Asymptotic Distribution of Misspecified Random Effects Estimator for a Dynamic Panel Model with Fixed Effects When Both n and T are Large

Jinyong Hahn
UCLA

Guido Kuersteiner
MIT

Myeong Hyeon Cho
Korea University

Abstract

We consider a dynamic panel AR(1) model with fixed effects when both n and T are large. It is shown that the MLE *motivated* by the random effects assumption is asymptotically unbiased even when the assumption is violated.

1   Introduction

In this paper, we consider estimation of a dynamic panel data model with fixed effects. We consider estimation of the autoregressive parameter β_0 in a dynamic panel model with fixed effects

$$y_{it} = \alpha_i + \beta_0 y_{i,t-1} + \varepsilon_{it}, \quad t = 1, 2, \ldots, T; \; i = 1, \ldots, n. \tag{1}$$

The model has additive individual time invariant intercepts (fixed effects) along with a parameter common to every individual. The total number of parameters is therefore equal to n plus the dimension of the common parameter, say K. When n is large relative to T, a maximum likelihood estimator (MLE) of all $n + K$ parameters would lead to a severely biased estimate of the common parameter of interest. See Neyman and Scott (1948), and Nickell (1981). Some of the recent literature proposed to remove such bias, adopting an asymptotic framework where n and T grow to infinity at the same rate. See Hahn and Kuersteiner (2002), Alvarez and Arellano (2003), or Hahn and Newey (2003).

In this paper, we consider the asymptotic properties of a random effects MLE. To be exact, we will consider the MLE computed under the (possibly) incorrect assumption that α_i is a normally distributed random variable *independent* of y_{i0}. It is shown that such a random effects MLE has a zero asymptotic bias. Moreover, the asymptotic distribution is the same as the bias corrected MLE developed in Hahn and Kuersteiner (2002). Therefore, the random effects MLE is efficient according to the convolution theorem derived there. These results hold regardless of the correctness of the random effects assumption.

Alvarez and Arellano (2003) obtained a similar but slightly more restrictive result. They showed that the pseudo MLE based on the correlated random effects model, where α_i is specified to be normally distributed with arbitrary unknown correlation with y_{i0}, is asymptotically unbiased. Our result strengthens their by showing that the pseudo MLE remains asymptotically unbiased even when α_i is specified to be independent of y_{i0}.

2   Random Effects Mis-Specification

We consider a panel model with fixed effects

$$y_{it} = \alpha_i + \beta_0 y_{it-1} + \varepsilon_{it}, \quad i = 1, \ldots, n; \; t = 1, \ldots, T \tag{2}$$

where we observe the sample $y_{i0}, \ldots, y_{iT}$ for each individual. We first examine the sampling properties of the maximum likelihood estimator (fixed effects estimator) when both n and T are large. We impose the following regularity condition.

Condition 1 *(i)* $\varepsilon_{it} \sim \mathcal{N}\left(0, \sigma_0^2\right)$ *i.i.d.; (ii)* $0 < \lim \frac{n}{T} \equiv \rho < \infty$*; (iii)* $|\beta_0| < 1$*; and (iv)* $\frac{1}{n} \sum_{i=1}^{n} y_{i0}^2 = O(1)$ *and* $\frac{1}{n} \sum_{i=1}^{n} \alpha_i^2 = O(1)$.

Observe that the MLE b is such that

$$\sqrt{nT}\left(b - \beta_0\right) = \frac{\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(\varepsilon_{it} - \overline{\varepsilon}_i\right) \cdot \left(y_{it-1} - \overline{y}_{i-}\right)}{\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(y_{it-1} - \overline{y}_{i-}\right)^2}, \tag{3}$$

where $\overline{y}_{i-} = \frac{1}{T} \sum_{t=1}^{T} y_{it-1}$ and $\overline{\varepsilon}_i = \frac{1}{T} \sum_{t}^{T} \varepsilon_{it}$. Under our asymptotics, we can obtain the result that the fixed effects MLE is asymptotically biased:

$$\sqrt{nT} \left(b - \beta_0 \right) \to \mathcal{N} \left(-\sqrt{\rho} \left(1 + \beta_0 \right), 1 - \beta_0^2 \right).$$

See Hahn and Kuersteiner (2002) or Alvarez and Arellano (2003).

