Supplementary Material on "a general class of non-nested test statistics for models defined through moment restrictions"

Paulo M.D.C. Parente

Instituto Universitário de Lisboa (ISCTE-IUL)

Business Research Unit (BRU-IUL)

This version: March 2017

In this supplement we provide the proofs of the theorems presented in Sections 4 and 5 of the paper, investigate the behavior of the random variable $W^*_{i,j}(\delta_h)$ $i = 1, 2,$ $j = 3, 4$ (defined in subsection 5.1 of the paper) as some elements of δ_h approach infinity, and present the results of the Monte Carlo study for the tests based on the exponential tilting estimator. This supplement is organized as follows. In Section SM1 we prove the relevant theorems of section 4. Section SM2 provides the proofs of the theorems of section 5 and analyzes the limit of $W^*_{i,j}(\delta_h)$ $i = 1, 2,$ $j = 3, 4$ as some elements of δ_h approach infinity. Finally, SM3 presents the additional results obtained in the Monte Carlo study for the tests based on the exponential tilting estimator.

In what follows CR, CS, L, and T denote the c_r, Cauchy-Schwarz, Lyapunov and triangle inequalities respectively. Furthermore, 'with probability approaching one' is abbreviated as 'wpa1'. Unless stated otherwise 'LLN' corresponds to the Khinchin law of large numbers, 'UWL' denotes a uniform weak law of large numbers, as Lemma 2.4 of Newey and McFadden (1994) or a uniform weak law of large numbers at the true parameter as Lemma 4.3 of Newey and McFadden (1994) and 'CLT' refers to the Lindeberg-Lévy central limit theorem. NS refers to Newey and Smith (2004).

SM1 Proofs of results of section 4

SM1.1 Proofs of the results of subsection 4.1

The following Lemma generalizes Lemma A.1 of Ramalho and Smith (2004). Let $g_i(\beta) = g(z_i, \beta)$ and $\hat{g}(\beta) = \sum_{i=1}^{n} g_i(\beta)/n$.

Lemma SM1.1 *Let Assumptions 2.1, 2.2 and 4.1 (a) hold. Then* $n\hat{p}_i^{v_c} = 1 + o_p(1)$ *and*

$$n^{1/2}\left(\hat{p}_i^{v_c} - \frac{1}{n}\right) = \kappa_{v_c}\frac{1}{n}\hat{g}_i' n^{1/2}\hat{\lambda}(1 + o_p(1)) + O_p(n^{-3/2}),$$

uniformly $(i = 1, ..., n)$ *where* $\hat{g}_i \equiv g(z_i, \hat{\beta})$ *and* $\kappa_{v_c} = v_{c,1}(0)/v_c(0)$.

Proof: Let $b_i \equiv \sup_{\beta\in B}|g_i(\beta)|$. From the proof of Lemma A1 and Theorem 3.1 in NS we have $\max_{1\le i\le n} b_i = O_p(n^{-1/\alpha})$ and $\hat{\lambda} = O_p(n^{-1/2})$. Thus, $\sup_{\beta\in B, 1\le i\le n}|\hat{\lambda}' g_i(\beta)| = O_p(n^{-(1/2-1/\alpha)})$. A first order Taylor expansion of $v_c(\hat{\lambda}'\hat{g}_i)$ around zero yields $v_c(\hat{\lambda}'\hat{g}_i) = v_c(0) + v_{c,1}(\dot{\lambda}'\hat{g}_i)\hat{\lambda}'\hat{g}_i$, where $\dot{\lambda}$ is on a line joining $\hat{\lambda}$ and zero. Now $\max_{1\le i\le n}|v_{c,1}(\dot{\lambda}' g_i(\hat{\beta})) - v_{c,1}(0)| = o_p(1)$ as $\sup_{\beta\in B, 1\le i\le n}|\hat{\lambda}' g_i(\beta)| = o_p(1)$ and so $v_{c,1}(\dot{\lambda}' g_i(\hat{\beta}))\hat{\lambda}'\hat{g}_i = v_{c,1}(0)\hat{\lambda}'\hat{g}_i(1+o_p(1))$. Therefore $v_c(\hat{\lambda}'\hat{g}_i) = v_c(0) + v_{c,1}(0)\hat{\lambda}'\hat{g}_i(1+o_p(1))$ uniformly $(i = 1, ..., n)$. Similarly,

$$\left[\sum_{j=1}^{n} v_c(\hat{\lambda}'\hat{g}_j)\right]^{-1} = (v_c(0)n)^{-1}(1 + O_p(n^{-1}))$$

as $\sum_{j=1}^{n}\hat{g}_j/n = O_p(n^{-1/2})$ and $\hat{\lambda} = O_p(n^{-1/2})$ by Theorem 3.1 of NS. Hence $\hat{p}_i^{v_c} = [v_c(0) + v_{c,1}(0)\hat{\lambda}'\hat{g}_i(1 + o_p(1))](v_c(0)n)^{-1}(1 + O_p(n^{-1}))$. It follows by Lemma A.1 of NS that $n\hat{p}_i^{v_c} = 1 + (v_{c,1}(0)/v_c(0))o_p(1)$ and that

$$n^{1/2}(\hat{p}_i^{v_c} - 1/n) = (v_{c,1}(0)/v_c(0))n^{-1}\hat{g}_i' n^{1/2}\hat{\lambda}(1 + o_p(1)) + O_p(n^{-3/2}).$$

■

Proof of Theorem 4.1: By the mean value theorem $v_a(\hat{p}_i n) = v_a(1) + v_{a,1}(\hat{\sigma}_i)(\hat{p}_i n - 1)$, where $\hat{\sigma}_i = \alpha_i + (1-\alpha_i)\hat{p}_i n$ and $\alpha_i \in (0,1)$ and consequently

$$\mathcal{S}_v = \sum_{i=1}^{n} v_{a,1}(\hat{\sigma}_i)\sqrt{n}\left(\hat{p}_i - \frac{1}{n}\right)n\hat{p}_i\left[v_b\left(\frac{\hat{q}_i^{v_c}}{\hat{p}_i^{v_c}}\right) - \sum_{\ell=1}^{n} v\left(\frac{\hat{q}_\ell^{v_c}}{\hat{p}_\ell^{v_c}}\right)\hat{p}_\ell\right].$$

By Lemma SM1.1 $\mathcal{S}_v = \sum_{j=1}^{4} R_{j,n}$, where $R_{j,n}, j = 1, ..., 4$ are defined below.

Let us consider first

$$R_{1,n} \equiv \frac{1}{n}\sum_{i=1}^{n} v_{a,1}(\hat{\sigma}_i)\, n\hat{p}_i v_b\left(\frac{v_c(\hat{\eta}' h_i(\hat{\gamma}))}{n\hat{p}_i^{v_c}\hat{v}_{c,h}}\right)\hat{g}_i'\sqrt{n}\hat{\lambda}(1 + o_p(1)),$$

where $\hat{v}_{c,h} \equiv \sum_{i=1}^{n} v_c(\hat{\eta}' h_i(\hat{\gamma}))/n$. Now by Lemma SM1.1, $v_{a,1}(\hat{\sigma}_i) = v_{a,1}(1) + o_p(1)$. Additionally $n\hat{p}_i = 1 + o_p(1)$ and $n\hat{p}_i^{v_c} = 1 + o_p(1)$ uniformly in $i = 1, ..., n$ by Lemma SM1.1. Also $\hat{v}_{c,h} = \mathrm{E}_{\mathrm{P}_0}[v_c(\eta^{*\prime} h(z, \gamma^*))] + o_p(1)$ by a UWL. It follows using the fact that $\sqrt{n}\hat{\lambda} = O_p(1)$, and a UWL that

$$R_{1,n} = A_v\sqrt{n}\hat{\lambda}(1 + o_p(1)) + o_p(1). \tag{1}$$

Hence by Theorem 3.2 of NS we have $R_{1,n} \xrightarrow{d} \mathcal{N}(0, \sigma_0^2)$.

