Online appendix to "Estimation of a Nonlinear Panel Data Model with Semiparametric Individual Effects"

Wayne-Roy Gayle[a], Soiliou Daw Namoro[b]

[a]*Department of Economics, University of Virginia, Monroe Hall, McCormick Rd, Room 208A, Charlottesville, VA 22903*

[b]*Department of Economics, University of Pittsburgh, 230 S. Bouquet St., Pittsburgh, PA 15260*

Abstract

In this online appendix, we prove Lemma 6.5 and 6.10.

1. Proof of Lemma 6.5

Proof. The proof borrows from Newey (1994) and Severini and Wong (1992). Let $\omega_{it}(P_t) := \sigma_2^{-1}K_2(\sigma_2^{-1}(P_{it0}-P_t))$, and $\omega_i(P) := \prod_{t=1}^{T}\omega_{it}(P_t)$. Let $\hat{\omega}_i(P_{-t}) := \prod_{s\neq t}\hat{\omega}_{is}(P_s)$, $\omega_i(P_{-t}) := \prod_{s\neq t}\omega_{is}(P_s)$, $K_2(u_{-t}) := \prod_{s\neq t}K_2(u_s)$, $K_2^{(r)}(u_t)$ be the r-th derivative of $K_2(u_t)$ with respect to u_t, and the vector $(P_{-t},\tilde{P}_t)$ be P with P_t replaced by $\tilde{P}_t$. Let $\otimes$ denote the Kronecker product. To prove the theorem, we introduce the following notations:

$$
\begin{aligned}
G_i(P,\theta) &:= \tau_i(\rho(x_i,P,\theta)\otimes\rho(x_i,P,\theta))',\\
\hat{G}^{(r)}(P_t,\theta) &:= \int\frac{1}{N}\sum_{i=1}^{N}G_i(P,\theta)\hat{\omega}_i(P_{-t})dP_{-t}\hat{\omega}_{it}^{(r)}(P_t),\\
G^{(r)}(P_t,\theta) &:= \int\frac{1}{N}\sum_{i=1}^{N}G_i(P,\theta)\omega_i(P_{-t})dP_{-t}\omega_{it}^{(r)}(P_t),\\
G_0^{(r)}(P_t,\theta) &:= E[G_i^{(r)}(P_{i0},\theta)|P_t]f_{P_t}(P_t).
\end{aligned}
$$

Let *vec* denote the vectorization operator. Then we have:

$$\begin{aligned}
\hat{Q}_i(P,\theta) &= G_i(P,\theta)vec(\hat{\Sigma}^{-1}),\\
\hat{Q}^{(r)}(P_t,\theta) &= \hat{G}^{(r)}(P_t,\theta)vec(\hat{\Sigma}^{-1}),\\
Q_0^{(r)}(P_t,\theta) &= G_0^{(r)}(P_t,\theta)vec(\Sigma^{-1}).
\end{aligned}$$

Let $Q^{(r)}(P_t,\theta) := G^{(r)}(P_t,\theta)vec(\Sigma^{-1})$. Note that by Assumptions 7.2, 7.3, and $\|\tau_i x_i\| \leq \sup_i \|\tau_i x_i\| \leq \sup_{w\in\mathcal{W}} \|w\| < \infty$ (because $\mathcal{W}$ is compact),

$$\sup_{\theta,P,w_i} \left\| \frac{\partial^{k+j}}{\partial\beta^k \partial\varphi_t^j} G_i(P,\theta) \right\| < \infty. \tag{1.1}$$

Also

$$\begin{aligned}
&E\left[\sup_{\theta} \left\| \int \frac{\partial^{k+j}}{\partial\beta^k \partial\varphi_t^j} G_i(P,\theta)\omega_i(P_{-t})dP_{-t} \right\|\right]\\
&= E\left[\sup_{\theta} \left\| \int \frac{\partial^{k+j}}{\partial\beta^k \partial\varphi_t^j} G_i(P_{i,-t}+\sigma_2 u_{-t}, P_t,\theta)K_2(u_{-t})du_{-t}. \right\|\right]\\
&\leq \sup_{\theta,P,w_i} \left\| \frac{\partial^{k+j}}{\partial\beta^k \partial\varphi_t^j} G_i(P,\theta) \right\| < \infty.
\end{aligned} \tag{1.2}$$

Note also, that

$$\frac{\partial^{k+j}}{\partial\beta^k \partial\varphi_t^j} G^{(r)}(P_t,\theta) = \frac{1}{N}\sum_{i=1}^{N} \int \frac{\partial^{k+j}}{\partial\beta^k \partial\varphi_t^j} G_i(P,\theta)\omega_i(P_{-t})dP_{-t}\omega_{it}^{(r)}(P_t).$$