We now consider the maximum likelihood estimator computed under the possibly *incorrect* assumption that α_i is a normally distributed random variable *independent* of y_{i0}:

$$\alpha_i | y_{i0} \sim \mathcal{N} \left(\mu, \omega^2 \right). \tag{4}$$

The nature of misspecification is that α_i is assumed to be independent of y_{i0} when in fact they may have arbitrary correlation. It will be established that the maximum likelihood estimator of the misspecified model is efficient under the alternative asymptotic approximation we adopt in this paper.

The log likelihood under (4) is equal to

$$\mathcal{L} \left(\sigma^2, \beta, \lambda, \mu \right) = -nT \log \sigma^2 + n \log \lambda - n \log \left(T + \lambda \right) - \frac{n\mu^2}{\sigma^2} \lambda$$

$$- \frac{1}{\sigma^2} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(y_{it} - \beta y_{it-1} \right)^2 + \frac{1}{\sigma^2} \frac{1}{T + \lambda} \sum_{i=1}^{n} \left(T \left(\overline{y}_i - \beta \overline{y}_{i-} \right) + \lambda \mu \right)^2,$$

where $\overline{y}_i \equiv \frac{1}{T} \sum_{t=1}^{T} y_{it}$, $\overline{y}_{i-} \equiv \frac{1}{T} \sum_{t=1}^{T} y_{it-1}$, and $\lambda \equiv \sigma^2 / \omega^2$. Let $\widehat{\beta}_{RE}$ denote the (pseudo-) maximum likelihood estimator solving $\max_{\sigma, \beta, \lambda, \mu} \mathcal{L} \left(\sigma^2, \beta, \lambda, \mu \right)$. It is shown that $\widehat{\beta}_{RE}$ is consistent and asymptotically normal:

Theorem 1 *Under Condition 1, we have* $\sqrt{nT} \left(\widehat{\beta}_{RE} - \beta_0 \right) \to \mathcal{N} \left(0, 1 - \beta_0^2 \right).$

Proof. See Appendix C. ■

The asymptotic distribution of $\widehat{\beta}_{RE}$ is the same as the bias corrected maximum likelihood estimator developed in Hahn and Kuersteiner (2002), and the result holds regardless of the correctness of the random effects assumption. This surprising result seems to have at least two practical consequences.

Our result casts reasonable doubt on the finite sample properties of Hausman tests for this particular model. Consider an econometrician faced with a dynamic panel model with fixed effects. He is entertaining two estimators: A maximum likelihood estimator for the random effects model, and an IV estimator based on first differencing. Assume that he approximates the model by the usual T fixed $n \to \infty$ asymptotics. Under the null, *i.e.*, under assumption (4), both estimators are consistent but MLE is more efficient. Under the alternative that (4) is incorrect, there would be a statistically significant difference between the two. If the difference is rather small, he will opt to use the random effects MLE. If the difference is large, he will use the IV estimator. Now, suppose that the alternative asymptotic approximation considered in this paper is a better one. It then follows that the difference between the two will be asymptotically nonzero whether the null is correct or not. It is quite possible that the econometrician would conclude that the difference between the two estimators is statistically significant, based on which he will opt to report the IV estimator. But the IV estimator is asymptotically biased under the alternative asymptotic approximation as argued by Alvarez and Arellano (2003). It therefore follows that, with high probability, an econometrician would choose to use an inferior estimator due to the puzzling property of the random effects estimator.

