Supplemental Appendix: Detailed proofs

In order to prove Proposition 1, we use the following five lemmas. First, we recall the definition of stochastic equicontinuity.

Definition 1 *(Andrews (1994)) The empirical process* $\{\mathbf{h}_T(\cdot) : T \geq 1\}$ *is stochastically equicontinuous if* $\forall\ \varepsilon > 0$ *and* $\eta > 0, \exists\ \delta > 0$ *such that*

$$\limsup_{T\to\infty} P\left[\sup_{\|\boldsymbol{\theta}_1-\boldsymbol{\theta}_2\|<\delta} \|\mathbf{h}_T(\boldsymbol{\theta}_1) - \mathbf{h}_T(\boldsymbol{\theta}_2)\| > \eta\right] < \varepsilon \tag{2}$$

Lemma 1 *Under Assumptions 1 and 2,*

(i) $\frac{1}{T}\sum_{t=1}^{T} \hat{F}_i(\hat{\eta}_{it})\hat{F}_j(\hat{\eta}_{jt}) \xrightarrow{p} \int\int uvdC_{\eta_i,\eta_j}(u,v;\boldsymbol{\theta}_0)$ *as* $T\to\infty$

(ii) $\frac{1}{T}\sum_{t=1}^{T} 1\left\{\hat{F}_i(\hat{\eta}_{it}) \leq q, \hat{F}_j(\hat{\eta}_{jt}) \leq q\right\} \xrightarrow{p} C_{\eta_i,\eta_j}(q,q;\boldsymbol{\theta}_0)$ *as* $T\to\infty$

(iii) $\frac{1}{S}\sum_{s=1}^{S} \hat{G}_i(x_{is}(\boldsymbol{\theta}))\hat{G}_j(x_{js}(\boldsymbol{\theta})) \xrightarrow{p} \int\int uvdC_{\eta_i,\eta_j}(u,v;\boldsymbol{\theta})$ *for* $\forall\ \boldsymbol{\theta}\in\boldsymbol{\Theta}$ *as* $S\to\infty$

(iv) $\frac{1}{S}\sum_{s=1}^{S} 1\left\{\hat{G}_i(x_{is}(\boldsymbol{\theta})) \leq q, \hat{G}_j(x_{js}(\boldsymbol{\theta})) \leq q\right\} \xrightarrow{p} C_{\eta_i,\eta_j}(q,q;\boldsymbol{\theta})$ *for* $\forall\ \boldsymbol{\theta}\in\boldsymbol{\Theta}$ *as* $S\to\infty$

(v) $\frac{1}{S}\sum_{s=1}^{S} G_i(x_{is}(\boldsymbol{\theta}))G_j(x_{js}(\boldsymbol{\theta})) \xrightarrow{p} \int\int uvdC_{\eta_i,\eta_j}(u,v;\boldsymbol{\theta})$ *for* $\forall\ \boldsymbol{\theta}\in\boldsymbol{\Theta}$ *as* $S\to\infty$

(vi) $\frac{1}{S}\sum_{s=1}^{S} 1\{G_i(x_{is}(\boldsymbol{\theta})) \leq q, G_j(x_{js}(\boldsymbol{\theta})) \leq q\} \xrightarrow{p} C_{\eta_i,\eta_j}(q,q;\boldsymbol{\theta})$ *for* $\forall\ \boldsymbol{\theta}\in\boldsymbol{\Theta}$ *as* $S\to\infty$

Proof of Lemma 1. Under Assumption 1, parts (iii) and (iv) of Lemma 1 can be proven by Theorem 3 and Theorem 6 of Fermanian, Radulović and Wegkamp (2004). Under Assumption 2, Corollary 1 of Rémillard (2010) proves that the empirical copula process constructed by the standardized residuals $\hat{\boldsymbol{\eta}}_t$ weakly converges to the limit of that constructed by the innovations $\boldsymbol{\eta}_t$, which combined with Theorem 3 and Theorem 6 of Fermanian, Radulović and Wegkamp (2004) yields parts (i) and (ii) above. In the case where it is possible to simulate directly from the copula rather than the joint distribution, e.g. Clayton/Gaussian copula in Section 3 or where we only can simulate from the joint distribution but know the marginal distribution G_i in closed form, it is not necessary to estimate marginal distribution G_i. In this case, instead of (iii) and (iv), (v) and (vi) are used for the later proofs. (v) and (vi) are proven by the standard law of large numbers. ■

Lemma 2 *(Lemma 2.8 of Newey and McFadden (1994)) Suppose* $\boldsymbol{\Theta}$ *is compact and* $\mathbf{g}_0(\boldsymbol{\theta})$ *is continuous. Then* $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \|\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_0(\boldsymbol{\theta})\| \xrightarrow{p} 0$ *as* $T, S \to \infty$ *if and only if* $\mathbf{g}_{T,S}(\boldsymbol{\theta}) \xrightarrow{p} \mathbf{g}_0(\boldsymbol{\theta})$ *for any* $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ *as* $T, S \to \infty$ *and* $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ *is stochastically equicontinuous.*

Lemma 2 states that sufficient and necessary conditions for uniform convergence are pointwise convergence and stochastic equicontinuity. The following lemma shows that uniform convergence of the moment functions $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ implies uniform convergence of the objective function $Q_{T,S}(\boldsymbol{\theta})$.

