of the autoregressive parameter in the dynamic linear panel model and derive its asymptotic bias. We find that the same higher order bias correction approach proposed by Hahn and Kuersteiner (2002, *Econometrica* 70, 1639–1659) can be applied to the dynamic linear panel model even when time specific effects are present.

1. INTRODUCTION

One of the advantages of panel data is that they allow the possibility of controlling for unobserved individual heterogeneity. Failure to control for such heterogeneity can result in misleading inferences. Although it is intuitive to deal with the unobserved individual effect by treating each such effect as a separate parameter to be estimated, such estimators are typically subject to the incidental parameters problem noted by Neyman and Scott (1948). In a simple dynamic linear panel regression model, the fixed effect estimator of the autoregressive coefficient is severely biased when the cross-sectional dimension is large but the time series dimension is small.[1] See Nickell (1981), Kiviet (1995), Alvarez and Arellano (2003), and Phillips and Sul (2003b). Phillips and Sul (2003a) investigate the median unbiased estimation method for various dynamic linear panel regression models.

Adopting a perspective that such bias can be understood as a higher order time series bias, Hahn and Kuersteiner (2002, 2003) propose a method that reduces the bias of the fixed effects estimator. Using alternative asymptotics where both n and T are large, they establish that a simpler form of the higher order bias can be derived. The alternative asymptotics, where both n and T grow

We thank Peter Phillips and three anonymous referees for helpful comments. The first author gratefully acknowledges financial support from NSF grant SES-0313651. The second author appreciates the Faculty Development Awards of USC for research support. Address correspondence to Hyungsik Roger Moon, Department of Economics, KAP 300, University Park Campus, University of Southern California, Los Angeles, CA 90089, USA; e-mail: moonr@usc.edu

 0266-4666/06 $12.00

to infinity,[2] can be quite convenient, especially for nonlinear panel models. See Hahn and Newey (2004) and Hahn and Kuersteiner (2003).

The panel models considered by Hahn and Kuersteiner (2003) and Hahn and Newey (2004) do not include any *time effect*, mainly because of analytical difficulties in nonlinear models. As argued in Hahn and Kuersteiner (2003), the alternative asymptotics can be understood as a simpler form of higher order time series asymptotics with fixed n. When time effects are not present, higher order time series asymptotics is straightforward, although tedious. Unfortunately, time effects create an incidental parameter problem in the time series domain, because the number of time effects grows to infinity as $T \to \infty$, which explains the analytic difficulty there.[3]

In this paper, we make a contribution to understanding the bias of fixed effects estimators in models with time effects. We establish the asymptotic distribution of the fixed effect estimator (or Gaussian quasi maximum likelihood estimator [QMLE]) of the autoregressive parameter when both n and T are large. In particular, we derive the asymptotic bias of the fixed effect estimator and propose an estimator that corrects for the asymptotic bias. We find that the asymptotic bias is the same as the one in the panel model of Hahn and Kuersteiner (2002) without the time effect. It follows that the same higher order bias correction approach as in Hahn and Kuersteiner (2002) can be adopted even when time effects are present. We should stress that such robustness is limited only to linear models. For more general models, we expect that estimation of time effects would lead to biases in addition to the biases due to estimation of individual effects. Because these two biases need to be analyzed simultaneously, it is not trivial to extend the analysis of Hahn and Kuersteiner (2003) or Hahn and Newey (2004) to nonlinear models with both time effects and individual effects.

This paper is organized as follows. In Section 2 we introduce a linear dynamic panel regression model with both individual effects and time effects and assumptions. The main results are summarized in Theorems 1 and 2. Section 3 concludes the paper. All the technical proofs and derivations are collected in the Appendix.

2. MAIN RESULT

We consider a simple dynamic panel regression model with fixed individual effects and time specific effects,

$$y_{it} = \theta y_{it-1} + (I_m - \theta)\alpha_i + f_t + \varepsilon_{it} \qquad (i = 1, \dots, n; t = 1, \dots, T), \tag{1}$$

where y_{it} are m-dimensional observables, ε_{it} are mean zero scalar error terms, θ is an $m \times m$ matrix of parameters of interest, i denotes the cross-sectional unit, and t denotes the time index.[4] We denote n and T to be the dimensions of cross section and time series, respectively, of the panel. In model (1), the parameter α_i ($m \times 1$) signifies fixed individual effects, and f_t ($m \times 1$) represents time

https://doi.org/10.1017/S0266466606060245 Published online by Cambridge University Press

specific effects. The dynamic panel model in (1) extends the conventional dynamic panel model with fixed effects,

$$y_{it} = \theta y_{it-1} + (I_m - \theta)\alpha_i + \varepsilon_{it}, \tag{2}$$

where the dynamics of the panel data y_{it} does not include time effects. In many empirical applications, the time effect f_t is included to model a simple form of nonstationarity in the time series of y_{it} or to represent an aggregate shock (e.g., a common macro shock) that is common to all the cross-section units. In the latter case, when the common shock f_t is random, the cross-sectional observations y_{it} have cross-sectional dependence.[5]

Before we proceed, we introduce a set of regularity conditions that will be used in deriving the main results in the following section. These conditions are the same as Conditions 1–3 in Hahn and Kuersteiner (2002).

Condition 1. (i) ε_{it} is independent and identically distributed (i.i.d.) across i and strictly stationary in t for each i, $E[\varepsilon_{it}] = 0$ for all i and t, $E[\varepsilon_{it}\varepsilon_{is}'] = \Omega$ for $t = s$ and $E[\varepsilon_{it}\varepsilon_{is}'] = 0$ for $t \neq s$, and has finite eighth moments; (ii) both n and T tend to infinity jointly under the restriction $n/T \to c$, where $0 < c < \infty$; (iii) $\lim_{n\to\infty}\theta^n = 0$; (iv) $(1/n)\sum_{i=1}^n \|\alpha_i\|^2 = O(1)$; (v) $(1/n)\sum_{i=1}^n \|y_{i0}\|^2 = O(1)$.