Appendix

Maximizing it over (λ, μ), we obtain the concentrated likelihood

$$\mathcal{L}\left(\sigma^2, \beta\right) = -n\left(T-1\right)\log \sigma^2 - n \log \left\{ \frac{1}{n} \sum_{i=1}^{n} \left(\overline{y}_i - \beta \overline{y}_{i-}\right)^2 - \left(\frac{1}{n} \sum_{i=1}^{n} \left(\overline{y}_i - \beta \overline{y}_{i-}\right)\right)^2 \right\}$$

$$- \frac{1}{\sigma^2} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(y_{it} - \beta y_{it-1}\right)^2 + T \frac{1}{\sigma^2} \sum_{i=1}^{n} \left(\overline{y}_i - \beta \overline{y}_{i-}\right)^2 .$$

It can also be shown that

$$\widehat{\lambda}\left(\sigma^2, \beta\right) = \frac{\sigma^2}{-\frac{\sigma^2}{T} + \frac{1}{n}\sum_{i=1}^{n}\left(\overline{y}_i - \beta\overline{y}_{i-}\right)^2 - \left(\frac{1}{n}\sum_{i=1}^{n}\left(\overline{y}_i - \beta\overline{y}_{i-}\right)\right)^2},$$

$$\widehat{\mu}\left(\sigma^2, \beta\right) = \frac{1}{n}\sum_{i=1}^{n}\left(\overline{y}_i - \beta\overline{y}_{i-}\right),$$

where $\left(\widehat{\lambda}\left(\sigma^2, \beta\right), \widehat{\mu}\left(\sigma^2, \beta\right)\right) \equiv \mathrm{argmax}_{(\lambda,\mu)} \mathcal{L}\left(\sigma^2, \beta, \lambda, \mu\right)$. Let $\left(\widehat{\sigma}^2_{RE}, \widehat{\beta}_{RE}, \widehat{\lambda}_{RE}, \widehat{\mu}_{RE}\right)$ denote the maximum likelihood estimator.

A   Technical Lemmas

Lemma 1 $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\alpha_i + \varepsilon_{it}\right)^2 = \lim \frac{1}{n}\sum_{i=1}^{n}\alpha_i^2 + \sigma_0^2 + o_p\left(1\right)$, $\frac{1}{n}\sum_{i=1}^{n}\left(\alpha_i + \overline{\varepsilon}_i\right)^2 = \lim \frac{1}{n}\sum_{i=1}^{n}\alpha_i^2 + o_p\left(1\right)$.

Proof. With some abuse of notation, we can write

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\alpha_i + \varepsilon_{it}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\alpha_i^2 + \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\varepsilon_{it}^2 + 2\frac{1}{nT}\sum_{i=1}^{n}\alpha_i\sum_{t=1}^{T}\varepsilon_{it}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\alpha_i^2 + \sigma_0^2\frac{\chi^2\left(nT\right)}{nT} + \mathcal{N}\left(0, \frac{4\sigma_0^2}{n^2T}\sum_{i=1}^{n}\alpha_i^2\right),$$

which establishes the first claim. The second can be similarly established. ∎

Lemma 2 $\frac{1}{n}\sum_{i=1}^{n}\overline{y}_{i-}^2 = \frac{1}{\left(1-\beta_0\right)^2}\lim\frac{1}{n}\sum_{i=1}^{n}\alpha_i^2 + o_p\left(1\right)$.

Proof. We have

$$\overline{y}_{i-} \sim \mathcal{N}\left(\frac{1}{T}\frac{1-\beta_0^T}{1-\beta_0}y_{i0} + \frac{1}{T}\frac{\beta_0 T - \beta_0 - \beta_0^2 T + \beta_0^{T+1}}{\left(1-\beta_0\right)^2\beta_0}\alpha_i, \gamma_T^2\right) \tag{5}$$

independent across is, where

$$\gamma_T \equiv \sqrt{\sigma_0^2\frac{T\left(1-\beta_0^2\right) - 2\beta_0 + 2\beta_0^{T+1} - \beta_0^{2T} + 2\beta_0^T - 1}{T^2\left(1-\beta_0\right)^2\left(1-\beta_0^2\right)}} = o\left(1\right). \tag{6}$$

