Appendix

A Proofs of Theorems 4.1 and 4.2

Throughout the proofs, we use the following notations:

$$\mathcal{P}_n \equiv D_n P_n^L \left[P_n^{L\prime} D_n P_n^L\right]^{-1} P_n^{L\prime} D_n \quad U_n(\theta) \equiv Y_n - S_n(\lambda) Z_n \beta \quad H_n(\theta) \equiv S_n(\lambda_0) Z_n \beta_0 - S_n(\lambda) Z_n \beta$$
$$\mathcal{A}_n \equiv D_n P_n^L \left[P_n^{L\prime} D_n P_n^L\right]^{-1} \quad \widehat{\mathcal{A}}_n \equiv D_n \widehat{P}_n^L \left[\widehat{P}_n^{L\prime} D_n \widehat{P}_n^L\right]^{-1}$$

The proofs of the technical lemmas are given in Section 1 in the online supplementary material.

A.1 Proof of Theorem 4.1

Lemma A.1 *(i)* $\left\|\widehat{P}_n^{L\prime} D_n/\sqrt{n} - P_n^{L\prime} D_n/\sqrt{n}\right\| = O_P(\zeta_1(L)/\sqrt{n}) = o_P(1)$; *(ii)* $\left\|P_n^{L\prime} D_n P_n^L/n - P_n^{L\prime} \overline{D}_n P_n^L/n\right\| = O_P(\zeta_0(L) L^{1/2}/\sqrt{n}) = o_P(1)$; *(iii) for sufficiently large n,*

$$\underline{c}_d \underline{c}_P - o_P(1) \leq \tau_{\min}\left(P_n^{L\prime} D_n P_n^L/n\right) \leq \tau_{\max}\left(P_n^{L\prime} D_n P_n^L/n\right) \leq \bar{c}_P + o_P(1);$$

(iv) $\left\|\widehat{P}_n^{L\prime} D_n \widehat{P}_n^L/n - P_n^{L\prime} D_n P_n^L/n\right\| = O_P(\zeta_1(L)/\sqrt{n}) = o_P(1)$; *(v) for sufficiently large n,*

$$\underline{c}_d \underline{c}_P - o_P(1) \leq \tau_{\min}\left(\widehat{P}_n^{L\prime} D_n \widehat{P}_n^L/n\right) \leq \tau_{\max}\left(\widehat{P}_n^{L\prime} D_n \widehat{P}_n^L/n\right) \leq \bar{c}_P + o_P(1);$$

(vi) $\sup_{\theta \in \mathcal{L} \times \mathcal{B}} \left\|\left(P_n^{L\prime} D_n - \widehat{P}_n^{L\prime} D_n\right) U_n(\theta)/n\right\| = o_P(1)$; *(vii)* $\sup_{\theta \in \mathcal{L} \times \mathcal{B}} \|U_n(\theta)' \mathcal{A}_n\| = O_P(1)$; *(viii)* $\sup_{\theta \in \mathcal{L} \times \mathcal{B}} \left\|U_n(\theta)' \widehat{\mathcal{A}}_n\right\| = O_P(1)$.

Lemma A.2 *$Q_n(\theta)$ converges in probability to $Q_n^*(\theta)$ as $n \to \infty$ uniformly in $\theta \in \mathcal{L} \times \mathcal{B}$, where*

$$Q_n^*(\theta) = \frac{1}{2n} \mathrm{E}\left(U_n(\theta)' D_n U_n(\theta)\right) - \frac{1}{2n} \mathrm{E}\left(U_n(\theta)' D_n P_n^L\right) \left(P_n^{L\prime} \overline{D}_n P_n^L\right)^{-1} \mathrm{E}\left(P_n^{L\prime} D_n U_n(\theta)\right).$$

Proof of consistency

We prove consistency following Theorem 3.4 of White (1996). We have already shown the uniform convergence of $Q_n(\theta)$ to $Q_n^*(\theta)$ on $\mathcal{L} \times \mathcal{B}$ in Lemma A.2. Then, it suffices to show the *identifiable uniqueness* condition (White, 1996, Definition 3.3): for all $\epsilon > 0$,

$$\liminf_{n \to \infty} \min_{\theta \in \mathcal{L} \times \mathcal{B}: ||\theta - \theta_0|| \geq \epsilon} \left[Q_n^*(\theta) - Q_n^*(\theta_0)\right] > 0. \tag{A.1}$$