The individual effect α_i and the initial observations y_{i0} are assumed to be deterministic sequences. It is in principle possible to treat α_i and y_{i0} as random, but we can avoid specifying their joint distribution by focusing our attention on the distribution of y's conditional on α_i and y_{i0}. Therefore, the distribution of the y's is in fact a conditional distribution. The time specific effect can be either a deterministic or a random sequence. When it is random, it does not have to be stationary. The fixed effects and the initial conditions are deterministic. Condition 1(ii) means that we adopt the "large n, T" asymptotics. Finally, notice that Condition 1(iii) excludes a possibility of unit roots in the panel. Our analysis fails to carry over to the case when the largest characteristic root of θ is one, which suggests that our approximation may not be accurate when y_{it} has the largest root near unity, which was confirmed by the Monte Carlo study in Hahn and Kuersteiner (2002) for the simpler model without time effects. For a nonstationary dynamic panel model, see Moon and Phillips (2004), Moon and Perron (2004), and Phillips and Sul (2003a).

The next two conditions restrict the higher order serial dependence of the error term ε_{it} and its moments. For this, define $u_{it}^* \equiv \sum_{j=0}^{\infty}\theta^j\varepsilon_{it-j}$, which is well defined under Conditions 1(i) and (iii). Also, define $z_{it} = (I_m \otimes u_{it-1}^*)\varepsilon_{it}$.

Condition 2. (i) $\sum_{t_1,t_2,t_3=-\infty}^{\infty}|\mathrm{cum}_{j_1,\ldots,j_4}(u_{it_1}^*, \varepsilon_{it_2}, u_{it_3}^*, \varepsilon_{i0})| < \infty$, and (ii) $\sum_{t_1,t_2,t_3=-\infty}^{\infty}|\mathrm{cum}_{j_1,\ldots,j_4}(z_{it_1}, z_{it_2}, z_{it_3}, z_{i0})| < \infty$, for all i and $j_1,\ldots,j_4 \in \{1,\ldots,m\}$, where $\mathrm{cum}_{j_1,\ldots,j_4}(\cdot,\cdot,\cdot,\cdot)$ is defined as in Brillinger (1981).

The estimator we consider in this paper is a fixed effect estimator. Let $\bar{y}_{\cdot,t} = (1/n)\sum_{i=1}^n y_{it}$, $\bar{y}_{i,\cdot} = (1/T)\sum_{t=1}^T y_{it}$, $\bar{y} = (1/nT)\sum_{i=1}^n\sum_{t=1}^T y_{it}$, $\bar{y}_{i,-1} =$

https://doi.org/10.1017/S0266466606060245 Published online by Cambridge University Press

$(1/T)\sum_{t=1}^{T} y_{it-1}$, $\bar{y}_{-1} = (1/nT)\sum_{i=1}^{n}\sum_{t=1}^{T} y_{it-1}$. Similarly, we define $\bar{\alpha}$, $\bar{\varepsilon}_{\cdot,t}$, $\bar{\varepsilon}_{i,\cdot}$, and $\bar{\varepsilon}$. For notational simplicity, write $\tilde{y}_{it} = y_{it} - \bar{y}_{\cdot,t} - \bar{y}_{i,\cdot} + \bar{y}$, $\tilde{y}_{it-1,-1} = y_{it-1} - \bar{y}_{\cdot,t-1} - \bar{y}_{i,\cdot,-1} + \bar{y}_{-1}$, and $\tilde{\varepsilon}_{it} = \varepsilon_{it} - \bar{\varepsilon}_{\cdot,t} - \bar{\varepsilon}_{i,\cdot} + \bar{\varepsilon}$. The estimator is defined as

$$\hat{\theta}' = \left(\sum_{i=1}^{n}\sum_{t=1}^{T} \tilde{y}_{it-1,-1}\tilde{y}'_{it-1,-1}\right)^{-1}\left(\sum_{i=1}^{n}\sum_{t=1}^{T} \tilde{y}_{it-1,-1}\tilde{y}'_{it}\right).$$

It can be shown that $\hat{\theta}$ is the Gaussian maximum likelihood estimator (MLE) (see, e.g., Hsiao, 2003). The main purposes of the paper are (i) to find an asymptotic bias of the fixed effect estimator $\hat{\theta}$ as $n, T \to \infty$ with $n/T \to c$, where $0 < c < \infty$ and (ii) to consider an estimator that corrects for the asymptotic bias.

Because $\tilde{y}'_{it} = \tilde{y}'_{it-1,-1}\theta'_0 + \tilde{\varepsilon}'_{it}$ by definition, we have

$$\sqrt{nT}(\hat{\theta}' - \theta') = \left(\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} \tilde{y}_{it-1,-1}\tilde{y}'_{it-1,-1}\right)^{-1}\left(\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T} \tilde{y}_{it-1,-1}\tilde{\varepsilon}'_{it}\right).$$

The following theorem finds the limiting distribution of $\sqrt{nT}(\hat{\theta} - \theta)$.

THEOREM 1. *Suppose that Conditions 1 and 2 hold. Then,*

$$\sqrt{nT}\,\mathrm{vec}(\hat{\theta}' - \theta') \Rightarrow \mathcal{N}\big(-\sqrt{c}(I_m \otimes \Upsilon)^{-1}(I_m \otimes I_m - (I_m \otimes \theta))^{-1}\,\mathrm{vec}(\Omega), (I_m \otimes \Upsilon)^{-1} \times (\Omega \otimes \Upsilon + \mathcal{K})(I_m \otimes \Upsilon)^{-1}\big),$$

where

$$\Upsilon = \Omega + \theta\Omega\theta' + \theta^2\Omega(\theta')^2 + \ldots, \qquad \mathcal{K} = \sum_{t=-\infty}^{\infty} \mathcal{K}(t,0),$$

$$\mathcal{K}(t_1, t_2) = E[z_{it_1} z'_{it_2}] - E[\varepsilon_{it_1}\varepsilon'_{it_2}] \otimes E[u^*_{i0} u^{*\prime}_{i0}].$$

