Econometrica Supplementary Material

SUPPLEMENT TO "QUANTILE FACTOR MODELS"

(*Econometrica*, Vol. 89, No. 2, March 2021, 875–910)

LIANG CHEN
HSBC Business School, Peking University

JUAN J. DOLADO
Department of Economics, Universidad Carlos III de Madrid

JESÚS GONZALO
Department of Economics, Universidad Carlos III de Madrid

S.1. PROOF OF THEOREM 4

WE ONLY PROVE the asymptotic distribution of $\tilde{\lambda}_i$ since the proof for $\tilde{f}_t$ is symmetric. Without loss of generality, we assume that $\tilde{\mathsf{S}} = \mathsf{sgn}(\tilde{F}'F_0/T) = \mathbb{I}_r$ to simply the notation. Define $\varrho(u) = [\tau - K(u/h)]u$, so that we can write the objective function as

$$\mathbb{S}_{NT}(\theta) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\varrho\big(X_{it} - \lambda_i' f_t\big).$$

Let $\varrho^{(j)}(u) = (\partial/\partial u)^j \varrho(u)$ for $j = 1, 2, 3$. For fixed λ_i, f_t, define

$$\bar{\varrho}^{(j)}\big(X_{it} - \lambda_i' f_t\big) = \mathbb{E}\big[\varrho^{(j)}\big(X_{it} - \lambda_i' f_t\big)\big] \quad \text{and}$$
$$\tilde{\varrho}^{(j)}\big(X_{it} - \lambda_i' f_t\big) = \varrho^{(j)}\big(X_{it} - \lambda_i' f_t\big) - \bar{\varrho}^{(j)}\big(X_{it} - \lambda_i' f_t\big) \quad \text{for } j = 1, 2, 3.$$

When the functions defined above are evaluated at the true parameters, we suppress their arguments to further simplify the notations. For example, $\varrho_{it} = \varrho(X_{it} - \lambda_{0i}' f_{0t})$, $\bar{\varrho}_{it}^{(j)} = \bar{\varrho}^{(j)}(X_{it} - \lambda_{0i}' f_{0t})$. Moreover, define

$$\mathbb{S}^*_{NT}(\theta) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\big[\varrho\big(X_{it} - \lambda_i' f_t\big) - \varrho\big(X_{it} - \lambda_{0i}' f_{0t}\big)\big],$$

$$\bar{\mathbb{S}}^*_{NT}(\theta) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{E}\big[\varrho\big(X_{it} - \lambda_i' f_t\big) - \varrho\big(X_{it} - \lambda_{0i}' f_{0t}\big)\big],$$

$$\mathbb{U}_{NT}(\theta) = \mathbb{S}^*_{NT}(\theta) - \bar{\mathbb{S}}^*_{NT}(\theta).$$

Using $\bar{O}(1)$ to denote a sequence that is uniformly (over i and t) bounded. Then we can show that the following holds.

LEMMA S.1: *Under Assumptions* 1 *and* 2,

(i) *There exists a constant* $\bar{C} > 0$ *such that* $h^{j-1} \cdot \sup_u |\varrho^{(j)}(u)| \leq \bar{C}$ *for* $j = 1, 2, 3$.

Liang Chen: chenliang@phbs.pku.edu.cn
Juan J. Dolado: dolado@eco.uc3m.es
Jesús Gonzalo: jgonzalo@est-econ.uc3m.es

 https://doi.org/10.3982/ECTA15746

(ii) $\bar{\varrho}_{it}^{(1)} = \bar{O}(h^m)$, $\bar{\varrho}^{(2)}(X_{it} - \lambda_i' f_t) = \mathsf{f}_{it}(\lambda_i' f_t - \lambda_{0i}' f_{0t}) + \bar{O}(h^m)$, *and* $\bar{\varrho}^{(3)}(X_{it} - \lambda_i' f_t) = -\mathsf{f}_{it}^{(1)}(\lambda_i' f_t - \lambda_{0i}' f_{0t}) + \bar{O}(h^m)$.

(iii) $\mathbb{E}(\varrho_{it}^{(1)})^2 = \tau(1-\tau) + \bar{O}(h)$, *and* $h \cdot \mathbb{E}[(\varrho_{it}^{(2)})^2] = \bar{O}(1)$.

PROOF: Since the proof steps follow the standard ones when computing the means of kernel density estimators, they are omitted here to save space. Similar results can be found in Horowitz (1998) and Galvao and Kato (2016). *Q.E.D.*

LEMMA S.2: *Under Assumptions* 1 *and* 2, $d(\tilde{\theta}, \theta_0) = o_P(1)$ *as* $N, T \to \infty$ *and* $h \to 0$.

PROOF: By definition, we have $\mathbb{S}_{NT}(\tilde{\theta}) \le \mathbb{S}_{NT}(\theta_0)$. Adding and subtracting terms, we have

$$d^2(\tilde{\theta}, \theta_0) \lesssim \bar{\mathbb{M}}_{NT}^*(\tilde{\theta}) \le \mathbb{M}_{NT}(\tilde{\theta}) - \mathbb{S}_{NT}(\tilde{\theta}) + \mathbb{S}_{NT}(\theta_0) - \mathbb{M}_{NT}(\theta_0) - \mathbb{W}_{NT}(\tilde{\theta}).$$

It follows that

$$d^2(\tilde{\theta}, \theta_0) \lesssim \sup_{\theta \in \Theta^r} \left| \mathbb{M}_{NT}(\theta) - \mathbb{S}_{NT}(\theta) \right| + \sup_{\theta \in \Theta^r} \left| \mathbb{W}_{NT}(\theta) \right|.$$

It is easy to see that the first term on the RHS of the above inequality is $O_p(h)$, and the second term is $o_P(1)$ as proved in Lemma 1. Then, the desired result follows. *Q.E.D.*

LEMMA S.3: *Under Assumptions* 1 *and* 2, $d(\tilde{\theta}, \theta_0) = O_P(1/L_{NT})$ *as* $N, T \to \infty$.

PROOF: First, since $\varrho^{(1)}(u)$ is uniformly bounded, we have $|\varrho(X_{it} - c_1) - \varrho(X_{it} - c_2)| \lesssim |c_1 - c_2|$. Then similar to the proof of Lemma 3, we can show that

$$\mathbb{E}\Big[\sup_{\theta \in \Theta^r(\delta)} \left| \mathbb{U}_{NT}(\theta) \right| \Big] \lesssim \frac{\delta}{L_{NT}}. \tag{S.1}$$

Analogous to the proof of Theorem 1, the parameter space Θ^r can be partitioned into shells $S_j = \{\theta \in \Theta^r : 2^{j-1} < L_{NT} \cdot d(\theta, \theta_0) \le 2^j\}$. Conclude that, for a given integer V and for every $\eta > 0$,

$$P\big[L_{NT} \cdot d(\tilde{\theta}, \theta_0) > 2^V\big] \le \sum_{j > V, 2^{j-1} \le \eta L_{NT}} P\Big[\inf_{\theta \in S_j} \mathbb{S}_{NT}^*(\theta) \le 0\Big] + P\big[d(\tilde{\theta}, \theta_0) \ge \eta\big].$$

For arbitrarily small $\eta > 0$, the second probability on the RHS of the above equation converges to 0 as $N, T \to \infty$ by Lemma S.2.

Next, expanding $\mathbb{S}_{NT}^*(\theta)$ around θ_0 and taking expectations yield

$$\bar{\mathbb{S}}_{NT}^*(\theta) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \bar{\varrho}_{it}^{(1)} \cdot \left(\lambda_i' f_t - \lambda_{0i}' f_{0t}\right) + \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} 0.5 \bar{\varrho}^{(2)}\left(c_{it}^*\right) \cdot \left(\lambda_i' f_t - \lambda_{0i}' f_{0t}\right)^2,$$

where c_{it}^* lies between $\lambda_i' f_t$ and $\lambda_{0i}' f_{0t}$. It then follows from Lemma S.1 and Assumption 2 that

$$\bar{\mathbb{S}}_{NT}^*(\theta) \ge O\big(h^m\big) + 0.5 \underline{\mathsf{f}} \cdot d^2(\theta, \theta_0).$$

Thus, for each θ in S_j we have

$$-\bar{\mathbb{S}}^*_{NT}(\theta) \le -0.5\underline{\mathrm{f}} \cdot d^2_{NT}(\theta, \theta_0) + O\big(h^m\big) \le -\underline{\mathrm{f}} \cdot \frac{2^{2j-3}}{L^2_{NT}} + O\big(h^m\big).$$

Therefore, $\inf_{\theta\in S_j} \mathbb{S}^*_{NT}(\theta) \le 0$ implies that

$$\inf_{\theta\in S_j} \mathbb{U}_{NT}(\theta) \le -\underline{\mathrm{f}} \cdot \frac{2^{2j-3}}{L^2_{NT}} + O\big(h^m\big),$$

and it follows that

$$\sum_{j>V, 2^{j-1}\le\eta L_{NT}} P\Big[\inf_{\theta\in S_j} \mathbb{S}^*_{NT}(\theta) \le 0\Big] \le \sum_{j>V, 2^{j-1}\le\eta L_{NT}} P\Big[\sup_{\theta\in S_j}\big|\mathbb{U}_{NT}(\theta)\big| \ge \underline{\mathrm{f}} \cdot \frac{2^{2j-3}}{L^2_{NT}} + O\big(h^m\big)\Big].$$

By (S.1) and Markov's inequality, we have

$$P\Big[\sup_{\theta\in S_j}\big|\mathbb{U}_{NT}(\theta)\big| \ge \underline{\mathrm{f}} \cdot \frac{2^{2j-3}}{L^2_{NT}} + O\big(h^m\big)\Big] \lesssim \frac{2^j}{2^{2j} + O\big(L^2_{NT}\cdot h^m\big)}.$$

By Assumption 2(v), $O(L^2_{NT}\cdot h^m) = o(1)$. Thus, the above inequality implies that

$$\sum_{j>V, 2^{j-1}\le\eta L_{NT}} P\Big[\inf_{\theta\in S_j} \mathbb{S}^*_{NT}(\theta) \le 0\Big] \lesssim \sum_{j>V} 2^{-j}.$$

The RHS of the previous expression converges to 0 as $V\to\infty$, implying that $L_{NT}\cdot d(\tilde{\theta}, \theta_0) = O_P(1)$, or $d(\tilde{\theta}, \theta_0) = O_P(1/L_{NT})$. *Q.E.D.*

Define:

$$\mathbb{S}^*_{i,T}(\lambda, F) = \frac{1}{T}\sum_{t=1}^{T}\big[\varrho\big(X_{it} - \lambda' f_t\big) - \varrho\big(X_{it} - \lambda'_{0i} f_{0t}\big)\big],$$

$$\bar{\mathbb{S}}^*_{i,T}(\lambda, F) = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\big[\varrho\big(X_{it} - \lambda' f_t\big) - \varrho\big(X_{it} - \lambda'_{0i} f_{0t}\big)\big],$$

$$\mathbb{M}^*_{i,T}(\lambda, F) = \frac{1}{T}\sum_{t=1}^{T}\big[\rho_\tau\big(X_{it} - \lambda' f_t\big) - \rho_\tau\big(X_{it} - \lambda'_{0i} f_{0t}\big)\big],$$

$$\bar{\mathbb{M}}^*_{i,T}(\lambda, F) = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\big[\rho_\tau\big(X_{it} - \lambda' f_t\big) - \rho_\tau\big(X_{it} - \lambda'_{0i} f_{0t}\big)\big].$$

LEMMA S.4: *Under Assumptions* 1 *and* 2, *we have* $\|\tilde{\lambda}_i - \lambda_{0i}\| = o_P(1)$ *for each* i.

