Journal of the Korean Statistical Society (2007), **36**: *4, pp 523–542*

EFFICIENT ESTIMATION IN SEMIPARAMETRIC RANDOM EFFECT PANEL DATA MODELS WITH AR(p) ERRORS[†]

YOUNG KYUNG LEE[1]

ABSTRACT

In this paper we consider semiparametric random effect panel models that contain AR(p) disturbances. We derive the efficient score function and the information bound for estimating the slope parameters. We make minimal assumptions on the distribution of the random errors, effects, and the regressors, and provide semiparametric efficient estimates of the slope parameters. The present paper extends the previous work of Park *et al.* (2003) where AR(1) errors were considered.

AMS 2000 subject classifications. Primary 62G05; Secondary 62G20.
Keywords. Correlated errors, efficient score functions, information bounds, panel data models, random effects, semiparametric efficiency.

1. INTRODUCTION

In this paper we assume that one observes (X_{it}, Y_{it}) such that

$$Y_{it} = X_{it}^\top \beta + \alpha_i + \varepsilon_{it}, \quad i = 1, \dots, N\,; t = 1, \dots, T, \tag{1.1}$$

where X_{it} are d-variate covariates, β is a d-dimensional unknown parameter and α_i are unobservable random effects. For the errors ε_{it} we assume an AR(p) model

$$\varepsilon_{it} = \phi_1 \varepsilon_{i,t-1} + \phi_2 \varepsilon_{i,t-2} + \cdots + \phi_p \varepsilon_{i,t-p} + u_{it},$$

where u_{it} are *i.i.d.* random variables from $N(0, \sigma^2)$. Writing $X_i \equiv (X_{i1}^\top, \dots, X_{iT}^\top)^\top$, we also assume that (X_i, α_i) are *i.i.d.* $(dT+1)$-dimensional random variables having unknown density q. For stationarity of the error process $\{\varepsilon_{it}\}_{t=1}^T$, we assume throughout the paper that the roots of the characteristic equation

$$1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p = 0$$

Received June 2007; accepted August 2007.
[†]This study was supported by 2007 Research Grants from Kangwon National University.
[1]Department of Statistics, Kangwon National University, Chuncheon 200-701, Korea (e-mail: youngklee@kangwon.ac.kr)

lie outside the unit ball. We consider two models. The first one is the model where X_i, α_i and $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iT})^\top$ are independent. The second model is the one where X_i and α_i are allowed to be dependent while (X_i, α_i) is independent of ε_i.

The main contribution of the paper is to derive the semiparametric information bound for estimating β in the presence of the infinite dimensional nuisance parameter q as well as the finite dimensional nuisance parameters ϕ and σ^2. Also, a semiparametric efficient estimator of β is provided that achieves the information bound. We make minimal assumptions on the distributions of the random effects and the covariates.

The general theory of semiparametric efficient estimation has been well developed, see Bickel (1982), Begun *et al.* (1983) and Bickel *et al.* (1993), among others. Semiparametric efficient estimation for the model (1.1) has been discussed extensively in the literature. For example, Park and Simar (1994) considered the case where the errors are assumed to be *i.i.d.* $N(0, \sigma^2)$. Park *et al.* (1998) extended their results to the case where some of the covariates are dependent of the random effects. Park *et al.* (2003) worked on the model (1.1) where the errors follow an AR(1) process. The present paper extends the previous work of Park *et al.* (2003) by relaxing the AR(1) error assumption to AR(p). This extension turns out to be not straightforward and calculation of the information bound as well as construction of an efficient estimator appear to be much more involved than in the AR(1) case.

The panel model (1.1) with AR(p) errors has numerous applications. One example is productivity analysis of financial industries where market shocks may not be adjusted to immediately and induce a serial correlation pattern in firm's use of best-practice financial technologies. In this case our semiparametric efficient estimators provide an important tool for making robust inferences on the productivity gains due to the economic reforms.

There have been a great deal of effort to address the question of efficient estimation for other panel data models. These include the earlier work of Arellano and Bond (1991), Arellano and Bover (1995) and Ahn and Schmidt (1995) on dynamic panel models. They investigated a number of moment conditions which if correctly specified the generalized method of moment technique yields efficient estimates. A number of excellent surveys and monographs have been written on this subject, see Baltagi (1995) and Mátyás and Sevestre (1992) for example. Recently, Park *et al.* (2007) studied semiparametric efficiency for dynamic panel data models. For extensive discussion on other panel models and their applications, see Hsiao (2007) and the references therein.

2. INFORMATION BOUNDS

2.1. Model 1

In this model we assume that X_i, α_i and ε_i are independent. We consider semiparametric efficient estimation of β from the sample $\{(X_i, Y_i) \mid i = 1, \dots, N\}$ in the presence of the nuisance parameters $(\sigma^2, \phi, h(\cdot), g(\cdot))$.

Let $Y = (Y_1, \dots, Y_T)^\top$, $X = (X_1^\top, \dots, X_T^\top)^\top$ for the generic of observation (X_i, Y_i) and (α, ε) for the generic of $(\alpha_i, \varepsilon_i)$. Thus, in these notations, (X_t, Y_t) are generics for (X_{it}, Y_{it}), $i = 1, \dots, N$. Let $h(\cdot)$ denote the univariate density of the α and $g(\cdot)$ the dT-variate density of X. Define $C_{ts}(\phi) = \mathrm{Cov}(\varepsilon_t, \varepsilon_s)/\sigma^2$ for $1 \le t, s \le p$ and let $c^{ts}(\phi)$ denotes the $(t,s)^{th}$ component of the inverse of the $p \times p$ matrix $C(\phi) \equiv (C_{ts}(\phi))$. We note that $C_{ts}(\phi)$ and $c^{ts}(\phi)$ are functions of ϕ only. Let $\phi(B)$ be a transfer function defined by $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ where B is the backward shift operator. Then, the probability density function of (X, Y) can be written as

$$\begin{aligned} &p(x, y; \beta, \sigma^2, \phi, h, g) \\ &= (\sqrt{2\pi}\sigma)^{-T} |C(\phi)|^{-1/2} g(x) \int \exp\left[-\frac{1}{2\sigma^2} \left\{ \sum_{t=1}^{p} \sum_{s=1}^{p} c^{ts}(\phi)(y_t - x_t^\top \beta - u) \right.\right. \\ &\quad \left.\left. \times (y_s - x_s^\top \beta - u) + \sum_{t=p+1}^{T} \left(\phi(B)(y_t - x_t^\top \beta - u) \right)^2 \right\} \right] h(u)\, du. \qquad (2.1) \end{aligned}$$

Define $T(\phi) = \sum_{t=1}^{p} \sum_{s=1}^{p} c^{ts}(\phi) + (T - p)\{\phi(B)1\}^2$, where $\phi(B)1 = 1 - \phi_1 - \cdots - \phi_p$, and

$$W(\phi, \beta) = T(\phi)^{-1} \left[\sum_{t=1}^{p} \sum_{s=1}^{p} c^{ts}(\phi)(Y_t - X_t^\top \beta) + \{\phi(B)1\} \sum_{t=p+1}^{T} \phi(B)(Y_t - X_t^\top \beta) \right].$$