Lemma 3 *If* $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \|\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_0(\boldsymbol{\theta})\| \xrightarrow{p} 0$ *as* $T, S \to \infty$, *then* $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} |Q_{T,S}(\boldsymbol{\theta}) - Q_0(\boldsymbol{\theta})| \xrightarrow{p} 0$ *as* $T, S \to \infty$.

Proof of Lemma 3. By the triangle inequality and Cauchy-Schwarz inequality

$$\begin{aligned}
|Q_{T,S}(\boldsymbol{\theta}) - Q_0(\boldsymbol{\theta})| &\leq \left|[\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_0(\boldsymbol{\theta})]' \hat{\mathbf{W}}_T [\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_0(\boldsymbol{\theta})]\right| \\
&\quad + \left|\mathbf{g}_0(\boldsymbol{\theta})' \left(\hat{\mathbf{W}}_T + \hat{\mathbf{W}}_T'\right) [\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_0(\boldsymbol{\theta})]\right| + \left|\mathbf{g}_0(\boldsymbol{\theta})' \left(\hat{\mathbf{W}}_T - \mathbf{W}_0\right) \mathbf{g}_0(\boldsymbol{\theta})\right| \\
&\leq \|\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_0(\boldsymbol{\theta})\|^2 \left\|\hat{\mathbf{W}}_T\right\| + 2\|\mathbf{g}_0(\boldsymbol{\theta})\| \|\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_0(\boldsymbol{\theta})\| \left\|\hat{\mathbf{W}}_T\right\| \\
&\quad + \|\mathbf{g}_0(\boldsymbol{\theta})\|^2 \left\|\hat{\mathbf{W}}_T - \mathbf{W}_0\right\|
\end{aligned} \tag{3}$$

Then note that $\mathbf{g}_0(\boldsymbol{\theta})$ is bounded, $\hat{\mathbf{W}}_T$ is $O_p(1)$ and converges to $\mathbf{W}_0$ by Assumption 3(iv), and $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \|\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_0(\boldsymbol{\theta})\| = o_p(1)$ is given. So

$$\begin{aligned}
\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} |Q_{T,S}(\boldsymbol{\theta}) - Q_0(\boldsymbol{\theta})| &\leq \left(\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \|\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_0(\boldsymbol{\theta})\|\right)^2 O_p(1) \\
&\quad + 2O(1) \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \|\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_0(\boldsymbol{\theta})\| O_p(1) + o_p(1) = o_p(1)
\end{aligned} \tag{4}$$

■

Lemma 4 *Under Assumption 1, Assumption 2, and Assumption 3(iii),*

(i) $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ *is stochastic Lipschitz continuous, i.e.*

$\exists B_{T,S} = O_p(1)$ *such that for all* $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \boldsymbol{\Theta}$, $\|\mathbf{g}_{T,S}(\boldsymbol{\theta}_1) - \mathbf{g}_{T,S}(\boldsymbol{\theta}_2)\| \leq B_{T,S} \cdot \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|$

(ii) There exists $\delta > 0$ *such that*

$$\lim \sup_{T,S\to\infty} E\left(B_{T,S}^{2+\delta}\right) < \infty \text{ for some } \delta > 0$$

Proof of Lemma 4. Without loss of generality, assume that $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ is scalar. By Lemma 1, we know that

$$\tilde{\mathbf{m}}_S(\boldsymbol{\theta}) = \mathbf{m}_0(\boldsymbol{\theta}) + o_p(1) \tag{5}$$

Also, by Assumption 3(iii) and the fact that $\mathbf{m}(\boldsymbol{\theta})$ consists of a function of Lipschitz continuous $C_{ij}(\boldsymbol{\theta})$, $\mathbf{m}_0(\boldsymbol{\theta})$ is Lipschitz continuous, i.e. $\exists K$ such that

$$|\mathbf{m}_0(\boldsymbol{\theta}_1) - \mathbf{m}_0(\boldsymbol{\theta}_2)| \leq K\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\| \tag{6}$$

Then,

$$\begin{aligned} |\mathbf{g}_{T,S}(\boldsymbol{\theta}_1) - \mathbf{g}_{T,S}(\boldsymbol{\theta}_2)| &= |\tilde{\mathbf{m}}_S(\boldsymbol{\theta}_1) - \tilde{\mathbf{m}}_S(\boldsymbol{\theta}_2)| = |\mathbf{m}_0(\boldsymbol{\theta}_1) - \mathbf{m}_0(\boldsymbol{\theta}_2) + o_p(1)| \qquad (7) \\ &\leq |\mathbf{m}_0(\boldsymbol{\theta}_1) - \mathbf{m}_0(\boldsymbol{\theta}_2)| + |o_p(1)| \\ &\leq K\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\| + |o_p(1)| \\ &= \underbrace{\left(K + \frac{|o_p(1)|}{\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|}\right)}_{=O_p(1)} \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\| \end{aligned}$$

and let $B_{T,S} = K + \frac{|o_p(1)|}{\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|}$. Then for some $\delta > 0$

$$\lim \sup_{T,S\to\infty} E\left(B_{T,S}^{2+\delta}\right) = \lim \sup_{T,S\to\infty} E\left(K + \frac{|o_p(1)|}{\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|}\right)^{2+\delta} < \infty \tag{8}$$

■

Lemma 5 *(Theorem 2.1 of Newey and McFadden (1994)) Suppose that (i)* $Q_0(\boldsymbol{\theta})$ *is uniquely minimized at* $\boldsymbol{\theta}_0$*; (ii)* $\boldsymbol{\Theta}$ *is compact; (iii)* $Q_0(\boldsymbol{\theta})$ *is continuous (iv)* $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left|\hat{Q}_T(\boldsymbol{\theta}) - Q_0(\boldsymbol{\theta})\right| \xrightarrow{p} 0$*. Then* $\hat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}_0$

Proof of Proposition 1. We prove this proposition by checking the conditions of Lemma 5.