We first prove the result for $k=j=0$. Define $H^{(r)}(P_t,\theta) = G^{(r)}(P_t,\theta) - E[G^{(r)}(P_t,\theta)]$. For any $P_t, \tilde{P}_t \in \mathcal{K}$, $\beta, \tilde{\beta} \in B$, and $\varphi, \tilde{\varphi} \in \prod_{t=1}^{T} \mathcal{S}_{\mathcal{K}}$, by equation (1.2), and by $K_2^{(r)}$ Lipschitz, three i.i.d. random variables $\{(M_i^l,\ i=1,\cdots,N),\ l=1,2,3,\ \}$ exist with $(M_i^1,\ i=1,\cdots,N)$ not

depending on $(P_t, \tilde{P}_t)$, $(M_i^2,\ i = 1, \cdots, N)$ not depending on $(\beta, \tilde{\beta})$, and $(M_i^3,\ i = 1, \cdots, N)$ not depending on $(\varphi, \tilde{\varphi})$, such that $E[|M_i^l|] < \infty, l = 1, 2, 3$, and with probability approaching one,

$$\begin{aligned}
\sup_{\beta,\varphi} \left\| H^r(P_t, \beta, \varphi) - H^r(\tilde{P}_t, \beta, \varphi) \right\| &\leq \sigma_2^{-r-2} |P_t - \tilde{P}_t| \frac{1}{N} \sum_{i=1}^N M_i^1, \\
\sup_{P_t,\varphi} \left\| H^r(P_t, \beta, \varphi) - H^r(P_t, \tilde{\beta}, \varphi) \right\| &\leq \sigma_2^{-r-1} \|\beta - \tilde{\beta}\| \frac{1}{N} \sum_{i=1}^N M_i^2, \\
\sup_{P_t,\beta} \| H^r(P_t, \beta, \varphi) - H^r(P_t, \beta, \tilde{\varphi}) \| &\leq \sigma_2^{-r-1} \|\varphi - \tilde{\varphi}\|_{s,0} \frac{1}{N} \sum_{i=1}^N M_i^3.
\end{aligned} \tag{1.3}$$

Therefore, by the triangular inequality, for $\delta > 0$, for N sufficiently large, there exists $M < \infty$ such that with probability approaching one,

$$\begin{aligned}
&\sup_{|P_t - \tilde{P}_t| \leq \delta} \sup_{\|\beta - \tilde{\beta}\| \leq \delta} \sup_{\|\varphi - \tilde{\varphi}\|_{s,0} \leq \delta} \left\| H^r(P_t, \beta, \varphi) - H^r(\tilde{P}_t, \tilde{\beta}, \tilde{\varphi}) \right\| \\
&\leq \sup_{|P_t - \tilde{P}_t| \leq \delta} \left\| H^r(P_t, \beta, \varphi) - H^r(\tilde{P}_t, \beta, \varphi) \right\| + \sup_{\|\beta - \tilde{\beta}\| \leq \delta} \left\| H^r(P_t, \beta, \varphi) - H^r(P_t, \tilde{\beta}, \varphi) \right\| \\
&+ \sup_{\|\varphi - \tilde{\varphi}\|_{s,0} \leq \delta} \| H^r(P_t, \beta, \varphi) - H^r(P_t, \beta, \tilde{\varphi}) \| \\
&\leq \delta \left(\sigma_2^{-r-2} \frac{1}{N} \sum_{i=1}^N M_i^1 + \sigma_2^{-r-1} \frac{1}{N} \sum_{i=1}^N M_i^2 + \sigma_2^{-r-1} \frac{1}{N} \sum_{i=1}^N M_i^3 \right) \\
&\leq \delta \sigma_2^{-r-2} \frac{1}{N} \sum_{i=1}^N (M_i^1 + M_i^2 + M_i^3) \leq M \delta \sigma_2^{-r-2}.
\end{aligned} \tag{1.4}$$

Because $\mathcal{K} \times \mathcal{B} \times \prod_{t=1}^T \mathcal{S}_{\mathcal{K}}$ is compact, it can be covered by $C\delta_l^{-K-T-1}$ open balls of radios δ_l. Let $(P_{tl}, \beta_l, \varphi_l)$ be the center of these balls, $l = 1, \cdots, L \leq C_1 \delta_l^{-K-T-1}$. Then by the triangular inequality and equation (1.4), for $(P_{tl}(P_t), \beta_l(\beta), \varphi_l(\varphi))$ equal to the center of an

open ball containing (P_t, β, φ),

$$\begin{aligned}
\sup_{P_t,\beta,\varphi} \|H^{(r)}(P_t,\beta,\varphi)\| &\leq \sup_{P_t,\beta,\varphi} \|H^{(r)}(P_t,\beta,\varphi) - H^{(r)}(P_{tl}(P_t),\beta_l(\beta),\varphi_l(\varphi))\| \\
&+ \sup_{P_t,\beta,\varphi} \|H^{(r)}(P_{tl}(P_t),\beta_l(\beta),\varphi_l(\varphi))\| \\
&\leq M\delta\sigma_2^{-r-2} + \sup_l \|H^{(r)}(P_{tl}(P_t),\beta_l(\beta),\varphi_l(\varphi))\|. \qquad (1.5)
\end{aligned}$$