We note that $W(\phi, \beta)$ is a weighted average of $Y_t - X_t^\top \beta$. In fact, $W(\phi, \beta) = \sum_{t=1}^{T} w_t (Y_t - X_t^\top \beta)$, where $w_t \equiv w_t(\phi)$ is defined by

$$T(\phi) w_t = \begin{cases} \sum_{s=1}^{p} c^{ts}(\phi) - \{\phi(B)1\} \sum_{s=p-t+1}^{p} \phi_s, & \text{if } 1 \le t \le p, \\ \{\phi(B)1\}^2, & \text{if } p+1 \le t \le T-p, \\ \{\phi(B)1\} \left(1 - \sum_{s=1}^{T-t} \phi_s\right), & \text{if } T-p+1 \le t \le T. \end{cases}$$

In general, for any sequence $\{a_t\}$, we have

$$\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)a_t + \{\phi(B)1\}\sum_{t=p+1}^{T}\phi(B)a_t = T(\phi)\sum_{t=1}^{T} w_t a_t. \tag{2.2}$$

The identity will be also used later. It can be seen that $\sum_{t=1}^{T} w_t = 1$. Also, it can be verified that

$$\begin{aligned}
&\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)(Y_t - X_t^\top\beta - u)(Y_s - X_s^\top\beta - u) + \sum_{t=p+1}^{T}\Big\{\phi(B)(Y_t - X_t^\top\beta - u)\Big\}^2 \\
&= \sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)(Y_t - X_t^\top\beta)(Y_s - X_s^\top\beta) + \sum_{t=p+1}^{T}\Big\{\phi(B)(Y_t - X_t^\top\beta)\Big\}^2 \\
&\quad + T(\phi)\{u - W(\phi,\beta)\}^2 - T(\phi)W(\phi,\beta)^2.
\end{aligned} \tag{2.3}$$

Define $v^2 \equiv v^2(\phi) = \sigma^2 T(\phi)^{-1}$ and $f \equiv f(\,\cdot\,; v^2, h)$ by

$$f(w) = \int \varphi_v(w-u)h(u)\,du,$$

where φ_v denotes the density function of $N(0, v^2)$. Plugging the expression at (2.3) into the right hand side of (2.1), we get

$$\begin{aligned}
&p(X,Y;\beta,\sigma^2,\phi,h,g) \\
&= (\sqrt{2\pi}\sigma)^{-T}|C(\phi)|^{-1/2}g(x)\exp\Bigg[-\frac{1}{2\sigma^2}\Bigg\{\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)(Y_t - X_t^\top\beta)(Y_s - X_s^\top\beta) \\
&\quad + \sum_{t=p+1}^{T}\Big\{\phi(B)(Y_t - X_t^\top\beta)\Big\}^2\Bigg\}\Bigg]\exp\left\{\frac{W(\phi,\beta)^2}{2v^2}\right\}\sqrt{2\pi}v f(W(\phi,\beta)).
\end{aligned}$$

It is worthwhile to note that f is the density function of $W(\phi,\beta)$. This follows from

$$W(\phi,\beta) = \alpha + T(\phi)^{-1}\left[\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)\varepsilon_t + \{\phi(B)1\}\sum_{t=p+1}^{T} u_t\right], \tag{2.4}$$

and the fact that the variance of the terms in the bracket on the right hand side of (2.4) equals

$$\begin{aligned}
&\mathrm{Var}\left(\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)\varepsilon_t\right) + \{\phi(B)1\}^2\mathrm{Var}\left(\sum_{t=p+1}^{T} u_t\right) \\
&= \sigma^2\left[\sum_{t=1}^{p}\sum_{t'=1}^{p}\left\{\sum_{s=1}^{p} c^{ts}(\phi)\right\}\left\{\sum_{s'=1}^{p} c^{t's'}(\phi)\right\}C_{tt'}(\phi) + (T-p)\{\phi(B)1\}^2\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sigma^2 \left[\sum_{t=1}^{p} \sum_{s=1}^{p} c^{ts}(\phi) + (T-p)\{\phi(B)1\}^2 \right] \\
&= T(\phi)\sigma^2. \qquad (2.5)
\end{aligned}
$$

The second equality in the above equations holds since

$$
\sum_{t'=1}^{p} C_{tt'}(\phi)c^{t's'}(\phi) = (C(\phi)C(\phi)^{-1})_{ts'} = \begin{cases} 1, \text{ if } t = s', \\ 0, \text{ if } t \neq s'. \end{cases} \qquad (2.6)
$$

We now calculate the score functions for the parameters β, σ^2 and ϕ. Let L denote the log-likelihood of $(\beta, \sigma^2, \phi, h, g)$ for the generic observation (X, Y). Define

$$
\ell_\beta = \frac{\partial L}{\partial \beta}, \quad \ell_{\sigma^2} = \frac{\partial L}{\partial \sigma^2}, \quad \ell_{\phi_j} = \frac{\partial L}{\partial \phi_j}.
$$

To give mathematical expressions for these score functions, write $\tilde{X}(\phi) = \sum_{t=1}^{T} w_t X_t$ and

$$
Z_t(\phi, \beta) = Y_t - X_t^\top \beta - W(\phi, \beta).
$$

Define

$$
\begin{aligned}
\kappa_j(\phi) &= -\frac{\partial}{\partial \phi_j}\left(\frac{1}{2}\log|C|\right), \quad \xi_j^{ts}(\phi) = -\frac{\partial}{\partial \phi_j}\left(\frac{1}{2}c^{ts}(\phi)\right), \\
S_j(\phi) &= -\frac{\partial}{\partial \phi_j} T(\phi) = 2\sum_{t=1}^{p}\sum_{s=1}^{p} \xi_j^{ts}(\phi) + 2(T-p)\{\phi(B)1\}.
\end{aligned}
$$

Let $\tilde{w}_{t,j} \equiv \tilde{w}_{t,j}(\phi)$ for a given j be another set of weights defined by

$$
S_j(\phi)\tilde{w}_{t,j} = \begin{cases}
2\sum_{s=1}^{p} \xi_j^{ts}(\phi) - \sum_{s=p-t+1}^{p} \phi_s, & \text{if } 1 \le t \le p-j, \\
2\sum_{s=1}^{p} \xi_j^{ts}(\phi) - \sum_{s=p-t+1}^{p} \phi_s + \{\phi(B)1\}, & \text{if } p-j+1 \le t \le p, \\
2\{\phi(B)1\}, & \text{if } p+1 \le t \le T-p, \\
1 - \sum_{s=1}^{T-t} \phi_s + \{\phi(B)1\}, & \text{if } T-p+1 \le t \le T-j, \\
1 - \sum_{s=1}^{T-t} \phi_s, & \text{if } T-j+1 \le t \le T.
\end{cases}
$$

It can be also verified that $\sum_{t=1}^{T} \tilde{w}_{t,j} = 1$ for all $j = 1, \ldots, p$. Finally, define

$$
\tilde{W}_j(\phi, \beta) = \sum_{t=1}^{T} \tilde{w}_{t,j}(Y_t - X_t^\top \beta). \qquad (2.7)
$$