(i) $Q_0(\boldsymbol{\theta})$ is uniquely minimized at $\boldsymbol{\theta}_0$ by Assumption 3(i) and Assumption 3(iv).

(ii) $\boldsymbol{\Theta}$ is compact by Assumption 3(ii).

(iii) $Q_0(\boldsymbol{\theta})$ consists of linear combinations of rank correlations and quantile dependence measures that are functions of pair-wise copula functions. Therefore, $Q_0(\boldsymbol{\theta})$ is continuous by Assumption 3(iii).

(iv) The pointwise convergence of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ to $\mathbf{g}_0(\boldsymbol{\theta})$ and the stochastic Lipschitz continuity of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ are shown by Lemma 1 and by Lemma 4(i), respectively. By Lemma 2.9 of Newey and McFadden (1994), the stochastic Lipschitz continuity of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$ ensures the stochastic equicontinuity of $\mathbf{g}_{T,S}(\boldsymbol{\theta})$, and under Assumption 3, $\boldsymbol{\Theta}$ is compact and $\mathbf{g}_0(\boldsymbol{\theta})$ is continuous in $\boldsymbol{\theta}$. Therefore, $\mathbf{g}_{T,S}$ uniformly converges in probability to $\mathbf{g}_0$ by Lemma 2. This implies that $Q_{T,S}$ uniformly converges in probability to Q_0 by Lemma 3. ■

The proof of Proposition 2 uses the following three lemmas.

Lemma 6 *Let the dependence measures of interest include rank correlation and quantile dependence measures, and possibly linear combinations thereof. Then under Assumptions 1 and 2,*

$$\sqrt{T}\left(\hat{\mathbf{m}}_T - \mathbf{m}_0(\boldsymbol{\theta}_0)\right) \xrightarrow{d} N(0, \boldsymbol{\Sigma}_0) \ \ \text{as } T \ \to \ \infty \tag{9}$$

$$\sqrt{S}\left(\tilde{\mathbf{m}}_S(\boldsymbol{\theta}_0) - \mathbf{m}_0(\boldsymbol{\theta}_0)\right) \xrightarrow{d} N(0, \boldsymbol{\Sigma}_0) \ \ \text{as } S \ \to \ \infty \tag{10}$$

Proof of Lemma 6. Follows from Theorem 3 and Theorem 6 of Fermanian, Radulović and Wegkamp (2004) and Corollary 1, Proposition 2 and Proposition 4 of Rémillard (2010). ■

We use Theorem 7.2 of Newey & McFadden (1994) to establish the asymptotic normality of our estimator, and this relies on showing the stochastic equicontinuity of $\mathbf{v}_{T,S}(\boldsymbol{\theta})$ defined below.

Lemma 7 *Suppose that Assumptions 1, 2, and 3(iii) hold. Then when* $S/T \to \infty$ *or* $S/T \to k \in (0,\infty)$, $\mathbf{v}_{T,S}(\boldsymbol{\theta}) = \sqrt{T}\left[\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_0(\boldsymbol{\theta})\right]$ *is stochastically equicontinuous and when* $S/T \to 0$, $\mathbf{v}_{T,S}(\boldsymbol{\theta}) = \sqrt{S}\left[\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_0(\boldsymbol{\theta})\right]$ *is stochastically equicontinuous.*

Proof of Lemma 7. By Lemma 4(i), $\{\mathbf{g}_{\cdot,\cdot}(\boldsymbol{\theta}) : \theta \in \Theta\}$ is a type II class of functions in Andrews (1994). By Theorem 2 of Andrews (1994), $\{\mathbf{g}_{\cdot,\cdot}(\boldsymbol{\theta}) : \theta \in \Theta\}$ satisfies Pollard's entropy condition with envelope $1 \vee \sup_\theta \|\mathbf{g}_{\cdot,\cdot}(\boldsymbol{\theta})\| \vee B_{\cdot,\cdot}$, so Assumption A of Andrews (1994) is satisfied. Since $\mathbf{g}_{\cdot,\cdot}(\boldsymbol{\theta})$ is bounded and by the condition of $\limsup_{T,S\to\infty} E\left(B_{T,S}^{2+\delta}\right) < \infty$ for some $\delta > 0$ by Lemma 4(ii), the Assumption B of Andrews (1994) is also satisfied. Therefore, $\mathbf{v}_{T,S}(\boldsymbol{\theta})$ is stochastically equicontinuous by Theorem 1 of Andrews (1994). ■