Thus by equation (1.5), for $\varepsilon > 0$,

$$\begin{aligned}
&Pr\left\{\sup_{P_t,\beta,\varphi} |H^{(r)}(P_t,\beta,\varphi)| > M\varepsilon\right\} \\
&\leq Pr\left\{\sup_l \|H^{(r)}(P_{tl}(P_t),\beta_l(\beta),\varphi_l(\varphi))\| > M(\varepsilon - \delta\sigma_2^{-r-2})\right\} \\
&\leq \sum_{l=1}^{L} Pr\left\{\|H^{(r)}(P_{tl}(P_t),\beta_l(\beta),\varphi_l(\varphi))\| > M(\varepsilon - \delta\sigma_2^{-r-2})\right\}. \qquad (1.6)
\end{aligned}$$

By the bounded kernel, by equation (1.1), and for C_2^{-1} and N large enough

$$\begin{aligned}
&E\left[\left\|\left(\int G_i(P,\theta)\omega_i(P_{-t})dP_{-t}\omega_{it}^{(r)}(P_t)\right)^2\right\|\right] \\
&\leq \int\left(\int E\left[\|G_i(P,\theta)\|^2 |P_{i0}\right] f_P(P_{i0})\omega_i(P_{-t})^2\omega_{it}^{(r)}(P_t)^2 dP_{i0}\right) dP_{-t} \\
&= \sigma^{-2r-1}\int\int E\left[\|G_i(P,\theta)\|^2 |P+\sigma_2 u\right] f_P(P+\sigma_2 u)K_2(u_{-t})^2 K_2^{(r)}(u_t)^2 du dP_{-t} \\
&\leq \frac{1}{8}C_2^{-1}\sigma_2^{-1-2r}. \qquad (1.7)
\end{aligned}$$

Also, by the same argument,

$$\left\|\int G_i(P,\theta)\omega_i(P_{it})dP_{-t}\omega_{it}^{(r)}(P_t)\right\| \leq \frac{3}{4}C_2^{-1}\sigma_2^{-1-r}, \qquad (1.8)$$

for C_2^{-1} and N large enough. Note that by $\ln(N)/N\sigma_2 \to 0$, $\sigma_2^{-1} < C_3 N$ for C_3 large enough. Hence, taking $\delta_l = \varepsilon\sigma_2^{r+2}/2$ and $\varepsilon = ((\ln N)/N\sigma_2^{2r+1})^{1/2}$, by equations (1.6) - (1.8), by Bernstein's inequality, and by $\ln(N)/N\sigma_2 \to 0$, for M and N large enough,

$$
\begin{aligned}
& Pr\left\{\sup_{P_t,\beta,\varphi} |H^{(r)}(P_t,\beta,\varphi)| > M\varepsilon\right\} \\
& \leq 2\sum_{l=1}^{L} \exp\left(-\frac{1}{2}N^2M^2(\varepsilon-\delta_l\sigma_2^{-r-2})^2 / \left[\frac{1}{8}C_2^{-1}N\sigma_2^{-1-2r} + \frac{1}{3}\frac{3}{4}C_2^{-1}N\sigma_2^{-1-r}(\varepsilon-\delta_l\sigma_2^{-r-2})\right]\right) \\
& \leq C_4(\varepsilon\sigma_2^{r+2})^{-(K+T+1)} \exp\left(-NC_2M^2\varepsilon^2\sigma_2^{1+2r}/[1+\sigma_2^r\varepsilon]\right) \\
& \leq C_4(\ln(N))^{-(K+T+1)/2}N^{(K+T+1)/2}\sigma_2^{-3(K+T+1)/2} \exp\left(-C_2M^2\ln(N)/[1+(\ln(N)/N\sigma_2)^{1/2}]\right) \\
& \leq C(\ln(N))^{-(K+T+1)/2}N^{2(K+T+1)} \exp\left(-C_2M^2\ln(N)/[1+(\ln(N)/N\sigma_2)^{1/2}]\right) \\
& \leq C(\ln(N))^{-(K+T+1)/2} \exp\left([1+(\ln(N)/N\sigma_2)^{1/2}]^{-1}\right) \\
& \times \exp\left(-[C_2M^2 - 2(K+T+1)]\ln(N)\right),
\end{aligned}
$$

which obtains

$$
Pr\left\{\sup_{P_t,\beta,\varphi} |H^{(r)}(P_t,\beta,\varphi)| > M\varepsilon\right\} = o(1)
$$

as $N \to \infty$. Hence,

$$
\sup_{P_t,\beta,\varphi} \|G^{(r)}(P_t,\theta) - E[G^{(r)}(P_t,\theta)]\| = O_p(\ln(N)^{1/2}(N\sigma_2^{2r+1})^{-1/2}). \tag{1.9}
$$