Then, we obtain

$$\ell_\beta = \frac{1}{\sigma^2}\left[\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)X_s Z_t(\phi,\beta) + \sum_{t=p+1}^{T}\{\phi(B)Z_t(\phi,\beta)\}\{\phi(B)X_t\}\right]$$
$$-\tilde{X}(\phi)\frac{f'}{f}(W(\phi,\beta)), \tag{2.8}$$
$$\ell_{\sigma^2} = \frac{1}{2\sigma^4}\left[\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)Z_t(\phi,\beta)Z_s(\phi,\beta) + \sum_{t=p+1}^{T}\{\phi(B)Z_t(\phi,\beta)\}^2\right]$$
$$-\frac{T}{2\sigma^2} + \frac{1}{2\sigma^2 f(W(\phi,\beta))}\int\left\{\frac{W(\phi,\beta)-u}{v}\right\}^2$$
$$\times\varphi_v(W(\phi,\beta)-u)h(u)\,du, \tag{2.9}$$
$$\ell_{\phi_j} = \frac{1}{\sigma^2}\left[\sum_{t=1}^{p}\sum_{s=1}^{p}\xi_j^{ts}(\phi)(Y_t - X_t^\top\beta)(Y_s - X_s^\top\beta)\right.$$
$$\left.+\sum_{t=p+1}^{T}\left\{\phi(B)(Y_t - X_t^\top\beta)\right\}(Y_{t-j} - X_{t-j}^\top\beta)\right] + \kappa_j(\phi)$$
$$+\frac{S_j(\phi)}{2\sigma^2 f(W(\phi,\beta))}\int\left\{u^2 - 2u\tilde{W}_j(\phi,\beta)\right\}$$
$$\times\varphi_v(W(\phi,\beta)-u)h(u)\,du. \tag{2.10}$$

The equation (2.8) follows from the fact that

$$\frac{\partial}{\partial\beta}W(\phi,\beta) = -\sum_{t=1}^{T} w_t X_t = -\tilde{X}(\phi).$$

Derivation of (2.9) is based on the identity

$$\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)(Y_t - X_t^\top\beta)(Y_s - X_s^\top\beta)$$
$$+\sum_{t=p+1}^{T}\left\{\phi(B)(Y_t - X_t^\top\beta)\right\}^2 - T(\phi)W(\phi,\beta)^2$$
$$= \sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)Z_t(\phi,\beta)Z_s(\phi,\beta) + \sum_{t=p+1}^{T}\{\phi(B)Z_t(\phi,\beta)\}^2.$$

The result (2.10) can be obtained by using

$$\frac{\partial}{\partial\phi_j}W(\phi,\beta) = \frac{S_j(\phi)}{T(\phi)}\left\{W(\phi,\beta) - \tilde{W}_j(\phi,\beta)\right\},$$

$$\frac{\partial}{\partial \phi_j}\left[\frac{W(\phi,\beta)^2}{2\,v(\phi)^2}\right] = \frac{1}{2\sigma^2}S_j(\phi)W(\phi,\beta)\left\{W(\phi,\beta) - 2\,\tilde{W}_j(\phi,\beta)\right\},$$

$$\frac{\partial}{\partial \phi_j}\left[\frac{\{W(\phi,\beta)-u\}^2}{2\,v(\phi)^2}\right] = \frac{S_j(\phi)}{2\sigma^2}\{W(\phi,\beta)-u\}\left\{W(\phi,\beta) - 2\,\tilde{W}_j(\phi,\beta)+u\right\}.$$

Methods of finding efficient score functions and information bounds for estimating parametric components in semiparametric models are well explained in Bickel *et al.* (1993). We adopt their approach here.

Let $\mathcal{S}$ be the tangent space for the nuisance parameters σ^2, ϕ, h, g. Then, $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3 + \mathcal{S}_4$, where $\mathcal{S}_1 = [\ell_\phi], \mathcal{S}_2 = [\ell_{\sigma^2}]$ and writing $W = W(\phi,\beta)$,

$$\mathcal{S}_3 = \{a(W) : a \in L_2(P), Ea(W) = 0\},$$
$$\mathcal{S}_4 = \{b(X) : b \in L_2(P), Eb(X) = 0\}.$$

Here and below, $[\ell]$ means the closed linear span of ℓ. Then, the efficient score ℓ^* is given by

$$\ell^* = \ell_\beta - \Pi(\ell_\beta|\mathcal{S}) = \Pi(\ell_\beta|\mathcal{S}_1^\perp \cap \mathcal{S}_2^\perp \cap \mathcal{S}_3^\perp \cap \mathcal{S}_4^\perp),$$

where $\Pi(\,\cdot\,|\mathcal{S})$ is the projection operator onto the space $\mathcal{S}$. Let $Q_i = \Pi(\,\cdot\,|\mathcal{S}_i^\perp)$. By the Halperin's theorem (see, Bickel *et al.*, 1993, p. 443, Theorem 3), it follows that

$$(Q_1\,Q_2\,Q_3\,Q_4)^r \ell_\beta \to \ell^*$$

in L_2-sense as $r \to \infty$. The information bound for estimating β is given by

$$I = E\,\ell^*\ell^{*\top}.$$

In the following theorem we give explicit formulas for the efficient score function ℓ^* and the information bound I. For this, let $J(f)$ denote the Fisher information for the location of f, *i.e.*,

$$J(f) = \int f'(w)^2/f(w)\,dw.$$

Also, define

$$\begin{aligned}\Sigma_1(\phi) &= E\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}(\phi)(X_t-\tilde{X})(X_s-\tilde{X})^\top \\ &\quad +E\sum_{t=p+1}^{T}\{\phi(B)(X_t-\tilde{X})\}\{\phi(B)(X_t-\tilde{X})\}^\top, \\ \Sigma_2(\phi) &= \mathrm{Var}(\tilde{X}) \;=\; E(\tilde{X}-E\tilde{X})(\tilde{X}-E\tilde{X})^\top.\end{aligned}$$

We assume that $J(f) < \infty$, and that Σ_1 and Σ_2 exist and are non-singular. We need these conditions for the regularity of the semiparametric model.

THEOREM 2.1. *Assume that $J(f) < \infty$, and that $\Sigma_1(\phi)$ and $\Sigma_2(\phi)$ exist and are non-singular. Then, the efficient score function for estimating β in Model 1 is given by*

$$\ell^* = \frac{1}{\sigma^2}\left[\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)Z_t(\phi,\beta)X_s + \sum_{t=p+1}^{T}\{\phi(B)Z_t(\phi,\beta)\}\{\phi(B)X_t\}\right] - \frac{f'}{f}(W(\phi,\beta))\{\tilde{X}(\phi) - E\tilde{X}(\phi)\}.$$

Also, the information bound is given by $I = \sigma^{-2}\Sigma_1(\phi) + J(f)\Sigma_2(\phi)$.