Lemma 8 *(Theorem 7.2 of Newey & McFadden (1994)) Suppose that* $\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}\right)' \hat{\mathbf{W}}_T \mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}\right) \leq \inf_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \mathbf{g}_{T,S}(\boldsymbol{\theta})' \hat{\mathbf{W}}_T \mathbf{g}_{T,S}(\boldsymbol{\theta}) + o_p\left(T^{-1}\right)$, $\hat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}_0$ *and* $\hat{\mathbf{W}}_T \xrightarrow{p} \mathbf{W}_0$, $\mathbf{W}_0$ *is positive semi-definite, where there is* $\mathbf{g}_0(\boldsymbol{\theta})$ *such that (i)* $\mathbf{g}_0(\boldsymbol{\theta}_0) = 0$, *(ii)* $\mathbf{g}_0(\boldsymbol{\theta})$ *is differentiable at* $\boldsymbol{\theta}_0$ *with derivative* $\mathbf{G}_0$ *such that* $\mathbf{G}_0'\mathbf{W}_0\mathbf{G}_0$ *is nonsingular, (iii)* $\boldsymbol{\theta}_0$ *is an interior point of* $\boldsymbol{\Theta}$, *(iv)* $\sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0) \xrightarrow{d} N\left(0, \boldsymbol{\Sigma}_0\right)$, *(v)* $\exists\delta$ *such that* $\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\leq\delta} \sqrt{T}\left\|\mathbf{g}_{T,S}(\boldsymbol{\theta}) - \mathbf{g}_{T,S}(\boldsymbol{\theta}_0) - \mathbf{g}_0(\boldsymbol{\theta})\right\| / \left[1 + \sqrt{T}\left\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\right\|\right] \xrightarrow{p} 0$. *Then* $\sqrt{T}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) \xrightarrow{d} N\left(0, \left(\mathbf{G}_0'\mathbf{W}_0\mathbf{G}_0\right)^{-1} \mathbf{G}_0'\mathbf{W}_0\boldsymbol{\Sigma}_0\mathbf{W}_0\mathbf{G}_0 \left(\mathbf{G}_0'\mathbf{W}_0\mathbf{G}_0\right)^{-1}\right)$.

Proof of Proposition 2. We prove this proposition by checking conditions of Lemma 8.

(i) $\mathbf{g}_0(\boldsymbol{\theta}_0) = 0$ by construction of $\mathbf{g}_0(\boldsymbol{\theta}) = \mathbf{m}_0(\boldsymbol{\theta}_0) - \mathbf{m}_0(\boldsymbol{\theta})$

(ii) $\mathbf{g}_0(\boldsymbol{\theta})$ is differentiable at $\boldsymbol{\theta}_0$ with derivative $\mathbf{G}_0$ such that $\mathbf{G}_0'\mathbf{W}_0\mathbf{G}_0$ is nonsingular by Assumption 4(ii).

(iii) $\boldsymbol{\theta}_0$ is an interior point of $\boldsymbol{\Theta}$ by Assumption 4(i).

(iv) If $S/T \to \infty$ as $T, S \to \infty$,

$$\begin{aligned}
\sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0) &= \sqrt{T}\left(\hat{\mathbf{m}}_T - \tilde{\mathbf{m}}_S(\boldsymbol{\theta}_0)\right) \\
&= \sqrt{T}\left(\hat{\mathbf{m}}_T - \mathbf{m}_0(\boldsymbol{\theta}_0)\right) - \sqrt{T}\left(\tilde{\mathbf{m}}_S(\boldsymbol{\theta}_0) - \mathbf{m}_0(\boldsymbol{\theta}_0)\right) \\
&= \underbrace{\sqrt{T}\left(\hat{\mathbf{m}}_T - \mathbf{m}_0(\boldsymbol{\theta}_0)\right)}_{\xrightarrow{d} N(0,\Sigma_0) \text{ by Lemma 6}} - \underbrace{\frac{\sqrt{T}}{\sqrt{S}}}_{=o(1)} \times \underbrace{\sqrt{S}\left(\tilde{\mathbf{m}}_S(\boldsymbol{\theta}_0) - \mathbf{m}_0(\boldsymbol{\theta}_0)\right)}_{\xrightarrow{d} N(0,\Sigma_0) \text{ by Lemma 6}}
\end{aligned} \tag{11}$$

Therefore,

$$\sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0) \xrightarrow{d} N(0, \Sigma_0) \text{ as } T, S \to \infty.$$

If $S/T \to k \in (0, \infty)$ as $T, S \to \infty$,

$$\sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0) = \underbrace{\sqrt{T}\left(\hat{\mathbf{m}}_T - \mathbf{m}_0(\boldsymbol{\theta}_0)\right)}_{\xrightarrow{d} N(0,\Sigma_0) \text{ by Lemma 6}} - \underbrace{\frac{\sqrt{T}}{\sqrt{S}}}_{\to 1/\sqrt{k}} \times \underbrace{\sqrt{S}\left(\tilde{\mathbf{m}}_S(\boldsymbol{\theta}_0) - \mathbf{m}_0(\boldsymbol{\theta}_0)\right)}_{\xrightarrow{d} N(0,\Sigma_0) \text{ by Lemma 6}}$$

Therefore,

$$\sqrt{T}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0) \xrightarrow{d} N\left(0, \left(1 + \frac{1}{k}\right)\Sigma_0\right) \text{ as } T, S \to \infty.$$

If $S/T \to 0$ as $T, S \to \infty$,

$$\sqrt{S}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0) = \underbrace{\frac{\sqrt{S}}{\sqrt{T}}}_{=o(1)} \times \underbrace{\sqrt{T}\left(\hat{\mathbf{m}}_T - \mathbf{m}_0(\boldsymbol{\theta}_0)\right)}_{\xrightarrow{d} N(0,\Sigma_0) \text{ by Lemma 6}} - \underbrace{\sqrt{S}\left(\tilde{\mathbf{m}}_S(\boldsymbol{\theta}_0) - \mathbf{m}_0(\boldsymbol{\theta}_0)\right)}_{\xrightarrow{d} N(0,\Sigma_0) \text{ by Lemma 6}}$$