PROOF: Note that

$$\tilde{\lambda}_i = \arg\min_{\lambda\in\mathcal{A}} \mathbb{S}^*_{i,T}(\lambda, \tilde{F}).$$

First, we show that

$$\sup_{\lambda\in\mathcal{A}}\left|\mathbb{S}^*_{i,T}(\lambda,\tilde{F})-\bar{\mathbb{M}}^*_{i,T}(\lambda,F_0)\right|=o_P(1). \tag{S.2}$$

Adding and subtracting terms, we have

$$\begin{aligned}&\mathbb{S}^*_{i,T}(\lambda,\tilde{F})-\bar{\mathbb{M}}^*_{i,T}(\lambda,F_0)\\&\quad=\mathbb{S}^*_{i,T}(\lambda,\tilde{F})-\mathbb{M}^*_{i,T}(\lambda,\tilde{F})+\mathbb{M}^*_{i,T}(\lambda,\tilde{F})-\mathbb{M}^*_{i,T}(\lambda,F_0)+\mathbb{M}^*_{i,T}(\lambda,F_0)-\bar{\mathbb{M}}^*_{i,T}(\lambda,F_0),\end{aligned}$$

thus,

$$\begin{aligned}&\sup_{\lambda\in\mathcal{A}}\left|\mathbb{S}^*_{i,T}(\lambda,\tilde{F})-\bar{\mathbb{M}}^*_{i,T}(\lambda,F_0)\right|\\&\quad\le\sup_{\lambda\in\mathcal{A}}\left|\mathbb{S}^*_{i,T}(\lambda,\tilde{F})-\mathbb{M}^*_{i,T}(\lambda,\tilde{F})\right|+\sup_{\lambda\in\mathcal{A}}\left|\mathbb{M}^*_{i,T}(\lambda,\tilde{F})-\mathbb{M}^*_{i,T}(\lambda,F_0)\right|\\&\qquad+\sup_{\lambda\in\mathcal{A}}\left|\mathbb{M}^*_{i,T}(\lambda,F_0)-\bar{\mathbb{M}}^*_{i,T}(\lambda,F_0)\right|.\end{aligned}$$

It is easy to show that

$$\begin{aligned}&\sup_{\lambda\in\mathcal{A}}\left|\mathbb{S}^*_{i,T}(\lambda,\tilde{F})-\mathbb{M}^*_{i,T}(\lambda,\tilde{F})\right|\lesssim h,\\&\sup_{\lambda\in\mathcal{A}}\left|\mathbb{M}^*_{i,T}(\lambda,\tilde{F})-\mathbb{M}^*_{i,T}(\lambda,F_0)\right|\lesssim\sup_{\lambda\in\mathcal{A}}\|\lambda\|\cdot\frac{1}{T}\sum_{t=1}^{T}\|\tilde{f}_t-f_{0t}\|\\&\qquad\qquad\lesssim\|\tilde{F}-F_0\|/\sqrt{T}=O_P(1/L_{NT}),\\&\sup_{\lambda\in\mathcal{A}}\left|\mathbb{M}^*_{i,T}(\lambda,F_0)-\bar{\mathbb{M}}^*_{i,T}(\lambda,F_0)\right|=o_P(1).\end{aligned}$$

Then (S.2) follows as $h\to 0$.

Second, we can show that for any $\epsilon>0$, and $B_i(\epsilon)=\{\lambda\in\mathcal{A}:\|\lambda-\lambda_{0i}\|\le\epsilon\}$,

$$\inf_{\lambda\in B_i^C(\epsilon)}\bar{\mathbb{M}}^*_{i,T}(\lambda,F_0)>\bar{\mathbb{M}}^*_{i,T}(\lambda_{0i},F_0)=0, \tag{S.3}$$

for example, the proof of Proposition 3.1 of Galvao and Kato (2016).

Finally, given (S.2) and (S.3), the consistency of $\tilde{\lambda}_i$ follows from a standard consistency proof of M-estimators (see Theorem 2.1 of Newey and McFadden (1994)), so that the proof concludes. *Q.E.D.*

LEMMA S.5: *Under Assumptions* 1 *and* 2, *we have* $\|\tilde{\lambda}_i-\lambda_{0i}\|=O_P(T^{-1/2}h^{-1})$ *for each* i.

PROOF: For any fixed $\lambda_i\in\mathcal{A}$ and $f_t\in\mathcal{F}$, expanding $\varrho^{(1)}(X_{it}-\lambda_i'f_t)f_t$ gives

$$\begin{aligned}&\varrho^{(1)}(X_{it}-\lambda_i'f_t)f_t\\&\quad=\varrho^{(1)}(X_{it}-\lambda_{0i}'f_t)f_t-\varrho^{(2)}(X_{it}-\lambda_{0i}'f_t)f_tf_t'\cdot(\lambda_i-\lambda_{0i})\\&\qquad+0.5\varrho^{(3)}(X_{it}-\lambda_i^{*\prime}f_t)f_t\big[(\lambda_i-\lambda_{0i})'f_t\big]^2\end{aligned}$$

$$
\begin{aligned}
&= \varrho_{it}^{(1)} f_{0t} + \varrho^{(1)}\big(X_{it} - \lambda_{0i}' f_t^*\big)(f_t - f_{0t}) - \varrho^{(2)}\big(X_{it} - \lambda_{0i}' f_t^*\big) f_t^* \lambda_{0i}'(f_t - f_{0t}) \\
&\quad - \varrho_{it}^{(2)} f_t f_t' \cdot (\lambda_i - \lambda_{0i}) \\
&\quad + \varrho^{(3)}\big(X_{it} - \lambda_{0i}' f_t^*\big) f_t f_t' \cdot (\lambda_i - \lambda_{0i}) \lambda_{0i}'(f_t - f_{0t}) \\
&\quad + 0.5 \varrho^{(3)}\big(X_{it} - \lambda_i^{*'} f_t\big) f_t \big[(\lambda_i - \lambda_{0i})' f_t\big]^2,
\end{aligned}
$$

where λ_i^* lies between λ_i and λ_{0i} and f_t^* lies between f_t and f_{0t}. Taking expectations of both sides of the above equation, and setting $\lambda_i = \tilde{\lambda}_i$, $f_t = \tilde{f}_t$, it follows from Lemma 4 that

$$
\begin{aligned}
&\frac{1}{T} \sum_{t=1}^{T} \bar{\varrho}^{(1)}\big(X_{it} - \tilde{\lambda}_i' \tilde{f}_t\big) \tilde{f}_t \\
&\quad = \frac{1}{T} \sum_{t=1}^{T} \bar{\varrho}_{it}^{(1)} f_{0t} - \left(\frac{1}{T} \sum_{t=1}^{T} \bar{\varrho}_{it}^{(2)} \tilde{f}_t \tilde{f}_t' \right) (\tilde{\lambda}_i - \lambda_{0i}) \\
&\qquad + O_P\big(T^{-1/2} \|\tilde{F} - F_0\|\big) + O_P\big(\|\tilde{\lambda}_i - \lambda_{0i}\|\big) \cdot O_P\big(T^{-1/2} \|\tilde{F} - F_0\|\big) + O_P\big(\|\tilde{\lambda}_i - \lambda_{0i}\|^2\big).
\end{aligned}
$$

Lemma S.1, Lemma S.3, and Assumption 2 imply that

$$
\frac{1}{T} \sum_{t=1}^{T} \bar{\varrho}_{it}^{(2)} \tilde{f}_t \tilde{f}_t' = \frac{1}{T} \sum_{t=1}^{T} \bar{\varrho}_{it}^{(2)} f_{0t} f_{0t}' + o_P(1) = \Phi_i + o_P(1).
$$

Then, from Lemmas S.1, S.3, and S.4, we get

$$
\Phi_i(\tilde{\lambda}_i - \lambda_{0i}) + o_P\big(\|\tilde{\lambda}_i - \lambda_{0i}\|\big) = O\big(h^m\big) + O_P(1/L_{NT}) - \frac{1}{T} \sum_{t=1}^{T} \bar{\varrho}^{(1)}\big(X_{it} - \tilde{\lambda}_i' \tilde{f}_t\big) \tilde{f}_t. \quad \text{(S.4)}
$$

Note that we can write

$$
\begin{aligned}
&\frac{1}{T} \sum_{t=1}^{T} \bar{\varrho}^{(1)}\big(X_{it} - \tilde{\lambda}_i' \tilde{f}_t\big) \tilde{f}_t \\
&\quad = -\frac{1}{T} \sum_{t=1}^{T} \tilde{\varrho}_{it}^{(1)} f_{0t} - \frac{1}{T} \sum_{t=1}^{T} \big[\tilde{\varrho}^{(1)}\big(X_{it} - \tilde{\lambda}_i' \tilde{f}_t\big) \tilde{f}_t - \tilde{\varrho}_{it}^{(1)} f_{0t}\big] \\
&\quad = -\frac{1}{T} \sum_{t=1}^{T} \tilde{\varrho}_{it}^{(1)} f_{0t} - \frac{1}{T} \sum_{t=1}^{T} \big[\tilde{\varrho}^{(1)}\big(X_{it} - \tilde{\lambda}_i' \tilde{f}_t\big) \tilde{f}_t - \tilde{\varrho}^{(1)}\big(X_{it} - \tilde{\lambda}_i' f_{0t}\big) f_{0t}\big] \\
&\qquad - \frac{1}{T} \sum_{t=1}^{T} \big[\tilde{\varrho}^{(1)}\big(X_{it} - \tilde{\lambda}_i' f_{0t}\big) - \tilde{\varrho}_{it}^{(1)}\big] f_{0t}.
\end{aligned}
$$

The first term on the RHS of the above equation is $O_P(T^{-1/2})$ by Lemma S.1 and Lyapunov's CLT. As for the second term on the RHS of the above equation, we have that

$$
\frac{1}{T} \sum_{t=1}^{T} \big[\tilde{\varrho}^{(1)}\big(X_{it} - \tilde{\lambda}_i' \tilde{f}_t\big) \tilde{f}_t - \tilde{\varrho}^{(1)}\big(X_{it} - \tilde{\lambda}_i' f_{0t}\big) f_{0t}\big]
$$

$$= \frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}^{(1)}\big(X_{it} - \tilde{\lambda}_i' f_t^*\big)(\tilde{f}_t - f_{0t})$$

$$- \frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}^{(2)}\big(X_{it} - \tilde{\lambda}_i' f_t^*\big) f_t^* \tilde{\lambda}_i'(\tilde{f}_t - f_{0t}), \tag{S.5}$$

where f_t^* lies between $\tilde{f}_t$ and f_{0t}. The first term on the RHS of (S.5) is $O_P(1/L_{NT})$ because $\varrho^{(1)}$ is uniformly bounded and $T^{-1}\sum_{t=1}^{T}\|\tilde{f}_t - f_{0t}\| = O_P(1/L_{NT})$ by Lemma S.3. Similarly, the second term on the RHS of (S.5) is $O_P(1/(L_{NT}h))$ because $h\varrho^{(2)}(u)$ is uniformly bounded. Finally, we can show that (see, e.g., Lemma B.2 of Galvao and Kato (2016))

$$\frac{1}{T}\sum_{t=1}^{T}\big[\tilde{\varrho}^{(1)}\big(X_{it} - \tilde{\lambda}_i' f_{0t}\big) - \tilde{\varrho}_{it}^{(1)}\big] f_{0t} = O_P\big(\|\tilde{\lambda}_i - \lambda_{0i}\|\big) \cdot O_P(1/\sqrt{Th}) = o_P\big(\|\tilde{\lambda}_i - \lambda_{0i}\|\big).$$