PROOF. We suppress ϕ and β in the arguments of W, Z_t, c^{ts}, $\tilde{X}$, *etc.* First, we show that ℓ_β is perpendicular to $\mathcal{S}_4$ in $L_2(P)$-sense, *i.e.*,

$$\Pi(\ell_\beta|\mathcal{S}_4) = E(\ell_\beta|X) = 0 \quad \text{or} \quad Q_4\ell_\beta = \ell_\beta. \tag{2.11}$$

Note that

$$Z_t = \varepsilon_t - T(\phi)^{-1}\left[\sum_{s=1}^{p}\sum_{s'=1}^{p} c^{ss'}\varepsilon_s + \{\phi(B)1\}\sum_{s=p+1}^{T} u_s\right]. \tag{2.12}$$

Thus, $Z_t(\phi,\beta)$ is independent of X. Also, from (2.4) one can see that $W(\phi,\beta)$ is independent of X. These entail

$$E(Z_t|X) = E(Z_t) = 0,$$
$$E\left\{\frac{f'}{f}(W)\Big|X\right\} = E\left\{\frac{f'}{f}(W)\right\} = 0.$$

The orthogonality (2.11) follows from (2.8).

Next, we claim that

$$W \text{ is independent of } \{Z_t\}_{t=1}^{T}. \tag{2.13}$$

To show this, we go back to the expression (2.4). Let σ^2 and ϕ be fixed so that v^2 is fixed. For the time being, we treat α as a parameter rather than a random variable. Consider the family of distributions of W indexed by α. In fact, W has a $N(\alpha, v^2)$ distribution. It is a complete and sufficient statistic for α. Furthermore, from (2.12) the distribution of $\{Z_t\}$ does not depend on α, thus it is ancillary for α. This establishes (2.13).

The independence (2.13) implies $E(Z_t|W) = E(Z_t) = 0$. Since $E(\tilde{X}|W) = E\tilde{X}$, we obtain

$$\Pi(\ell_\beta|\mathcal{S}_3) = -(E\tilde{X})\frac{f'}{f}(W).$$

Thus, from (2.11) it follows that

$$\begin{aligned} Q_3Q_4\ell_\beta &= \ell_\beta - \Pi(\ell_\beta|\mathcal{S}_3) \\ &= \frac{1}{\sigma^2}\left[\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts} Z_t X_s + \sum_{t=p+1}^{T}\{\phi(B)Z_t\}\{\phi(B)X_t\}\right] \\ &\quad - \frac{f'}{f}(W)\{\tilde{X} - E\tilde{X}\}. \end{aligned}$$

Since $\{Z_t\} \stackrel{d}{=} \{-Z_t\}$, all third-order moments of Z_t's are zero. With the independence of X, W and $\{Z_t\}$, this entails

$$E\{(Q_3Q_4\ell_\beta)\ell_{\sigma^2}\} = 0 \quad \text{or} \quad Q_2Q_3Q_4\ell_\beta = Q_3Q_4\ell_\beta.$$

We prove

$$Q_3Q_4\ell_\beta \text{ is perpendicular to } \ell_{\phi_j} \text{ for all } 1 \le j \le p. \tag{2.14}$$

This implies $Q_1Q_2Q_3Q_4\ell_\beta = Q_3Q_4\ell_\beta$. Thus, the assertion (2.14) concludes the proof of the first part of the theorem. To show (2.14) we use a representation of ℓ_{ϕ_j} that is different from the one given at (2.10). Note that for any sequence $\{a_t\}$

$$2\sum_{t=1}^{p}\sum_{s=1}^{p}\xi_j^{ts} a_t + \sum_{t=p+1}^{T}\phi(B)a_t + \{\phi(B)1\}\sum_{t=p+1}^{T} a_{t-j} = S_j \sum_{t=1}^{T}\tilde{w}_{t,j}a_t. \tag{2.15}$$

Using (2.15) and the fact $\sum_{t=1}^{T}\tilde{w}_{t,j}Z_t = \tilde{W}_j - W$, one can verify

$$\begin{aligned} &\sum_{t=1}^{p}\sum_{s=1}^{p}\xi_j^{ts}(Y_t - X_t^\top\beta)(Y_s - X_s^\top\beta) + \sum_{t=p+1}^{T}\left\{\phi(B)(Y_t - X_t^\top\beta)\right\}(Y_{t-j} - X_{t-j}^\top\beta) \\ &= \sum_{t=1}^{p}\sum_{s=1}^{p}\xi_j^{ts} Z_t Z_s + \sum_{t=p+1}^{T}\{\phi(B)Z_t\}Z_{t-j} + \frac{1}{2}S_j W(2\tilde{W}_j - W). \end{aligned}$$

Plugging this into (2.10), we get

$$\ell_{\phi_j} = \frac{1}{\sigma^2}\left[\sum_{t=1}^{p}\sum_{s=1}^{p}\xi_j^{ts} Z_t Z_s + \sum_{t=p+1}^{T}\{\phi(B)Z_t\}Z_{t-j}\right] + \kappa_j(\phi) + \frac{1}{2\sigma^2}S_j W(2\tilde{W}_j - W)$$
$$+ \frac{S_j(\phi)}{2\sigma^2 f(W(\phi,\beta))}\int\left\{u^2 - 2u\tilde{W}_j(\phi,\beta)\right\}\varphi_v(W(\phi,\beta)-u)h(u)\,du.$$

By similar arguments for proving (2.13), one can show that $(W, \tilde{W}_j)$ is independent of $\{Z_t\}_{t=1}^T$. Using again the fact $\{Z_t\} \overset{d}{=} \{-Z_t\}$ and the independence of X, $(W, \tilde{W}_j)$ and $\{Z_t\}$, one can conclude (2.14).

Next, we prove the second part of the theorem. By the independence between $\{Z_t\}, \{X_t\}$ and W, we get

$$E\ell^*\ell^{*\top} = J(f)\mathrm{Var}(\tilde{X}) + \frac{1}{\sigma^4}E\left[\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}Z_t X_s + \sum_{t=p+1}^{T}\{\phi(B)Z_t\}\{\phi(B)X_t\}\right]$$
$$\times\left[\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}Z_t X_s + \sum_{t=p+1}^{T}\{\phi(B)Z_t\}\{\phi(B)X_t\}\right]^\top.$$

Define

$$\tilde{\varepsilon} \equiv \tilde{\varepsilon}(\phi) = T(\phi)^{-1}\left[\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}\varepsilon_t + \{\phi(B)1\}\sum_{t=p+1}^{T}u_t\right].$$

From (2.4) we can write $Z_t = \varepsilon_t - \tilde{\varepsilon}$. Recall that

$$\mathrm{Var}(\tilde{\varepsilon}) = T(\phi)^{-1}\sigma^2,$$

see (2.5). Also, we find that for $1 \le t' \le p$,

$$E(\varepsilon_{t'}\tilde{\varepsilon}) = T(\phi)^{-1}\sigma^2\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}C_{tt'} = T(\phi)^{-1}\sigma^2.$$

Thus, we obtain

$$E(Z_t Z_{t'}) = \sigma^2\{C_{tt'} - T(\phi)^{-1}\}, \quad 1 \le t, t' \le p. \tag{2.16}$$

Since $\phi(B)Z_t = u_t - \{\phi(B)1\}\tilde{\varepsilon}$ and $E(\tilde{\varepsilon}u_{t'}) = T(\phi)^{-1}\sigma^2\{\phi(B)1\}$ for $p+1 \le t' \le T$, it follows that

$$EZ_t\{\phi(B)Z_{t'}\} = -\frac{\sigma^2}{T(\phi)}\{\phi(B)1\}, \quad 1 \le t \le p;\ p+1 \le t' \le T, \tag{2.17}$$

$$E\{\phi(B)Z_t\}\{\phi(B)Z_{t'}\} = E(u_t u_{t'}) - \frac{\sigma^2}{T(\phi)}\{\phi(B)1\}^2, \quad p+1 \le t, t' \le T. \tag{2.18}$$