Let us consider now

$$R_{2,n} \equiv \frac{1}{n}\sum_{i=1}^{n} v_{a,1}\left(\hat{\sigma}_i\right) n\hat{p}_i v_b \left(\frac{v_c\left(\hat{\eta}' h_i\left(\hat{\gamma}\right)\right)}{n\hat{p}_i^{v_c}\hat{v}_{c,h}}\right) O_p(n^{-1/2}).$$

Using the same arguments as above $n^{-1}\sum_{i=1}^{n} v_{a,1}\left(\hat{\sigma}_i\right) n\hat{p}_i v_b(v_c\left(\hat{\eta}' h_i\left(\hat{\gamma}\right)\right)/[n\hat{p}_i^{v_c}\hat{v}_{c,h}]) = O_p(1)$. It follows that $R_{2,n} = O_p(n^{-1/2})$. Now define

$$R_{3,n} \equiv \frac{1}{n}\sum_{i=1}^{n} v_{a,1}\left(\hat{\sigma}_i\right) n\hat{p}_i \hat{g}_i' \sqrt{n}\hat{\lambda}(1+o_p(1)) \sum_{\ell=1}^{n} v_b\left(\frac{\hat{q}_\ell^{v_c}}{\hat{p}_\ell^{v_c}}\right)\hat{p}_\ell.$$

Note that $n^{-1}\sum_{i=1}^{n} v_{a,1}\left(\hat{\sigma}_i\right) n\hat{p}_i \hat{g}_i' = o_p(1)$ by a UWL, $\sqrt{n}\hat{\lambda}(1+o_p(1)) = O_p(1)$ by Theorem 3.2 of NS and that

$$\sum_{\ell=1}^{n} v_b\left(\frac{\hat{q}_\ell^{v_c}}{\hat{p}_\ell^{v_c}}\right)\hat{p}_\ell = \frac{1}{n}\sum_{\ell=1}^{n} v_b\left(\frac{v_c\left(\hat{\eta}' h_\ell\left(\hat{\gamma}\right)\right)}{n\hat{p}_\ell^{v_c}\hat{v}_{c,h}}\right) n\hat{p}_\ell$$

converges in probability to $\mathrm{E}_{\mathrm{P}_0}[v_b(v_c\left(\eta^{*\prime} h\left(z,\gamma^*\right)\right)/\mathrm{E}_{\mathrm{P}_0}\left[v_c\left(\eta^{*\prime} h\left(z,\gamma^*\right)\right)\right])] + o_p\left(1\right)$ by a UWL. Hence $R_{3,n} = o_p(1)$. Finally, consider

$$R_{4,n} \equiv \frac{1}{n}\sum_{i=1}^{n} v_{a,1}\left(\hat{\sigma}_i\right) O_p(n^{-1/2}) \sum_{\ell=1}^{n} v_b\left(\frac{\hat{q}_\ell^{v_c}}{\hat{p}_\ell^{v_c}}\right)\hat{p}_\ell.$$

Since by a UWL $n^{-1}\sum_{i=1}^{n} v_{a,1}\left(\hat{\sigma}_i\right) = O_p(n^{-1/2})$ and $\sum_{\ell=1}^{n} v_b\left(\hat{q}_\ell^{v_c}/\hat{p}_\ell^{v_c}\right)\hat{p}_\ell = O_p(1)$ we have $R_{4,n} = o_p(1)$. Hence $\mathcal{S}_v \xrightarrow{d} \mathcal{N}\left(0,\sigma_0^2\right)$.

The fact that $\tilde{\mathcal{S}}_v = \mathcal{S}_v + o_p(1)$ follows from the arguments above and the fact that $n\hat{p}_i^{v_c} = 1 + o_p\left(1\right)$ by Lemma SM1.1.

Concerning the Lagrange multiplier statistic $\mathcal{LM}_v = \hat{A}_v\sqrt{n}\hat{\lambda}$, note that $\mathcal{S}_v = R_{1.n} + o_p(1)$ and using (1) one obtains $R_{1.n} - \mathcal{LM}_v = [A_{0,v} - \hat{A}_v]\sqrt{n}\hat{\lambda} + A_v\sqrt{n}\hat{\lambda}o_p(1) + o_p(1)$. Because $A_{0,v} - \hat{A}_v = o_p(1)$ and $\sqrt{n}\hat{\lambda} = O_p(1)$ it follows that $R_{1.n} - \mathcal{LM}_v = o_p(1)$.

Finally we consider the statistic $\mathcal{J}_v = -\hat{A}_v\hat{\Omega}_g^{-1}\sqrt{n}\hat{g}(\hat{\beta})$. Note that NS proved in the proof of Theorem 3.2 (p. 240) that $\hat{g}(\hat{\beta}) = -\Omega_g\hat{\lambda} + o_p(n^{1/2})$. Therefore $\mathcal{J}_v - \mathcal{LM}_v = \left[\hat{A}_v\hat{\Omega}_g^{-1}\Omega_g - \hat{A}_v\right]\sqrt{n}\hat{\lambda} + o_p(1)$, as $\hat{A}_v\hat{\Omega}_g^{-1} = O_p(1)$ and $\sqrt{n}\hat{\lambda} = O_p(1)$. Since $\hat{\Omega}_g^{-1} = \Omega_g^{-1} + o_p(1)$ and $\hat{A}_v = A_{0,v} + o_p(1)$, it follows that $\mathcal{J}_v - \mathcal{LM}_v = o_p(1)$. ■

Proof of Theorem 4.2: We prove here only consistency of $\hat{\sigma}_3^2$ for σ_0^2 as the proof for the other estimators $\hat{\sigma}_j^2, j = 1,2,4$ is simpler. First note that $\hat{P} \xrightarrow{p} P_{0,g}$ by a UWL and the Slutsky theorems under Assumptions 2.1 and 2.3. Now note that

$$\hat{A}_{v,3} = \frac{1}{n}\sum\nolimits_{i=1}^{n} v_b\left(\hat{q}_i^{v_c}/\hat{p}_i^{v_c}\right)\hat{g}_i' n\hat{p}_i = \frac{1}{n}\sum\nolimits_{i=1}^{n} v_b\left(\frac{v_c(\hat{\eta}' h_i(\hat{\gamma}))}{\frac{1}{n}\sum_{i=1}^{n} v_c(\hat{\eta}' h_i(\hat{\gamma}))}\frac{1}{n\hat{p}_i}\right)\hat{g}_i' n\hat{p}_i$$