Therefore,

$$\sqrt{S}\mathbf{g}_{T,S}(\boldsymbol{\theta}_0) \xrightarrow{d} N(0, \Sigma_0) \text{ as } T, S \to \infty$$

(v) We established the stochastic equicontinuity of $\mathbf{v}_{T,S}(\boldsymbol{\theta}) = \sqrt{T}\left[\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_0(\boldsymbol{\theta})\right]$ when $S/T \to \infty$ or $S/T \to k$ by Lemma 7, i.e. for $\forall \varepsilon > 0$, $\eta > 0, \exists \delta$ such that

$$\begin{aligned} & \limsup_{T\to\infty} P\left[\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|<\delta} \left\|\mathbf{v}_{T,S}(\boldsymbol{\theta}) - \mathbf{v}_{T,S}(\boldsymbol{\theta}_0)\right\| > \eta\right] \\ = & \limsup_{T\to\infty} P\left[\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|<\delta} \sqrt{T}\left\|\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_{T,S}(\boldsymbol{\theta}_0) - \boldsymbol{g}_0(\boldsymbol{\theta})\right\| > \eta\right] < \varepsilon \end{aligned} \tag{12}$$

and from the following inequality

$$\sqrt{T}\left\|\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_{T,S}(\boldsymbol{\theta}_0) - \boldsymbol{g}_0(\boldsymbol{\theta})\right\| / \left[1 + \sqrt{T}\left\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\right\|\right] \le \sqrt{T}\left\|\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_{T,S}(\boldsymbol{\theta}_0) - \boldsymbol{g}_0(\boldsymbol{\theta})\right\| \tag{13}$$

we know that

$$\begin{aligned} & \limsup_{T\to\infty} P\left[\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|<\delta} \sqrt{T}\left\|\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_{T,S}(\boldsymbol{\theta}_0) - \boldsymbol{g}_0(\boldsymbol{\theta})\right\| / \left[1 + \sqrt{T}\left\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\right\|\right] > \eta\right] \\ \le & \limsup_{T\to\infty} P\left[\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|<\delta} \sqrt{T}\left\|\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_{T,S}(\boldsymbol{\theta}_0) - \boldsymbol{g}_0(\boldsymbol{\theta})\right\| > \eta\right] < \varepsilon \end{aligned} \tag{14}$$

Similarly, it can be shown that when $S/T \to 0$,

$$\limsup_{S\to\infty} P\left[\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|<\delta} \sqrt{S}\left\|\boldsymbol{g}_{T,S}(\boldsymbol{\theta}) - \boldsymbol{g}_{T,S}(\boldsymbol{\theta}_0) - \boldsymbol{g}_0(\boldsymbol{\theta})\right\| / \left[1 + \sqrt{S}\left\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\right\|\right] > \eta\right] < \varepsilon. \tag{15}$$

■

Proof of Proposition 3. First, we prove the consistency of the numerical derivatives $\hat{\mathbf{G}}_{T,S}$. This part of the proof is similar to that of Theorem 7.4 in Newey and McFadden (1994). We will consider one-sided derivatives first, with the same arguments applying to two-sided derivatives. First we consider the case where $S/T \to \infty$ or $S/T \to k > 0$ as $T, S \to \infty$. We know that $\left\|\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right\| = O_p\left(T^{-1/2}\right)$ by the conclusion of Proposition 2. Also, by assumption we have $\varepsilon_{T,S} \to 0$ and $\varepsilon_{T,S}\sqrt{T} \to \infty$, so

$$\left\|\hat{\boldsymbol{\theta}}_{T,S} + \mathbf{e}_k\varepsilon_{T,S} - \boldsymbol{\theta}_0\right\| \le \left\|\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right\| + \left\|\mathbf{e}_k\varepsilon_{T,S}\right\| = O_p\left(T^{-1/2}\right) + O\left(\varepsilon_{T,S}\right) = O_p\left(\varepsilon_{T,S}\right)$$

(Recall that $\mathbf{e}_k$ is the k^{th} unit vector.) In the proof of Proposition 2, it is shown that $\exists\delta$ such that

$$\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\leq\delta}\sqrt{T}\left\|\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)-\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)-\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right\|/\left[1+\sqrt{T}\left\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\right\|\right]=o_p\left(1\right)$$

Substituting $\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}$ for $\boldsymbol{\theta}$, then for T,S large, it follows that

$$\sqrt{T}\left\|\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)-\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)-\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)\right\|/\left[1+\sqrt{T}\left\|\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}-\boldsymbol{\theta}_0\right\|\right]\leq o_p\left(1\right)$$

$$\begin{aligned}
&\text{so}\quad\left\|\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)-\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)-\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)\right\| \\
\leq&\quad\left[1+\sqrt{T}\underbrace{\left\|\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}-\boldsymbol{\theta}_0\right\|}_{=O_p\left(\varepsilon_{T,S}\right)}\right]o_p\left(\frac{1}{\sqrt{T}}\right) \\
=&\quad\sqrt{T}O_p\left(\varepsilon_{T,S}\right)o_p\left(\frac{1}{\sqrt{T}}\right)=O_p\left(\varepsilon_{T,S}\right)o_p\left(1\right) \\
=&\quad o_p\left(\varepsilon_{T,S}\right) \qquad (16)
\end{aligned}$$