From (2.16)–(2.18) and (2.6) we get

$$E\Big(\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}Z_tX_s\Big)\Big(\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}Z_tX_s\Big)^\top$$
$$= \sigma^2\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}E(X_sX_t^\top) - \frac{\sigma^2}{T(\phi)}E\Big(\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}X_s\Big)\Big(\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}X_s\Big)^\top, \quad (2.19)$$

$$E\Big(\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}Z_tX_s\Big)\Big(\sum_{t=p+1}^{T}\{\phi(B)Z_t\}\{\phi(B)X_t\}\Big)^\top$$
$$= -\frac{\sigma^2}{T(\phi)}E\Big(\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}X_s\Big)\Big(\{\phi(B)1\}\sum_{t=p+1}^{T}\{\phi(B)X_t^\top\}\Big), \quad (2.20)$$

$$E\Big(\sum_{t=p+1}^{T}\{\phi(B)Z_t\}\{\phi(B)X_t\}\Big)\Big(\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}Z_tX_s\Big)^\top$$
$$= -\frac{\sigma^2}{T(\phi)}E\Big(\{\phi(B)1\}\sum_{t=p+1}^{T}\{\phi(B)X_t\}\Big)\Big(\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}X_s^\top\Big), \quad (2.21)$$

$$E\Big(\sum_{t=p+1}^{T}\{\phi(B)Z_t\}\{\phi(B)X_t\}\Big)\Big(\sum_{t=p+1}^{T}\{\phi(B)Z_t\}\{\phi(B)X_t\}\Big)^\top$$
$$= \sigma^2E\sum_{t=p+1}^{T}\{\phi(B)X_t\}\{\phi(B)X_t\}^\top - \frac{\sigma^2}{T(\phi)}\{\phi(B)1\}^2E\Big(\sum_{t=p+1}^{T}\phi(B)X_t\Big)$$
$$\times\Big(\sum_{t=p+1}^{T}\phi(B)X_t^\top\Big). \quad (2.22)$$

Putting (2.19)–(2.22) together and using (2.2), we obtain

$$E\Big[\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}Z_tX_s + \sum_{t=p+1}^{T}\{\phi(B)Z_t\}\{\phi(B)X_t\}\Big]$$
$$\times\Big[\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}Z_tX_s + \sum_{t=p+1}^{T}\{\phi(B)Z_t\}\{\phi(B)X_t\}\Big]^\top$$
$$= \sigma^2\Big[\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}E(X_sX_t^\top) + \sum_{t=p+1}^{T}E\{\phi(B)X_t\}\{\phi(B)X_t\}^\top\Big] - \sigma^2T(\phi)E(\tilde{X}\tilde{X}^\top)$$
$$= \sigma^2\Sigma_1.$$

This concludes the proof of the theorem. □

2.2. *Model 2*

In this model we allow α_i and X_i to be dependent, but continue to assume that (α_i, X_i) is independent of ε_i. Let q denote the $(dT+1)$-variate density function of (α_i, X_i). Define

$$\eta(w,x) = \int \varphi_v(w-u)q(u,x)\,du.$$

The function η is the joint density function of $(W(\phi,\beta), X)$. Then, the log-likelihood of $(\beta, \sigma^2, \phi, q)$ for the generic observation (X,Y) equals

$$\begin{aligned} &L(\beta,\sigma^2,\phi,h,g;X,Y) \\ &= -\frac{T}{2}\log(2\pi\sigma^2) - \frac{1}{2}\log|C(\phi)| - \frac{1}{2\sigma^2}\left\{\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)(Y_t - X_t^{\top}\beta)(Y_s - X_s^{\top}\beta)\right. \\ &\quad \left. + \sum_{t=p+1}^{T}\left\{\phi(B)(Y_t - X_t^{\top}\beta)\right\}^2\right\} + \frac{W(\phi,\beta)^2}{2v^2} + \frac{1}{2}\log(2\pi v^2) \\ &\quad + \log f(W(\phi,\beta),X). \end{aligned}$$

The following theorem gives the efficient score function and the information bound for estimating β in Model 2.

THEOREM 2.2. *Assume that $\Sigma_1(\phi)$ exists and is non-singular. Then, the efficient score function for estimating β in Model 2 is given by*

$$\ell^* = \frac{1}{\sigma^2}\left[\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)Z_t(\phi,\beta)X_s + \sum_{t=p+1}^{T}\{\phi(B)Z_t(\phi,\beta)\}\{\phi(B)X_t\}\right].$$

Also, the information bound is given by $I = \sigma^{-2}\Sigma_1(\phi)$.

PROOF. The theorem follows by the same arguments used in the proof of Theorem 2.1 with the fact that $(W(\phi,\beta),X)$ is independent of $\{Z_t\}$. In this case, the score functions are as given by (2.8)–(2.10), but with $h(u)$ and $f(W(\phi,\beta))$ being replaced by $q(u,X)$ and $\eta(W(\phi,\beta),X)$, respectively. Also, $\eta'(W(\phi,\beta),X)$ is substituted for $f'(W(\phi,\beta))$, where $\eta'(w,x) = (\partial/\partial w)\eta(w,x)$. The tangent space for the nuisance parameters σ^2, ϕ, q is given by $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3$, where $\mathcal{S}_1 = [\ell_\phi], \mathcal{S}_2 = [\ell_{\sigma^2}]$ and $\mathcal{S}_3 = \{a(W,X) : a \in L_2(P), Ea(W,X) = 0\}$. Thus, one can verify that

$$\ell_\beta - \Pi(\ell_\beta|\mathcal{S}_3) = \frac{1}{\sigma^2}\left[\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)Z_t(\phi,\beta)X_s + \sum_{t=p+1}^{T}\{\phi(B)Z_t(\phi,\beta)\}\{\phi(B)X_t\}\right],$$

and that this is perpendicular to ℓ_{σ^2} and ℓ_{ϕ_j} for all $1 \le j \le p$. □

3. EFFICIENT ESTIMATION

3.1. Model 1

To construct an efficient estimator of β, we need preliminary estimators of β and ϕ that are $\sqrt{N}$-consistent. Let $\tilde{\beta}$ be the OLS estimator obtained by regressing $Y_{it} - \overline{Y}_i$ on $X_{it} - \overline{X}_i$, where $\overline{X}_i = T^{-1}(X_{i1} + \cdots + X_{iT})$ and $\overline{Y}_i = T^{-1}(Y_{i1} + \cdots + Y_{iT})$. Then,

$$\begin{aligned}\tilde{\beta} &= \left[\sum_{i=1}^{N}\sum_{t=1}^{T}(X_{it} - \overline{X}_i)(X_{it} - \overline{X}_i)^\top\right]^{-1}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}(X_{it} - \overline{X}_i)(Y_{it} - \overline{Y}_i)\right] \\ &= \beta + \left[\sum_{i=1}^{N}\sum_{t=1}^{T}(X_{it} - \overline{X}_i)(X_{it} - \overline{X}_i)^\top\right]^{-1}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}(X_{it} - \overline{X}_i)(\varepsilon_{it} - \overline{\varepsilon}_i)\right].\end{aligned}$$

It is easy to see that $\tilde{\beta}$ is $\sqrt{N}$-consistent.