By a UWL $\sum_{i=1}^{n} v_c(\hat{\eta}' h_i(\hat{\gamma}))/n \xrightarrow{p} \mathrm{E}_{\mathrm{P}_0}(v_c(\eta^{*\prime} h(z,\gamma^*)))$. Also by Lemma SM1.1, $n\hat{p}_i = 1 + o_p(1)$. Hence, by a UWL and continuity of $v_b\left(\cdot\right)$ it follows that $\hat{A}_{v,3} = \mathrm{E}_{\mathrm{P}_0}(v_b(v_c(\eta^{*\prime} h(z,\gamma^*))/\mathrm{E}_{\mathrm{P}_0}[v_c(\eta^{*\prime} h(z,\gamma^*))])g(z,\beta_0)) + o_p(1)$. Hence consistency of $\hat{\sigma}_3^2$ for σ_0^2 follows from the Slutsky theorem. ■

SM1.2 Proofs of the results of subsection 4.2

Proof of Theorem 4.2: Let us first consider $\mathcal{S}_v$. Note that

$$\begin{aligned}\mathcal{S}_v/\sqrt{n} &= \frac{1}{n}\sum_{i=1}^n v_a\left(\hat{p}_i n\right) n\hat{p}_i v_b\left(\frac{n\hat{q}_i^{v_c}}{n\hat{p}_i^{v_c}}\right) - \\ &\quad \frac{1}{n}\sum_{i=1}^n v_a\left(\hat{p}_i n\right) n\hat{p}_i \frac{1}{n}\sum_{\ell=1}^n n\hat{p}_\ell v_b\left(\frac{n\hat{q}_\ell^{v_c}}{n\hat{p}_\ell^{v_c}}\right).\end{aligned}$$

Now notice that $n\hat{q}_i^{v_c} = v_c(\hat{\eta}'h_i(\hat{\gamma}))/[\sum_{i=1}^n v_c(\hat{\eta}'h_i(\hat{\gamma}))/n]$, $n\hat{p}_i^{v_c} = v_c(\hat{\lambda}'g_i(\hat{\beta})) \Big/ \left[\sum_{i=1}^n v_c(\hat{\lambda}'g_i(\hat{\beta}))/n\right]$ and $n\hat{p}_i = \rho_1(\hat{\lambda}'g_i(\hat{\beta})) \Big/ \left[\sum_{i=1}^n \rho_1(\hat{\lambda}'g_i(\hat{\beta}))/n\right]$. A UWL implies that $\sum_{i=1}^n v_c(\hat{\eta}'h_i(\hat{\gamma})/n = \mathrm{E}_{\mathrm{P}_0}[v_c(\eta^{*\prime}h(z,\gamma^*))] + o_p(1)$, $\sum_{i=1}^n v_c(\hat{\lambda}'g_i(\hat{\beta}))/n = \mathrm{E}_{\mathrm{P}_0}[v_c(g(z,\beta^*))] + o_p(1)$, $\sum_{i=1}^n \rho_1(\hat{\lambda}'g_i(\hat{\beta}))/n = \mathrm{E}_{\mathrm{P}_0}[\rho_1(g(z,\beta^*))] + o_p(1)$. It follows by a UWL that

$$\frac{1}{n}\sum_{i=1}^n v_a\left(\hat{p}_i n\right) n\hat{p}_i v_b\left(\frac{n\hat{q}_i^{v_c}}{n\hat{p}_i^{v_c}}\right) = \mathrm{E}_{\mathrm{P}_0}\left[v_a(\rho_1^{g,z})\rho_1^{g,z} v_b\left(\frac{v_c^{h,z}}{v_c^{g,z}}\right)\right] + o_p(1).$$

Also

$$\frac{1}{n}\sum_{i=1}^n v_a\left(\hat{p}_i n\right) n\hat{p}_i = \mathrm{E}_{\mathrm{P}_0}\left[v_a(\rho_1^{g,z})\rho_1^{g,z}\right] + o_p(1)$$

and

$$\frac{1}{n}\sum_{\ell=1}^n n\hat{p}_\ell v_b\left(\frac{n\hat{q}_\ell^{v_c}}{n\hat{p}_\ell^{v_c}}\right) = \mathrm{E}_{\mathrm{P}_0}[\rho_1^{g,z} v_b\left(\frac{v_c^{h,z}}{v_c^{g,z}}\right)] + o_p(1).$$

Because $\mathrm{E}_{\mathrm{P}_0}[v_a(\rho_1^{g,z})\rho_1^{g,z} v_b(v_c^{h,z}/v_c^{g,z})] \neq \mathrm{E}_{\mathrm{P}_0}[v_a(\rho_1^{g,z})\rho_1^{g,z}]\mathrm{E}_{\mathrm{P}_0}[\rho_1^{g,z} v_b(v_c^{h,z}/v_c^{g,z})]$, $\mathcal{S}_v \xrightarrow{p} \pm\infty$.

Consider now $\tilde{\mathcal{S}}_v$. Note that

$$\begin{aligned}\tilde{\mathcal{S}}_v/\sqrt{n} &= \frac{1}{n}\sum_{i=1}^n v_a\left(\hat{p}_i n\right) n\hat{p}_i v_b\left(n\hat{q}_i^{v_c}\right) \\ &\quad -\frac{1}{n}\sum_{i=1}^n v_a\left(\hat{p}_i n\right) n\hat{p}_i \frac{1}{n}\sum_{\ell=1}^n v_b\left(n\hat{q}_\ell^{v_c}\right) n\hat{p}_\ell.\end{aligned}$$

Using similar arguments to those described above we have $\sum_{i=1}^n v_a\left(\hat{p}_i n\right) n\hat{p}_i v_b\left(n\hat{q}_i^{v_c}\right)/n = \mathrm{E}_{\mathrm{P}_0}[v_a(\rho_1^{g,z})\rho_1^{g,z} v_b(v_c^{h,z})] + o_p(1)$, $\sum_{i=1}^n v_a\left(\hat{p}_i n\right) n\hat{p}_i/n = \mathrm{E}_{\mathrm{P}_0}[v_a(\rho_1^{g,z})\rho_1^{g,z}] + o_p(1)$ and $\sum_{\ell=1}^n n\hat{p}_\ell v_b\left(n\hat{q}_\ell^{v_c}\right)/n = \mathrm{E}_{\mathrm{P}_0}[\rho_1^{g,z} v_b(v_c^{h,z})] + o_p(1)$. Since $\mathrm{E}_{\mathrm{P}_0}[v_a(\rho_1^{g,z})\rho_1^{g,z} v_b(v_c^{h,z})] \neq \mathrm{E}_{\mathrm{P}_0}[v_a(\rho_1^{g,z})\rho_1^{g,z}]\mathrm{E}_{\mathrm{P}_0}[\rho_1^{g,z} v_b(v_c^{h,z})]$, $\tilde{\mathcal{S}}_v \xrightarrow{p} \pm\infty$.

Concerning the Lagrange multiplier statistic $\mathcal{LM}_v = \hat{A}_v\sqrt{n}\hat{\lambda}$, note that $\mathcal{LM}_v/\sqrt{n} = \hat{A}_v\hat{\lambda} = A_v^*\lambda^* + o_p(1)$. Since $A_v^*\lambda^* \neq 0$ it follows that $\mathcal{LM}_v \xrightarrow{p} \pm\infty$. Finally we consider the statistic $\mathcal{J}_v/\sqrt{n} = -\hat{A}_v\hat{\Omega}_g^{-1}\hat{g}(\hat{\beta})$. Note that $\mathcal{J}_v/\sqrt{n} \xrightarrow{p} -A_v^*\Omega_g^{*-1}\mathrm{E}_{\mathrm{P}_0}[g_i(\beta^*)]$. Given that $A_v^*\Omega_g^{*-1}\mathrm{E}_{\mathrm{P}_0}[g_i(\beta^*)] \neq 0$, $\mathcal{J}_v \xrightarrow{p} \pm\infty$. ■

SM1.3 Proofs of the results of subsection 4.3

Let $\{z_{in}\}_{i=1}^n$ be a triangular array which we assume to be row wise independent and identically distributed (iid). Let $g_{in}(\beta) = g(z_{i,n},\beta)$, $\hat{g}(\beta) = \sum_{i=1}^n g_{i,n}(\beta)/n$, $\hat{g}_{in} \equiv g(z_{i,n},\hat{\beta})$ and $h_{in}(\gamma) = h(z_{i,n},\gamma)$.