4

based on which we obtain

$$
\frac{1}{n} \sum_{i=1}^{n} \overline{y}_{i-}^2 = \frac{1}{n} \sum_{i=1}^{n} E\left[\overline{y}_{i-}^2\right] + o_p(1)
$$

$$
= \frac{1}{n} \sum_{i=1}^{n} \left(\left(E\left[\overline{y}_{i-}\right]\right)^2 + \mathrm{Var}\left(\overline{y}_{i-}\right)\right) + o_p(1) = \frac{1}{\left(1-\beta_0\right)^2} \lim \frac{1}{n} \sum_{i=1}^{n} \alpha_i^2 + o_p(1).
$$

∎

Lemma 3 $\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(\alpha_i + \varepsilon_{it}\right) y_{it-1} = \frac{1}{1-\beta_0} \lim \frac{1}{n} \sum_{i=1}^{n} \alpha_i^2.$

Proof. We have

$$
\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(\alpha_i + \varepsilon_{it}\right) y_{it-1} = \frac{1}{n} \sum_{i=1}^{n} \alpha_i \overline{y}_{i-} + \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(\varepsilon_{it} - \overline{\varepsilon}_i\right) \left(y_{it-1} - \overline{y}_{i-}\right) + \frac{1}{n} \sum_{i=1}^{n} \overline{\varepsilon}_i \overline{y}_{i-}. \quad (7)
$$

Using equation (5) in the proof of Lemma 2, we can show that

$$
\frac{1}{n} \sum_{i=1}^{n} \alpha_i \overline{y}_{i-} = \frac{1}{1-\beta_0} \lim \frac{1}{n} \sum_{i=1}^{n} \alpha_i^2 + o_p(1) \quad (8)
$$

By Hahn and Kuersteiner (2002, Lemma 6), we also have

$$
\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(\varepsilon_{it} - \overline{\varepsilon}_i\right) \left(y_{it-1} - \overline{y}_{i-}\right) = o_p(1). \quad (9)
$$

Finally,

$$
\frac{1}{n} \sum_{i=1}^{n} \overline{\varepsilon}_i \overline{y}_{i-} = o_p(1), \quad (10)
$$

because

$$
\left| \frac{1}{n} \sum_{i=1}^{n} \overline{\varepsilon}_i \overline{y}_{i-} \right|^2 \leq \left| \frac{1}{n} \sum_{i=1}^{n} \overline{\varepsilon}_i^2 \right| \left| \frac{1}{n} \sum_{i=1}^{n} \overline{y}_{i-}^2 \right| = \left| \frac{\sigma_0^2}{nT} \chi^2(n) \right| \left| O_p(1) \right| = o_p(1),
$$

where the first equality is based on Lemma 2. The conclusion follows from (7), (8), (9), and (10). ∎

Lemma 4 $\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} y_{it-1}^2 = \frac{1}{\left(1-\beta_0\right)^2} \cdot \lim \frac{1}{n} \sum_{i=1}^{n} \alpha_i^2 + \frac{\sigma_0^2}{1-\beta_0^2} + o_p(1).$

Proof. We have

$$
\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} y_{it-1}^2 = \frac{1}{nT} \sum_{i=1}^{n} \left(T\overline{y}_{i-}^2 + \sum_{t=1}^{T} \left(y_{it-1} - \overline{y}_{i-}\right)^2 \right) = \frac{1}{n} \sum_{i=1}^{n} \overline{y}_{i-}^2 + \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(y_{it-1} - \overline{y}_{i-}\right)^2.
$$

The conclusion follows from Lemma 2 and Hahn and Kuersteiner (2002, Lemma 7). ∎

B Consistency of $\left(\widehat{\beta}_{RE}, \widehat{\sigma}_{RE}^2\right)$

The pointwise limit of $\mathcal{L}\left(\sigma^2, \beta\right) / nT$ is equal to the limit of

$$
-\log \sigma^2 - \frac{1}{\sigma^2} \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(y_{it} - \beta y_{it-1}\right)^2 + \frac{1}{\sigma^2} \frac{1}{n} \sum_{i=1}^{n} \left(\overline{y}_i - \beta \overline{y}_{i-}\right)^2. \quad (11)
$$