On the other hand, since $\mathbf{g}_0\left(\boldsymbol{\theta}\right)$ is differentiable at $\boldsymbol{\theta}_0$ with derivative $\mathbf{G}_0$ by Assumption 4(ii), a Taylor expansion of $\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)$ around $\boldsymbol{\theta}_0$ is

$$\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)=\mathbf{g}_0\left(\boldsymbol{\theta}_0\right)+\mathbf{G}_0\cdot\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}-\boldsymbol{\theta}_0\right)+o\left(\left\|\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}-\boldsymbol{\theta}_0\right\|\right)$$

with $\mathbf{g}_0\left(\boldsymbol{\theta}_0\right)=0$. Then divide by $\varepsilon_{T,S}$,

$$\begin{aligned}
\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)/\varepsilon_{T,S} &= \mathbf{G}_0\cdot\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}-\boldsymbol{\theta}_0\right)/\varepsilon_{T,S}+o\left(\varepsilon_{T,S}^{-1}\left\|\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}-\boldsymbol{\theta}_0\right\|\right) \\
\text{so}\quad\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)/\varepsilon_{T,S}-\mathbf{G}_0\mathbf{e}_k &= \mathbf{G}_0\cdot\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)/\varepsilon_{T,S}+o\left(\varepsilon_{T,S}^{-1}\left\|\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}-\boldsymbol{\theta}_0\right\|\right)
\end{aligned}$$

The triangle inequality implies that

$$\begin{aligned}
\left\|\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)/\varepsilon_{T,S}-\mathbf{G}_0\mathbf{e}_k\right\| &\leq \left\|\mathbf{G}_0\cdot\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)/\varepsilon_{T,S}\right\|+o\left(\varepsilon_{T,S}^{-1}\left\|\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}-\boldsymbol{\theta}_0\right\|\right) \\
&= \frac{1}{\sqrt{T}\varepsilon_{T,S}}\left\|\mathbf{G}_0\cdot\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)\right\|+\varepsilon_{T,S}^{-1}\left\|\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}-\boldsymbol{\theta}_0\right\|o\left(1\right) \\
&= o\left(1\right)O_p\left(1\right)+\varepsilon_{T,S}^{-1}O_p\left(\varepsilon_{T,S}\right)o\left(1\right) \qquad (17) \\
&= o_p\left(1\right)
\end{aligned}$$

Combining the inequalities in equations (16) and (17) gives

$$\begin{aligned}
\left(\frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)-\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)}{\varepsilon_{T,S}}-\mathbf{G}_0\mathbf{e}_k\right) &= \left(\frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)-\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)-\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)}{\varepsilon_{T,S}}\right) \\
&\quad +\left(\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)/\varepsilon_{T,S}-\mathbf{G}_0\mathbf{e}_k\right) \\
\left\|\frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)-\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)}{\varepsilon_{T,S}}-\mathbf{G}_0\mathbf{e}_k\right\| &\leq \left\|\frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)-\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)-\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)}{\varepsilon_{T,S}}\right\| \\
&\quad +\left\|\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)/\varepsilon_{T,S}-\mathbf{G}_0\mathbf{e}_k\right\| \\
&\leq o_p(1)
\end{aligned}$$

Then,

$$\frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)-\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)}{\varepsilon_{T,S}}\xrightarrow{p}\mathbf{G}_0\mathbf{e}_k$$

and the same arguments can be applied to the two-sided derivative:

$$\frac{\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}+\mathbf{e}_k\varepsilon_{T,S}\right)-\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}-\mathbf{e}_k\varepsilon_{T,S}\right)}{2\varepsilon_{T,S}}\xrightarrow{p}\mathbf{G}_0\mathbf{e}_k$$

This holds for each column $k=1,2...,p$. Thus $\hat{\mathbf{G}}_{T,S}\xrightarrow{p}\mathbf{G}_0$.

In the case where $S/T\to 0$ as $T,S\to\infty$, the proof for the consistency of $\hat{\mathbf{G}}_{T,S}$ is done in the similar way using the following facts:

$$\left\|\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right\|=O_p\left(S^{-1/2}\right) \tag{18}$$

and $\exists\delta$

$$\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\leq\delta}\sqrt{S}\left\|\mathbf{g}_{T,S}\left(\boldsymbol{\theta}\right)-\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)-\mathbf{g}_0\left(\boldsymbol{\theta}\right)\right\|/\left[1+\sqrt{S}\left\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\right\|\right]=o_p(1) \tag{19}$$

Next, we show the consistency of $\hat{\boldsymbol{\Sigma}}_{T,B}$. If $\boldsymbol{\mu}_t$ and $\boldsymbol{\sigma}_t$ are known constant, or if $\boldsymbol{\phi}_0$ is known, then the result follows from Theorems 5 and 6 of Fermanian, Radulović and Wegkamp (2004). When $\boldsymbol{\phi}_0$ is estimated, the result is obtained by combining the results in Fermanian, *et al.* with those of Rémillard (2010): For simplicity, assume that only one dependence measure is used. Let $\hat{\tau}_{ij}$ and $\hat{\tau}_{ij}^{(b)}$ be the sample quantile dependence constructed from the standardized residuals $\left\{\hat{\eta}_t^i, \hat{\eta}_t^j\right\}_{t=1}^T$ and from the bootstrap counterpart $\left\{\hat{\eta}_t^{(b)i}, \hat{\eta}_t^{(b)j}\right\}_{t=1}^T$. Also, define the corresponding estimates, $\ddot{\tau}_{ij}$ and $\ddot{\tau}_{ij}^{(b)}$, using the true innovations $\left\{\eta_t^i, \eta_t^j\right\}_{t=1}^T$ and the bootstrapped true innovations $\left\{\eta_t^{(b)i}, \eta_t^{(b)j}\right\}_{t=1}^T$ (where the same bootstrap time indices are used for both $\left\{\hat{\eta}_t^{(b)i}, \hat{\eta}_t^{(b)j}\right\}_{t=1}^T$ and $\left\{\eta_t^{(b)i}, \eta_t^{(b)j}\right\}_{t=1}^T$). Define true τ as τ_0. Theorem 5 of Fermanian, Radulović and Wegkamp (2004) shows that

$$\sqrt{T}\left(\ddot{\tau}_{ij}-\tau_0\right)=\sqrt{T}\left(\ddot{\tau}_{ij}^{(b)}-\ddot{\tau}_{ij}\right)+o_p(1) \tag{20}$$