According to Theorem 1, as $n, T \to \infty$ with $n/T \to c$, the fixed effects estimator $\hat{\theta}$ has a normal limiting distribution with an asymptotic bias $-(1/T)(I_m \otimes \Upsilon)^{-1}(I_m \otimes I_m - (I_m \otimes \theta))^{-1}\,\mathrm{vec}(\Omega)$. This bias is the same as the bias found by Hahn and Kuersteiner (2002) with the linear dynamic panel regression with fixed effects in (2). Unlike the conventional model in (2), our model (1) assumes incidental parameters in both cross-section and time series. Theorem 1 shows that the incidental parameters in the time series, f_t, do not contribute to the asymptotic bias of the fixed effect estimator and it is the α_i, the cross-sectional incidental parameters, that cause the asymptotic bias.

To understand the different roles of the two incidental parameters, it is useful to consider a simple case where y_{it} is univariate and ε_{it} are i.i.d. First, notice

https://doi.org/10.1017/S0266466606060245 Published online by Cambridge University Press

that QMLE estimation eliminates the individual effect α_i through the following time series filtering:

$$y_{it} - \bar{y}_{i,\cdot} = \theta(y_{it-1} - \bar{y}_{i,\cdot,-1}) + \left(f_t - T^{-1}\sum_{s=1}^{T} f_s\right) + (\varepsilon_{it} - \bar{\varepsilon}_{i,\cdot}) \tag{3}$$

and then eliminates the time effect f_t through the following cross-sectional filtering:[6]

$$\begin{aligned} y_{it} - \bar{y}_{i,\cdot} - (\bar{y}_{\cdot,t} - \bar{y}) &= \theta(y_{it-1} - \bar{y}_{i,\cdot,-1} - (\bar{y}_{\cdot,t-1} - \bar{y}_{-1})) \\ &\quad + (\varepsilon_{it} - \bar{\varepsilon}_{i,\cdot} - (\bar{\varepsilon}_{\cdot,t} - \bar{\varepsilon})). \end{aligned} \tag{4}$$

The covariance between the filtered regressor $y_{it-1} - \bar{y}_{i,\cdot,-1} - (\bar{y}_{\cdot,t-1} - \bar{y}_{-1})$ and the filtered error term $\varepsilon_{it} - \bar{\varepsilon}_{i,\cdot} - (\bar{\varepsilon}_{\cdot,t} - \bar{\varepsilon})$ can be shown[7] to consist of the following four covariances: (i) the correlation between y_{it-1} and ε_{it}, (ii) the correlation between $\bar{y}_{i,\cdot,-1}$ and $\bar{\varepsilon}_{i,\cdot}$, (iii) the correlation between $\bar{y}_{\cdot,t-1}$ and $\bar{\varepsilon}_{\cdot,t}$, and (iv) the correlation between $\bar{y}_{-1}$ and $\bar{\varepsilon}$. The second correlation is generated by the time series filtering in (3), and the third and fourth correlations are due to the cross-sectional filtering eliminating f_t in (4). Now, because of the weak exogeneity of y_{it-1}, it is easy to see that the first correlation is zero. Second, according to Hahn and Kuersteiner (2002), the second correlation times $\sqrt{nT}$, the convergence rate of the fixed effects estimator, does not vanish even though $T \to \infty$ and remains as a bias in the limit distribution, if $n, T \to \infty$ with $n/T \to c$, where $0 < c < \infty$. As for the third correlation, we observe that because the time series of y_{it-1} is weakly exogeneous and the cross section is independent, the time series of the cross section aggregate $\bar{y}_{\cdot,t-1}$ is also weakly exogeneous with respect to the time series of the cross section aggregate $\bar{\varepsilon}_{\cdot,t}$. Therefore, we expect that the third correlation is zero. Finally, we expect that the fourth correlation is negligible in large n and T samples. As a consequence, the additional filtering in (4) to eliminate the time effect does not have the same effect of the time series filtering in (3), which is why the asymptotic bias in Theorem 1 is identical to that in Hahn and Kuersteiner (2002) where only α_i are assumed present.

In view of the limiting distribution of $\sqrt{nT}\,\mathrm{vec}(\hat{\theta}' - \theta')$ in Theorem 1, to fix the asymptotic bias in $\sqrt{nT}\,\mathrm{vec}(\hat{\theta}' - \theta')$, we can use the same bias correction formula as in Hahn and Kuersteiner (2002). Define the following bias corrected estimator:

$$\begin{aligned} \mathrm{vec}(\tilde{\theta}') &= \left[I_m \otimes \left(\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{y}_{it-1,-1} \tilde{y}_{it-1,-1}' \right)^{-1} \right] \\ &\quad \times \left[\begin{array}{l} \dfrac{1}{nT} \displaystyle\sum_{i=1}^{n} \sum_{t=1}^{T} (I_m \otimes \tilde{y}_{it-1,-1}) \tilde{y}_{it}' \\ \qquad + \dfrac{1}{T} ((I_m \otimes I_m) - (I_m \otimes \hat{\theta}))^{-1} \mathrm{vec}(\hat{\Omega}) \end{array} \right], \end{aligned}$$

https://doi.org/10.1017/S0266466606060245 Published online by Cambridge University Press

where

$$\hat{\Upsilon} = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{y}_{it-1,-1} \tilde{y}'_{it-1,-1},$$

$$\text{vec}(\hat{\Omega}) = ((I_m \otimes I_m) - (\hat{\theta} \otimes \hat{\theta}))\text{vec}(\hat{\Upsilon}).$$

We can easily see that the bias corrected estimator $\tilde{\theta}$ is asymptotically centered at zero:[8]

THEOREM 2. *Under Conditions 1 and 2, we have*

$$\sqrt{nT}\,\text{vec}(\tilde{\theta}' - \theta') \Rightarrow \mathcal{N}(0, (I_m \otimes \Upsilon)^{-1}(\Omega \otimes \Upsilon + \mathcal{K})(I_m \otimes \Upsilon)^{-1}).$$