Next, we describe a $\sqrt{N}$-consistent estimator of ϕ. Let $\rho_j = \mathrm{Corr}(\varepsilon_t, \varepsilon_{t-j})$ and define for $1 \le j \le p$,

$$r_j = \frac{\rho_j - \rho_{j+1}}{1 - \rho_1}.$$

Then, since the autocorrelation function satisfies the difference equation

$$\rho_j = \phi_1\rho_{j-1} + \phi_2\rho_{j-2} + \cdots + \phi_p\rho_{j-p}, \quad j \ge 1,$$

it follows that

$$\begin{aligned} r_j &= r_{j-1}\phi_1 + \cdots + r_1\phi_{j-1} + \phi_j + (-1)\phi_{j+1} \\ &\quad +(-r_1)\phi_{j+2} + \cdots + (-r_{p-j-1})\phi_p, \qquad 1 \le j \le p.\end{aligned}$$

Thus, writing $r = (r_1, \ldots, r_p)^\top$ and

$$R = \begin{pmatrix} 1 & -1 & -r_1 & \cdots & -r_{p-4} & -r_{p-3} & -r_{p-2} \\ r_1 & 1 & -1 & \cdots & -r_{p-5} & -r_{p-4} & -r_{p-3} \\ r_2 & r_1 & 1 & \cdots & -r_{p-6} & -r_{p-5} & -r_{p-4} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ r_{p-3} & r_{p-4} & r_{p-5} & \cdots & 1 & -1 & -r_1 \\ r_{p-2} & r_{p-3} & r_{p-4} & \cdots & r_1 & 1 & -1 \\ r_{p-1} & r_{p-2} & r_{p-3} & \cdots & r_2 & r_1 & 1 \end{pmatrix},$$

we have

$$r = R\phi \quad \text{or} \quad \phi = R^{-1}r.$$

Let $\tilde{r}_j$ be estimators of r_j and $\tilde{R}$ be the corresponding estimator of R. We propose as the preliminary estimator of ϕ

$$\tilde{\phi} = \tilde{R}^{-1}\tilde{r}.$$

It is clear that if $\tilde{r}_j$ are $\sqrt{N}$-consistent, then $\tilde{\phi}$ is also a $\sqrt{N}$-consistent estimator of ϕ. To define $\tilde{r}_j$, let

$$\gamma_{ij}(\beta) = \frac{1}{(T-j)} \sum_{t=j+1}^{T} (Y_{it} - X_{it}^{\top}\beta)(Y_{i,t-j} - X_{i,t-j}^{\top}\beta).$$

Then, we have $E\{\gamma_{ij}(\beta)|\alpha_i\} = \alpha_i^2 + \rho_j \mathrm{Var}(\varepsilon_t)$, so that

$$\begin{aligned} E\{\gamma_{ij}(\beta) - \gamma_{i,j+1}(\beta)|\alpha_i\} &= (\rho_j - \rho_{j+1})\mathrm{Var}(\varepsilon_t), \\ E\{\gamma_{i0}(\beta) - \gamma_{i1}(\beta)|\alpha_i\} &= (1 - \rho_1)\mathrm{Var}(\varepsilon_t), \end{aligned} \tag{3.1}$$

for all i. This shows that $\tilde{r}_j$ defined by

$$\tilde{r}_j = \frac{\sum_{i=1}^{N}\{\gamma_{ij}(\tilde{\beta}) - \gamma_{i,j+1}(\tilde{\beta})\}}{\sum_{i=1}^{N}\{\gamma_{i0}(\tilde{\beta}) - \gamma_{i1}(\tilde{\beta})\}}$$

is a $\sqrt{N}$-consistent estimator of r_j.

Once an estimator of ϕ is obtained, one can construct an estimator of $\rho \equiv (\rho_1, \ldots, \rho_p)^{\top}$ from the Yule-Walker equations for AR(p) process. Let $\tilde{\rho}$ be the solution of the following system of linear equations with respect to ρ:

$$\rho_j = \tilde{\phi}_1\rho_{j-1} + \tilde{\phi}_2\rho_{j-2} + \cdots + \tilde{\phi}_p\rho_{j-p}, \quad 1 \le j \le p.$$

From (3.1) and the identity

$$\mathrm{Var}(\varepsilon_t) = \sigma^2(1 - \rho_1\phi_1 - \cdots - \rho_p\phi_p)^{-1},$$

we have

$$\sigma^2 = (1 - \rho_1)^{-1}(1 - \rho_1\phi_1 - \cdots - \rho_p\phi_p)E\{\gamma_{i0}(\beta) - \gamma_{i1}(\beta)\}.$$

This suggests the following estimator of σ^2 which is also $\sqrt{N}$-consistent:

$$\tilde{\sigma}^2 = (1 - \tilde{\rho}_1)^{-1}(1 - \tilde{\rho}_1\tilde{\phi}_1 - \cdots - \tilde{\rho}_p\tilde{\phi}_p)N^{-1}\sum_{i=1}^{N}\{\gamma_{i0}(\tilde{\beta}) - \gamma_{i1}(\tilde{\beta})\}.$$

Now, we define an estimator of the information matrix I. Let $\tilde{X}.(\tilde{\phi}) = N^{-1}\sum_{i=1}^{N}\tilde{X}_i(\tilde{\phi})$. Define

$$\hat{\Sigma}_1 = \frac{1}{N}\sum_{i=1}^{N}\left\{\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}(\tilde{\phi})(X_{it}-\tilde{X}_i(\tilde{\phi}))(X_{is}-\tilde{X}_i(\tilde{\phi}))^\top \right.$$
$$\left. + \sum_{t=p+1}^{T}\{\tilde{\phi}(B)(X_{it}-\tilde{X}_i(\tilde{\phi}))\}\{\tilde{\phi}(B)(X_{it}-\tilde{X}_i(\tilde{\phi}))\}^\top\right\},$$
$$\hat{\Sigma}_2 = \frac{1}{N}\sum_{i=1}^{N}\{\tilde{X}_i(\tilde{\phi})-\tilde{X}.(\tilde{\phi})\}\{\tilde{X}_i(\tilde{\phi})-\tilde{X}.(\tilde{\phi})\}^\top,$$
$$\hat{J}(f) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{f'}{f}\right)^2(W_i(\tilde{\phi},\tilde{\beta})).$$

We estimate f by a kernel estimator:

$$\hat{f}(w) = \frac{1}{N}\sum_{i=1}^{N}K_b(w-W_i(\tilde{\phi},\tilde{\beta}))+c,$$

where $K_b(u) = (1/b)K(u/b)$ and K is a probability density function such that $|K^{(j)}/K|$ are bounded for $j = 1, 2, 3$. An example of K satisfying this condition is given by $K(u) = e^{-u}(1+e^{-u})^{-2}$. The bandwidth b and the constant c tends to zero at some appropriate rates described below. We define

$$\hat{I} = \tilde{\sigma}^{-2}\hat{\Sigma}_1 + \hat{J}(\hat{f})\hat{\Sigma}_2.$$

The efficient estimator of β is now defined by

$$\hat{\beta} = \tilde{\beta} + \frac{1}{N}\hat{I}^{-1}\sum_{i=1}^{N}\left[\tilde{\sigma}^{-2}\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}(\tilde{\phi})Z_{it}(\tilde{\phi},\tilde{\beta})X_{is}\right.$$
$$\left. +\tilde{\sigma}^{-2}\sum_{t=p+1}^{T}\{\tilde{\phi}(B)Z_{it}(\tilde{\phi},\tilde{\beta})\}\{\tilde{\phi}(B)X_{it}\} - \{\tilde{X}_i(\tilde{\phi})-\tilde{X}.(\tilde{\phi})\}\frac{\hat{f}'}{\hat{f}}(W_i(\tilde{\phi},\tilde{\beta}))\right].$$

The following theorem tells that $\hat{\beta}$ is a semiparametric efficient estimator of β.