Combining the above results yields

$$\frac{1}{T}\sum_{t=1}^{T}\bar{\varrho}^{(1)}\big(X_{it} - \tilde{\lambda}_i'\tilde{f}_t\big)\tilde{f}_t = O_P\bigg(\frac{1}{L_{NT}h}\bigg) + o_P\big(\|\tilde{\lambda}_i - \lambda_{0i}\|\big), \tag{S.6}$$

and the desired result follows from (S.4), (S.6), and Assumption 2. *Q.E.D.*

To derive the asymptotic distribution of $\tilde{\lambda}_i$, it is essential to obtain the stochastic expansion of $\tilde{f}_t$. To do so, define

$$\mathbb{P}_{NT}(\theta) = b\Bigg[\frac{1}{2N}\sum_{p=1}^{r}\sum_{q>p}^{r}\Bigg(\sum_{i=1}^{N}\lambda_{ip}\lambda_{iq}\Bigg)^2 + \frac{1}{2T}\sum_{p=1}^{r}\sum_{q>p}^{r}\Bigg(\sum_{t=1}^{T}f_{tp}f_{tq}\Bigg)^2 + \frac{1}{8T}\sum_{k=1}^{r}\Bigg(\sum_{t=1}^{T}f_{tk}^2 - T\Bigg)^2\Bigg]$$

for some $b > 0$. Define

$$\mathcal{S}^*(\theta) = \Bigg[\underbrace{\ldots, -\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}\bar{\varrho}^{(1)}\big(X_{it} - \lambda_i' f_t\big)f_t', \ldots,}_{1\times Nr}\underbrace{\ldots, -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\bar{\varrho}^{(1)}\big(X_{it} - \lambda_i' f_t\big)\lambda_i', \ldots}_{1\times Tr}\Bigg]',$$

$$\mathcal{S}(\theta) = \mathcal{S}^*(\theta) + \partial\mathbb{P}_{NT}(\theta)/\partial\theta, \qquad \mathcal{H}(\theta) = \partial\mathcal{S}^*(\theta)/\partial\theta' + \partial^2\mathbb{P}_{NT}(\theta)/\partial\theta\partial\theta',$$

and let $\mathcal{H} = \mathcal{H}(\theta_0)$.

Expanding $\mathcal{S}(\tilde{\theta})$ around $\mathcal{S}(\theta_0)$ gives

$$\mathcal{S}(\tilde{\theta}) = \mathcal{S}(\theta_0) + \mathcal{H}\cdot(\tilde{\theta} - \theta_0) + 0.5\mathcal{R}(\tilde{\theta}), \tag{S.7}$$

where

$$\mathcal{R}(\tilde{\theta}) = \Bigg(\sum_{j=1}^{M}\partial\mathcal{H}\big(\theta^*\big)/\partial\theta_j\cdot(\tilde{\theta}_j - \theta_{0j})\Bigg)(\tilde{\theta} - \theta_0),$$

and θ^* lies between $\tilde{\theta}$ and θ_0.

Further, define

$$\mathcal{H}_d = \begin{pmatrix} \mathcal{H}_d^{\Lambda} & 0 \\ 0 & \mathcal{H}_d^{F} \end{pmatrix}, \qquad \mathcal{H}_d^{\Lambda} = \frac{\sqrt{T}}{\sqrt{N}} \operatorname{diag}[\Phi_{T,1}, \ldots, \Phi_{T,i}, \ldots, \Phi_{T,N}],$$

$$\mathcal{H}_d^{F} = \frac{\sqrt{N}}{\sqrt{T}} \operatorname{diag}[\Psi_{N,1}, \ldots, \Psi_{N,t}, \ldots, \Psi_{N,T}],$$

where

$$\Phi_{T,i} = \frac{1}{T}\sum_{t=1}^{T} \bar{\varrho}_{it}^{(2)} f_{0t} f_{0t}', \qquad \Psi_{N,t} = \frac{1}{N}\sum_{i=1}^{N} \bar{\varrho}_{it}^{(2)} \lambda_{0i}\lambda_{0i}'.$$

The following lemma is important for the stochastic expansion of $\tilde{f}_t$.

LEMMA S.6: *The matrix $\mathcal{H}$ is invertible and $\|\mathcal{H}^{-1} - \mathcal{H}_d^{-1}\|_{\max} = O(1/T)$.*

PROOF: To simplify the notations, we consider the case $r = 2$, but the proof can be easily generalized to the case $r > 2$. Note that $\lambda_{0i} = (\lambda_{0i,1}, \lambda_{0i,2})'$ and $f_{0t} = (f_{0t,1}, f_{0t,2})'$, and $\mathbb{P}_{NT}(\theta)$ simplifies to

$$\mathbb{P}_{NT}(\theta) = b\Bigg[\frac{1}{2N}\left(\sum_{i=1}^{N} \lambda_{i1}\lambda_{i2}\right)^2 + \frac{1}{2T}\left(\sum_{t=1}^{T} f_{t1} f_{t2}\right)^2 + \frac{1}{8T}\left(\sum_{t=1}^{T} f_{t1}^2 - T\right)^2 + \frac{1}{8T}\left(\sum_{t=1}^{T} f_{t2}^2 - T\right)^2\Bigg].$$

First, define

$$\gamma_1' = \big[\mathbf{0}_{1\times 2N}, (f_{01,1}, 0), \ldots, (f_{0t,1}, 0), \ldots, (f_{0T,1}, 0)\big]/\sqrt{T},$$
$$\gamma_2' = \big[\mathbf{0}_{1\times 2N}, (0, f_{01,2}), \ldots, (0, f_{0t,2}), \ldots, (0, f_{0T,2})\big]/\sqrt{T},$$
$$\gamma_3' = \big[\mathbf{0}_{1\times 2N}, (f_{01,2}, f_{01,1}), \ldots, (f_{0t,2}, f_{0t,1}), \ldots, (f_{0T,2}, f_{0T,1})\big]/\sqrt{T},$$
$$\gamma_4' = \big[(\lambda_{01,2}, \lambda_{01,1}), \ldots, (\lambda_{0i,2}, \lambda_{0i,1}), \ldots, (\lambda_{0N,2}, \lambda_{0N,1}), \mathbf{0}_{1\times 2T}\big]/\sqrt{N},$$

and note that $\partial^2\mathbb{P}_{NT}(\theta_0)/\partial\theta\partial\theta' = b(\sum_{k=1}^{4} \gamma_k\gamma_k')$.

Second, define

$$\omega_1' = \big[\underbrace{(\lambda_{01,1}, 0)/\sqrt{N}, \ldots, (\lambda_{0N,1}, 0)/\sqrt{N}}_{\omega_{1\Lambda}'}, \underbrace{(-f_{01,1}, 0)/\sqrt{T}, \ldots, (-f_{0T,1}, 0)/\sqrt{T}}_{\omega_{1F}'}\big],$$

$$\omega_2' = \big[\underbrace{(0, \lambda_{01,2})/\sqrt{N}, \ldots, (0, \lambda_{0N,2})/\sqrt{N}}_{\omega_{2\Lambda}'}, \underbrace{(0, -f_{01,2})/\sqrt{T}, \ldots, (0, -f_{0T,2})/\sqrt{T}}_{\omega_{2F}'}\big],$$

$$\omega_3' = \big[\underbrace{(\lambda_{01,2}, 0)/\sqrt{N}, \ldots, (\lambda_{0N,2}, 0)/\sqrt{N}}_{\omega_{3\Lambda}'}, \underbrace{(0, -f_{01,1})/\sqrt{T}, \ldots, (0, -f_{0T,1})/\sqrt{T}}_{\omega_{3F}'}\big],$$

$$\omega_4' = \big[\underbrace{(0, \lambda_{01,1})/\sqrt{N}, \ldots, (0, \lambda_{0N,1})/\sqrt{N}}_{\omega_{4\Lambda}'}, \underbrace{(-f_{01,2}, 0)/\sqrt{T}, \ldots, (-f_{0T,2}, 0)/\sqrt{T}}_{\omega_{4F}'}\big],$$

and $\omega = [\omega_1, \omega_2, \omega_3, \omega_4]$. It is easy to check that $\omega_p' \omega_q = 0$ for $p \neq q$. Moreover, we have

$$\omega\omega' = \sum_{k=1}^{4} \omega_k \omega_k' = \begin{pmatrix} \sum_{k=1}^{4} \omega_{k\Lambda}\omega_{k\Lambda}' & -(NT)^{-1/2}\{f_{0t}\lambda_{0i}'\}_{i\le N, t\le T} \\ -(NT)^{-1/2}\{\lambda_{0i}f_{0t}'\}_{t\le T, i\le N,} & \sum_{k=1}^{4} \omega_{kF}\omega_{kF}' \end{pmatrix}, \qquad \text{(S.8)}$$

where $\{f_{0t}\lambda_{0i}'\}_{i\le N, t\le T}$ denotes a $2N \times 2T$ matrix whose $\{i, t\}$th block is $f_{0t}\lambda_{0i}'$. Further, it is easy to see that under our normalizations,

$$\omega'\omega = \begin{pmatrix} \sigma_{N1}+1 & 0 & 0 & 0 \\ 0 & \sigma_{N2}+1 & 0 & 0 \\ 0 & 0 & \sigma_{N2}+1 & 0 \\ 0 & 0 & 0 & \sigma_{N1}+1 \end{pmatrix}.$$

Next, we project γ_k onto ω, and write $\gamma_k = \omega\beta_k + \zeta_k$ for $k = 1, \ldots, 4$, where $\beta_k = (\omega'\omega)^{-1}\omega'\gamma_k$. In particular,

$$\beta_1 = \begin{pmatrix} -\frac{1}{\sigma_{N1}+1} \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \beta_2 = \begin{pmatrix} 0 \\ -\frac{1}{\sigma_{N2}+1} \\ 0 \\ 0 \end{pmatrix},$$

$$\beta_3 = \begin{pmatrix} 0 \\ 0 \\ -\frac{1}{\sigma_{N2}+1} \\ -\frac{1}{\sigma_{N1}+1} \end{pmatrix}, \qquad \beta_4 = \begin{pmatrix} 0 \\ 0 \\ \frac{\sigma_{N2}}{\sigma_{N2}+1} \\ \frac{\sigma_{N1}}{\sigma_{N1}+1} \end{pmatrix}.$$

Define $B_N = \sum_{k=1}^{4} \beta_k\beta_k'$. It is easy to show that there exists $\underline{\rho} > 0$ such that $\rho_{\min}(B_N) > \underline{\rho}$ for all large N as long as $\sigma_{N1} - \sigma_{N2}$ is bounded below by a positive constant for all large N, which is true under our assumption that $\sigma_{N1} \to \sigma_1$, $\sigma_{N2} \to \sigma_2$, and $\sigma_1 > \sigma_2$. It then follows that

$$\begin{aligned} \partial^2\mathbb{P}_{NT}(\theta_0)/\partial\theta\partial\theta' &= b\left(\sum_{k=1}^{4} \gamma_k\gamma_k'\right) = b \cdot \omega\left(\sum_{k=1}^{4} \beta_k\beta_k'\right)\omega' + b\left(\sum_{k=1}^{4} \zeta_k\zeta_k'\right) \\ &= b\underline{\rho} \cdot \omega\omega' + b \cdot \omega(B_N - \underline{\rho} \cdot \mathbb{I}_4)\omega' + b\left(\sum_{k=1}^{4} \zeta_k\zeta_k'\right). \end{aligned} \qquad \text{(S.9)}$$