To assess the effectiveness of bias correction as discussed in Theorem 2, we conducted a small-scale Monte Carlo study for the simple case when y_{it} is a scalar and the error term ε_{it} is i.i.d. with constant variance, which is summarized in Table 1. Note that the limiting distribution of $\sqrt{nT}(\hat{\theta} - \theta)$ in Theorem 1 simplifies to $\mathcal{N}(-\sqrt{c}(1+\theta), 1-\theta^2)$. For this simple situation, we may use the bias correction formula

$$\tilde{\theta} = \hat{\theta} + \frac{1}{T}(1 + \hat{\theta})$$

as in Hahn and Kuersteiner (2002). This estimator may be intuitively understood by observing that the bias of $\sqrt{nT}(\hat{\theta} - \theta)$ is approximately equal to $-\sqrt{n/T}(1+\theta)$, which may be estimated by $-\sqrt{n/T}(1+\hat{\theta})$, and as a consequence, the asymptotic distribution of $\sqrt{nT}(\hat{\theta} - \theta) - (-\sqrt{n/T}(1+\hat{\theta})) = \sqrt{nT}(\hat{\theta} + (1/T)(1+\hat{\theta}) - \theta)$ is centered at zero. For this bias corrected estimator, we find that the performance is almost identical whether the time effect is present in the model or not.

3. CONCLUSION

This paper investigates a simple dynamic linear panel regression model with both fixed effects and time effects. Using large n and T asymptotics, we approximate the distribution of the fixed effect estimator of the autoregressive parameter in the dynamic linear panel model and derive its asymptotic bias. As main results, we find that the asymptotic bias is the same as the one in the panel model of Hahn and Kuersteiner (2002) without the time effect, and we show that the same higher order bias correction approach proposed by Hahn and Kuersteiner (2002) can be applied to the dynamic linear panel model with both fixed effects and time effects. However, as mentioned in the introduction, we stress that the robustness of the bias correction approach of Hahn and Kuersteiner (2002) to the time effects model is limited only to a linear model so far.

TABLE 1. Performance of bias corrected maximum likelihood estimators

			$y_{it} = \theta y_{it-1} + \alpha_i(1-\theta) + \varepsilon_{it}$				$y_{it} = \theta y_{it-1} + \alpha_i(1-\theta) + f_t + \varepsilon_{it}$			
			Mean bias		RMSE		Mean bias		RMSE	
θ	T	N	*MLE*	*BCMLE*	*MLE*	*BCMLE*	*MLE*	*BCMLE*	*MLE*	*BCMLE*
0	5	100	−0.200183	−0.040219	0.205075	0.066876	−0.200264	−0.040317	0.205199	0.067136
0.3	5	100	−0.275611	−0.070733	0.279555	0.090312	−0.275699	−0.070838	0.279680	0.090566
0.6	5	100	−0.363652	−0.116382	0.366854	0.130049	−0.363722	−0.116467	0.366954	0.130247
0.9	5	100	−0.465102	−0.178123	0.467710	0.187699	−0.465128	−0.178154	0.467752	0.187784
0	5	200	−0.201328	−0.041594	0.203771	0.056170	−0.201271	−0.041525	0.203727	0.056184
0.3	5	200	−0.276146	−0.071375	0.278150	0.081819	−0.276084	−0.071301	0.278098	0.081801
0.6	5	200	−0.363556	−0.116267	0.365196	0.123444	−0.363508	−0.116210	0.365156	0.123425
0.9	5	200	−0.464653	−0.177584	0.465986	0.182544	−0.464662	−0.177595	0.466004	0.182588
0	10	100	−0.099987	−0.009985	0.104884	0.036244	−0.099999	−0.009999	0.104949	0.036435
0.3	10	100	−0.135512	−0.019064	0.139248	0.040068	−0.135521	−0.019073	0.139301	0.040260
0.6	10	100	−0.179708	−0.037679	0.182324	0.050648	−0.179720	−0.037692	0.182371	0.050816
0.9	10	100	−0.245161	−0.079677	0.246808	0.085608	−0.245230	−0.079753	0.246901	0.085765
0	10	200	−0.100514	−0.010566	0.103072	0.027235	−0.100536	−0.010589	0.103109	0.027318
0.3	10	200	−0.135533	−0.019087	0.137504	0.031865	−0.135560	−0.019116	0.137542	0.031945
0.6	10	200	−0.179086	−0.036994	0.180467	0.044381	−0.179113	−0.037024	0.180502	0.044443
0.9	10	200	−0.244303	−0.078733	0.245162	0.081901	−0.244330	−0.078763	0.245192	0.081940
0	20	100	−0.049880	−0.002373	0.054840	0.024050	−0.049902	−0.002397	0.054928	0.024222
0.3	20	100	−0.066671	−0.005004	0.070375	0.024183	−0.066697	−0.005032	0.070455	0.024364
0.6	20	100	−0.086287	−0.010602	0.088688	0.023990	−0.086325	−0.010641	0.088764	0.024169
0.9	20	100	−0.120618	−0.031649	0.121735	0.036058	−0.120669	−0.031702	0.121799	0.036154
0	20	200	−0.050403	−0.002923	0.052881	0.017051	−0.050396	−0.002916	0.052884	0.017080
0.3	20	200	−0.066847	−0.005190	0.068707	0.017459	−0.066844	−0.005186	0.068711	0.017490
0.6	20	200	−0.086115	−0.010420	0.087326	0.018446	−0.086116	−0.010422	0.087333	0.018478
0.9	20	200	−0.120352	−0.031370	0.120916	0.033674	−0.120357	−0.031375	0.120924	0.033693

Note: Simulations are based on 5,000 replications. The fixed effects α_i, the time effects f_t, and the innovations ε_{it} are assumed to have independent standard normal distributions. The initial observations y_{i0} are assumed to be generated by the stationary distribution $\mathcal{N}(\alpha_i, 1/(1-\theta^2))$. The BCMLE is the bias corrected MLE and is equal to $MLE + (1 + MLE)/MLE$. The RMSE is the root mean square error.

https://doi.org/10.1017/S0266466606060245 Published online by Cambridge University Press

NOTES

1. Various alternative methods, including the generalized method of moments (GMM) approach with first differenced data, have been proposed. For a survey on recent developments of these methods, one can refer to Arellano and Honoré (2001).