Lemma SM1.2 *Under Assumption 4.6 the following result holds* $\sup_{\beta\in\mathcal{B}}\|\hat{g}(\beta)-\mathrm{E}_{\mathrm{P}_{0n}}[g_{in}(\beta)]\| = o_p(1)$, *and* $\{\mathrm{E}_{\mathrm{P}_{0n}}[g_{in}(\beta)]\}_{n=1}^{\infty}$ *is uniformly equicontinuous in* $\beta\in\mathcal{B}$.

Proof: We use the UWL corresponding to Theorem 4 of Andrews (1992) together with the Weak Law of Large Numbers for Triangular Arrays (Davidson, 1994, 19.9 Corollary p.301). Note that by the UWL $\sup_{\beta\in\mathcal{B}}|\hat{g}(\beta)-\mathrm{E}_{\mathrm{P}_n}[g_{i,n}(\beta)]| = o_p(1)$, where the UWL applies because the four sufficient conditions are satisfied. In particular the total boundedness condition (BD) holds by Assumption 4.6 (c). Assumption 4.6 (e) implies both the pointwise convergence condition (P-WLLN) (by the LLN) and the domination condition (DM). The termwise stochastic equicontinuity (TSE) condition is satisfied because

$$\mathrm{E}_{\mathrm{P}_{0n}}[\sup_{\beta,\beta'\in\mathcal{B}:\|\beta-\beta'\|\leq d}\|g_{in}(\beta)-g_{in}(\beta')\|] \leq \mathrm{E}_{\mathrm{P}_{0n}}[\sup_{\beta\in\mathcal{B}}\left\|\frac{\partial g_{i,n}(\beta)}{\partial\beta}\right\|]d \leq Cd, \tag{2}$$

where the first inequality holds by a mean-value expansion (which relies on Assumption 4.6 (d)) and the second holds by Assumption 4.6 (e). In addition to guaranteeing TSE, equation (2) also shows the uniform equicontinuity of $\{\mathrm{E}_{\mathrm{P}_{0n}}[g(z_{in},\beta)]\}_{n=1}^{\infty}$. ■

Lemma SM1.3 *Under Assumptions 4.6, 4.7 and 4.8 the following results hold:*

1. $\frac{1}{n}\sum_{i=1}^{n}\hat{g}_{in}\hat{g}_{in}' - \Omega_{0,g} = o_p(1)$ *if* $\hat{\beta}\xrightarrow{p}\beta_0$;
2. $\frac{1}{n}\sum_{i=1}^{n}\|\hat{g}_{in}\|^2 - \mathrm{E}_{\mathrm{P}_{0n}}[\|g_{in}(\beta_0)\|^2] = o_p(1)$ *if* $\hat{\beta}\xrightarrow{p}\beta_0$;
3. $\frac{1}{n}\sum_{i=1}^{n}\frac{\partial g_{in}(\hat{\beta})}{\partial\beta'} - D_{0,g} = o_p(1)$ *if* $\hat{\beta}\xrightarrow{p}\beta_0$;
4. $\frac{1}{n}\sum_{i=1}^{n}v_c\left(\hat{\eta}'h_{in}\left(\hat{\gamma}\right)\right) - \mathrm{E}_{\mathrm{P}_{0n}}[v_c(\eta^{*\prime}h_{in}\left(\gamma^*\right))] = o_p(1)$ *if* $\hat{\beta}\xrightarrow{p}\beta_0$, $\hat{\gamma}\xrightarrow{p}\gamma^*$ *and* $\hat{\eta}\xrightarrow{p}\eta^*$;
5. $\frac{1}{n}\sum_{i=1}^{n}v_b\left(\frac{v_c\left(\hat{\eta}'h_{in}(\hat{\gamma})\right)}{\sum_{i=1}^{n}v_c(\hat{\eta}'h_{in}(\hat{\gamma}))/n}\right) - \mathrm{E}_{\mathrm{P}_{0n}}[v_b(\frac{v_c(\eta^{*\prime}h_{in}(\gamma^*))}{\mathrm{E}_{\mathrm{P}_{0n}}[v_c(\eta^{*\prime}h_{in}(\gamma^*))]})] = o_p(1)$ *if* $\hat{\beta}\xrightarrow{p}\beta_0$, $\hat{\gamma}\xrightarrow{p}\gamma^*$ *and* $\hat{\eta}\xrightarrow{p}\eta^*$;
6. $\frac{1}{n}\sum_{i=1}^{n}v_b\left(\frac{v_c\left(\hat{\eta}'h_{in}(\hat{\gamma})\right)}{\sum_{i=1}^{n}v_c(\hat{\eta}'h_{in}(\hat{\gamma}))/n}\right))\hat{g}_{in} - A_{0,v} = o_p(1)$, *if* $\hat{\beta}\xrightarrow{p}\beta_0$, $\hat{\gamma}\xrightarrow{p}\gamma^*$ *and* $\hat{\eta}\xrightarrow{p}\eta^*$;

Proof: We prove results (1), (5) and (6) as the proofs of the remaining results are similar. Proof of 1: Using a proof similar to that Lemma SM1.2 we have $\sup_{\beta\in\mathcal{B}}\|\sum_{i=1}^{n}g_{i,n}(\beta)g_{i,n}(\beta)'/n - \mathrm{E}_{\mathrm{P}_{0n}}[g_{i,n}(\beta)g_{i,n}(\beta)']\| = o_p(1)$ which relies on Assumptions 4.6 (c), (d) and (e) and CS. Now we use the fact that $\mathrm{E}_{\mathrm{P}_{0n}}[g_{i,n}(\beta_0)g_{i,n}(\beta_0)'] \to \Omega_{0,g}$ by Assumption 4.6 (f).