Next consider $E\left[G^{(r)}(P_t,\theta)\right]-G_0^{(r)}(P_t,\theta)$:

$$
\begin{aligned}
&E\left[G^{(r)}(P_t,\theta)\right]=E\left[\int G_i(P,\theta)\omega_i(P_{-t})dP_{-t}\omega_{it}^{(r)}(P_t)\right]\\
&=E\left[E\left[\int G_i(P_{i,-t0}-\sigma_2 u_{-t},P_t,\theta)K_2(u_{-t})du_{-t}\middle|P_{it0}\right]\omega_{it}^{(r)}(P_t)\right]\\
&=\int E\left[\int G_i(P_{i0}-\sigma_2 u,\theta)K_2(u_{-t})du_{-t}\middle|P_t+\sigma_2 u_t\right]f_{P_t}(P_t+\sigma_2 u_t)K_2^{(r)}(u_t)/\sigma_2^r du_t\\
&=\int\left\{\frac{\partial^r}{\partial P_t^r}E\left[\int G_i(P_{i0}-\sigma_2 u,\theta)K_2(u_{-t})du_{-t}\middle|P_t+\sigma_2 u_t\right]f_{P_t}(P_t+\sigma_2 u_t)\right\}K_2(u_t)du_t\\
&=G_0^{(r)}(P_t,\theta)+O_p(\sigma_2^2) \qquad (1.10)
\end{aligned}
$$

uniformly in $\Theta\times\mathcal{W}$. It follows from equations (1.9), (1.10), and the triangular inequality that

$$
\sup_{\theta,P_t}\left\|G^{(r)}(P_t,\theta)-G_0^{(r)}(P_t,\theta)\right\|=O_p\left(\ln(N)^{1/2}(N\sigma_2^{2r+1})^{-1/2}+\sigma_2^2\right). \qquad (1.11)
$$

Next consider $\|\hat{G}^{(r)}(P_t,\theta)-G^{(r)}(P_t,\theta)\|$. For $t=1,\cdots,T$, define the parameters $\alpha_t\in(0,1)$. By the mean value expansion, $\hat{G}^{(r)}(P_t,\theta)=G^{(r)}(P_t,\theta)+T_{1t}+T_{2t}$, where $T_{1t}=\sum_{i=1}^N T_{1it}/N$, $T_{2t}=\sum_{i=1}^N T_{2it}/N$,

$$
T_{1it}=\int G_i(P,\theta)\prod_{s\neq t}\frac{1}{\sigma_2}K\left(\frac{P_{is0}+\alpha_s a_{is}-P_s}{\sigma_2}\right)dP_{-t}\frac{1}{\sigma_2^{r+2}}K_2^{(r+1)}\left(\frac{P_{it0}+\alpha_t a_{it}-P_t}{\sigma_2}\right)a_{it},\text{ and}
$$

$$
\begin{aligned}
T_{2it}=&\int G_i(P,\theta)\left(\sum_{s\neq t}\frac{1}{\sigma_2^2}K^{(1)}\left(\frac{P_{is0}+\alpha_s a_{is}-P_s}{\sigma_2}\right)a_{is}\prod_{s'\neq s}\frac{1}{\sigma_2}K\left(\frac{P_{is'0}+\alpha_{s'}a_{is'}-P_{s'}}{\sigma_2}\right)dP_{-t}\right)\\
&\times\frac{1}{\sigma_2^{r+1}}K^{(r)}\left(\frac{P_{it0}+\alpha_t a_{it}-P_t}{\sigma_2}\right).
\end{aligned}
$$

Under the assumptions of the theorem, change of variables, mean value expansion, and

integration by parts obtain

$$
\begin{aligned}
&\|E[T_{1it}]\| = \\
&= \left\| E\left[\int\int E[G_i(P,\theta)|P_{i0},a_i] \prod_{s\neq t} \frac{1}{\sigma_2} K\left(\frac{P_{is0}+\alpha_s a_{is}-P_s}{\sigma_2} \right) \right.\right. \\
&\times \left.\left. \sigma_2^{-r-2} K_2\left(\frac{P_{it0}+\alpha_t a_{it}-P_t}{\sigma_2} \right) f_{P|a}(P_{i0}|a_i) a_{it} dP_{i0} dP_{-t} \right] \right\| \\
&= \left\| E\left[\int\int \frac{\partial^{r+1}}{\partial u_t^{r+1}} \Big[E[G_i(P,\theta)|\{P_s-\alpha_s a_{is}+\sigma_2 u_s\},a_i] \right.\right. \\
&\times \left.\left. f_{P|a}(\{P_s-\alpha_s a_{is}+\sigma_2 u_s\}|a_i) \Big] a_{it} K_2(u) du dP_{-t} \right] \right\| \\
&= O_p(\sup_w \|\hat{P}(w)-P_0(w)\|)
\end{aligned}
$$

uniformly over $\Theta \times \mathcal{W}$, so that $\sup_{\theta,P_t} \|E[T_{1t}]\| = O_p(\sup_w \|\hat{P}(w)-P_0(w)\|)$. Similar calculations show $\sup_{\theta,P_t} \|E[T_{2t}]\| = O_p(\sup_w \|\hat{P}(w)-P_0(w)\|)$. Consider

$$
\|E[(T_{1t}+T_{2t})'(T_{1t}+T_{2t})]\| \leq \|E[T_{1t}'T_{1t}]\| + \|E[T_{2t}'T_{2t}]\| + \|E[T_{1t}'T_{2t}]\| + \|E[T_{2t}'T_{1t}]\|.
$$