2. In other words, the alternative asymptotics are joint asymptotics, not sequential asymptotics. For discussion on the difference between the two, see, for example, Phillips and Moon (1999).

3. The GMM approach based on fixed T assumption can easily handle models with time effects. See, for example, Holtz-Eakin, Newey, and Rosen (1988), and Ahn, Lee, and Schmidt (2001). These papers, in fact, consider more general models where the time effects are individually heterogeneous.

4. The panel vector autoregression (VAR) model may be understood as a completion of the univariate dynamic panel AR(1) model with additional regressors. If we write $y_{it} = (Y_{it}, X'_{it+1})'$, then the first component of the model (1) can be rewritten as

$$Y_{it} = c_i + \beta Y_{it-1} + \gamma' X_{it} + g_t + e_{it},$$

where c_i, g_t, and (β, γ') denote the first components of $(I_m - \theta)\alpha_i$, f_t, and the first row of θ, respectively. This implies that, under the special circumstances where X_{it} follows a first-order VAR, we can regard model (1) as a completion of this model. Under this interpretation, model (1) encompasses panel models with further regressors such as this model.

5. Recently Bai and Ng (2004), Moon and Perron (2004), and Phillips and Sul (2003a) have used a dynamic factor model to model cross-sectional dependence. The time effect model in (1) may correspond to a special case of the factor model with known homogeneous factor loading coefficients.

6. The filtering sequence is chosen for convenience of explanation.

7. See equation (A.3) in the Appendix.

8. The proof of Theorem 2 is straightforward and we omit it.

REFERENCES

Ahn, S.C., Y.H. Lee, & P. Schmidt (2001) GMM estimation of linear panel data models with time-varying individual effects. *Journal of Econometrics* 101, 219–255.

Alvarez, J. & M. Arellano (2003) The time series and cross-section asymptotics of dynamic panel data estimators. *Econometrica* 71, 1121–1160.

Arellano, M. & B. Honoré (2001) Panel data models: Some recent developments. In J. Heckman & E. Leamer (eds.), *Handbook of Econometrics*, vol. 5, pp. 3229–3296. Elsevier Science.

Bai, J. & S. Ng (2004) A PANIC attack on unit roots and cointegration. *Econometrica* 72, 1127–1177.

Brillinger, D. (1981) *Time Series: Data Analysis and Theory.* Holden-Day.

Hahn, J. & G. Kuersteiner (2002) Asymptotically unbiased inference for a dynamic panel model with fixed effects when both n and T are large. *Econometrica* 70, 1639–1659.

Hahn, J. & G. Kuersteiner (2003) Bias Reduction for Dynamic Nonlinear Panel Models with Fixed Effects. Mimeo, UCLA.

Hahn, J. & W. Newey (2004) Jackknife and analytical bias reduction for nonlinear panel models. *Econometrica* 72, 1295–1319.

Holtz-Eakin, D., W. Newey, & H. Rosen (1988) Estimating vector autoregressions with panel data. *Econometrica* 56, 1371–1395.

Hsiao, C. (2003) *Analysis of Panel Data*, 2nd ed. Cambridge University Press.

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

Moon, H.R. & B. Perron (2004) Testing for a unit root in panels with dynamic factors. *Journal of Econometrics* 122, 81–126.

https://doi.org/10.1017/S0266466606060245 Published online by Cambridge University Press

Moon, H.R. & P.C.B. Phillips (2004) GMM estimation of autoregressive roots near unity with panel data. *Econometrica* 72, 467–522.
Neyman, J. & E. Scott (1948) Consistent estimates based on partially consistent observations. *Econometrica* 16, 1–31.
Nickell, S.J. (1981) Biases in dynamic models with fixed effects. *Econometrica* 49, 1417–1426.
Phillips, P.C.B. & H.R. Moon (1999) Linear regression limit theory for nonstationary panel data. *Econometrica* 67, 1057–1111.
Phillips, P.C.B. & D. Sul (2003a) Dynamic panel estimation and homogeneity testing under cross section dependence. *Econometric Journal* 6, 217–259.
Phillips, P.C.B. & D. Sul (2003b) Bias in Dynamic Panel Estimation with Fixed Effects, Incidental Trends and Cross Section Dependence. Mimeo, Yale University and University of Auckland.

APPENDIX

Before we start the proof of Theorem 1, we introduce the following notation. Define $\Theta_{t-1} = (I_m + \theta + \cdots + \theta^{t-2})(I_m - \theta) = (I_m - \theta^{t-1})$, $F_{t-1} = f_{t-1} + \theta f_{t-2} + \cdots + \theta^{t-2} f_1$, and $u_{it-1} = \varepsilon_{it-1} + \theta\varepsilon_{it-2} + \cdots + \theta^{t-2}\varepsilon_{i1}$, where $t = 2,\ldots,T$. For notational convenience, define $\Theta_0 = 0$, $F_0 = 0$, and $u_{i0} = 0$. We let $\bar{\Theta}_T = (1/T)\sum_{t=1}^{T}\Theta_{t-1} = I_m - (1/T)(I_m - \theta^T)(I_m - \theta)^{-1}$, $\bar{F}_{-1} = (1/T)\sum_{t=1}^{T} F_{t-1}$, $\bar{u}_{i,\cdot,-1} = (1/T)\sum_{t=1}^{T} u_{it-1}$, $\bar{u}_{\cdot,t-1} = (1/n)\sum_{i=1}^{n} u_{it-1}$, $\bar{u}_{-1} = (1/nT)\sum_{i=1}^{n}\sum_{t=1}^{T} u_{it-1}$, and $\bar{y}_{\cdot,0} = (1/n)\sum_{i=1}^{n} y_{i0}$. By definition, then, we have