Observe that

$$\begin{aligned} Q_n^*(\theta) - Q_n^*(\theta_0) &= \frac{1}{2n} \mathrm{E}\left(H_n(\theta)' D_n H_n(\theta)\right) + \frac{1}{n} \mathrm{E}\left(H_n(\theta)' D_n \mathcal{E}_n\right) \\ &\quad - \frac{1}{2n} \mathrm{E}\left(H_n(\theta)' D_n P_n^L\right) \left(P_n^{L\prime} \overline{D}_n P_n^L\right)^{-1} \mathrm{E}\left(P_n^{L\prime} D_n H_n(\theta)\right) \\ &\quad - \frac{1}{n} \mathrm{E}\left(H_n(\theta)' D_n P_n^L\right) \left(P_n^{L\prime} \overline{D}_n P_n^L\right)^{-1} \mathrm{E}\left(P_n^{L\prime} D_n \mathcal{E}_n\right). \end{aligned}$$

Furthermore, by the law of iterated expectations,

$$\begin{aligned}\frac{1}{n}\mathrm{E}\left(H_n(\theta)'D_n\mathcal{E}_n\right) &= \frac{1}{n}\mathrm{E}\left(H_n(\theta)'D_n\Pi_0(V_n)\right)\\ &= \frac{1}{n}\mathrm{E}\left(H_n(\theta)'D_nP_n^LB_{0L}\right)+\frac{1}{n}\mathrm{E}\left(H_n(\theta)'D_n\left[\Pi_0(V_n)-P_n^LB_{0L}\right]\right)\\ &= \frac{1}{n}\mathrm{E}\left(H_n(\theta)'D_nP_n^LB_{0L}\right)+O(L^{-\rho}).\end{aligned}$$

Similarly, we obtain

$$\frac{1}{n}\mathrm{E}\left(H_n(\theta)'D_nP_n^L\right)\left(P_n^{L\prime}\overline{D}_nP_n^L\right)^{-1}\mathrm{E}\left(P_n^{L\prime}D_n\mathcal{E}_n\right)=\frac{1}{n}\mathrm{E}\left(H_n(\theta)'D_nP_n^LB_{0L}\right)+O(L^{-\rho}).$$

Consequently,

$$\begin{aligned}Q_n^*(\theta)-Q_n^*(\theta_0) &= \frac{1}{2}\left[H_n(\theta)'\overline{D}_nH_n(\theta)/n-\left(H_n(\theta)'\overline{D}_nP_n^L/n\right)\left(P_n^{L\prime}\overline{D}_nP_n^L/n\right)^{-1}\left(P_n^{L\prime}\overline{D}_nH_n(\theta)/n\right)\right]+o(1)\\ &= \frac{1}{2}\left(\theta_0-\theta\right)'\Lambda_n(\lambda)\left(\theta_0-\theta\right)+o(1).\end{aligned}$$

Then, since $\Lambda_n(\lambda)$ is positive definite for any $\lambda\in\mathcal{L}$ as $n\to\infty$ by Assumption 7.2, this implies (A.1). ■

Lemma A.3 $\Gamma_n'\left[I_n-\widehat{\mathcal{P}}_n\right]D_n\Pi_0(\widehat{V}_n)/\sqrt{n}=o(1)$.

Lemma A.4 $\Gamma_n'\left[I_n-\widehat{\mathcal{P}}_n\right]D_n\widetilde{\mathcal{E}}_n/\sqrt{n}=\Gamma_n^{\perp\prime}D_n\widetilde{\mathcal{E}}_n/\sqrt{n}+o_P(1)$.

Lemma A.5 $\Gamma_n'\left[I_n-\widehat{\mathcal{P}}_n\right]\left(D_n\Pi_0(V_n)-D_n\Pi_0(\widehat{V}_n)\right)/\sqrt{n}=-\left[\Gamma_n^{\perp\prime}\overline{D}_n\dot{\Pi}_0(V_n)W_n/n\right]\psi_n+o_P(1)$.