Concerning (5) and (6), write $\hat{a}_n = \sum_{i=1}^{n}v_c\left(\hat{\eta}'h_{in}\left(\hat{\gamma}\right)\right)/n$ and $a_n = \mathrm{E}_{\mathrm{P}_{0n}}[v_c(\eta^{*\prime}h_{in}\left(\gamma^*\right))]$, we know by Assumption 4.8 (c) that $a_n\in\mathcal{A}$, $n\geq 1$. Let $\psi = (\beta',\gamma',\eta',a')'$. Using a proof similar to that of Lemma SM1.2 we have $\sup_{\psi\in\mathcal{B}\times\mathcal{G}\times\mathcal{H}\times\mathcal{A}}||\sum_{i=1}^{n}v_b\left(v_c\left(\eta'h_{in}\left(\gamma\right)\right)/a\right))g\left(z,\beta\right)/n - \mathrm{E}_{\mathrm{P}_{0n}}[v_b\left(v_c\left(\eta'h_{in}\left(\gamma\right)\right)/a\right))g_{in}\left(\beta\right)]|| = o_p(1)$ using Assumptions 4.6 (c), (d) , (e), 4.7 (a), 4.8 (a), (b) and (c) and CS and consequently by result (4) we have $\sum_{i=1}^{n}v_b\left(v_c\left(\hat{\eta}'h_{in}\left(\hat{\gamma}\right)\right)/\hat{a}_n\right)\hat{g}_{in}/n - \mathrm{E}_{\mathrm{P}_{0n}}[v_b(v_c(\eta^{*\prime}h_{in}\left(\gamma^*\right))/a_n)g_{in}\left(\beta_0\right)] = o_p(1)$ which proves (5). The conclusion (6) follows from the fact that $\mathrm{E}_{\mathrm{P}_{0n}}[v_b(v_c(\eta^{*\prime}h_{in}\left(\gamma^*\right))/a_n)g_{in}\left(\beta_0\right)] \to A_{0,v}$ by Assumption 4.8 (d). ■

Lemma SM1.4 *If Assumption 4.6 is satisfied, then* $\sqrt{n}\hat{g}(\beta_{0,n}) \xrightarrow{D} \mathcal{N}(\delta_g, \Omega_{0,g})$.

Proof: First we use the Cramér Wold device to show that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}[g_{in}(\beta_{0,n}) - \mathrm{E}_{\mathrm{P}_{0n}}(g_{in}(\beta_{0,n}))] \xrightarrow{D} \mathcal{N}(0, \Omega_{0,g}).$$

That is, we show that for a fixed $\lambda \neq 0$

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\lambda'[g_{in}(\beta_{0,n}) - \mathrm{E}_{\mathrm{P}_{0n}}(g_{in}(\beta_{0,n}))]}{\sqrt{\lambda' B_n \lambda}} \xrightarrow{D} \mathcal{N}(0,1), \tag{3}$$

where $B_n = \mathrm{E}_{\mathrm{P}_{0n}}([g_{in}(\beta_{0,n}) - \mathrm{E}_{\mathrm{P}_{0n}}(g_{in}(\beta_{0,n}))][g_{in}(\beta_{0,n}) - \mathrm{E}_{\mathrm{P}_{0n}}(g_{in}(\beta_{0,n}))]') \to \Omega_{0,g}$. Note first that

$$B_n = \mathrm{E}_{\mathrm{P}_{0n}}[g_{in}(\beta_{0,n})g_{in}(\beta_{0,n})'] - n\mathrm{E}_{\mathrm{P}_{0n}}(g_{in}(\beta_{0,n}))\mathrm{E}_{\mathrm{P}_{0n}}(g_{in}(\beta_{0,n}))'/n.$$

Now $n\mathrm{E}_{\mathrm{P}_{0n}}(g_{in}(\beta_{0,n}))\mathrm{E}_{\mathrm{P}_{0n}}(g_{in}(\beta_{0,n}))' = \delta_g\delta_g'$. Additionally, $\|\mathrm{E}_{\mathrm{P}_{0n}}[g_{in}(\beta_{0,n})g_{in}(\beta_{0,n})'] - \mathrm{E}_{\mathrm{P}_{0n}}[g_{in}(\beta_0)g_{in}(\beta_0)']\| \to 0$ as $\beta_{0,n} \to \beta_0$ because

$$\mathrm{E}_{\mathrm{P}_{0n}}[\sup_{\beta_a,\beta_b\in\mathcal{B}:\|\beta_a-\beta_b\|\leq d}\|g_{in}(\beta_a)g_{in}(\beta_a)' - g_{in}(\beta_b)g_{in}(\beta_b)'\|] \leq Cd,$$

by a mean-value expansion (which holds by Assumption 4.6 (d)), Assumption 4.6 (e) and CS. It follows from Assumption 4.6 (f) that $B_n \to \Omega_{0,g}$. Also note that B_n is positive definite for n large enough because $\Omega_{0,g}$ is positive definite. Now for $a = 2+\delta$ we have by CR, L, CS and Assumption 4.6 (e)

$$\begin{aligned}\mathrm{E}_{\mathrm{P}_{0n}}\left[|\lambda'[g_{in}(\beta_{0,n}) - \mathrm{E}_{\mathrm{P}_{0n}}(g_{in}(\beta_{0,n}))]|^a\right] &\leq 2^{a-1}\left[\mathrm{E}_{\mathrm{P}_{0n}}\|\lambda' g_{in}(\beta_{0,n})\|^a + |\mathrm{E}_{\mathrm{P}_{0n}}(\lambda' g_{in}(\beta_{0,n}))|^a\right] \\ &\leq 2^a\left[\|\lambda\|^a \mathrm{E}_{\mathrm{P}_{0n}}[\|g_{in}(\beta_{0,n})\|^a\right] < 2^a\|\lambda\|^a C.\end{aligned}$$

Therefore

$$\frac{1}{n^{a/2}}\sum_{i=1}^{n}\mathrm{E}_{\mathrm{P}_{0n}}\left[|\lambda'[g_{in}(\beta_{0,n}) - \mathrm{E}_{\mathrm{P}_{0n}}(g_{in}(\beta_{0,n}))]|^a\right] \leq \frac{2^a\|\lambda\|^a C}{n^{a/2-1}} \to 0.$$

Hence by the Lyapunov CLT (Serfling, 1980, p.31-32, Corollary) it follows that (3) holds. Now note that

$$\sqrt{n}\hat{g}(\beta_{0,n}) = \sqrt{n}\left[\hat{g}(\beta_{0,n}) - \mathrm{E}_{\mathrm{P}_{0n}}(\hat{g}(\beta_{0,n}))\right] + \sqrt{n}\mathrm{E}_{\mathrm{P}_{0n}}(\hat{g}(\beta_{0,n}))$$

and the first term converges to $\mathcal{N}(0, \Omega_{0,g})$ while the second is equal to δ_g which proves the result. ■.

Lemmata SM1.5 to SM1.7 correspond to versions of Lemmata A1 to A3 of NS for iid triangular arrays and the proofs are similar to those Lemmata given in NS (with β_0 replaced by $\beta_{0,n}$ in those proofs) and therefore are omitted.

Lemma SM1.5 *If Assumption 4.6 is satisfied, then for any* $1/\alpha < \zeta < 1/2$ *and* $\Lambda_n = \{\lambda : \|\lambda\| \leq n^{-\zeta}\}$ *and with wpa*1 $\Lambda_n \subseteq \hat{\Lambda}_n(\beta)$ *for all* $\beta \in \mathcal{B}$.

Lemma SM1.6 *If Assumption 4.6 is satisfied,* $\bar{\beta} \in \mathcal{B}$, $\bar{\beta} - \beta_{0,n} \xrightarrow{p} 0$ *and* $\hat{g}(\bar{\beta}) = O_p(n^{-1/2})$, *then* $\bar{\lambda} = \arg\max_{\lambda\in\hat{\Lambda}_n(\bar{\beta})}\hat{P}_n^g(\bar{\beta},\lambda)$ *exists wpa*1, *and* $\bar{\lambda} = O_p(n^{-1/2})$, $\sup_{\lambda\in\hat{\Lambda}_n(\bar{\beta})}\hat{P}_n^g(\bar{\beta},\lambda) \leq O_p(n^{-1})$.