$$y_{it-1} = \Theta_{t-1}\alpha_i + F_{t-1} + u_{it-1} + \theta^{t-1} y_{i0},$$

$$\bar{y}_{\cdot,t-1} = \Theta_{t-1}\bar{\alpha} + F_{t-1} + \bar{u}_{\cdot,t-1} + \theta^{t-1}\bar{y}_{\cdot,0},$$

$$\bar{y}_{i,-1} = \bar{\Theta}_T\alpha_i + \bar{F}_{-1} + \bar{u}_{i,\cdot,-1} + \left(\frac{1}{T}\sum_{t=1}^{T}\theta^{t-1}\right) y_{i0},$$

$$\bar{y}_{-1} = \bar{\Theta}_T\bar{\alpha} + \bar{F}_{-1} + \bar{u}_{-1} + \left(\frac{1}{T}\sum_{t=1}^{T}\theta^{t-1}\right)\bar{y}_{\cdot,0}.$$

Therefore, we have

$$\tilde{y}_{it-1,-1} = \tilde{u}_{it-1,-1} + (\Theta_{t-1} - \bar{\Theta}_T)(\alpha_i - \bar{\alpha}) + \left(\theta^{t-1} - \left(\frac{1}{T}\sum_{t=1}^{T}\theta^{t-1}\right)\right)(y_{i0} - \bar{y}_{\cdot,0}), \tag{A.1}$$

where $\tilde{u}_{it-1,-1} = u_{it-1} - \bar{u}_{\cdot,t-1} - \bar{u}_{i,\cdot,-1} + \bar{u}_{-1}$.

It will be convenient to define another process such that $Y_{i0} = y_{i0}$ and

$$Y_{it} = \theta Y_{it-1} + (I_m - \theta)\alpha_i + \varepsilon_{it}. \tag{A.2}$$

Note that

$$\tilde{y}_{it-1,-1} = Y_{it-1} - \overline{Y}_{i,-1} - (\overline{Y}_{\cdot,t-1} - \overline{Y}_{-1}),$$

$$\tilde{\varepsilon}_{it} = \varepsilon_{it} - \bar{\varepsilon}_{i,\cdot} - (\bar{\varepsilon}_{\cdot,t} - \bar{\varepsilon}),$$

so that

$$\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{y}_{it-1,-1} \tilde{\varepsilon}'_{it} = \frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - \overline{Y}_{i,-1})(\varepsilon_{it} - \bar{\varepsilon}_{i,\cdot})' - \sqrt{\frac{n}{T}} \sum_{t=1}^{T} (\overline{Y}_{\cdot,t-1} - \overline{Y}_{-1})(\bar{\varepsilon}_{\cdot,t} - \bar{\varepsilon})' \quad \textbf{(A.3)}$$

and

$$\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{y}_{it-1,-1} \tilde{y}'_{it-1,-1} = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it-1} - \overline{Y}_{i,-1})(Y_{it-1} - \overline{Y}_{i,-1})' - \frac{1}{T} \sum_{t=1}^{T} (\overline{Y}_{\cdot,t-1} - \overline{Y}_{-1})(\overline{Y}_{\cdot,t-1} - \overline{Y}_{-1})'. \quad \textbf{(A.4)}$$

This is because

$$\begin{aligned} \tilde{y}_{it-1,-1} \tilde{\varepsilon}'_{it} &= \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} ((Y_{it-1} - \overline{Y}_{i,-1}) - (\overline{Y}_{\cdot,t-1} - \overline{Y}_{-1}))((\varepsilon_{it} - \bar{\varepsilon}_{i,\cdot}) - (\bar{\varepsilon}_{\cdot,t} - \bar{\varepsilon}))' \\ &= \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it-1} - \overline{Y}_{i,-1})(\varepsilon_{it} - \bar{\varepsilon}_{i,\cdot})' - \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it-1} - \overline{Y}_{i,-1})(\bar{\varepsilon}_{\cdot,t} - \bar{\varepsilon})' \\ &\quad - \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (\overline{Y}_{\cdot,t-1} - \overline{Y}_{-1})(\varepsilon_{it} - \bar{\varepsilon}_{i,\cdot})' \\ &\quad + \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (\overline{Y}_{\cdot,t-1} - \overline{Y}_{-1})(\bar{\varepsilon}_{\cdot,t} - \bar{\varepsilon})'. \end{aligned}$$

But because

$$\begin{aligned} \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (\overline{Y}_{\cdot,t-1} - \overline{Y}_{-1})(\varepsilon_{it} - \bar{\varepsilon}_{i,\cdot})' &= \frac{1}{T} \sum_{t=1}^{T} (\overline{Y}_{\cdot,t-1} - \overline{Y}_{-1}) \left(\frac{1}{n} \sum_{i=1}^{n} (\varepsilon_{it} - \bar{\varepsilon}_{i,\cdot})' \right) \\ &= \frac{1}{T} \sum_{t=1}^{T} (\overline{Y}_{\cdot,t-1} - \overline{Y}_{-1})(\bar{\varepsilon}_{\cdot,t} - \bar{\varepsilon})' \end{aligned}$$

and

$$\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (\overline{Y}_{\cdot,t-1} - \overline{Y}_{-1})(\bar{\varepsilon}_{\cdot,t} - \bar{\varepsilon})' = \frac{1}{T} \sum_{t=1}^{T} (\overline{Y}_{\cdot,t-1} - \overline{Y}_{-1})(\bar{\varepsilon}_{\cdot,t} - \bar{\varepsilon})',$$

https://doi.org/10.1017/S0266466606060245 Published online by Cambridge University Press

we have the desired simplification. We note that Conditions 1 and 2 are identical to Conditions 1–3 in Hahn and Kuersteiner (2002). We also note that Y_{it} is the same process considered there. Therefore, we can conclude that their Lemmas 6 and 7 are satisfied for Y_{it}.