Proof of normality

By the first-order condition and the mean value theorem,

$$\mathbf{0}_{(k_1+k_2+2)}=\sqrt{n}\partial_\theta Q_n(\theta_0)+\partial_{\theta\theta'}Q_n(\overline{\theta}_n)\sqrt{n}(\widehat{\theta}_n-\theta_0)$$

with $\overline{\theta}_n\in\left[\widehat{\theta}_n,\theta_0\right]$. By Lemmas A.3, A.4, and A.5, we obtain

$$\begin{aligned}\sqrt{n}\partial_\theta Q_n(\theta_0) &= -\Gamma_n'D_n\left[I_n-\widehat{\mathcal{P}}_n\right]\left(D_nY_n-D_nS_n(\lambda_0)Z_n\beta_0\right)/\sqrt{n}\\ &= -\left(\Gamma_n^{\perp\prime}D_n\widetilde{\mathcal{E}}_n/\sqrt{n}-\left[\Gamma_n^{\perp\prime}\overline{D}_n\dot{\Pi}_0(V_n)W_n/n\right]\psi_n\right)+o_P(1).\end{aligned}$$

Now, we show that $\Gamma_n^{\perp\prime}D_n\widetilde{\mathcal{E}}_n/\sqrt{n}\xrightarrow{d}N\left(\mathbf{0}_{(k_1+k_2+2)},\Psi_1\right)$. To demonstrate this, we show that

$$n^{-1/2}\mathbf{c}'\Psi_{n1}^{-1/2}\Gamma_n^{\perp\prime}D_n\widetilde{\mathcal{E}}_n\equiv\Xi_n\xrightarrow{d}N\left(0,1\right),\tag{A.2}$$

where $\mathbf{c}$ is an arbitrary $(k_1+k_2+2)\times 1$ vector satisfying $\|\mathbf{c}\|=1$. Furthermore, we define $\Xi_n\equiv\sum_{j=1}^n\xi_j=\sum_{j=1}^n\left(\xi_{1j}+\xi_{2j}\right)$ with

$$\begin{aligned}\xi_{1j} &\equiv n^{-1/2}\mathbf{c}'\Psi_{n1}^{-1/2}\varsigma_jd_j\widetilde{\varepsilon}_j,\\ \xi_{2j} &\equiv -n^{-1/2}\mathbf{c}'\Psi_{n1}^{-1/2}\Gamma_n'\overline{D}_nP_n^L\left(P_n^{L\prime}\overline{D}_nP_n^L\right)^{-1}p^L(v_j)d_j\widetilde{\varepsilon}_j,\end{aligned}$$

where ς_j is the jth column of Γ_n', and $\widetilde{\varepsilon}_j \equiv \varepsilon_j - \pi_0(v_j)$. Then, by Assumptions 3.1 and 7.1 and the law of iterated expectations,

$$\begin{aligned}\sum_{j=1}^n \mathrm{E}\left[\xi_{1j}^4\right] &= n^{-2}\sum_{j=1}^n \mathrm{E}\left\{\left(\mathbf{c}'\Psi_{n1}^{-1/2}\varsigma_j\right)^4 d_j \mathrm{E}\left[\widetilde{\varepsilon}_j^4|d_j\right]\right\} \\ &\leq cn^{-2}\sum_{j=1}^n \left\{\mathbf{c}'\Psi_{n1}^{-1/2}\varsigma_j\varsigma_j'\Psi_{n1}^{-1/2}\mathbf{c}\mathbf{c}'\Psi_{n1}^{-1/2}\varsigma_j\varsigma_j'\Psi_{n1}^{-1/2}\mathbf{c}\right\} \\ &\leq cn^{-2}\sum_{j=1}^n \mathbf{c}'\Psi_{n1}^{-1}\Psi_{n1}^{-1}\mathbf{c} = O(n^{-1}) = o(1)\end{aligned}$$

as $n \to \infty$. Similarly

$$\begin{aligned}\sum_{j=1}^n \mathrm{E}\left[\xi_{2j}^4\right] &\leq c\zeta_0^2(L)n^{-2}\left\{\mathbf{c}'\Psi_{n1}^{-1/2}\Gamma_n'\overline{D}_nP_n^L\left(P_n^{L\prime}\overline{D}_nP_n^L\right)^{-1}P_n^{L\prime}\overline{D}_nP_n^L\left(P_n^{L\prime}\overline{D}_nP_n^L\right)^{-1}P_n^{L\prime}\overline{D}_n\Gamma_n\Psi_{n1}^{-1/2}\mathbf{c}\right\} \\ &\leq c\zeta_0^2(L)n^{-1}\mathbf{c}'\Psi_{n1}^{-1}\mathbf{c} = O\left(\zeta_0^2(L)/n\right) = o(1)\end{aligned}$$