THEOREM 3.1. *Assume the conditions given in Theorem 2.1. Furthermore, assume that* $E(e^{t\|\tilde{X}_1\|}) < \infty$ *for some* $t > 0$ *and that* $\int u^2 h(u)\,du < \infty$. *If* $b \to 0$, $c \to 0$ *and* $Nb^2s^6 \to \infty$ *as* $N \to \infty$, *then*

$$N^{1/2}(\hat{\beta}-\beta) \xrightarrow{d} N(0, I^{-1}).$$

PROOF. We only give a sketch of the proof. In the proof we write W_i, Z_{it}, $\tilde{X}_i$ and $\tilde{X}.$ for $W_i(\phi,\beta)$, $Z_{it}(\phi,\beta)$, $\tilde{X}_i(\phi)$ and $\tilde{X}.(\phi)$, respectively, when ϕ and β are the true parameters. Define

$$f_N(w;\tau^2) = K_b * \int \varphi_\tau(w-u)h(u)\,du + c.$$

Note that $f_N(w;v^2) = K_b * f(w) + c$. Also define

$$\begin{aligned} J_N(f) :&= \int \left\{ \frac{f_N'}{f_N}(w;v^2) \right\}^2 f(w)\,dw, \\ I_N :&= \sigma^{-2}\Sigma_1 + J_N(f)\Sigma_2. \end{aligned}$$

From the standard theory of kernel density estimation it follows that as $N \to \infty$

$$E\left[\frac{f_N'}{f_N}(W_1;v^2) - \frac{f'}{f}(W_1)\right]^2 \to 0, \tag{3.2}$$

See, for example, Bickel and Ritov (1987), Park (1990), or Park and Simar (1994) for the proofs of these results. From (3.2) one can show

$$I_N \to I, \tag{3.3}$$

$$\{\tilde{X}. - E(\tilde{X}_1)\}N^{-1/2}\sum_{i=1}^{N}\frac{f_N'}{f_N}(W_i;v^2) \xrightarrow{p} 0, \tag{3.4}$$

$$N^{-1/2}\sum_{i=1}^{N}\{\tilde{X}_i - E(\tilde{X}_1)\}\left[\frac{f_N'}{f_N}(W_i;v^2) - \frac{f'}{f}(W_i)\right] \xrightarrow{p} 0. \tag{3.5}$$

Now, define

$$\begin{aligned} &Q_N(\zeta,\theta) \\ &= N^{-1/2}\sum_{i=1}^{N}\Bigg[\sigma^{-2}\sum_{t=1}^{p}\sum_{s=1}^{p}c^{ts}(\theta)Z_{it}(\theta,\zeta)X_{is} \\ &+\sigma^{-2}\sum_{t=p+1}^{T}\{\theta(B)Z_{it}(\theta,\zeta)\}\{\theta(B)X_{it}\} - \{\tilde{X}_i(\theta) - \tilde{X}.(\theta)\}\frac{f_N'}{f_N}\Big(W_i(\theta,\zeta);\frac{\sigma^2}{T(\theta)}\Big)\Bigg]. \end{aligned}$$

Then, by the central limit theorem and (3.3)–(3.5) we have

$$I_N^{-1}Q_N(\beta,\phi) \xrightarrow{d} N(0,I^{-1}).$$

□

Note that if we replace $\tilde{\sigma}^2$ by σ^2 and $\hat{f}$ by $f_N(\cdot;\tilde{\sigma}^2/T(\tilde{\phi}))$ in the definition of $\hat{\beta}$, then

$$\hat{\beta} = \tilde{\beta} + N^{-1/2}\hat{I}^{-1}Q_N(\tilde{\beta},\tilde{\phi}).$$

Thus, if we define

$$R_N(\zeta,\theta) = N^{-1/2}\sum_{i=1}^{N}\{\tilde{X}_i(\theta)-\tilde{X}_\cdot(\theta)\}\left[\frac{\hat{f}'}{\hat{f}}(W_i(\theta,\zeta)) - \frac{f_N'}{f_N}\Big(W_i(\theta,\zeta);\frac{\sigma^2}{T(\theta)}\Big)\right],$$

then the proof of the theorem is reduced to the proofs of

$$\hat{I} - I_N \xrightarrow{p} 0, \tag{3.6}$$

$$R_N(\tilde{\beta},\tilde{\phi}) \xrightarrow{p} 0, \tag{3.7}$$

$$\Big|Q_N(\tilde{\beta},\tilde{\phi}) - Q_N(\beta,\phi) - (\tilde{\beta}-\beta)^\top\frac{\partial}{\partial\beta}Q_N(\beta,\phi) - (\tilde{\phi}-\phi)^\top\frac{\partial}{\partial\phi}Q_N(\beta,\phi)\Big| \xrightarrow{p} 0, \tag{3.8}$$

$$N^{-1/2}\frac{\partial}{\partial\beta}Q_N(\beta,\phi) + I_N \xrightarrow{p} 0, \tag{3.9}$$

$$N^{-1/2}\frac{\partial}{\partial\phi}Q_N(\beta,\phi) \xrightarrow{p} 0. \tag{3.10}$$

The proofs of (3.6)–(3.10) can be done similarly as in Park *et al.* (2003). For instance, to prove (3.10) one may use $E(\tilde{W}_j|W) = W$ and

$$\frac{\partial}{\partial\phi_j}W(\phi,\beta) = \frac{S_j(\phi)}{T(\phi)}\left\{W(\phi,\beta) - \tilde{W}_j(\phi,\beta)\right\},$$

$$E\left[\frac{\partial}{\partial\phi_j}\{\tilde{X}_i(\phi) - \tilde{X}_\cdot(\phi)\}\right] = 0,$$

where W_j is defined at (2.7).