By the triangular, and Cauchy-Schwarz inequalities,

$$
\|E[T_{1t}'T_{1t}]\| \leq \frac{1}{N}\sum_{i=1}^{N} E[\|T_{1it}\|^2]^{1/2} \frac{1}{N}\sum_{j=1}^{N} E[\|T_{1jt}\|^2]^{1/2}. \tag{1.12}
$$

Now,

$$
\begin{aligned}
&E[\|T_{1it}\|^2] \\
&\leq E\left\{\int\int E\left[\|G_i(P_t,\{P_{is0}+\alpha_s a_{is}-\sigma_2 u_s : s\neq t\},\theta)\|^2 | P_{it0}, a_i\right] f_{P|a}(P_{it0}|a_i)\right. \\
&\left.\times \prod_{s\neq t} K_2(u_s)^2 du_{-t}\sigma_2^{-2r-4} K_2^{(r+1)}\left(\frac{P_{it0}+\alpha_t a_{it}-P_t}{\sigma_2}\right)^2 a_{it}^2 dP_{it0}\right\} \\
&= \sigma^{-2r-3} E\left\{\int\int E\left[\|G_i(P_t,\{P_{is0}+\alpha_s a_{is}-\sigma_2 u_s : s\neq t\},\theta)\|^2 | P_t-\alpha_t a_{it}+\sigma_2 u_t, a_i\right]\right. \\
&\left.\times\ f_{P|a}(P_t-\alpha_t a_{it}+\sigma_2 u_t|a_i)\prod_{s\neq t} K_2(u_s)^2 du_{-t} K_2^{(r+1)}(u_t)^2 a_{it}^2 du_t\right\} \\
&= O_p(\sigma_2^{-2r-3}\sup_w \|\hat{P}(w)-P_0(w)\|^2),
\end{aligned}
$$

uniformly over $\Theta\times P$, so that $\sup_{\theta,P_t}\|E[T_{1t}'T_{1t}]\| = O_p(\sigma_2^{-2r-3}\sup_w\|\hat{P}(w)-P_0(w)\|^2)$. Similar calculations show the last three terms on the RHS of equation 1.12 obtain the same uniform rate of convergence. The results of Newey (1994) and Newey and McFadden (1994) show that, under Assumption 7.3, $\sup_w\|\hat{P}(w)-P_0(w)\| = O_p((\ln(N)/N\sigma_1^{K+L})^{1/2}+\sigma_1^m)$. Therefore, by the triangular inequality,

$$
\begin{aligned}
&\sup_{P_t}\sup_{\theta}\|\hat{G}^{(r)}(P_t,\theta)-G^{(r)}(P_t,\theta)\| = \\
&O_p((\ln(N)/(N\sigma_1^{K+L}))^{1/2}+\sigma_1^m+\ln(N)/(N\sigma_2^{2r+3}\sigma_1^{K+L})+\sigma_1^{2m}/\sigma_2^{2r+3}). \qquad (1.13)
\end{aligned}
$$

Equations (1.11), (1.13), and the triangular inequality imply

$$
\begin{aligned}
&\sup_{P_t,\theta}\left\|\hat{Q}^{(r)}(P_t,\theta)-Q_0^{(r)}(P_t,\theta)\right\| \\
&\leq \sup_{P_t,\theta}\left\|\hat{Q}^{(r)}(P_t,\theta)-Q^{(r)}(P_t,\theta)\right\|+\sup_{P_t,\theta}\left\|Q^{(r)}(P_t,\theta)-Q_0^{(r)}(P_t,\theta)\right\| \\
&\leq \left\|\hat{\Sigma}^{-1}\right\|\sup_{P_t,\theta}\left\|\hat{G}^{(r)}(P_t,\theta)-G^{(r)}(P_t,\theta)\right\|+\sup_{P_t,\theta}\left\|G^{(r)}(P_t,\theta)\right\|\left\|\hat{\Sigma}^{-1}-\Sigma^{-1}\right\| \\
&+\|\Sigma^{-1}\|\sup_{P_t,\theta}\left\|G^{(r)}(P_t,\theta)-G_0^{(r)}(P_t,\theta)\right\| \\
&=O_p\left(\ln(N)^{1/2}(N\sigma_2^{2r+1})^{-1/2}+\sigma_2^2+\ln(N)^{1/2}(N\sigma_1^{K+L})^{-1/2}+\sigma_1^m\right. \\
&\left.+(\ln(N)/(N\sigma_2^{2r+3}\sigma_1^{K+L}))+\sigma_1^{2m}/\sigma_2^{2r+3}+\left\|\hat{\Sigma}^{-1}-\Sigma^{-1}\right\|\right).
\end{aligned}
$$