Corollary 1 and Proposition 4 of Rémillard (2010) shows, under Assumption 2, that

$$\begin{aligned}\sqrt{T}\left(\hat{\tau}_{ij}-\ddot{\tau}_{ij}\right) &= o_p(1) &(21)\\ \text{and}\quad \sqrt{T}\left(\hat{\tau}_{ij}^{(b)}-\ddot{\tau}_{ij}^{(b)}\right) &= o_p(1) &(22)\end{aligned}$$

Combining those three equations, we obtain

$$\sqrt{T}\left(\hat{\tau}^{ij}-\tau_0\right)=\sqrt{T}\left(\hat{\tau}_{ij}^{(b)}-\hat{\tau}^{ij}\right)+o_p(1), \text{ as } T,B\rightarrow\infty \tag{23}$$

and so $\hat{\mathbf{\Sigma}}_{T,B}$ defined in equation (11) in the paper is a consistent estimator of $\mathbf{\Sigma}_0$. ■

Proof of Proposition 4. First consider $S/T\rightarrow\infty$ or $S/T\rightarrow k>0$. A Taylor expansion of $\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}\right)$ around $\boldsymbol{\theta}_0$ yields

$$\sqrt{T}\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}\right)=\sqrt{T}\mathbf{g}_0\left(\boldsymbol{\theta}_0\right)+\mathbf{G}_0\cdot\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)+o\left(\sqrt{T}\left\|\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right\|\right) \tag{24}$$

and since $\mathbf{g}_0\left(\boldsymbol{\theta}_0\right)=0$ and $\sqrt{T}\left\|\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right\|=O_p(1)$

$$\sqrt{T}\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}\right)=\mathbf{G}_0\cdot\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)+o_p(1) \tag{25}$$

Then consider the following expansion of $\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)$ around $\boldsymbol{\theta}_0$

$$\sqrt{T}\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)=\sqrt{T}\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)+\hat{\mathbf{G}}_{T,S}\cdot\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)+\mathbf{R}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) \tag{26}$$

where the remaining term is captured by $\mathbf{R}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)$. Combining equations (25) and (26) we obtain

$$\sqrt{T}\left[\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)-\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)-\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}\right)\right]=\left(\hat{\mathbf{G}}_{T,S}-\mathbf{G}_0\right)\cdot\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)+\mathbf{R}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)+o_p\left(1\right)$$

Lemma 7 shows the stochastic equicontinuity of $\mathbf{v}_{T,S}\left(\boldsymbol{\theta}\right)$, which implies (see proof of Proposition 2) that

$$\sqrt{T}\left[\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)-\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)-\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}\right)\right]=o_p\left(1\right)$$

By Proposition 3, $\hat{\mathbf{G}}_{T,S}-\mathbf{G}_0=o_p\left(1\right)$, which implies $\mathbf{R}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)=o_p\left(1\right)$. Thus, we obtain the expansion of $\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)$ around $\boldsymbol{\theta}_0$:

$$\sqrt{T}\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)=\sqrt{T}\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)+\hat{\mathbf{G}}_{T,S}\cdot\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)+o_p\left(1\right) \tag{27}$$

The remainder of the proof is the same as in standard GMM applications: From the proof of Proposition 2, we have $\sqrt{T}\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)\xrightarrow{d}N\left(0,\boldsymbol{\Sigma}_0\right)$ and rewrite this as $-\boldsymbol{\Sigma}_0^{-1/2}\sqrt{T}\mathbf{g}_{T,S}\left(\boldsymbol{\theta}_0\right)\equiv\mathbf{u}_{T,S}\xrightarrow{d}\mathbf{u}\sim N\left(0,\mathbf{I}\right)$, and from Proposition 2, we have $\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{T,S}-\boldsymbol{\theta}_0\right)=\left(\mathbf{G}_0'\mathbf{W}_0\mathbf{G}_0\right)^{-1}\mathbf{G}_0'\mathbf{W}_0\boldsymbol{\Sigma}_0^{1/2}\mathbf{u}_{T,S}+o_p\left(1\right)$. By these two equations and Proposition 3, equation (27) becomes

$$\begin{aligned}\sqrt{T}\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) &= -\hat{\boldsymbol{\Sigma}}_{T,B}^{1/2}\mathbf{u}_{T,S}+\hat{\mathbf{G}}_{T,S}\left(\hat{\mathbf{G}}_{T,S}'\hat{\mathbf{W}}_T\hat{\mathbf{G}}_{T,S}\right)^{-1}\hat{\mathbf{G}}_{T,S}'\hat{\mathbf{W}}_T\hat{\boldsymbol{\Sigma}}_{T,B}^{1/2}\mathbf{u}_{T,S}+o_p\left(1\right) \quad (28)\\ &= -\hat{\boldsymbol{\Sigma}}_{T,B}^{1/2}\hat{\mathbf{R}}\mathbf{u}_{T,S}+o_p\left(1\right)\end{aligned}$$