LEMMA 1 (Hahn and Kuersteiner, 2002, Lem. 6). *Under Conditions 1 and 2, we have*

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}(I_m \otimes (Y_{it}-\overline{Y}_{i,-1}))(\varepsilon_{it}-\bar{\varepsilon}_{i,\cdot})$$

$$\Rightarrow \mathcal{N}(-\sqrt{c}(I_m \otimes I_m - (I_m \otimes \theta))^{-1}\operatorname{vec}(\Omega), (\Omega \otimes \Upsilon + \mathcal{K})).$$

LEMMA 2 (Hahn and Kuersteiner 2002, Lem. 7). *Under Conditions 1 and 2, we have*

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(Y_{it-1}-\overline{Y}_{i,-1})(Y_{it-1}-\overline{Y}_{i,-1})' = \Upsilon + o_p(1).$$

In light of Lemmas 1 and 2, we can show that

$$\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{y}_{it-1,-1}\tilde{\varepsilon}_{it}' \Rightarrow \mathcal{N}(-\sqrt{c}(I_m \otimes I_m - (I_m \otimes \theta))^{-1}\operatorname{vec}(\Omega), (\Omega \otimes \Upsilon + \mathcal{K}))$$

and

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{y}_{it-1,-1}\tilde{y}_{it-1,-1}' = \Upsilon + o_p(1)$$

by proving

$$\sqrt{\frac{n}{T}}\sum_{t=1}^{T}(\overline{Y}_{\cdot,t-1}-\overline{Y}_{-1})(\bar{\varepsilon}_{\cdot,t}-\bar{\varepsilon})' = o_p(1), \tag{A.5}$$

$$\frac{1}{T}\sum_{t=1}^{T}(\overline{Y}_{\cdot,t-1}-\overline{Y}_{-1})(\overline{Y}_{\cdot,t-1}-\overline{Y}_{-1})' = o_p(1). \tag{A.6}$$

Because

$$\sqrt{\frac{n}{T}}\sum_{t=1}^{T}(\overline{Y}_{\cdot,t-1}-\overline{Y}_{-1})(\bar{\varepsilon}_{\cdot,t}-\bar{\varepsilon})' = \frac{1}{n\sqrt{nT}}\sum_{t=1}^{T}\sum_{i=1}^{n}(Y_{it-1}-\overline{Y}_{i,-1})(\varepsilon_{it}-\bar{\varepsilon}_{i,\cdot})'$$

$$+\frac{1}{n\sqrt{nT}}\sum_{i\neq j}^{n}\sum_{t=1}^{T}(Y_{it-1}-\overline{Y}_{i,-1})(\varepsilon_{jt}-\bar{\varepsilon}_{j,\cdot})'$$

https://doi.org/10.1017/S0266466606060245 Published online by Cambridge University Press

and the first term is of order $O_p(n^{-1}) = o_p(1)$ by Lemma 1, it suffices to prove that

$$\frac{1}{n\sqrt{nT}} \sum_{i \neq j}^{n} \sum_{t=1}^{T} (Y_{it-1} - \bar{Y}_{i,-1})(\varepsilon_{jt} - \bar{\varepsilon}_{j,\cdot})' = o_p(1) \tag{A.7}$$

to establish (A.5). On the other hand, because

$$Y_{it-1} - \bar{Y}_{i,-1} = (\Theta_{t-1} - \bar{\Theta}_T)\alpha_i + (u_{it-1} - \bar{u}_{i,\cdot,-1}) + \left(\theta^{t-1} - \frac{1}{T}\sum_{t=1}^{T}\theta^{t-1}\right) y_{i0},$$

we can establish (A.7) by showing that

$$\frac{1}{n\sqrt{nT}} \sum_{i \neq j}^{n} \sum_{t=1}^{T} (u_{it-1} - \bar{u}_{i,\cdot,-1})\varepsilon_{jt}' = o_p(1), \tag{A.8}$$

$$\frac{1}{n\sqrt{nT}} \sum_{i \neq j}^{n} \sum_{t=1}^{T} (\Theta_{t-1} - \bar{\Theta}_T)\alpha_i \varepsilon_{jt}' = o_p(1), \tag{A.9}$$

$$\frac{1}{n\sqrt{nT}} \sum_{i \neq j}^{n} \sum_{t=1}^{T} \left(\theta^{t-1} - \left(\frac{1}{T}\sum_{t=1}^{T}\theta^{t-1}\right)\right) y_{i0}\varepsilon_{jt}' = o_p(1). \tag{A.10}$$

As for (A.6), we note that

$$\bar{Y}_{\cdot,t-1} - \bar{Y}_{-1} = (\bar{u}_{\cdot,t-1} - \bar{u}_{-1}) + (\Theta_{t-1} - \bar{\Theta}_T)\bar{\alpha} + \left(\theta^{t-1} - \left(\frac{1}{T}\sum_{t=1}^{T}\theta^{t-1}\right)\right)\bar{y}_{\cdot,0}$$

and establish (A.6) by showing that

$$\frac{1}{T}\sum_{t=1}^{T} (\bar{u}_{\cdot,t-1} - \bar{u}_{-1})(\bar{u}_{\cdot,t-1} - \bar{u}_{-1})' = o_p(1), \tag{A.11}$$

$$\frac{1}{T}\sum_{t=1}^{T} (\Theta_{t-1} - \bar{\Theta}_T)\bar{\alpha}\bar{\alpha}'(\Theta_{t-1} - \bar{\Theta}_T)' = o_p(1), \tag{A.12}$$

$$\frac{1}{T}\sum_{t=1}^{T} \left(\theta^{t-1} - \left(\frac{1}{T}\sum_{t=1}^{T}\theta^{t-1}\right)\right)\bar{y}_{\cdot,0}\bar{y}_{\cdot,0}'\left(\theta^{t-1} - \left(\frac{1}{T}\sum_{t=1}^{T}\theta^{t-1}\right)\right)' = o_p(1), \tag{A.13}$$

noting that the cross-product terms will all be of order $o_p(1)$ by Cauchy–Schwartz.