By equation (1.2), the proofs for $j=1$, $k=1,2$, $k+r\leq 1$ follow the same arguments. Also, By the same arguments, it can be shown that

$$
\begin{aligned}
&\sup_{\tilde{P}_t}\left\|\hat{f}_{P_t}(\tilde{P}_t)-f_{P_t}(\tilde{P}_t)\right\|= \\
&O_p\left(\ln(N)^{1/2}(N\sigma_2)^{-1/2}+\sigma_2^2+\ln(N)^{1/2}(N\sigma_1^{K+L})^{-1/2}+\sigma_1^m\right. \\
&\left.+\ln(N)/(N\sigma_2^3\sigma_1^{K+L})+\sigma_1^{2m}/\sigma_2^3\right), \quad \text{and}
\end{aligned}
$$

$$
\begin{aligned}
&\sup_{\tilde{P}_t,\tilde{P}_s}\left\|\hat{f}_{t,s}(\tilde{P}_t,\tilde{P}_s)-f_{t,s}(\tilde{P}_t,\tilde{P}_s)\right\|= \\
&O_p\left(\ln(N)^{1/2}(N\sigma_2^2)^{-1/2}+\sigma_2^2+\ln(N)^{1/2}(N\sigma_1^{K+L})^{-1/2}+\sigma_1^m\right. \\
&\left.+\ln(N)/(N\sigma_2^4\sigma_1^{K+L})+\sigma_1^{2m}/\sigma_2^4\right).
\end{aligned}
$$

□

2. Proof of Theorem 6.10

Proof. By the mean value expansion,

$$\begin{aligned} 0 &= \frac{\partial}{\partial\beta}\hat{Q}(\hat{\beta},\hat{\varphi}(\hat{\beta})) \\ &= \frac{\partial}{\partial\beta}\hat{Q}(\beta_0,\hat{\varphi}(\beta_0)) + \frac{\partial^2}{\partial\beta^2}\hat{Q}(\bar{\beta},\hat{\varphi}(\bar{\beta}))[\hat{\beta}-\beta_0], \end{aligned}$$

where $\bar{\beta}$ lies between $\hat{\beta}$ and β_0. Also, because $Q_{\varphi_t,0}(\beta_0,\varphi_0(\beta_0)) = 0$, $t = 1,\cdots,T$, mean value expansion of $\hat{Q}(\beta_0,\hat{\varphi}(\beta_0))$ obtains

$$\begin{aligned} \hat{Q}(\beta_0,\hat{\varphi}(\beta_0)) = \hat{Q}(\beta_0,\varphi_0(\beta_0)) &+ \sum_{t=1}^{T}(\hat{Q}_{\varphi_t}(\beta_0,\varphi_0(\beta_0)) - Q_{\varphi_t,0}(\beta_0,\varphi_0(\beta_0)))(\hat{\varphi}_t(\beta_0)-\varphi_{t0}(\beta_0)) \\ &+ \frac{1}{2}\sum_{t=1}^{T}\sum_{s=1}^{T}\hat{Q}_{\varphi_t\varphi_s}(\beta_0,\bar{\varphi}(\beta_0))(\hat{\varphi}_t(\beta_0)-\varphi_{t0}(\beta_0))(\hat{\varphi}_s(\beta_0)-\varphi_{s0}(\beta_0)), \end{aligned}$$

where $\hat{Q}_{\varphi_t\varphi_s}(\beta_0,\bar{\varphi}(\beta_0)) = \partial^2\hat{Q}(\beta_0,\bar{\varphi}(\beta_0))/\partial\varphi_t\partial\varphi_s$. Under the conditions of the theorem, and by $\hat{Q}_{\varphi_t\varphi_s}(\beta_0,\bar{\varphi}(\beta_0)) = 2\hat{\sigma}_*^{ts}\hat{f}_{t,s}(P_t,P_s)$, the above equation implies

$$\frac{\partial}{\partial\beta}\hat{Q}(\beta_0,\hat{\varphi}(\beta_0)) = \frac{\partial}{\partial\beta}\hat{Q}(\beta_0,\varphi_0(\beta_0)) + o_p(N^{-1/2}).$$

Thus

$$\sqrt{N}(\hat{\beta}-\beta_0) = -\left[\frac{\partial^2}{\partial\beta^2}\hat{Q}(\bar{\beta},\hat{\varphi}(\bar{\beta}))\right]^{+}\left\{\sqrt{N}\frac{\partial}{\partial\beta}\hat{Q}(\beta_0,\varphi_0(\beta_0)) + o_P(1)\right\}$$

under the conditions of the theorem. By standard arguments we have that

$$\int\frac{\partial}{\partial\beta}\hat{Q}_i(P,\theta_0)\omega_i(P)dP = \frac{\partial}{\partial\beta}\hat{Q}_i(\hat{P}_i,\theta_0) + O_p(\sigma_2^2). \tag{2.1}$$