Lemma 5 $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it}-\beta y_{it-1}\right)^{2}=\left(\lim\frac{1}{n}\sum_{i=1}^{n}\alpha_{i}^{2}\right)\left(1-\frac{\beta-\beta_{0}}{1-\beta_{0}}\right)^{2}+\sigma_{0}^{2}\left(1+\frac{(\beta-\beta_{0})^{2}}{1-\beta_{0}^{2}}\right)+o_{p}\left(1\right).$

Proof. We have

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it}-\beta y_{it-1}\right)^{2}=\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\alpha_{i}+\varepsilon_{it}\right)^{2}-2\left(\beta-\beta_{0}\right)\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\alpha_{i}+\varepsilon_{it}\right)y_{it-1}$$

$$+\left(\beta-\beta_{0}\right)^{2}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}y_{it-1}^{2}. \tag{12}$$

Conclusion follows from combination of (12), Lemmas 1, 3, and 4. ∎

Lemma 6 $\frac{1}{n}\sum_{i=1}^{n}\left(\overline{y}_{i}-\beta\overline{y}_{i-}\right)^{2}=\left(1-\frac{\beta-\beta_{0}}{1-\beta_{0}}\right)^{2}\lim\frac{1}{n}\sum_{i=1}^{n}\alpha_{i}^{2}+o_{p}\left(1\right).$

Proof. Note that

$$\frac{1}{n}\sum_{i=1}^{n}\left(\overline{y}_{i}-\beta\overline{y}_{i-}\right)^{2}=\frac{1}{n}\sum_{i=1}^{n}\left(\alpha_{i}+\overline{\varepsilon}_{i}\right)^{2}+\left(\beta-\beta_{0}\right)^{2}\frac{1}{n}\sum_{i=1}^{n}\overline{y}_{i-}^{2}-2\left(\beta-\beta_{0}\right)\frac{1}{n}\sum_{i=1}^{n}\left(\alpha_{i}+\overline{\varepsilon}_{i}\right)\overline{y}_{i-}.$$

Combining this result with Lemma 1, and Lemma 2, equations (8), and (10), we obtain the desired conclusion. ∎

Theorem 2 $\left(\widehat{\beta}_{RE},\widehat{\sigma}_{RE}^{2}\right)=\left(\beta_{0},\sigma_{0}^{2}\right)+o_{p}\left(1\right).$

Proof. Combining (11) with Lemmas 5, and 6, we conclude that the pointwise limit of $\mathcal{L}\left(\sigma^{2},\beta\right)/nT$ is equal to $-\log\sigma^{2}-\frac{\sigma_{0}^{2}}{\sigma^{2}}\left(1+\frac{(\beta-\beta_{0})^{2}}{1-\beta_{0}^{2}}\right)$, which is maximized at $\left(\beta_{0},\sigma_{0}^{2}\right)$. Because the convergence is uniform, as is verified in Appendix D, $\left(\widehat{\beta},\widehat{\sigma}^{2}\right)$ should be consistent for $\left(\beta_{0},\sigma_{0}^{2}\right)$. Conclusion follows. ∎

C   Proof of Theorem 1

We have

$$\sqrt{nT}\left(\widehat{\beta}_{RE}-\beta_{0}\right)=\frac{\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\widetilde{y}_{it-1}\widetilde{\varepsilon}_{it}+\frac{T\widehat{\lambda}_{RE}}{T+\widehat{\lambda}_{RE}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\widetilde{\overline{y}}_{i-}\widetilde{\alpha}_{i}+\frac{T\widehat{\lambda}_{RE}}{T+\widehat{\lambda}_{RE}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\widetilde{\overline{y}}_{i-}\widetilde{\overline{\varepsilon}}_{i}}{\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\widetilde{y}_{it-1}^{2}+\frac{1}{nT}\frac{T\widehat{\lambda}_{RE}}{T+\widehat{\lambda}_{RE}}\sum_{i=1}^{n}\widetilde{\overline{y}}_{i-}^{2}}.$$