where $\hat{\mathbf{R}}\equiv\left(\mathbf{I}-\hat{\boldsymbol{\Sigma}}_{T,B}^{-1/2}\hat{\mathbf{G}}_{T,S}\left(\hat{\mathbf{G}}_{T,S}'\hat{\mathbf{W}}_T\hat{\mathbf{G}}_{T,S}\right)^{-1}\hat{\mathbf{G}}_{T,S}'\hat{\mathbf{W}}_T\hat{\boldsymbol{\Sigma}}_{T,B}^{1/2}\right)$. The test statistic is

$$\begin{aligned}T\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right)'\hat{\mathbf{W}}_T\mathbf{g}_{T,S}\left(\hat{\boldsymbol{\theta}}_{T,S}\right) &= \mathbf{u}_{T,S}'\hat{\mathbf{R}}'\hat{\boldsymbol{\Sigma}}_{T,B}^{1/2\prime}\hat{\mathbf{W}}_T\hat{\boldsymbol{\Sigma}}_{T,B}^{1/2}\hat{\mathbf{R}}\mathbf{u}_{T,S}+o_p\left(1\right) \quad (29)\\ &= \mathbf{u}'\mathbf{R}_0'\boldsymbol{\Sigma}_0^{1/2\prime}\mathbf{W}_0\boldsymbol{\Sigma}_0^{1/2}\mathbf{R}_0\mathbf{u}+o_p\left(1\right)\end{aligned}$$

where $\mathbf{R}_0\equiv\left(\mathbf{I}-\boldsymbol{\Sigma}_0^{-1/2}\mathbf{G}_0\left(\mathbf{G}_0'\mathbf{W}_0\mathbf{G}_0\right)^{-1}\mathbf{G}_0'\mathbf{W}_0\boldsymbol{\Sigma}_0^{1/2}\right)$. When $\hat{\mathbf{W}}_T=\hat{\boldsymbol{\Sigma}}_{T,B}^{-1}$, $\hat{\mathbf{R}}$ is symmetric and idempotent with $rank\left(\hat{\mathbf{R}}\right)=tr\left(\hat{\mathbf{R}}\right)=m-p$, and the test statistic converges to a χ^2_{m-p} random variable, as usual. In general, the asymptotic distribution is a sample-dependent combination of m independent standard Normal variables, namely that of $\mathbf{u}'\mathbf{R}_0'\boldsymbol{\Sigma}_0^{1/2\prime}\mathbf{W}_0\boldsymbol{\Sigma}_0^{1/2}\mathbf{R}_0\mathbf{u}$ where $\mathbf{u} \sim N(0, \mathbf{I})$.

When $S/T \to 0$, a similar proof can be given using Taylor expansion of $\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}\right)$

$$\sqrt{S}\mathbf{g}_0\left(\hat{\boldsymbol{\theta}}_{T,S}\right) = \sqrt{S}\mathbf{g}_0(\boldsymbol{\theta}_0) + \mathbf{G}_0 \cdot \sqrt{S}\left(\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right) + o\left(\sqrt{S}\left\|\hat{\boldsymbol{\theta}}_{T,S} - \boldsymbol{\theta}_0\right\|\right) \tag{30}$$

■

Supplemental Appendix: Additional tables

Table S1: Summary statistics on the daily stock returns

	Bank of America	Bank of N.Y.	Citi Group	Goldman Sachs	JP Morgan	Morgan Stanley	Wells Fargo
Mean	0.038	0.015	-0.020	0.052	0.041	0.032	0.047
Std dev	3.461	2.797	3.817	2.638	2.966	3.814	2.965
Skewness	1.048	0.592	1.595	0.984	0.922	4.982	2.012
Kurtosis	28.190	18.721	43.478	18.152	16.006	119.757	30.984

Notes: This table presents some summary statistics of the seven daily equity returns data used in the empirical analysis.

Table S2: Parameter estimates for the conditional mean and variance models

	Bank of America	Bank of N.Y.	Citi Group	Goldman Sachs	JP Morgan	Morgan Stanley	Wells Fargo
Constant (ϕ_0)	0.038	0.017	-0.019	0.058	0.043	0.031	0.051
$r_{i,t-1}$	0.020	-0.151	0.053	-0.156	-0.035	0.004	-0.078
$r_{m,t-1}$	-0.053	-0.011	0.029	0.282	-0.141	0.063	-0.099
Constant (ω)	0.009	0.069	0.019	0.034	0.014	0.036	0.008
$\sigma^2_{i,t-1}$	0.931	0.895	0.901	0.953	0.926	0.922	0.926
$\varepsilon^2_{i,t-1}$	0.031	0.017	0.036	0.000	0.025	0.002	0.021
$\varepsilon^2_{i,t-1} \cdot \mathbf{1}_{\{\varepsilon_{i,t-1} \leq 0\}}$	0.048	0.079	0.123	0.077	0.082	0.135	0.108
$\varepsilon^2_{m,t-1}$	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$\varepsilon^2_{m,t-1} \cdot \mathbf{1}_{\{\varepsilon_{m,t-1} \leq 0\}}$	0.068	0.266	0.046	0.012	0.064	0.077	0.013

Notes: This table presents the estimated models for the conditional mean (top panel) and conditional variance (lower panel).