Lemma SM1.7 *If Assumption 4.6 is satisfied, then* $\left\|\hat{g}(\hat{\beta})\right\| = O_p(n^{-1/2})$.

The proof of the following Lemma follows the same steps of the proof of Theorem 3.1 of NS, but since there are some small differences we present it below.

Lemma SM1.8 *If Assumption 4.6 is satisfied, then* $\hat{\beta} \xrightarrow{p} \beta_0$, $\hat{\beta} - \beta_{0,n} \xrightarrow{p} 0$, $g(\hat{\beta}) = O_p(n^{-1/2})$, $\hat{\lambda} = \arg\max_{\lambda \in \hat{\Lambda}_n(\hat{\beta})} \sum_{i=1}^n \rho(\lambda' g_{in}(\hat{\beta}))/n$ *exists wpa*1, *and* $\hat{\lambda} = O_p(n^{-1/2})$.

Proof: Let $g_n(\beta) = \mathrm{E}_{\mathrm{P}_{0n}}[g(z_{in}, \beta)]$. By Lemma SM1.7 $\hat{g}(\hat{\beta}) = o_p(1)$, and by Lemma SM1.2 $\sup_{\beta \in \mathcal{B}} \|\hat{g}(\beta) - g_n(\beta)]\| \xrightarrow{p} 0$ and $\{g_n(\beta)\}_{n=1}^{\infty}$ is uniformly equicontinuous. Additionally, since $\lim_{n\to\infty} g_n(\beta) = g(\beta)$ for each $\beta \in \mathcal{B}$, we have $\sup_{\beta\in\mathcal{B}} \|g_n(\beta) - g(\beta)\| \to 0$ (see Rudin, 1976, Exercise 16, p.168). Hence by T $g(\hat{\beta}) \xrightarrow{p} 0$. Since $g(\beta) = 0$ has a unique zero at β_0, $g(\beta)$ must be bounded away from zero outside any neighborhood of β_0. Therefore, $\hat{\beta}$ must be inside any neighborhood of β_0 wpa1, i.e. $\hat{\beta} \xrightarrow{p} \beta_0$, giving the first conclusion. The second conclusion follows from the inequality $\left\|\hat{\beta} - \beta_{0,n}\right\| \leq \left\|\hat{\beta} - \beta_0\right\| + \|\beta_0 - \beta_{0,n}\|$, the first conclusion and the fact that $\|\beta_0 - \beta_{0,n}\| \to 0$ by Assumption 4.6 (a). The third conclusion is due to Lemma SM1.7. Also, note that by the second and third conclusions the hypotheses of Lemma SM1.6 are satisfied for $\bar{\beta} = \hat{\beta}$, so that the last conclusion follows from Lemma SM1.6. ■

Lemma SM1.9 *If Assumption 4.6 is satisfied, then*

$$\sqrt{n}\begin{pmatrix} \hat{\beta} - \beta_{0,n} \\ \hat{\lambda} \end{pmatrix} \xrightarrow{d} \mathcal{N}(\begin{pmatrix} -H_{0,g}\delta_g \\ -P_{0,g}\delta_g \end{pmatrix}, \begin{pmatrix} \Sigma_{0,g} & 0 \\ 0 & P_{0,g} \end{pmatrix}).$$

Proof: Let $\hat{\theta} = (\hat{\beta}', \hat{\lambda}')'$ and $\theta_{0,n} = \left(\beta_{0,n}', 0'\right)'$. Note that since $\beta_0 \in int\,(\mathcal{B})$ and $\beta_{0,n} \to \beta_0$, then $\beta_{0,n} \in int\,(\mathcal{B})$ for n large enough. Using arguments similar to those of NS in the proof of their Theorem 3.2 (which in our case are based on a first order Taylor expansion of the first order conditions of the GEL objective function around $\theta_{0,n}$ and require the fact that $\beta_{0,n} \to \beta_0$, Lemma SM1.8 and the Lemma SM1.3) we have

$$\sqrt{n}(\hat{\theta} - \theta_{0,n}) = -(H_{0,g}', -P_{0,g})\sqrt{n}\hat{g}(\beta_{0,n}) + o_p(1). \tag{4}$$

Now apply the CLT given by Lemma SM1.4. ■

The following Lemma generalizes the results of Lemma SM1.1 for triangular arrays and its proof is similar and therefore omitted, though available upon request.

Lemma SM1.10 *Let Assumptions 4.6, 4.7 and 4.8 hold. Then* $n\hat{p}_i^{v_c} = 1 + o_p(1)$ *and*

$$n^{1/2}\left(\hat{p}_i^{v_c} - \frac{1}{n}\right) = \kappa_{v_c}\frac{1}{n}\hat{g}_{in}' n^{1/2}\hat{\lambda}(1 + o_p(1)) + O_p(n^{-3/2}),$$

uniformly $(i = 1, ..., n)$ *where* $\kappa_{v_c} = v_{c,1}(0)/v_c(0)$.

Proof of Theorem 4.3: Using the similar arguments to those used in the proof of Theorem 4.1, which require SM1.3 rather than Lemma 4.3 of Newey and McFadden (1994) and Lemma SM1.10 we have $\mathcal{S}_v = A_{0,v}\sqrt{n}\hat{\lambda}(1+o_p(1))+o_p(1)$. Now, given that by Lemma SM1.9 $\sqrt{n}\hat{\lambda} \xrightarrow{d} \mathcal{N}(-P_{0,g}\delta_g, P_{0,g})$, it follows that $\mathcal{N}\left(-A_{0,v}P_{0,g}\delta_g, \sigma_0^2\right)$.

The demonstration of the asymptotic equivalence of the statistics $\mathcal{S}_v$, $\tilde{\mathcal{S}}_v$, $\mathcal{LM}_v$ and $\mathcal{J}_v$ is similar to the proof of asymptotic equivalence of these statistics given in the proof of Theorem 4.2. ■

SM2 Proofs of the results of section 5 and discussion

In this section we provide the proofs of the theorems presented in section 5 of the paper and investigate the behavior of the random variable $W^*_{i,j}(\delta_h)$ $i=1,2$, $j=3,4$ (defined in subsection 5.1 of the paper) as some elements of δ_h approach infinity.

SM2.1 Proofs of the results of section 5

We start by compiling a number of Lemmata without presenting their proofs either because the proofs are very similar to those given in NS or to those provided in the previous sections. Let $g_{in}(\beta)=g(z_{i,n},\beta)$, $\hat{g}(\beta)=\sum_{i=1}^n g_{i,n}(\beta)/n$, $\hat{g}_{in}=g_{in}(\hat{\beta})$ and $h_{in}(\gamma)=h(z_{i,n},\gamma)$, $\hat{h}_{in}\equiv h(z_{i,n},\hat{\gamma})$, $\hat{h}(\gamma)=\sum_{i=1}^n h_{in}(\gamma)/n$, $s_{in}(\varphi)=s(z_{i,n},\varphi)$ and $\hat{s}(\varphi)=\sum_{i=1}^n s_{in}(\varphi)/n$.

The proofs of Lemmata SM2.1 to SM2.3 are similar to the proofs of Lemmata SM1.2 to SM1.4 above.