Now let $\underline{b} = \min\{\underline{\mathfrak{f}}, b\underline{\rho}\}$. Then it follows from (S.9) that

$$\mathcal{H} = \partial\mathcal{S}^*(\theta_0)/\partial\theta' + \partial^2\mathbb{P}_{NT}(\theta_0)/\partial\theta\partial\theta'$$

$$= \partial \mathcal{S}^*(\theta_0)/\partial\theta' + \underline{b}\cdot\omega\omega' + \underbrace{(b\underline{\rho} - \underline{b})\cdot\omega\omega'}_{\geq 0} + \underbrace{b\cdot\omega(B_N - \underline{\rho} I_4)\omega'}_{\geq 0} + \underbrace{b\left(\sum_{k=1}^{4}\zeta_k\zeta_k'\right)}_{\geq 0}$$

$$\geq \partial \mathcal{S}^*(\theta_0)/\partial\theta' + \underline{b}\cdot\omega\omega'.$$

Moreover, we can write

$$\begin{aligned}
&\partial \mathcal{S}^*(\theta_0)/\partial\theta' \\
&= \begin{pmatrix} (NT)^{-1/2}\operatorname{diag}\left[\left\{\sum_{t=1}^{T}\bar{\varrho}_{it}^{(2)} f_{0t}f_{0t}'\right\}_{i\leq N}\right] & (NT)^{-1/2}\left\{\bar{\varrho}_{it}^{(2)} f_{0t}\lambda_{0i}'\right\}_{i\leq N, t\leq T} \\ (NT)^{-1/2}\left\{\bar{\varrho}_{it}^{(2)}\lambda_{0i}f_{0t}'\right\}_{t\leq T, i\leq N} & (NT)^{-1/2}\operatorname{diag}\left[\left\{\sum_{i=1}^{N}\bar{\varrho}_{it}^{(2)}\lambda_{0i}\lambda_{0i}'\right\}_{t\leq N}\right]\end{pmatrix} \\
&= \underbrace{\underline{b}\begin{pmatrix}\operatorname{diag}\left[\left\{(NT)^{-1/2}\sum_{t=1}^{T} f_{0t}f_{0t}'\right\}_{i\leq N}\right] & \mathbf{0}_{2N\times 2T} \\ \mathbf{0}_{2T\times 2N} & \operatorname{diag}\left[\left\{(NT)^{-1/2}\sum_{i=1}^{N}\lambda_{0i}\lambda_{0i}'\right\}_{t\leq N}\right]\end{pmatrix}}_{I} \\
&\quad + \underbrace{\underline{b}\begin{pmatrix}\mathbf{0}_{2N\times 2N} & (NT)^{-1/2}\left\{f_{0t}\lambda_{0i}'\right\}_{i\leq N, t\leq T} \\ (NT)^{-1/2}\left\{\lambda_{0i}f_{0t}'\right\}_{t\leq T, i\leq N} & \mathbf{0}_{2T\times 2T}\end{pmatrix}}_{II} \\
&\quad + \underbrace{\begin{pmatrix}(NT)^{-1/2}\operatorname{diag}\left[\left\{\sum_{t=1}^{T}(\bar{\varrho}_{it}^{(2)} - \underline{b}) f_{0t}f_{0t}'\right\}_{i\leq N}\right] & (NT)^{-1/2}\left\{(\bar{\varrho}_{it}^{(2)} - \underline{b}) f_{0t}\lambda_{0i}'\right\}_{i\leq N, t\leq T} \\ (NT)^{-1/2}\left\{(\bar{\varrho}_{it}^{(2)} - \underline{b})\lambda_{0i}f_{0t}'\right\}_{t\leq T, i\leq N} & (NT)^{-1/2}\operatorname{diag}\left[\left\{\sum_{i=1}^{N}(\bar{\varrho}_{it}^{(2)} - \underline{b})\lambda_{0i}\lambda_{0i}'\right\}_{t\leq N}\right]\end{pmatrix}}_{III}.
\end{aligned}$$

Note that by our Assumptions 1 and 2, there exists a constant $\underline{c} > 0$ such that

$$I = \underline{b}\begin{pmatrix}\sqrt{T/N}\cdot\mathbb{I}_{2N} & \mathbf{0}_{2N\times 2T} \\ \mathbf{0}_{2T\times 2N} & \sqrt{N/T}\cdot\mathbb{I}_T\otimes\operatorname{diag}(\sigma_{N1},\sigma_{N2})\end{pmatrix} \geq \underline{c}\cdot\mathbb{I}_{2(N+T)}. \tag{S.10}$$

From (S.8), we have

$$II + \underline{b}\cdot\omega\omega' = \underline{b}\cdot\begin{pmatrix}\sum_{k=1}^{4}\omega_{k\Lambda}\omega_{k\Lambda}' & \mathbf{0}_{2N\times 2T} \\ \mathbf{0}_{2T\times 2N} & \sum_{k=1}^{4}\omega_{kF}\omega_{kF}'\end{pmatrix} \geq 0. \tag{S.11}$$

For the last term, we have, for N, T large enough,

$$III = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \big(\bar{\varrho}_{it}^{(2)} - \underline{b}\big) \mu_{it} \mu_{it}' \geq 0, \tag{S.12}$$

where $\mu_{it} = [\underbrace{\mathbf{0}_{1\times 2}, \ldots, f_{0t}', \ldots, \mathbf{0}_{1\times 2}}_{1\times 2N}, \underbrace{\mathbf{0}_{1\times 2}, \ldots, \lambda_{0i}', \ldots, \mathbf{0}_{1\times 2}}_{1\times 2T}]'$, because Assumption 1 and Lemma S.1 imply that $\bar{\varrho}_{it}^{(2)} \geq \underline{\mathrm{f}}$ for all i, t. It then follows from (S.10), (S.11), and (S.12) that

$$\mathcal{H} \geq \partial \mathcal{S}^*(\theta_0)/\partial \theta' + \underline{b} \cdot \omega\omega' = I + II + III + \underline{b} \cdot \omega\omega' \geq \underline{c} \cdot \mathbb{I}_{2(N+T)},$$

and thus

$$\mathcal{H}^{-1} \leq \underline{c}^{-1} \cdot \mathbb{I}_{2(N+T)}. \tag{S.13}$$

Finally, write $\mathcal{H} = \mathcal{H}_d + \mathcal{C}$, where

$$\mathcal{C} = \begin{pmatrix} \mathbf{0}_{2N\times 2N} & (NT)^{-1/2}\big\{\bar{\varrho}_{it}^{(2)} f_{0t} \lambda_{0i}'\big\}_{i\leq N, t\leq T} \\ (NT)^{-1/2}\big\{\bar{\varrho}_{it}^{(2)} \lambda_{0i} f_{0t}'\big\}_{t\leq T, i\leq N} & \mathbf{0}_{2T\times 2T} \end{pmatrix} + b\left(\sum_{k=1}^{4} \gamma_k \gamma_k'\right).$$

Note that $\mathcal{H}^{-1} - \mathcal{H}_d^{-1} = -\mathcal{H}_d^{-1}\mathcal{C}\mathcal{H}_d^{-1} + \mathcal{H}_d^{-1}\mathcal{C}\mathcal{H}^{-1}\mathcal{C}\mathcal{H}_d^{-1}$, and thus $\|\mathcal{H}^{-1} - \mathcal{H}_d^{-1}\|_{\max} \leq \|\mathcal{H}_d^{-1}\mathcal{C}\mathcal{H}_d^{-1}\|_{\max} + \|\mathcal{H}_d^{-1}\mathcal{C}\mathcal{H}^{-1}\mathcal{C}\mathcal{H}_d^{-1}\|_{\max}$. Inequality (S.13) implies that $\mathcal{H}_d^{-1}\mathcal{C}\mathcal{H}^{-1}\mathcal{C}\mathcal{H}_d^{-1} \leq \underline{c}^{-1}\mathcal{H}_d^{-1}\mathcal{C}^2\mathcal{H}_d^{-1}$, and thus the jth diagonal element of $\mathcal{H}_d^{-1}\mathcal{C}\mathcal{H}^{-1}\mathcal{C}\mathcal{H}_d^{-1}$ is smaller than the jth diagonal element of $\underline{c}^{-1}\mathcal{H}_d^{-1}\mathcal{C}^2\mathcal{H}_d^{-1}$. It then follows that $\|\mathcal{H}_d^{-1}\mathcal{C}\mathcal{H}^{-1}\mathcal{C}\mathcal{H}_d^{-1}\|_{\max} \leq \underline{c}^{-1}\|\mathcal{H}_d^{-1}\mathcal{C}^2\mathcal{H}_d^{-1}\|_{\max}$ and therefore

$$\big\|\mathcal{H}^{-1} - \mathcal{H}_d^{-1}\big\|_{\max} \leq \big\|\mathcal{H}_d^{-1}\mathcal{C}\mathcal{H}_d^{-1}\big\|_{\max} + \underline{c}^{-1}\big\|\mathcal{H}_d^{-1}\mathcal{C}^2\mathcal{H}_d^{-1}\big\|_{\max}$$

because the entry with the largest absolute value of a positive semidefinite matrix is always on the diagonal. Since $\mathcal{H}_d^{-1}$ is a block-diagonal matrix whose elements are all $O(1)$ by Assumption 2, and both $\|\mathcal{C}\|_{\max}$ and $\|\mathcal{C}^2\|_{\max}$ can be shown to be $O(1/T)$, then the desired result follows. *Q.E.D.*

Since $\partial \mathbb{P}_{NT}(\tilde{\theta})/\partial\theta = \partial \mathbb{P}_{NT}(\theta_0)/\partial\theta = 0$, (S.7) implies that

$$\tilde{\theta} - \theta_0 = \mathcal{H}^{-1}\mathcal{S}^*(\tilde{\theta}) - \mathcal{H}^{-1}\mathcal{S}^*(\theta_0) - 0.5\mathcal{H}^{-1}\mathcal{R}(\tilde{\theta}). \tag{S.14}$$

Define

$$\mathcal{S}_{NT}^*(\theta) = \Bigg[\underbrace{\ldots, -\frac{1}{\sqrt{NT}} \sum_{t=1}^{T} \varrho^{(1)}\big(X_{it} - \lambda_i' f_t\big) f_t', \ldots}_{1\times Nr}, \underbrace{\ldots, -\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \varrho^{(1)}\big(X_{it} - \lambda_i' f_t\big) \lambda_i', \ldots}_{1\times Tr}\Bigg]',$$

$\tilde{\mathcal{S}}^*(\theta) = \mathcal{S}_{NT}^*(\theta) - \mathcal{S}^*(\theta)$, and $\mathcal{D} = \mathcal{H}^{-1} - \mathcal{H}_d^{-1}$. Note that by the first-order conditions, $\mathcal{S}_{NT}^*(\tilde{\theta}) = 0$. As a result, we can write

$$\mathcal{H}^{-1}\mathcal{S}^*(\tilde{\theta}) = \mathcal{H}_d^{-1}\mathcal{S}^*(\tilde{\theta}) + \mathcal{D}\mathcal{S}^*(\tilde{\theta}) = -\mathcal{H}_d^{-1}\tilde{\mathcal{S}}^*(\tilde{\theta}) + \mathcal{D}\mathcal{S}^*(\tilde{\theta})$$

$$= -\mathcal{H}_d^{-1}\tilde{\mathcal{S}}^*(\theta_0) - \mathcal{H}_d^{-1}\big(\tilde{\mathcal{S}}^*(\tilde{\theta}) - \tilde{\mathcal{S}}^*(\theta_0)\big) + \mathcal{D}\mathcal{S}^*(\tilde{\theta}) \tag{S.15}$$

$$= -\mathcal{H}_d^{-1}\tilde{\mathcal{S}}^*(\theta_0) - \mathcal{H}_d^{-1}\big(\tilde{\mathcal{S}}^*(\tilde{\theta}) - \tilde{\mathcal{S}}^*(\theta_0)\big) - \mathcal{D}\tilde{\mathcal{S}}^*(\theta_0) - \mathcal{D}\big(\tilde{\mathcal{S}}^*(\tilde{\theta}) - \tilde{\mathcal{S}}^*(\theta_0)\big). \tag{S.16}$$

Next, let $\mathcal{R}(\tilde{\theta})_j$ denote the vector containing the $(j-1)r+1$th to the jrth elements of $\mathcal{R}(\tilde{\theta})$ for $j = 1, \ldots, N+T$, and let $\bar{O}_P(\cdot)$ denote a stochastic order that is uniformly in i and t.[1] Then, by the result of Lemma S.3, it can be shown that

$$\mathcal{R}(\tilde{\theta})_i = \bar{O}_P(1)\|\tilde{\lambda}_i - \lambda_{0i}\|^2 + \bar{O}_P(1/\sqrt{T})\|\tilde{\lambda}_i - \lambda_{0i}\| + \bar{O}_P(1/T) \tag{S.17}$$

for $i = 1, \ldots, N$, and

$$\mathcal{R}(\tilde{\theta})_{N+t} = \bar{O}_P(1)\|\tilde{f}_t - f_{0t}\|^2 + \bar{O}_P(1/\sqrt{T})\|\tilde{f}_t - f_{0t}\| + \bar{O}_P(1/T) \tag{S.18}$$

for $t = 1, \ldots, T$.