3.2. Model 2

We claim that a GLS within estimator of β is semiparametric efficient. To define the GLS within estimator, let

$$X_{it}^*(\phi) = \begin{cases} X_{it} - \tilde{X}_i(\phi), & \text{if } 1 \le t \le p, \\ \tilde{\phi}(B)(X_{it} - \tilde{X}_i(\phi)), & \text{if } p+1 \le t \le T. \end{cases}$$

Let $\tilde{Y}(\phi) = \sum_{t=1}^T w_t(\phi)Y_t$ and define

$$Y_{it}^*(\phi) = \begin{cases} Y_{it} - \tilde{Y}_i(\phi), & \text{if } 1 \le t \le p, \\ \tilde{\phi}(B)(Y_{it} - \tilde{Y}_i(\phi)), & \text{if } p+1 \le t \le T. \end{cases}$$

Then the GLS within estimator of β is defined by

$$\hat{\beta}_{GLS} = \left[\sum_{i=1}^{N}\Big\{\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\tilde{\phi})X_{it}^{*}(\tilde{\phi})X_{is}^{*}(\tilde{\phi})^{\top} + \sum_{t=p+1}^{T} X_{it}^{*}(\tilde{\phi})X_{it}^{*}(\tilde{\phi})^{\top}\Big\}\right]^{-1} \times \left[\sum_{i=1}^{N}\Big\{\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\tilde{\phi})X_{it}^{*}(\tilde{\phi})Y_{is}^{*}(\tilde{\phi}) + \sum_{t=p+1}^{T} X_{it}^{*}(\tilde{\phi})Y_{it}^{*}(\tilde{\phi})\Big\}\right].$$

Using (2.2) one may verify

$$\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\tilde{\phi})X_{it}^{*}(\tilde{\phi}) + \{\tilde{\phi}(B)1\}\sum_{t=p+1}^{T} X_{it}^{*}(\tilde{\phi}) = 0.$$

This entails

$$\tilde{\beta}_{GLS} = \beta + \left[\frac{1}{N}\sum_{i=1}^{N}\Big\{\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\tilde{\phi})X_{it}^{*}(\tilde{\phi})X_{is}^{*}(\tilde{\phi})^{\top} + \sum_{t=p+1}^{T} X_{it}^{*}(\tilde{\phi})X_{it}^{*}(\tilde{\phi})^{\top}\Big\}\right]^{-1} \times \left[\frac{1}{N}\sum_{i=1}^{N}\Big\{\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\tilde{\phi})X_{it}^{*}(\tilde{\phi})\varepsilon_{is} + \sum_{t=p+1}^{T} X_{it}^{*}(\tilde{\phi})\{\tilde{\phi}(B)\varepsilon_{it}\Big\}\right]. \qquad (3.11)$$

Define β_{GLS} to be the right hand side of (3.11) with $\tilde{\phi}$ being replaced by the true ϕ, *i.e.*,

$$\beta_{GLS} = \beta + \left[\frac{1}{N}\sum_{i=1}^{N}\Big\{\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)X_{it}^{*}(\phi)X_{is}^{*}(\phi)^{\top} + \sum_{t=p+1}^{T} X_{it}^{*}(\phi)X_{it}^{*}(\phi)^{\top}\Big\}\right]^{-1} \times \left[\frac{1}{N}\sum_{i=1}^{N}\Big\{\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)X_{it}^{*}(\phi)\varepsilon_{is} + \sum_{t=p+1}^{T} X_{it}^{*}(\phi)u_{it}\right].$$

Note that

$$\mathrm{Var}\left(\sum_{t=1}^{p}\sum_{s=1}^{p} c^{ts}(\phi)X_{it}^{*}(\phi)\varepsilon_{is} + \sum_{t=p+1}^{T} X_{it}^{*}(\phi)u_{it}\,\Big|X_1,\ldots,X_T\right) = \sigma^2\left[\sum_{t=1}^{p}\sum_{t'=1}^{p}\Big\{\sum_{s=1}^{p}\sum_{s'=1}^{p} c^{ts}(\phi)c^{t's'}(\phi)C_{ss'}(\phi)\Big\}X_{it}^{*}(\phi)X_{it'}^{*}(\phi)^{\top}\right.$$

$$+ \sum_{t=p+1}^{p} X_{it}^*(\phi) X_{it}^*(\phi)^\top \Bigg]$$

$$= \sigma^2 \left[\sum_{t=1}^{p} \sum_{t'=1}^{p} c^{tt'}(\phi) X_{it}^*(\phi) X_{it'}^*(\phi)^\top + \sum_{t=p+1}^{p} X_{it}^*(\phi) X_{it}^*(\phi)^\top \right].$$

Thus, the conditional variance of β_{GLS} equals $\Sigma_1(\phi)^{-1}\sigma^2$, so that

$$\sqrt{N}(\beta_{GLS} - \beta) \xrightarrow{d} N(0, \sigma^2 \Sigma_1^{-1}(\phi)).$$

From the consistency of $\tilde{\phi}$ it can be verified that $\sqrt{N}(\tilde{\beta}_{GLS} - \beta_{GLS}) \to 0$ in probability. Thus, we have the following theorem, which implies in view of Theorem 2.2 that $\tilde{\beta}_{GLS}$ is semiparametric efficient for Model 2.

THEOREM 3.2. *Under the conditions of Theorem 2.2, we have*

$$\sqrt{N}(\tilde{\beta}_{GLS} - \beta) \xrightarrow{d} N(0, \sigma^2 \Sigma_1^{-1}(\phi)).$$

REFERENCES

AHN, S. C. AND SCHMIDT, P. (1995). "Efficient estimation of models for dynamic panel data", *Journal of Econometrics*, **68**, 5–27.

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

ARELLANO, M. AND BOVER, O. (1995). "Another look at the instrumental variable estimation of error-components models", *Journal of Econometrics*, **68**, 29–51.

BALTAGI, B. (1995). *Econometric Analysis of Panel Data*, John Wiley & Sons, New York.

BEGUN, J. M., HALL, W. J., HUANG, W. M. AND WELLNER, J. A. (1983). "Information and asymptotic efficiency in parametric-nonparametric models", *The Annals of Statistics*, **11**, 432–452.

BICKEL, P. J. (1982). "On adaptive estimation", *The Annals of Statistics*, **10**, 647–671.

BICKEL, P. J. AND RITOV, Y. (1987). "Efficient estimation in the errors in variables model", *The Annals of Statistics*, **15**, 513–540.

BICKEL, P. J., KLAASSEN, C. A. J., RITOV, Y. AND WELLNER, J. A. (1993). *Efficient and Adaptive Estimation in Non- and Semi-parametric Models*, Johns Hopkins University Press, Baltimore.

HSIAO, C. (2007). "Panel data analysis–advantages and challenges (with discussion)", *Test*, **16**, 1–57.

MÁTYÁS, L. AND SEVESTRE, P. (1992). *The Econometrics of Panel Data: Handbook of Theory and Applications*, Kluwer Academic Publishers, London.

PARK, B. U. (1990). "Efficient estimation in the two sample semiparametric location-scale model", *Probability Theory and Related Fields*, **86**, 21–39.

PARK, B. U. AND SIMAR, L. (1994). "Efficient semiparametric estimation in a stochastic frontier models", *Journal of the American Statistical Association*, **89**, 929–936.

PARK, B. U., SICKLES, R. C. AND SIMAR, L. (1998). "Stochastic frontiers: A semiparametric approach", *Journal of Econometrics*, **84**, 273–301.

PARK, B. U., SICKLES, R. C. AND SIMAR, L. (2003). "Semiparametric-efficient estimation of AR(1) panel data models", *Journal of Econometrics*, **117**, 279–309.

PARK, B. U., SICKLES, R. C. AND SIMAR, L. (2007). "Semiparametric efficient estimation of dynamic panel data models", *Journal of Econometrics,* **136,** 281–301.