Also, under the conditions of the theorem and because $\rho(x_i, P_{i0}, \theta_0) = 0$,

$$\begin{aligned}
\frac{\partial}{\partial\beta}\hat{Q}_i(\hat{P}_i,\theta_0) &= 2\tau_i\left[\frac{\partial}{\partial\beta}\rho(x_i,\hat{P}_i,\theta_0)\right]'\hat{\Sigma}^{-1}\rho(x_i,\hat{P}_i,\theta_0)\\
&= 2\tau_i\left[\frac{\partial}{\partial\beta}\rho(x_i,P_{i0},\theta_0)\right]'\hat{\Sigma}^{-1}\left[\frac{\partial}{\partial P}\rho(x_i,P_{i0},\theta_0)\right](\hat{P}_i-P_{i0})+O_p(\|\hat{P}_i-P_{i0}\|_{s,0}^2)\\
&= 2\tau_i\left[\frac{\partial}{\partial\beta}\rho(x_i,P_{i0},\theta_0)\right]'\Sigma^{-1}\left[\frac{\partial}{\partial P}\rho(x_i,P_{i0},\theta_0)\right](\hat{P}_i-P_{i0})\\
&+ O_p(\|\hat{\Sigma}^{-1}-\Sigma^{-1}\|\|\hat{P}_i-P_{i0}\|_{s,0}+\|\hat{P}_i-P_{i0}\|_{s,0}^2).\\
&= 2\tau_i\left[\frac{\partial}{\partial\beta}\rho(x_i,P_{i0},\theta_0)\right]'\Sigma^{-1}\left[\frac{\partial}{\partial P}\rho(x_i,P_{i0},\theta_0)\right](\hat{P}_i-P_{i0})+o_p(1/\sqrt{N}).
\end{aligned}$$

Recall $h_{i0}(P_i) = \partial\Delta\varphi_0(P_i,\beta_0)/\partial\beta - \Delta x_i$, $h_{i0} = h_{i0}(P_{i0})$, $R(P_i) =$diag$[(-\varphi_0'(P_{it-1}),\varphi_0'(P_{it})), t = 2,\cdots,T]$ and $R_i = R(P_{i0})$. Then under the conditions of the theorem,

$$\sqrt{N}(\hat{\beta}-\beta_0) = -\left[\frac{1}{2}\frac{\partial^2}{\partial\beta^2}\hat{Q}(\bar{\beta},\hat{\varphi}(\bar{\beta}))\right]^+\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\tau_i h_{i0}'\Sigma^{-1}R_i(\hat{P}_i-P_{i0})+o_P(1). \quad (2.2)$$

Let $\hat{\gamma}_{it} := \hat{\gamma}_t(w_{it})$ and $\hat{\gamma}_i := (\hat{\gamma}_{i1},\cdots,\hat{\gamma}_{iT})$. Define γ_{it0} and γ_{i0} analogously. Define $g(x_i,\hat{\gamma}_i) = \tau_i h_{i0}'\Sigma^{-1}R_i(\hat{P}_i-P_{i0})$. The rest of this section of the proof involves checking conditions (i)-(iv) of Theorem 8.11 of Newey and McFadden (1994). Notice that $g(x_i,\gamma_{i0}) = 0$, implying that $E[g(x_i,\gamma_{i0})] = 0$ and $E[\|g(x_i,\gamma_{i0})\|^2] = 0$. Linearizing $g(x_i,\hat{\gamma}_i)$ around γ_{i0} gives $D(w_i,\hat{\gamma}-\gamma_0) := \tau_i h_{i0}\Sigma^{-1}R_i\underline{f}^{-1}(w_i)G_i[\hat{\gamma}(w_i)-\gamma_0(w_i)]$, where

$$\begin{aligned}
\underline{f}^{-1}(w_i) &:= \text{diag}\left(f^{-1}(w_{it}), t = 1,\cdots,T\right)\\
G_i &:= \text{diag}((-P_{it0}\ 1), t = 1,\cdots,T)\\
\hat{\gamma}(w_i) &:= (\hat{\gamma}(w_{i1}),\cdots,\hat{\gamma}(w_{iT}))'.
\end{aligned}$$

Conditions (i) and (ii) of Theorem 8.11 of Newey and McFadden (1994) are satisfied by noting that boundedness of $\tau_i \Delta x_i$ (recall that $\mathcal{W}$ is compact), of γ_0 and its first two derivatives on $\mathcal{K}$, and of Σ^{-1} gives

$$\|g(x_i, \hat{\gamma}_i) - D(w_i, \hat{\gamma} - \gamma_0)\| \leq b(w) \|\hat{\gamma}(w_i) - \gamma_0(w_i)\|^2,$$

with $E[\tau(w)b(w)] < \infty$, and

$$D(w, \gamma) = \tau_i h_{i0}' \Sigma^{-1} R_i \underline{f}^{-1}(w_i) G_i \gamma \leq c(w) \|\gamma\|$$

with $E[\tau(w)c(w)^2] < \infty$. Condition (iii) is also immediately satisfied by observing that

$$\int D(w, \gamma) f_w(w) \mathrm{d}w = \int \nu(w)\gamma(w) \mathrm{d}w,$$

where $\nu(w_i) := \tau_i h_{i0}' \Sigma^{-1} R_i G_i$. Condition (iv) of Newey and McFadden (1994) follows from continuity of $v(w)$ on $\mathcal{W}$, the compactness of $\mathcal{B}$ and $\mathcal{W}$, the uniform boundedness conditions of P_{i0}, φ_{t0} on $\mathcal{K}$ and $\mathcal{S}_{\mathcal{K}}$, and the nonsingularity of Σ. Therefore, by Theorem 8.11 of Newey and McFadden (1994),