Write $\mathcal{D}_{j,s}$ as the $r \times r$ matrix containing the $(j-1)r+1$ to jr rows and $(s-1)r+1$ to sr columns of $\mathcal{D}$. Note that Lemma S.1 and Lemma S.6 imply that $\|\mathcal{H}^{-1}\mathcal{S}^*(\theta_0)\|_{\max} = \bar{O}(h^m)$. Then, from (S.14) to (S.18), we can write

$$\begin{aligned}
\tilde{f}_t - f_{0t} = {} & (\Psi_{N,t})^{-1}\frac{1}{N}\sum_{j=1}^{N}\tilde{\varrho}_{jt}^{(1)}\lambda_{0j} + \frac{1}{\sqrt{NT}}\sum_{j=1}^{N}\sum_{s=1}^{T}\mathcal{D}_{N+t,j}\cdot\tilde{\varrho}_{js}^{(1)}\cdot f_{0s} \\
& + \frac{1}{\sqrt{NT}}\sum_{j=1}^{N}\sum_{s=1}^{T}\mathcal{D}_{N+t,N+s}\cdot\tilde{\varrho}_{js}^{(1)}\cdot\lambda_{0j} \\
& + (\Psi_{N,t})^{-1}\frac{1}{N}\sum_{j=1}^{N}\big\{\tilde{\varrho}^{(1)}\big(X_{jt} - \tilde{\lambda}_j'\tilde{f}_t\big)\tilde{\lambda}_j - \tilde{\varrho}_{jt}^{(1)}\lambda_{0j}\big\} \\
& + \frac{1}{\sqrt{NT}}\sum_{j=1}^{N}\sum_{s=1}^{T}\mathcal{D}_{N+t,j}\big\{\tilde{\varrho}^{(1)}\big(X_{js} - \tilde{\lambda}_j'\tilde{f}_s\big)\tilde{f}_s - \tilde{\varrho}_{js}^{(1)}f_{0s}\big\} \\
& + \frac{1}{\sqrt{NT}}\sum_{j=1}^{N}\sum_{s=1}^{T}\mathcal{D}_{N+t,N+s}\big\{\tilde{\varrho}^{(1)}\big(X_{js} - \tilde{\lambda}_j'\tilde{f}_s\big)\tilde{\lambda}_j - \tilde{\varrho}_{js}^{(1)}\lambda_{0j}\big\} \\
& - 0.5(\Psi_{N,t})^{-1}\mathcal{R}(\tilde{\theta})_{N+t} - 0.5\sum_{j=1}^{N}\mathcal{D}_{N+t,j}\mathcal{R}(\tilde{\theta})_j \\
& - 0.5\sum_{s=1}^{T}\mathcal{D}_{N+t,N+s}\mathcal{R}(\tilde{\theta})_{N+s} + \bar{O}\big(h^m\big).
\end{aligned} \tag{S.19}$$

[1]For example, $Z_{it} = \bar{O}_P(1)$ means that $\max_{i\le N, t\le T}\|Z_{it}\| = O_P(1)$.

LEMMA S.7: *Let* $c_1, \dots, c_T$ *be a sequence of uniformly bounded constants. Then, under Assumptions* 1 *and* 2, *we have*

$$\frac{1}{T}\sum_{t=1}^{T} c_t(\tilde{f}_t - f_{0t}) = O_P\left(\frac{1}{Th}\right).$$

PROOF: Define $d_j = \sqrt{NT} \cdot T^{-1}\sum_{t=1}^{T} c_t \mathcal{D}_{N+t,j}$ for $j = 1, \dots, N+T$. Lemma S.6 implies that $\max_{1\le j\le N+T}\|d_j\|$ is bounded. From (S.19), we have

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T} c_t(\tilde{f}_t - f_{0t}) \\
&= \frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T} c_t(\Psi_{N,t})^{-1}\tilde{\varrho}_{jt}^{(1)}\lambda_{0j} + \frac{1}{NT}\sum_{j=1}^{N}\sum_{s=1}^{T} d_j\tilde{\varrho}_{js}^{(1)} f_{0s} + \frac{1}{NT}\sum_{j=1}^{N}\sum_{s=1}^{T} d_{N+s}\tilde{\varrho}_{js}^{(1)}\lambda_{0j} \\
&\quad + \frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T} c_t(\Psi_{N,t})^{-1}\big\{\tilde{\varrho}^{(1)}\big(X_{jt} - \tilde{\lambda}_j'\tilde{f}_t\big)\tilde{\lambda}_j - \tilde{\varrho}_{jt}^{(1)}\lambda_{0j}\big\} \\
&\quad + \frac{1}{NT}\sum_{j=1}^{N}\sum_{s=1}^{T} d_j\big\{\tilde{\varrho}^{(1)}\big(X_{js} - \tilde{\lambda}_j'\tilde{f}_s\big)\tilde{f}_s - \tilde{\varrho}_{js}^{(1)} f_{0s}\big\} \\
&\quad + \frac{1}{NT}\sum_{j=1}^{N}\sum_{s=1}^{T} d_{N+s}\big\{\tilde{\varrho}^{(1)}\big(X_{js} - \tilde{\lambda}_j'\tilde{f}_s\big)\tilde{\lambda}_j - \tilde{\varrho}_{js}^{(1)}\lambda_{0j}\big\} - 0.5\frac{1}{T}\sum_{t=1}^{T} c_t(\Psi_{N,t})^{-1}\mathcal{R}(\tilde{\theta})_{N+t} \\
&\quad - 0.5\frac{1}{\sqrt{NT}}\sum_{j=1}^{N} d_j\mathcal{R}(\tilde{\theta})_j - 0.5\frac{1}{\sqrt{NT}}\sum_{s=1}^{T} d_{N+s}\mathcal{R}(\tilde{\theta})_{N+s} + \bar{O}\big(h^m\big).
\end{aligned} \tag{S.20}$$

First, by Lyapunov's CLT, it is easy to see that the first three terms on the RHS of (S.20) are all $O_P(1/\sqrt{NT})$.

Next, it follows from Lemma S.3, (S.17), (S.18), and Assumption 2(v) that the last four terms on the RHS of (S.20) are all $O_P(1/L_{NT}^2)$. Finally, we will show that the remaining three terms on the RHS of (S.20) are all $O_P(1/(Th))$, from which the desired result follows.

Define

$$\mathbb{V}_{NT}(\theta) = \frac{1}{NT}\sum_{j=1}^{N}\sum_{s=1}^{T} d_j\big\{\tilde{\varrho}^{(1)}\big(X_{js} - \lambda_j' f_s\big) f_s - \tilde{\varrho}_{js}^{(1)} f_{0s}\big\},$$

and $\Delta_{NT}(\theta_a, \theta_b) = \sqrt{NT}h[\mathbb{V}_{NT}(\theta_a) - \mathbb{V}_{NT}(\theta_b)]$. Note that

$$\Delta_{NT}(\theta_a, \theta_b) = \underbrace{\frac{h}{\sqrt{NT}}\sum_{j=1}^{N}\sum_{s=1}^{T} d_j \cdot \tilde{\varrho}^{(1)}\big(X_{js} - \lambda_{aj}' f_{as}\big)\cdot(f_{as} - f_{bs})}_{\Delta_{1,NT}(\theta_a,\theta_b)}$$

$$+ \underbrace{\frac{h}{\sqrt{NT}} \sum_{j=1}^{N} \sum_{s=1}^{T} d_j \cdot \left[\tilde{\varrho}^{(1)}\left(X_{js} - \lambda'_{aj} f_{as}\right) - \tilde{\varrho}^{(1)}\left(X_{js} - \lambda'_{bj} f_{bs}\right)\right] \cdot f_{bs}}_{\Delta_{2,NT}(\theta_a,\theta_b)}.$$

Lemma S.1 implies that

$$h\left|d_j \cdot \tilde{\varrho}^{(1)}\left(X_{js} - \lambda'_{aj} f_{as}\right) \cdot (f_{as} - f_{bs})\right| \lesssim \left|\lambda'_{aj} f_{as} - \lambda'_{bj} f_{bs}\right|,$$
$$h\left|\varrho^{(1)}\left(X_{js} - \lambda'_{aj} f_{as}\right) - \varrho^{(1)}\left(X_{js} - \lambda'_{bj} f_{bs}\right)\right| \lesssim \left|\lambda'_{aj} f_{as} - \lambda'_{bj} f_{bs}\right|.$$

By Hoeffding's inequality and Lemma 2.2.1 of van der Vaart and Wellner (1996), we can show that for $d(\theta_a, \theta_b)$ sufficiently small,

$$\left\|\Delta_{1,NT}(\theta_a, \theta_b)\right\|_{\psi_2} \lesssim d(\theta_a, \theta_b) \quad \text{and} \quad \left\|\Delta_{2,NT}(\theta_a, \theta_b)\right\|_{\psi_2} \lesssim d(\theta_a, \theta_b).$$

Thus,

$$\left\|\Delta_{NT}(\theta_a, \theta_b)\right\|_{\psi_2} \le \left\|\Delta_{1,NT}(\theta_a, \theta_b)\right\|_{\psi_2} + \left\|\Delta_{2,NT}(\theta_a, \theta_b)\right\|_{\psi_2} \lesssim d(\theta_a, \theta_b).$$

Therefore, similar to the proof of Lemma 3, we can show that for sufficiently small $\delta > 0$,

$$\mathbb{E}\Big[\sup_{\theta \in \Theta^r(\delta)} \left|\mathbb{V}_{NT}(\theta)\right|\Big] \lesssim \frac{\delta}{L_{NT} h}. \tag{S.21}$$

It then follows from (S.21) and Lemma S.3 that $\mathbb{V}_{NT}(\tilde{\theta}) = O_P(1/(L_{NT}^2 h)) = O_P(1/(Th))$, for example, the fifth term on the right of (S.20) is $O_P(1/(Th))$. Similar results can be obtained for the fourth and sixth terms on the right of (S.20), and thus the desired result follows. *Q.E.D.*

LEMMA S.8: *Under Assumptions* 1 *and* 2, *for each* i *we have*

$$\frac{1}{T}\sum_{t=1}^{T} \tilde{\varrho}_{it}^{(1)}(\tilde{f}_t - f_{0t}) = O_P\left(\frac{1}{Th}\right) \quad \textit{and} \quad \frac{1}{T}\sum_{t=1}^{T} \tilde{\varrho}_{it}^{(2)} f_{0t}(\tilde{f}_t - f_{0t})' = O_P\left(\frac{1}{Th^2}\right).$$