Lemma SM2.1 *Suppose Assumption 5.1 holds. Under a sequence $\{\mathrm{P}_n\}_{n=1}^{\infty}\in\mathcal{P}$ the following results hold:*

1. $\sup_{\beta\in\mathcal{B}}\|\hat{g}(\beta)-\mathrm{E}_{\mathrm{P}_n}[g_{in}(\beta)]\|=o_p(1)$, *and* $\{\mathrm{E}_{\mathrm{P}_n}[g_{in}(\beta)]\}_{n=1}^{\infty}$ *is uniformly equicontinuous in* $\beta\in\mathcal{B}$.
2. $\sup_{\gamma\in\mathcal{G}}\left\|\hat{h}(\gamma)-\mathrm{E}_{\mathrm{P}_n}[h_{in}(\gamma)]\right\|=o_p(1)$, *and* $\{\mathrm{E}_{\mathrm{P}_n}[h_{in}(\gamma)]\}_{n=1}^{\infty}$ *is uniformly equicontinuous in* $\gamma\in\mathcal{G}$.

Lemma SM2.2 *Suppose Assumption 5.1 holds. Under a sequence $\{\mathrm{P}_n\}_{n=1}^{\infty}\in\mathrm{Seq}(\beta_0,\gamma^*,\eta^*,\delta_h,\Omega,D_g,D_h,A_v)$ we have:*

1. $\frac{1}{n}\sum_{i=1}^n s_{in}(\hat{\varphi})s_{in}(\hat{\varphi})'-\Omega=o_p(1)$ *if* $\hat{\varphi}\xrightarrow{p}\varphi^*$;
2. $\frac{1}{n}\sum_{i=1}^n \|s_{in}(\hat{\varphi})\|^2-\mathrm{E}_{\mathrm{P}_n}[\|s_{in}(\varphi^*)\|^2]=o_p(1)$ *if* $\hat{\varphi}\xrightarrow{p}\varphi^*$;
3. $\frac{1}{n}\sum_{i=1}^n \frac{\partial g_{in}(\hat{\beta})}{\partial\beta'}-D_g=o_p(1)$ *if* $\hat{\beta}\xrightarrow{p}\beta_0$;
4. $\frac{1}{n}\sum_{i=1}^n \frac{\partial h_{in}(\hat{\gamma})}{\partial\gamma'}-D_h=o_p(1)$ *if* $\hat{\gamma}\xrightarrow{p}\gamma^*$;
5. $\frac{1}{n}\sum_{i=1}^n v_c(\hat{\eta}'h_{in}(\hat{\gamma}))-\mathrm{E}_{\mathrm{P}_n}[v_c(\eta^{*\prime}h_{in}(\gamma^*))]=o_p(1)$ *if* $\hat{\beta}\xrightarrow{p}\beta_0$, $\hat{\gamma}\xrightarrow{p}\gamma^*$ *and* $\hat{\eta}\xrightarrow{p}\eta^*$;

6. $\frac{1}{n}\sum_{i=1}^{n} v_b\left(\frac{v_c(\hat{\eta}' h_{in}(\hat{\gamma}))}{\sum_{i=1}^{n} v_c(\hat{\eta}' h_{in}(\hat{\gamma}))/n}\right) - \mathrm{E}_{\mathrm{P}_n}[v_b(\frac{v_c(\eta^{*\prime} h_{in}(\gamma^*))}{\mathrm{E}_{\mathrm{P}_n}[v_c(\eta^{*\prime} h_{in}(\gamma^*))]})] = o_p(1)$ *if* $\hat{\beta} \xrightarrow{p} \beta_0$, $\hat{\gamma} \xrightarrow{p} \gamma^*$ *and* $\hat{\eta} \xrightarrow{p} \eta^*$;

7. $\frac{1}{n}\sum_{i=1}^{n} v_b\left(\frac{v_c(\hat{\eta}' h_{in}(\hat{\gamma}))}{\sum_{i=1}^{n} v_c(\hat{\eta}' h_{in}(\hat{\gamma}))/n}\right))\hat{g}_{in} - A_v = o_p(1)$, *if* $\hat{\beta} \xrightarrow{p} \beta_0$, $\hat{\gamma} \xrightarrow{p} \gamma^*$ *and* $\hat{\eta} \xrightarrow{p} \eta^*$.

Lemma SM2.3 *Suppose Assumption 5.1 holds. Under a sequence* $\{\mathrm{P}_n\}_{n=1}^{\infty} \in \mathrm{Seq}(\beta_0, \gamma^*, \eta^*, \delta_h, \Omega, D_g, D_h, A_v)$ *with* $\|\delta_h\|_{\infty} < \infty$ *and satisfying* $\sqrt{n}\mathrm{E}_{\mathrm{P}_n}(g_{in}(\beta_{0,\mathrm{P}_n})) \to \delta_g$ *with* $\|\delta_g\|_{\infty} < \infty$, *we have* $\sqrt{n}\hat{s}(\varphi_n^*) \xrightarrow{D} \mathcal{N}(\delta, \Omega)$, *where* $\varphi_n^* = (\beta_{0,\mathrm{P}_n}', \gamma_{\mathrm{P}_n}^{*\prime})'$ *and* $\delta = (\delta_g', \delta_h')'$.

The proofs of Lemmata SM2.4 to SM2.6 are similar to the proofs of Lemma A1 to A3 of NS. The proofs of these Lemmata of NS required a LLN, a UWL and a CLT. In our framework these are replaced by the LLN for triangular arrays in Davidson (1994, 19.9 Corollary p.301), Lemmata SM2.2 and SM2.3 respectively. The proof of part (2) of Lemma SM2.5 is similar to that of Lemma A2 of NS, but uses the assumptions that $H_n \subset \mathcal{H}$, $n \geq 1$ and $\mathcal{H}$ is a convex set.

Lemma SM2.4 *Suppose Assumption 5.1 holds. Under a sequence* $\{\mathrm{P}_n\}_{n=1}^{\infty} \in \mathcal{P}$ *for any* $1/(2+\delta) < \zeta < 1/2$, *the following results hold:*

1. $\sup_{\beta \in \mathcal{B}, \lambda \in \Lambda_n, 1 \leq i \leq n} |\lambda' g_{in}(\beta)| \xrightarrow{p} 0$, *where* $\Lambda_n = \{\lambda : \|\lambda\| \leq n^{-\zeta}\}$ *and wpa*1 $\Lambda_n \subseteq \hat{\Lambda}_n(\beta)$ *for all* $\beta \in \mathcal{B}$.

2. $\sup_{\gamma \in \mathcal{G}, \eta \in H_n, 1 \leq i \leq n} |\eta' h_{in}(\gamma)| \xrightarrow{p} 0$, *where* $H_n = \{\eta : \|\eta\| \leq n^{-\zeta}\}$.

Lemma SM2.5 *Suppose Assumption 5.1 holds. Under a sequence* $\{\mathrm{P}_n\}_{n=1}^{\infty} \in \mathrm{Seq}(\beta_0, \gamma^*, \eta^*, \delta_h, \Omega, D_g, D_h, A_v)$:

1. *if* $\sqrt{n}\mathrm{E}_{\mathrm{P}_n}(g_{in}(\beta_{0,\mathrm{P}_n})) \to \delta_g$, *with* $\|\delta_g\|_{\infty} < +\infty$, $\bar{\beta} \in \mathcal{B}$, $\bar{\beta} - \beta_{0,\mathrm{P}_n} \xrightarrow{p} 0$ *and* $\hat{g}(\bar{\beta}) = O_p(n^{-1/2})$, *then* $\bar{\lambda} = \arg\max_{\lambda \in \hat{\Lambda}_n(\bar{\beta})} \hat{P}_g(\bar{\beta}, \lambda)$ *exists wpa*1, $\bar{\lambda} = O_p(n^{-1/2})$ *and* $\sup_{\lambda \in \hat{\Lambda}_n(\bar{\beta})} \hat{P}_g(\bar{\beta}, \lambda) \leq \rho_0 + O_p(1/n)$.