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} g(w_i, \hat{\gamma}) \xrightarrow{d} N(0, W), \tag{2.3}$$

where $W = \mathrm{Var}(\delta(w))$, $\delta(w) = \nu(w)q - E[\nu(w)q]$, $q = (q_1, \cdots, q_T)'$. Application of the law of iterated expectations gives $E[\nu(w)q] = 0$. Also, straightforward calculations show $\nu(w_i)q_i = \tau_i h_{i0} \Sigma^{-1} R_i \varepsilon_i$, where $\varepsilon_i = (y_i - P_{i0})$. Therefore,

$$W = E[\tau_i h_{i0}' \Sigma^{-1} R_i \varepsilon_i \varepsilon_i' R_i' \Sigma^{-1} h_{i0}].$$

Under the conditions of the theorem, Theorems 7.5, 7.8, and 7.9 obtain

$$\begin{aligned}
\frac{\partial^2}{\partial\beta^2}\hat{Q}(\bar{\beta},\hat{\varphi}(\bar{\beta})) &= \frac{\partial^2}{\partial\beta^2}Q_0(\bar{\beta},\hat{\varphi}(\bar{\beta})) + o_p(N^{-1/4}),\\
&= \frac{\partial^2}{\partial\beta^2}Q_0(\bar{\beta},\varphi_0(\bar{\beta})) + \frac{\partial^3}{\partial\beta^2\partial\varphi}Q_0(\bar{\beta},\bar{\varphi}(\bar{\beta}))(\hat{\varphi}(\bar{\beta})-\varphi_0(\bar{\beta}))) + o_p(N^{-1/4}),\\
&= \frac{\partial^2}{\partial\beta^2}Q_0(\beta_0,\varphi_0(\beta_0)) + \frac{\partial^3}{\partial\beta^3}Q_0(\tilde{\beta},\varphi_0(\tilde{\beta}))(\bar{\beta}-\beta_0) + o_p(N^{-1/4}),\\
&= \frac{\partial^2}{\partial\beta^2}Q_0(\beta_0,\varphi_0(\beta_0)) + o_p(1). \qquad (2.4)
\end{aligned}$$

Also, because $\Delta\varphi_0(P_{i0},\beta_0) - \Delta x_i\beta_0 = 0$,

$$\begin{aligned}
\frac{1}{2}\frac{\partial^2}{\partial\beta^2}Q_0(\beta_0,\varphi_0(\beta_0)) &= E\left[\tau_i h_{i0}\Sigma^{-1}h_{i0}\right]\\
&\quad + E\left[\tau_i\left(\Delta\varphi_0(P_{i0},\beta_0)-\Delta x_i\beta_0\right)'\Sigma^{-1}\otimes I_k\left(\frac{\partial^2}{\partial\beta^2}\Delta\varphi_0(P_{i0},\beta_0)\right)\right],\\
&= E\left[\tau_i h_{i0}\Sigma^{-1}h_{i0}\right].
\end{aligned}$$

Therefore,

$$\frac{1}{2}\frac{\partial^2}{\partial\beta^2}\hat{Q}(\bar{\beta},\hat{\varphi}(\bar{\beta})) = E\left[\tau_i h_{i0}\Sigma^{-1}h_{i0}\right] + o_p(1). \qquad (2.5)$$

From equations (2.2), (2.3), and (2.5), along with the continuous mapping theorem and the Slutsky theorem, we conclude

$$\sqrt{N}(\hat{\beta}-\beta_0) \xrightarrow{d} N(0,V),$$

where

$$V := E[\tau_i h_{i0}'\Sigma^{-1}h_{i0}]^{+}E[\tau_i h_{i0}'\Sigma^{-1}R_i\varepsilon_i\varepsilon_i'R_i'\Sigma^{-1}h_{i0}]E[\tau_i h_{i0}'\Sigma^{-1}h_{i0}]^{+}.$$

□

References

Newey, W. K., 1994. Kernel estimation of partial means and a general variance estimator. Econometric Theory 10 (2), 233–253.

Newey, W. K., McFadden, D., 1994. Large Sample Estimation and Hypothesis Testing. Elsevier Science Publishers.

Severini, T. A., Wong, W. H., 1992. Profile likelihood and conditionally parametric models. The Annals of Statistics 20 (4), 1768–1802.