PROOF: To save space, we only prove the second result, since the proof of the first result is similar. Using (S.15) and (S.19), we can write

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T} \tilde{\varrho}_{it}^{(2)} f_{0t}(\hat{f}_t - f_{0t})' \\
&\quad = \frac{1}{NT}\sum_{t=1}^{T}\sum_{j=1}^{N} \tilde{\varrho}_{it}^{(2)} \tilde{\varrho}_{jt}^{(1)} f_{0t} \lambda'_{0j} (\Psi_{N,t})^{-1} \\
&\qquad + \frac{1}{NT}\sum_{t=1}^{T}\sum_{j=1}^{N} \tilde{\varrho}_{it}^{(2)} f_{0t} \cdot \left\{\tilde{\varrho}^{(1)}\left(X_{jt} - \tilde{\lambda}'_j \tilde{f}_t\right)\tilde{\lambda}'_j - \tilde{\varrho}_{jt}^{(1)} \lambda'_{0j}\right\}(\Psi_{N,t})^{-1} \\
&\qquad - \frac{1}{\sqrt{NT}}\sum_{j=1}^{N}\sum_{s=1}^{T}\left(\frac{1}{T}\sum_{t=1}^{T} \tilde{\varrho}_{it}^{(2)} f_{0t} \tilde{f}'_s \mathcal{D}'_{N+t,j}\right)\bar{\varrho}^{(1)}\left(X_{js} - \tilde{\lambda}'_j \tilde{f}_s\right)
\end{aligned}$$

$$- \frac{1}{\sqrt{NT}} \sum_{j=1}^{N} \sum_{s=1}^{T} \left(\frac{1}{T} \sum_{t=1}^{T} \tilde{\varrho}_{it}^{(2)} f_{0t} \tilde{\lambda}_j' \mathcal{D}_{N+t,N+s}' \right) \bar{\varrho}^{(1)}\big(X_{js} - \tilde{\lambda}_j' \tilde{f}_s\big)$$

$$+ \frac{1}{2T} \sum_{t=1}^{T} \tilde{\varrho}_{it}^{(2)} f_{0t} \mathcal{R}(\tilde{\theta})_{N+t}' (\Psi_{N,t})^{-1} + \frac{1}{2T} \sum_{j=1}^{N} \sum_{t=1}^{T} \tilde{\varrho}_{it}^{(2)} f_{0t} \mathcal{R}(\tilde{\theta})_j' \mathcal{D}_{N+t,j}'$$

$$+ \frac{1}{2T} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{\varrho}_{it}^{(2)} f_{0t} \mathcal{R}(\tilde{\theta})_{N+s}' \mathcal{D}_{N+t,N+s}' + O\big(h^{m-1}\big). \tag{S.22}$$

First, we can write

$$\frac{1}{NT} \sum_{t=1}^{T} \sum_{j=1}^{N} \tilde{\varrho}_{it}^{(2)} \tilde{\varrho}_{jt}^{(1)} f_{0t} \lambda_{0j}' (\Psi_{N,t})^{-1}$$

$$= \frac{1}{NT} \sum_{t=1}^{T} \tilde{\varrho}_{it}^{(2)} \tilde{\varrho}_{it}^{(1)} f_{0t} \lambda_{0i}' (\Psi_{N,t})^{-1} + \frac{1}{NT} \sum_{t=1}^{T} \sum_{j=1, j\neq i}^{N} \tilde{\varrho}_{it}^{(2)} \tilde{\varrho}_{jt}^{(1)} f_{0t} \lambda_{0j}' (\Psi_{N,t})^{-1}.$$

Since $h\varrho_{it}^{(2)}(\cdot)$ is uniformly bounded by Lemma S.1, $\max_{t\leq T} \|(\Psi_{N,t})^{-1}\| = O(1)$ for large N by Assumption 2, the first term on the RHS of the above equation is $O_P((Nh)^{-1})$. Using Lyapunov's CLT and Lemma S.1, the second term on the RHS of the above equation can be shown to be $O_P((NTh)^{-1/2})$. Thus, the first term on the RHS of (S.22) is $O_P((Th)^{-1})$.

Second, consider the second term on the RHS of (S.22), which can be written as

$$O_P\left(\frac{1}{Nh}\right) + \frac{1}{NT} \sum_{t=1}^{T} \sum_{j=1, j\neq i}^{N} \tilde{\varrho}_{it}^{(2)} f_{0t} \cdot \big\{ \bar{\varrho}^{(1)}\big(X_{jt} - \tilde{\lambda}_j' \tilde{f}_t\big) \tilde{\lambda}_j' - \tilde{\varrho}_{jt}^{(1)} \lambda_{0j}' \big\} (\Psi_{N,t})^{-1}.$$

Similar to the proof of Lemma S.7, the second term of the above expression can be shown to be $O_P(1/(Th^2))$. So the second term on the RHS of (S.22) is $O_P(1/(Th^2))$.

Next, for the third term on the RHS of (S.22), its p, qth element is given by

$$\frac{1}{NT} \sum_{j=1}^{N} \sum_{s=1}^{T} \chi_{i,j} \cdot \bar{\varrho}^{(1)}\big(X_{js} - \tilde{\lambda}_j' \tilde{f}_s\big) \tilde{f}_s,$$

where $\chi_{i,j} = T^{-1} \sum_{t=1}^{T} (\sqrt{NT} \mathcal{D}_{N+t,j,q}) \cdot f_{0t,p} \tilde{\varrho}_{it}^{(2)}$, and $\mathcal{D}_{N+t,j,q}$ is the qth row of $\mathcal{D}_{N+t,j}$. Therefore,

$$\left\| \frac{1}{NT} \sum_{j=1}^{N} \sum_{s=1}^{T} \chi_{i,j} \bar{\varrho}^{(1)}\big(X_{js} - \tilde{\lambda}_j' \tilde{f}_s\big) \tilde{f}_s \right\|$$

$$\leq \sqrt{\frac{1}{N} \sum_{j=1}^{N} \|\chi_{i,j}\|^2} \cdot \sqrt{\frac{1}{NT} \sum_{j=1}^{N} \sum_{s=1}^{T} \big[\bar{\varrho}^{(1)}\big(X_{js} - \tilde{\lambda}_j' \tilde{f}_s\big)\big]^2 \|\tilde{f}_s\|^2}.$$

Since $|\sqrt{NT} f_{0t,p} \mathcal{D}_{N+t,j,q}|$ is uniformly is bounded by Lemma S.6, it can be shown that $\mathbb{E}\|\chi_{i,j}\|^2 = O((Th)^{-1})$. Moreover,

$$\big[\bar{\varrho}^{(1)}\big(X_{js} - \tilde{\lambda}_j' \tilde{f}_s\big)\big]^2 = \big[\bar{\varrho}_{it}^{(1)} + \bar{\varrho}^{(2)}\big(X_{js} - \lambda_j^{*\prime} f_s^*\big) \cdot \big(\lambda_{0i}' f_{0t} - \tilde{\lambda}_j' \tilde{f}_s\big)\big]^2 \lesssim O\big(h^{2m}\big) + \big(\lambda_{0i}' f_{0t} - \tilde{\lambda}_j' \tilde{f}_s\big)^2,$$

thus,

$$\sqrt{\frac{1}{NT}\sum_{j=1}^{N}\sum_{s=1}^{T}\big[\bar{\varrho}^{(1)}\big(X_{js}-\tilde{\lambda}_j'\tilde{f}_s\big)\big]^2\|\tilde{f}_s\|^2} \lesssim O\big(h^m\big)+d(\tilde{\theta},\theta_0)=O_P(1/L_{NT})$$

by Lemma S.3. So, the third term on the RHS of (S.22) is $O_P(T^{-1}h^{-1/2})$, while the fourth term on the RHS of (S.22) can be shown to be $O_P(T^{-1}h^{-1/2})$ in the same way.

Finally, it follows from Lemma S.3 and (S.18) that the fifth term on the RHS of (S.22) is $O_P((L_{NT})^{-2}h^{-1})=O_P((Th)^{-1})$. The p,qth element of the sixth term on the RHS of (S.22) can be written as $(2\sqrt{NT})^{-1}\sum_{j=1}^{N}\chi_{i,j}\mathcal{R}(\tilde{\theta})_j$, which is bounded by

$$\frac{\sqrt{N}}{2\sqrt{T}}\sqrt{\frac{1}{N}\sum_{j=1}^{N}\|\chi_{i,j}\|^2}\sqrt{\frac{1}{N}\sum_{j=1}^{N}\big\|\mathcal{R}(\tilde{\theta})_j\big\|^2}=O_P\big((Th)^{-1/2}\big)O_P\big(L_{NT}^{-2}\big)=O_P\big(T^{-3/2}h^{-1/2}\big).$$

The same bound for the seventh term on the RHS of (S.22) can be obtained using the same argument. Thus, combining the above results, we get

$$\frac{1}{T}\sum_{t=1}^{T}\bar{\varrho}_{it}^{(2)}f_{0t}(\tilde{f}_t-f_{0t})'=O_P\bigg(\frac{1}{Th^2}\bigg).$$

Q.E.D.

PROOF OF THEOREM 4: From the expansion in the proof of Lemma S.5, we have

$$\begin{aligned}\Phi_{T,i}(\tilde{\lambda}_i-\lambda_{0i}) = &-\frac{1}{T}\sum_{t=1}^{T}\bar{\varrho}^{(1)}\big(X_{it}-\tilde{\lambda}_i'\tilde{f}_t\big)\tilde{f}_t+\frac{1}{T}\sum_{t=1}^{T}\bar{\varrho}_{it}^{(1)}f_{0t}+\frac{1}{T}\sum_{t=1}^{T}\bar{\varrho}_{it}^{(1)}(\tilde{f}_t-f_{0t})\\ &-\frac{1}{T}\sum_{t=1}^{T}\bar{\varrho}_{it}^{(2)}f_{0t}(\tilde{f}_t-f_{0t})'\lambda_{0i}+O_P\big(T^{-1}\|\tilde{F}-F_0\|^2\big)+o_P\big(\|\tilde{\lambda}_i-\lambda_{0i}\|\big).\end{aligned}$$

Then, it follows from Lemmas S.1, S.3, and S.7 that

$$\Phi_{T,i}(\tilde{\lambda}_i-\lambda_{0i})=-\frac{1}{T}\sum_{t=1}^{T}\bar{\varrho}^{(1)}\big(X_{it}-\tilde{\lambda}_i'\tilde{f}_t\big)\tilde{f}_t+O\big(h^m\big)+O_P\bigg(\frac{1}{Th}\bigg)+o_P\big(\|\tilde{\lambda}_i-\lambda_{0i}\|\big).$$

Note that

$$\begin{aligned}&-\frac{1}{T}\sum_{t=1}^{T}\bar{\varrho}^{(1)}\big(X_{it}-\tilde{\lambda}_i'\tilde{f}_t\big)\tilde{f}_t\\ &=\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}^{(1)}\big(X_{it}-\tilde{\lambda}_i'\tilde{f}_t\big)\tilde{f}_t\\ &=\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(1)}\tilde{f}_t-\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(2)}\cdot\big(\tilde{\lambda}_i'\tilde{f}_t-\lambda_{0i}'f_{0t}\big)\tilde{f}_t+0.5\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(3)}(*)\big(\tilde{\lambda}_i'\tilde{f}_t-\lambda_{0i}'f_{0t}\big)^2\tilde{f}_t,\end{aligned}$$

where $\tilde{\varrho}_{it}^{(3)}(*)=\tilde{\varrho}_{it}^{(3)}(c_{it}^*)$ and c_{it}^* is between $\lambda_{0i}'f_{0t}$ and $\tilde{\lambda}_i'\tilde{f}_t$.