We first show (A.8). Note that

$$\begin{aligned}\operatorname{vec}\left(\frac{1}{n\sqrt{nT}} \sum_{i \neq j}^{n} \sum_{t=1}^{T} (u_{it-1} - \bar{u}_{i,\cdot,-1})\varepsilon_{jt}'\right) &= \frac{1}{n}\left(\frac{1}{\sqrt{nT}} \sum_{i \neq j}^{n} \sum_{t=1}^{T} (I_m \otimes u_{it-1})\varepsilon_{jt}\right) \\ &\quad - \frac{1}{n}\left(\sqrt{\frac{T}{n}} \sum_{i \neq j}^{n} (I_m \otimes \bar{u}_{i,\cdot,-1})\varepsilon_{jt}\right).\end{aligned}$$

https://doi.org/10.1017/S0266466606060245 Published online by Cambridge University Press

The first term on the right has mean zero and variance equal to

$$\frac{1}{n^3T}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{i\neq j}^{n}E[(I_m\otimes u_{it-1})\varepsilon_{jt}\varepsilon_{js}'(I_m\otimes u_{is-1}')]$$

$$=\frac{n(n-1)}{n^3}\frac{1}{T}\sum_{t=1}^{T}E[(I_m\otimes u_{it-1})\Omega(I_m\otimes u_{it-1})']=O\left(\frac{1}{n}\right)=o(1).$$

The second term on the right also has mean zero and variance equal to

$$\frac{Tn(n-1)}{n^3}E[(I_m\otimes \bar{u}_{i,\cdot,-1})\Omega(I_m\otimes \bar{u}_{i,\cdot,-1})']$$

and is of the same order as $(T/n)E[\bar{u}_{i,\cdot,-1}\bar{u}_{i,\cdot,-1}']$. But because

$$TE[\bar{u}_{i,\cdot,-1}\bar{u}_{i,\cdot,-1}']=\frac{1}{T}E\left[\left(\sum_{t=1}^{T}u_{it-1}\right)\left(\sum_{t=1}^{T}u_{it-1}\right)'\right]=\Upsilon+o(1),$$

the second term is of order $o_p(1)$ also. Therefore, we obtain (A.8).

As for (A.9), we note that

$$\mathrm{vec}\left(\frac{1}{n\sqrt{nT}}\sum_{i\neq j}^{n}\sum_{t=1}^{T}(\Theta_{t-1}-\bar{\Theta}_T)\alpha_i\varepsilon_{jt}'\right)=\mathrm{vec}\left(\frac{1}{n\sqrt{nT}}\sum_{i\neq j}^{n}\sum_{t=1}^{T}(I_m\otimes(\Theta_{t-1}-\bar{\Theta}_T)\alpha_i)\varepsilon_{jt}\right)$$

has mean zero and variance equal to

$$\frac{n-1}{n^3T}\sum_{i,j=1}^{n}\sum_{t=1}^{T}(I_m\otimes(\Theta_{t-1}-\bar{\Theta}_T)\alpha_i)\Omega(I_m\otimes(\Theta_{t-1}-\bar{\Theta}_T)\alpha_j)'=O\left(\frac{1}{T}\right)=o(1),$$

from which we obtain (A.9). Likewise, we can establish (A.10).

As for (A.11), we note that

$$\frac{1}{T}\sum_{t=1}^{T}(\bar{u}_{\cdot,t-1}-\bar{u}_{-1})(\bar{u}_{\cdot,t-1}-\bar{u}_{-1})'=\frac{1}{T}\sum_{t=1}^{T}\bar{u}_{\cdot,t-1}\bar{u}_{\cdot,t-1}'-\bar{u}_{-1}\bar{u}_{-1}'$$

and that

$$E\left[\left\|\frac{1}{T}\sum_{t=1}^{T}\bar{u}_{\cdot,t-1}\bar{u}_{\cdot,t-1}'\right\|\right]\leq\frac{1}{nT}\sum_{t=1}^{T}E\left[\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}u_{it-1}\right\|^2\right]=O\left(\frac{1}{n}\right),$$

$$\bar{u}_{-1}=O_p\left(\frac{1}{\sqrt{nT}}\right),$$

from which we obtain

$$\frac{1}{T}\sum_{t=1}^{T}(\bar{u}_{\cdot,t-1}-\bar{u}_{-1})(\bar{u}_{\cdot,t-1}-\bar{u}_{-1})'=o_p(1).$$

As for (A.12), we note that

$$
\begin{aligned}
&\operatorname{vec}\left(\frac{1}{T}\sum_{t=1}^{T}(\Theta_{t-1}-\bar{\Theta}_T)\bar{\alpha}\bar{\alpha}'(\Theta_{t-1}-\bar{\Theta}_T)'\right)\\
&\quad=\frac{1}{T}\left[\sum_{t=1}^{T}(\Theta_{t-1}-\bar{\Theta}_T)\otimes(\Theta_{t-1}-\bar{\Theta}_T)\right]\operatorname{vec}(\bar{\alpha}\bar{\alpha}')\\
&\quad=\frac{1}{T}\left[\sum_{t=1}^{T}\left(-\theta^{t-1}+\frac{1}{T}(I_m-\theta^T)(I_m-\theta)^{-1}\right)\right.\\
&\qquad\qquad\left.\otimes\left(-\theta^{t-1}+\frac{1}{T}(I_m-\theta^T)(I_m-\theta)^{-1}\right)\right]\operatorname{vec}(\bar{\alpha}\bar{\alpha}')\\
&\quad=O\left(\frac{1}{T}\right)=o(1)
\end{aligned}
$$

and that (A.13) can be similarly established.

https://doi.org/10.1017/S0266466606060245 Published online by Cambridge University Press