By Hahn and Kuersteiner (2002, Lemmas 6 and 7), we have

$$\frac{1}{\sqrt{nT}}\sum_{i}\sum_{t=1}^{T}\widetilde{y}_{it-1}\widetilde{\varepsilon}_{it}\rightarrow-\sqrt{\frac{n}{T}}\sigma_{0}^{2}\frac{1}{1-\beta_{0}}+\mathcal{N}\left(0,\frac{\sigma_{0}^{4}}{1-\beta_{0}^{2}}\right), \tag{13}$$

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\widetilde{y}_{it-1}^{2}=\frac{\sigma_{0}^{2}}{1-\beta_{0}^{2}}+o_{p}\left(1\right). \tag{14}$$

Using consistency of $\left(\widehat{\beta},\widehat{\sigma}^{2}\right)$ along with equations (5) and (6), we also obtain $\widehat{\mu}_{RE}=\lim\frac{1}{n}\sum_{i=1}^{n}\alpha_{i}+o_{p}\left(1\right)$, $\frac{T\widehat{\lambda}_{RE}}{T+\widehat{\lambda}_{RE}}=\frac{\sigma_{0}^{2}}{\lim\frac{1}{n}\sum_{i=1}^{n}(\alpha_{i}-\overline{\alpha})^{2}}+o_{p}\left(1\right)$, $\frac{1}{nT}\sum_{i=1}^{n}\widetilde{\overline{y}}_{i-}^{2}=o_{p}\left(1\right)$, $\frac{T\widehat{\lambda}_{RE}}{T+\widehat{\lambda}_{RE}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\widetilde{\overline{y}}_{i-}\widetilde{\alpha}_{i}=\sqrt{\frac{n}{T}}\frac{\sigma_{0}^{2}}{1-\beta_{0}}+o_{p}\left(1\right)$, $\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\widetilde{\overline{y}}_{i-}\widetilde{\overline{\varepsilon}}_{i}=o_{p}\left(1\right)$, from which we can infer

$$\frac{1}{nT}\frac{T\widehat{\lambda}_{RE}}{T+\widehat{\lambda}_{RE}}\sum_{i=1}^{n}\widetilde{\overline{y}}_{i-}^{2}=o_{p}\left(1\right),$$

$$\frac{T\widehat{\lambda}}{T+\widehat{\lambda}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\widetilde{\overline{y}}_{i-}\widetilde{\alpha}_{i}+\frac{T\widehat{\lambda}}{T+\widehat{\lambda}}\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\widetilde{\overline{y}}_{i-}\widetilde{\overline{\varepsilon}}_{i}=\sqrt{\frac{n}{T}}\frac{\sigma_{0}^{2}}{1-\beta_{0}}+o_{p}\left(1\right).$$

6

Combined with (13) and (14), we obtain the conclusion.

D   Uniform Convergence

Let $\widehat{Q}_n\left(\sigma^2,\beta\right) \equiv \frac{1}{nT}\mathcal{L}\left(\sigma^2,\beta\right)$ and $Q\left(\sigma^2,\beta\right) \equiv -\log\sigma^2 - \frac{\sigma_0^2}{\sigma^2}\left(1 + \frac{(\beta-\beta_0)^2}{1-\beta_0^2}\right)$. It was already shown that $\widehat{Q}_n\left(\sigma^2,\beta\right) \rightarrow Q\left(\sigma^2,\beta\right)$ pointwise. Therefore, the convergence is uniform as long as

1. $\widehat{Q}_n$ is maximized over $\left(\sigma^2,\beta\right) \in \Theta$, a compact set, and $\left(\sigma_0^2,\beta_0\right) \in \Theta$;

2. $\widehat{Q}_n\left(\sigma^2,\beta\right)$ satisfies stochastic equicontinuity.