First, by Lemma S.8, we have

$$\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(1)}\tilde{f}_t=\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(1)}f_{0t}+\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(1)}(\tilde{f}_t-f_{0t})=\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(1)}f_{0t}+O_P\left(\frac{1}{Th}\right).$$

Second,

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(2)}\cdot\left(\tilde{\lambda}_i'\tilde{f}_t-\lambda_{0i}'f_{0t}\right)\tilde{f}_t\\
&\quad=\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(2)}\tilde{f}_t\cdot(\tilde{f}_t-f_{0t})'\tilde{\lambda}_i+\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(2)}\tilde{f}_tf_{0t}'\cdot(\tilde{\lambda}_i-\lambda_{0i})\\
&\quad=\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(2)}f_{0t}\cdot(\tilde{f}_t-f_{0t})'\tilde{\lambda}_i+\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(2)}(\tilde{f}_t-f_{0t})\cdot(\tilde{f}_t-f_{0t})'\tilde{\lambda}_i\\
&\qquad+\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(2)}f_{0t}f_{0t}'\cdot(\tilde{\lambda}_i-\lambda_{0i})\\
&\qquad+\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(2)}(\tilde{f}_t-f_{0t})f_{0t}'\cdot(\tilde{\lambda}_i-\lambda_{0i}).
\end{aligned}\tag{S.23}$$

It then follows from Lemma S.3 that the second term on the RHS of (S.23) is $O_P((Th)^{-1})$, and Lemma S.8 implies that the first term is $O_P(T^{-1}h^{-2})$. It is easy to show that the last two terms on the right of (S.23) are both $o_P(\|\tilde{\lambda}_i-\lambda_{0i}\|)$. Thus, we have

$$\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(2)}\cdot\left(\tilde{\lambda}_i'\tilde{f}_t-\lambda_{0i}'f_{0t}\right)\tilde{f}_t=O_P\left(T^{-1}h^{-2}\right)+o_P\left(\|\tilde{\lambda}_i-\lambda_{0i}\|\right).\tag{S.24}$$

Next, it is also easy to show that

$$\left\|\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(3)}(*)\left(\tilde{\lambda}_i'\tilde{f}_t-\lambda_{0i}'f_{0t}\right)^2\tilde{f}_t\right\|\lesssim\|\tilde{\lambda}_i-\lambda_{0i}\|^2\frac{1}{T}\sum_{t=1}^{T}\left|\tilde{\varrho}_{it}^{(3)}(*)\right|+\frac{1}{T}\sum_{t=1}^{T}\left|\tilde{\varrho}_{it}^{(3)}(*)\right|\cdot\|\tilde{f}_t-f_{0t}\|^2.$$

Therefore, from Lemma S.1, Lemma S.3, and Lemma S.5, we have

$$\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(3)}(*)\left(\tilde{\lambda}_i'\tilde{f}_t-\lambda_{0i}'f_{0t}\right)^2\tilde{f}_t=O_P\left(\|\tilde{\lambda}_i-\lambda_{0i}\|\right)\cdot O_P\left(T^{-1/2}h^{-3}\right)+O_P\left(T^{-1}h^{-2}\right),$$

which is $o_P(\|\tilde{\lambda}_i-\lambda_{0i}\|)+O_P(T^{-1}h^{-2})$ under the assumption that $\sqrt{T}h^3\to\infty$.

Finally, combining all the above results, we get

$$\Phi_{T,i}(\tilde{\lambda}_i-\lambda_{0i})=\frac{1}{T}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(1)}f_{0t}+o_P\left(\|\tilde{\lambda}_i-\lambda_{0i}\|\right)+O_P\left(\frac{1}{Th^2}\right)+O\left(h^m\right),\tag{S.25}$$

and from Lemma S.1 it is easy to show that

$$\frac{1}{T}\sum_{t=1}^{T}\bar{\varrho}_{it}^{(2)} f_{0t}f_{0t}' \to \Phi_i > 0 \quad \text{and} \quad \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\tilde{\varrho}_{it}^{(1)} f_{0t} \xrightarrow{d} \mathcal{N}\big(0, \tau(1-\tau)\mathbb{I}_r\big). \tag{S.26}$$

Since our assumption implies that $\sqrt{T}h^2 \to \infty$ and $\sqrt{T}h^m \to 0$, the desired results follow from (S.25) and (S.26). *Q.E.D.*

S.2. ESTIMATING THE COVARIANCE MATRICES

In this subsection, we show that the estimators of the covariance matrices proposed in Remark 4.4 are consistent. Recall that

$$V_{f_t} = \tau(1-\tau)\Psi_t^{-1}\Sigma_\Lambda\Psi_t^{-1}, \qquad \tilde{V}_{f_t} = \tau(1-\tau)\tilde{\Psi}_t^{-1}\tilde{\Sigma}_\Lambda\tilde{\Psi}_t^{-1},$$

where $\tilde{\Psi}_t = (Nb)^{-1}\sum_{i=1}^{N} l(\tilde{u}_{it}/b)\cdot\tilde{\lambda}_i\tilde{\lambda}_i'$ and $\tilde{\Sigma}_\Lambda = \tilde{\Lambda}'\tilde{\Lambda}/N$. It suffices to show that $\|\tilde{\Sigma}_\Lambda - \Sigma_\Lambda\| = o_P(1)$ and $\|\tilde{\Psi}_t - \Psi_t\| = o_P(1)$.

First, from Lemma 2 and Lemma S.2,

$$\begin{aligned}
\|\tilde{\Sigma}_\Lambda - \Sigma_\Lambda\| &\le \big\|\tilde{\Lambda}'\tilde{\Lambda}/N - \Lambda_0'\Lambda_0/N + \Lambda_0'\Lambda_0/N - \Sigma_\Lambda\big\| \\
&\le \big\|\tilde{\Lambda}'\tilde{\Lambda}/N - \Lambda_0'\Lambda_0/N\big\| + o(1) \\
&\le 2\cdot\frac{\|\Lambda_0\|}{\sqrt{N}}\cdot\frac{\|\tilde{\Lambda}-\Lambda_0\|}{\sqrt{N}} + \left(\frac{\|\tilde{\Lambda}-\Lambda_0\|}{\sqrt{N}}\right)^2 + o(1) = o_P(1).
\end{aligned}$$

Second,

$$\begin{aligned}
\tilde{\Psi}_t &= \frac{1}{Nb}\sum_{i=1}^{N} l(u_{it}/b)\cdot\lambda_{0i}\lambda_{0i}' + \frac{1}{Nb}\sum_{i=1}^{N}\big[l(\tilde{u}_{it}/b) - l(u_{it}/b)\big]\cdot\lambda_{0i}\lambda_{0i}' \\
&\quad + \frac{1}{Nb}\sum_{i=1}^{N} l(\tilde{u}_{it}/b)\cdot\big(\tilde{\lambda}_i\tilde{\lambda}_i' - \lambda_{0i}\lambda_{0i}'\big).
\end{aligned} \tag{S.27}$$

For the first term on the RHS of (S.27), following the standard proof of kernel density estimator, we have

$$\frac{1}{Nb}\sum_{i=1}^{N} l(u_{it}/b)\cdot\lambda_{0i}\lambda_{0i}' = \frac{1}{N}\sum_{i=1}^{N}\mathsf{f}_{it}(0)\lambda_{0i}\lambda_{0i}' + O_P\big(b^2\big) + O_P(1/\sqrt{Nb}).$$

For the third term on the RHS of (S.27), by Lemma S.3 one can show that

$$\left\|\frac{1}{Nb}\sum_{i=1}^{N} l(\tilde{u}_{it}/b)\cdot\big(\tilde{\lambda}_i\tilde{\lambda}_i' - \lambda_{0i}\lambda_{0i}'\big)\right\| \lesssim \frac{1}{Nb}\sum_{i=1}^{N}\|\tilde{\lambda}_i\tilde{\lambda}_i' - \lambda_{0i}\lambda_{0i}'\| = O_P\big(1/\sqrt{Nb^2}\big).$$

As regards the second term on the RHS of (S.27), we have

$$\frac{1}{Nb}\sum_{i=1}^{N}\big[l(\tilde{u}_{it}/b) - l(u_{it}/b)\big]\cdot\lambda_{0i}\lambda_{0i}'$$

$$= \frac{1}{Nb^2}\sum_{i=1}^{N} l^{(1)}(u_{it}/b)\cdot\lambda_{0i}\lambda_{0i}'\cdot(\tilde{u}_{it}-u_{it}) + \frac{1}{Nb^3}\sum_{i=1}^{N} l^{(2)}\big(u_{it}^*/b\big)\cdot\lambda_{0i}\lambda_{0i}'\cdot(\tilde{u}_{it}-u_{it})^2$$

$$= \frac{1}{Nb^2}\sum_{i=1}^{N} l^{(1)}(u_{it}/b)\cdot\lambda_{0i}\lambda_{0i}'\lambda_{0i}\cdot(\tilde{f}_t - f_{0t}) + \frac{1}{Nb^2}\sum_{i=1}^{N} l^{(1)}(u_{it}/b)\cdot\lambda_{0i}\lambda_{0i}'(\tilde{\lambda}_i-\lambda_{0i})\cdot\tilde{f}_t$$

$$+ \frac{1}{Nb^3}\sum_{i=1}^{N} l^{(2)}\big(u_{it}^*/b\big)\cdot\lambda_{0i}\lambda_{0i}'\cdot(\tilde{u}_{it}-u_{it})^2,$$

where u_{it}^* is between u_{it} and $\tilde{u}_{it}$. It is easy to show that

$$\frac{1}{Nb^2}\sum_{i=1}^{N} l^{(1)}(u_{it}/b)\cdot\lambda_{0i}\lambda_{0i}'\lambda_{0i} = O_P(1),$$

$$\left\|\frac{1}{Nb^2}\sum_{i=1}^{N} l^{(1)}(u_{it}/b)\cdot\lambda_{0i}\lambda_{0i}'(\tilde{\lambda}_i-\lambda_{0i})\cdot\tilde{f}_t\right\|$$

$$\lesssim \frac{1}{b^{3/2}}\sqrt{\frac{1}{Nb}\sum_{i=1}^{N}\big[l^{(1)}(u_{it}/b)\big]^2}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\|\tilde{\lambda}_i-\lambda_{0i}\|^2}$$

$$= O_P\big(b^{-3/2}\big)O_P(1)O_P\big(L_{NT}^{-1}\big) = O_P\big(1/\sqrt{Nb^3}\big),$$

and

$$\left\|\frac{1}{Nb^3}\sum_{i=1}^{N} l^{(2)}\big(u_{it}^*/b\big)\cdot\lambda_{0i}\lambda_{0i}'\cdot(\tilde{u}_{it}-u_{it})^2\right\| \lesssim \frac{1}{b^3}\cdot\frac{1}{N}\sum_{i=1}^{N}(\tilde{u}_{it}-u_{it})^2 = O_P\big(1/Nb^3\big).$$

Thus, it follows that

$$\left\|\frac{1}{Nb}\sum_{i=1}^{N}\big[l(\tilde{u}_{it}/b)-l(u_{it}/b)\big]\cdot\lambda_{0i}\lambda_{0i}'\right\| = O_P\big(N^{-1}\big) + O_P\big(1/\sqrt{Nb^3}\big) + O_P\big(1/Nb^3\big).$$

Finally, it follows from the definition of Ψ_t and the condition $b\to 0$, $Nb^3\to\infty$ that $\tilde{\Psi}_t \xrightarrow{p} \Psi_t$ and thus $\tilde{V}_{f_t} \xrightarrow{p} V_{f_t}$. The proof for $\tilde{\Phi}_i$ is similar and therefore is omitted.