holds as $n \to \infty$ by Assumption 5.2. Then, we obtain

$$\begin{aligned}\sum_{j=1}^n \mathrm{E}\left[\xi_j^4\right] &= \sum_{j=1}^n \mathrm{E}\left[(\xi_{1j}+\xi_{2j})^4\right] \\ &\leq 8\left\{\sum_{j=1}^n \mathrm{E}\left[\xi_{1j}^4\right] + \sum_{j=1}^n \mathrm{E}\left[\xi_{2j}^4\right]\right\} = o(1)\end{aligned}$$

by the c_r-inequality. Therefore, (A.2) is verified by Lyapunov's central limit theorem, and we obtain

$$n^{-1/2}\Psi_{n1}^{-1/2}\Gamma_n^{\perp\prime}D_n\widetilde{\mathcal{E}}_n \xrightarrow{d} N\left(0, I_{(k_1+k_2+2)}\right)$$

by the Cramer–Wold device. Noting the consistency of $\overline{\theta}_n$ for θ_0 and, thus, the convergence in probability of $\partial_{\theta\theta'}Q_n(\overline{\theta}_n)$ to Λ as $n \to \infty$ by the continuous mapping theorem, the proof is completed by applying Slutsky's theorem. ■

A.2 Proof of Theorem 4.2

Lemma A.6 $\left\|B_{0L} - \widehat{B}_{nL}\right\| = O_P\left(\sqrt{L}/\sqrt{n}\right)$.

Proof of Theorem 4.2

First, note that

$$\begin{aligned}\max_{1\leq i\leq n}\left|\partial_v\pi_0(\widehat{v}_i) - \partial_v p^L(\widehat{v}_i)'\widehat{B}_{nL}\right| &\leq \sup_{v\in[-c_v,c_v]}\left|\partial_v\pi_0(v) - \partial_v p^L(v)'B_{0L}\right| + \sup_{v\in[-c_v,c_v]}\left|\partial_v p^L(v)'\left(B_{0L}-\widehat{B}_{nL}\right)\right| \\ &= O(L^{-(\rho-1)}) + O_P(\zeta_1(L)\sqrt{L}/\sqrt{n}) = o_P(1)\end{aligned}$$

as $n \to \infty$ by Lemma A.6 and the assumptions made in the theorem. This leads to $\left\| \widehat{\Psi}_{n2} - \Psi_{n2} \right\| = o_P(1)$ in a straightforward way by the continuity of $\partial_v \pi_0(\cdot)$.

Next, by the same argument as in Newey (2009, A.12), we can obtain $\left\| \widehat{\Psi}_{n1} - \widehat{\Gamma}_n^{\perp\prime} \Sigma_n^* \widehat{\Gamma}_n^{\perp} / n \right\| = o_P(1)$. Further,

$$
\begin{aligned}
\left\| \widehat{\Gamma}_n^{\perp\prime} \Sigma_n^* \widehat{\Gamma}_n^{\perp} / n - \Psi_{n1} \right\| &\leq \left\| \left(\widehat{\Gamma}_n^{\perp} - \Gamma_n^{\perp} \right)' \Sigma_n^* \left(\widehat{\Gamma}_n^{\perp} - \Gamma_n^{\perp} \right) / n \right\| + 2 \left\| \left(\widehat{\Gamma}_n^{\perp} - \Gamma_n^{\perp} \right)' \Sigma_n^* \Gamma_n^{\perp} / n \right\| \\
&= \| J_{n1} \| + 2 \| J_{n2} \|, \text{ say.}
\end{aligned}
$$

For J_{n1}, we have $\| J_{n1} \| \leq \bar{c}_{\varepsilon,2} \left\| \left(\widehat{\Gamma}_n^{\perp} - \Gamma_n^{\perp} \right) / \sqrt{n} \right\|^2 = o_P(1)$. For J_{n2}, by the matrix Cauchy–Schwarz inequality,

$$
\begin{aligned}
\| J_{n2} \| &\leq \left\| \Gamma_n^{\perp\prime} \Sigma_n^* \Gamma_n^{\perp} / n \right\|^{1/2} \| J_{n1} \|^{1/2} \\
&\leq c \bar{c}_{\varepsilon,2}^{1/2} \| J_{n1} \|^{1/2} = o_P(1).
\end{aligned}
$$

This implies that $\left\| \widehat{\Psi}_{n1} - \Psi_{n1} \right\| = o_P(1)$ by the triangle inequality. The rest of the proof is straightforward and is omitted. ■