See Newey (1991) for details.

We will simply assume that Θ is compact. As for stochastic equicontinuity, we will show that $\widehat{Q}_n\left(\sigma^2,\beta\right)$ satisfies Lipschitz condition. By Newey (1991, Corollary 2.2), this will be sufficient for uniform convergence. We will show that the derivatives of $\widehat{Q}_n$ are $O_p\left(1\right)$: By the usual mean-value-theorem type argument, this will establish Lipschitz.

We note that

$$\frac{\partial}{\partial\sigma^2}\widehat{Q}_n\left(\sigma^2,\beta\right) = -\frac{T-1}{T\sigma^2} + \frac{1}{(\sigma^2)^2}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it}-\beta y_{it-1}\right)^2 - \frac{1}{(\sigma^2)^2}\frac{1}{n}\sum_{i=1}^{n}\left(\overline{y}_i - \beta\overline{y}_{i-}\right)^2,$$

and

$$\frac{\partial}{\partial\beta}\widehat{Q}_n\left(\sigma^2,\beta\right) = \frac{2n}{T}\frac{1}{\frac{1}{n}\sum_{i=1}^{n}\left(\overline{y}_i - \beta\overline{y}_{i-}\right)^2 - \left(\frac{1}{n}\sum_{i=1}^{n}\left(\overline{y}_i - \beta\overline{y}_{i-}\right)\right)^2}$$
$$\times\left\{\frac{1}{n}\sum_{i=1}^{n}\left(\overline{y}_i - \beta\overline{y}_{i-}\right)\overline{y}_{i-} - \left(\frac{1}{n}\sum_{i=1}^{n}\left(\overline{y}_i - \beta\overline{y}_{i-}\right)\right)\left(\frac{1}{n}\sum_{i=1}^{n}\overline{y}_{i-}\right)\right\}$$
$$+ \frac{2}{\sigma^2}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it}-\beta y_{it-1}\right)y_{it-1} - \frac{2}{\sigma^2}\frac{1}{n}\sum_{i=1}^{n}\left(\overline{y}_i - \beta\overline{y}_{i-}\right)\overline{y}_{i-}.$$

Therefore, Newey's (1991) Assumption 3A would be satisfied if $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it}-\beta y_{it-1}\right)^2$, $\frac{1}{n}\sum_{i=1}^{n}\left(\overline{y}_i - \beta\overline{y}_{i-}\right)^2$, $\frac{1}{n}\sum_{i=1}^{n}\left(\overline{y}_i - \beta\overline{y}_{i-}\right)\overline{y}_{i-}$, $\frac{1}{n}\sum_{i=1}^{n}\left(\overline{y}_i - \beta\overline{y}_{i-}\right)$, $\frac{1}{n}\sum_{i=1}^{n}\overline{y}_{i-}$, and $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it}-\beta y_{it-1}\right)y_{it-1}$ are all $O_p\left(1\right)$, which follow from the compactness of Θ and Lemma 6.

References

[1] ALVAREZ, J., AND M. ARELLANO (2003), "The Time Series and Cross-Section Asymptotics of Dynamic Panel Data Estimators", *Econometrica* 71, pp. 1121 – 1159.

[2] HAHN, J., AND G. KUERSTEINER (2002), "Asymptotically Unbiased Inference for a Dynamic Panel Model with Fixed Effects when Both n and T are Large", *Econometrica* 70, pp. 1639 – 1657.

[3] HAHN, J., AND W. NEWEY (2003), "Jackknife and Analytical Bias Reduction for Nonlinear Panel Models", forthcoming in *Econometrica*.

[4] NEYMAN, J., AND E. SCOTT (1948), "Consistent Estimates Based on Partially Consistent Observations", *Econometrica* 16, pp. 1 -31.

[5] NICKEL, S. (1981), "Biases in Dynamic Models with Fixed Effects", *Econometrica* 49, pp. 1417 - 1426.