TABLE S.I

ESTIMATION OF QFM BY PCA[a]

	M1				M2				M3			
	$\hat{r}_{IC}$	f_{1t}	f_{2t}	f_{3t}	$\hat{r}_{IC}$	f_{1t}	f_{2t}	f_{3t}	$\hat{r}_{IC}$	f_{1t}	f_{2t}	f_{3t}
(50,50)	7.19	0.982	0.968	0.323	7.44	0.949	0.910	0.302	7.99	0.982	0.968	0.339
(50,100)	4.47	0.981	0.964	0.122	5.48	0.946	0.894	0.134	7.99	0.982	0.967	0.178
(50,200)	2.32	0.981	0.961	0.016	3.53	0.946	0.892	0.035	8.00	0.982	0.964	0.088
(100,50)	7.44	0.991	0.984	0.380	7.60	0.974	0.947	0.364	7.99	0.990	0.982	0.398
(100,100)	5.49	0.991	0.983	0.182	6.33	0.974	0.947	0.188	7.99	0.991	0.981	0.227
(100,200)	2.98	0.991	0.982	0.046	4.28	0.974	0.947	0.061	7.99	0.990	0.980	0.120
(200,50)	7.58	0.996	0.992	0.419	7.67	0.985	0.967	0.398	7.98	0.994	0.990	0.455
(200,100)	6.63	0.996	0.992	0.252	7.03	0.987	0.973	0.232	7.98	0.995	0.990	0.284
(200,200)	4.33	0.996	0.991	0.097	5.60	0.987	0.973	0.105	7.97	0.995	0.990	0.149

[a]Simulation results from 1000 repetitions. The DGP is fully described in Section 5.2 of the paper, and corresponds to: $X_{it} = \lambda_{1i} f_{1t} + \lambda_{2i} f_{2t} + (\lambda_{3i} f_{3t}) \cdot e_{it}$, $f_{1t} = 0.8 f_{1,t-1} + \epsilon_{1t}$, $f_{2t} = 0.5 f_{2,t-1} + \epsilon_{2t}$, $f_{3t} = |g_t|$, $\lambda_{1i}, \lambda_{2i}, \epsilon_{1t}, \epsilon_{2t}, g_t \sim$ i.i.d. $\mathcal{N}(0,1)$, and $\lambda_{3i} \sim$ i.i.d. $U[1,2]$. $e_{it} = \beta e_{i,t-1} + v_{it} + \rho \cdot \sum_{j=i-J, j\neq i}^{i+J} v_{jt}$. M1: $v_{it} \sim$ i.i.d. $\mathcal{N}(0,1)$, $\beta = \rho = 0$; M2: $v_{it} \sim$ i.i.d. t_3, $\beta = \rho = 0$; M3: $v_{it} \sim$ i.i.d. $\mathcal{N}(0,1)$, $\beta = \rho = 0.2$, $J = 3$. For each model, the first column reports the averages of $\hat{r}_{IC}$, which is the estimated number of factors using IC_{P1} of BN (2002), while the second to the fourth columns report the average R^2 in the regression of (each of) the true factors on the PCA factors $\hat{F}_{PCA}$, where the number of PCA factors is given by $\hat{r}_{IC}$.

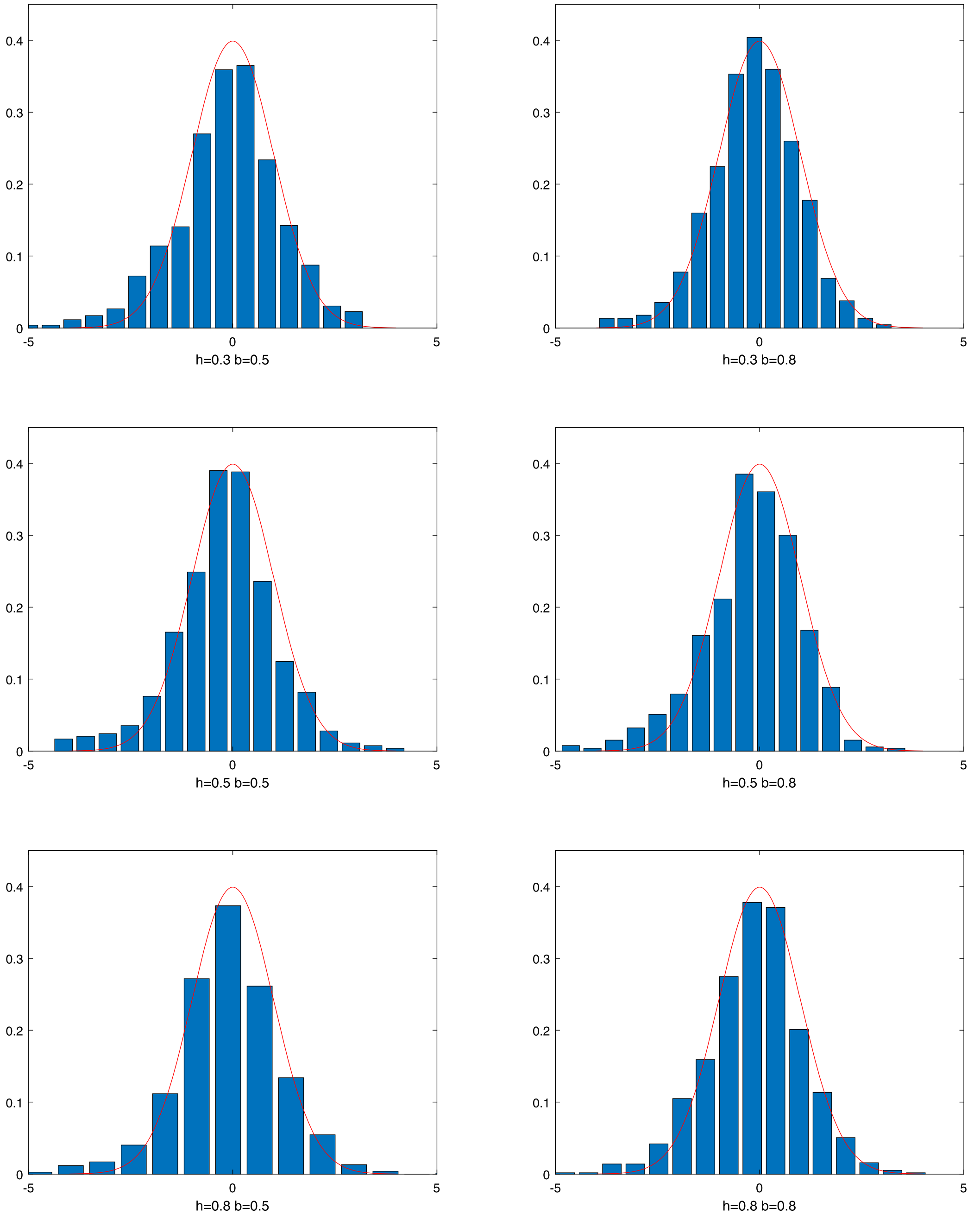

FIGURE S.1.—Normal approximations of the estimated QFA factors using SQR for $N = T = 50$.

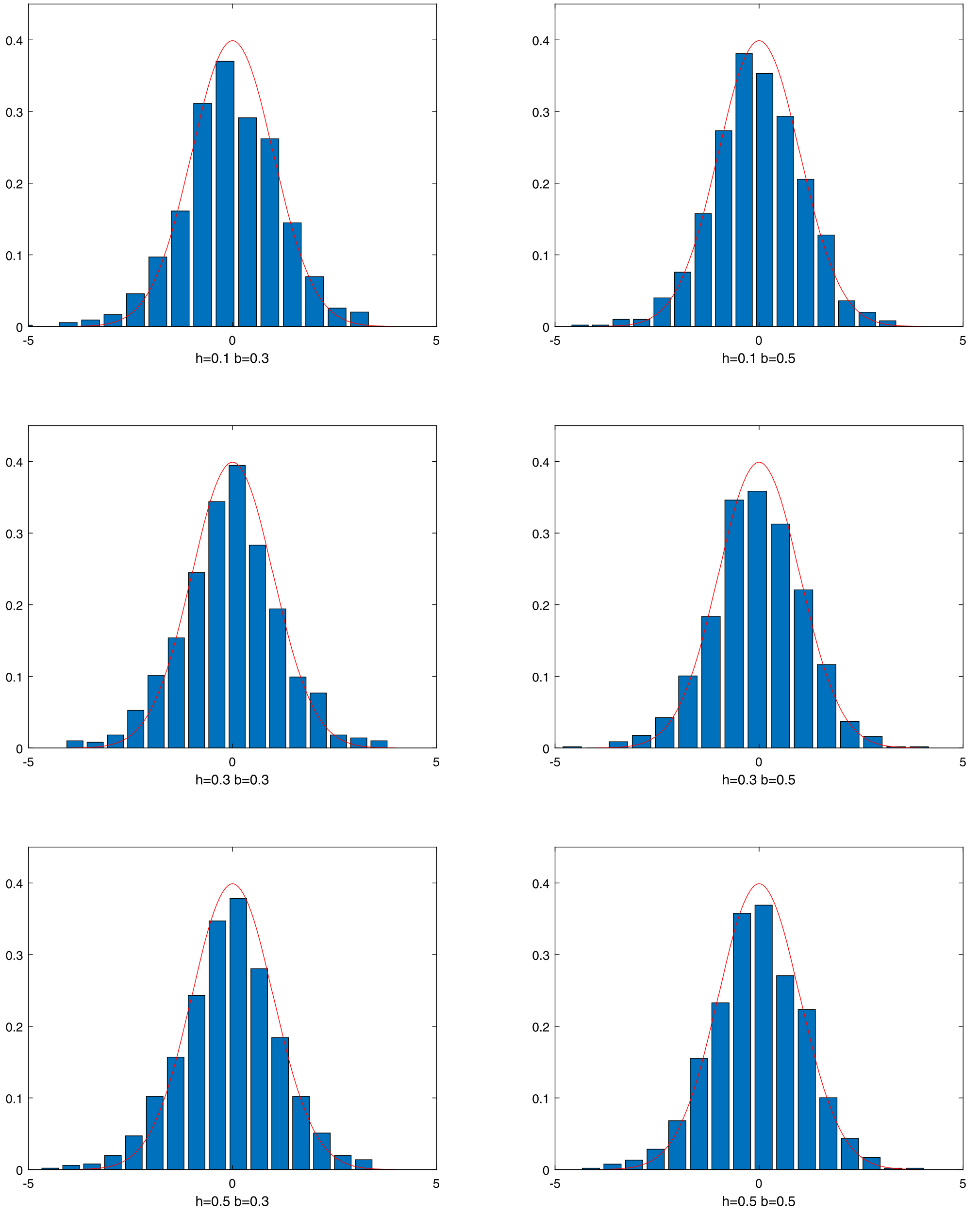

FIGURE S.2.—Normal approximations of the estimated factors using SQR for $N = T = 200$.

REFERENCES

GALVAO, A. F., AND K. KATO (2016): "Smoothed Quantile Regression for Panel Data," *Journal of Econometrics*, 193 (1), 92–112. [2,4,6]

HOROWITZ, J. (1998): "Bootstrap Methods for Median Regression Models," *Econometrica*, 66 (6), 1327–1352. [2]

NEWEY, W. K., AND D. MCFADDEN (1994): "Large Sample Estimation and Hypothesis Testing," *Handbook of Econometrics*, 4, 2111–2245. [4]

VAN DER VAART, A., AND J. WELLNER (1996): *Weak Convergence and Empirical Processes*. New York: Springer. [13]

Co-editor Ulrich K. Müller handled this manuscript.

Manuscript received 8 October, 2017; final version accepted 20 October, 2020; available online 22 October, 2020.