2. *if* $\|\delta_h\|_{\infty} < +\infty$, $\bar{\gamma} \in \mathcal{G}$, $\bar{\gamma} - \gamma_{\mathrm{P}_n}^* \xrightarrow{p} 0$ *and* $\hat{h}(\bar{\gamma}) = O_p(n^{-1/2})$, *then* $\bar{\eta} = \arg\max_{\eta \in \mathcal{H}} \hat{P}_h(\bar{\gamma}, \eta)$ *exists wpa*1, $\bar{\eta} = O_p(n^{-1/2})$ *and* $\sup_{\eta \in \mathcal{H}} \hat{P}_h(\bar{\gamma}, \eta) \leq \rho_0 + O_p(1/n)$.

Lemma SM2.6 *Suppose Assumption 5.1 holds. Under a sequence* $\{\mathrm{P}_n\}_{n=1}^{\infty} \in \mathrm{Seq}(\beta_0, \gamma^*, \eta^*, \delta_h, \Omega, D_g, D_h, A_v)$:

1. *if* $\sqrt{n}\mathrm{E}_{\mathrm{P}_n}(g_{in}(\beta_{0,\mathrm{P}_n})) \to \delta_g$, *with* $\|\delta_g\|_{\infty} < +\infty$, *then* $\left\|\hat{g}(\hat{\beta})\right\| = O_p(n^{-1/2})$.

2. *if* $\|\delta_h\|_{\infty} < +\infty$, *then* $\left\|\hat{h}(\hat{\gamma})\right\| = O_p(n^{-1/2})$.

The proof of the following Lemma is similar to that of Lemma SM1.8.

Lemma SM2.7 *Suppose Assumption 5.1 holds. Under a sequence* $\{\mathrm{P}_n\}_{n=1}^{\infty} \in \mathrm{Seq}(\beta_0, \gamma^*, \eta^*, \delta_h, \Omega, D_g, D_h, A_v)$:

1. *if* $\sqrt{n}\mathrm{E}_{\mathrm{P}_n}(g_{in}(\beta_{0,\mathrm{P}_n})) \to \delta_g$ *with* $\|\delta_g\|_{\infty} < +\infty$, *then* $\hat{\beta} \xrightarrow{p} \beta_0$, $\hat{\beta} - \beta_{0,\mathrm{P}_n} \xrightarrow{p} 0$, $\hat{\lambda} = \arg\max_{\lambda \in \hat{\Lambda}_n(\hat{\beta})} \sum_{i=1}^{n} \rho(\lambda' g_{in}(\hat{\beta}))/n$ *exists wpa*1, *and* $\hat{\lambda} = O_p(n^{-1/2})$.

2. *if* $\|\delta_h\|_\infty < +\infty$, *then* $\hat{\gamma} \xrightarrow{p} \gamma^*$, $\hat{\gamma} - \gamma^*_{\mathrm{P}_n} \xrightarrow{p} 0$, $\hat{\eta} = \arg\max_{\eta\in\mathcal{H}} \sum_{i=1}^n \rho(\eta' h_{in}(\hat{\gamma}))/n$ *exists wpa*1, *and* $\hat{\eta} = O_p(n^{-1/2})$.

The proof of the following Lemma is similar to that of Lemma SM1.9.

Lemma SM2.8 *Suppose Assumption 5.1 holds. Under a sequence* $\{\mathrm{P}_n\}_{n=1}^\infty \in \mathrm{Seq}\,(\beta_0, \gamma^*, \eta^*, \delta_h, \Omega, D_g, D_h, A_v)$*:*

1. *if* $\sqrt{n}\mathrm{E}_{\mathrm{P}_n}(g_{in}(\beta_{0,\mathrm{P}_n})) \to \delta_g$ *with* $\|\delta_g\|_\infty < +\infty$, *then*

$$\sqrt{n}\begin{pmatrix} \hat{\beta} - \beta_{0,\mathrm{P}_n} \\ \hat{\lambda} \end{pmatrix} \xrightarrow{d} \mathcal{N}(\begin{pmatrix} -H_{0,g}\delta_g \\ -P_{0,g}\delta_g \end{pmatrix}, \begin{pmatrix} \Sigma_{0,g} & 0 \\ 0 & P_{0,g} \end{pmatrix}).$$

2. *if* $\|\delta_h\|_\infty < +\infty$, *then*

$$\sqrt{n}\begin{pmatrix} \hat{\gamma} - \gamma^*_{\mathrm{P}_n} \\ \hat{\eta} \end{pmatrix} \xrightarrow{d} \mathcal{N}(\begin{pmatrix} -H_{0,h}\delta_h \\ -P_{0,h}\delta_h \end{pmatrix}, \begin{pmatrix} \Sigma_{0,h} & 0 \\ 0 & P_{0,h} \end{pmatrix}).$$

The proof of the following Lemma is similar to that of Lemma SM1.1.

Lemma SM2.9 *Suppose Assumption 5.1 holds. Under a sequence* $\{\mathrm{P}_n\}_{n=1}^\infty \in \mathrm{Seq}\,(\beta_0, \gamma^*, \eta^*, \delta_h, \Omega, D_g, D_h, A_v)$*:*

1. *if* $\sqrt{n}\mathrm{E}_{\mathrm{P}_n}(g_{in}(\beta_{0,\mathrm{P}_n})) \to \delta_g$ *with* $\|\delta_g\|_\infty < +\infty$, *we have* $n\hat{p}_i^{v_c} = 1 + o_p(1)$ *and*

$$n^{1/2}\left(\hat{p}_i^{v_c} - \frac{1}{n}\right) = \kappa_{v_c}\frac{1}{n}\hat{g}'_{in}n^{1/2}\hat{\lambda}(1 + o_p(1)) + O_p(n^{-3/2}),$$

uniformly $(i = 1, ..., n)$ *where* $\kappa_{v_c} = v_{c,1}(0)/v_c(0)$.

2. *if* $\|\delta_h\|_\infty < +\infty$, *we have* $n\hat{q}_i^{v_c} = 1 + o_p(1)$ *and*

$$n^{1/2}\left(\hat{q}_i^{v_c} - \frac{1}{n}\right) = \kappa_{v_c}\frac{1}{n}\hat{h}'_{in}n^{1/2}\hat{\eta}(1 + o_p(1)) + O_p(n^{-3/2}),$$

uniformly $(i = 1, ..., n)$ *where* $\kappa_{v_c} = v_{c,1}(0)/v_c(0)$.

The proofs of the following two Lemmata are similar to those of Theorem 4.3 and Theorem 4.2.

Lemma SM2.10 *Suppose Assumption 5.1 holds. Under a sequence* $\{\mathrm{P}_n\}_{n=1}^\infty \in \mathrm{Seq}\,(\beta_0, \gamma^*, \eta^*, \delta_h, \Omega, D_g, D_h, A_v)$ *that satisfies* $\sqrt{n}\mathrm{E}_{\mathrm{P}_n}(g_{in}(\beta_{0,\mathrm{P}_n})) \to \delta_g$ *with* $\|\delta_g\|_\infty < +\infty$ *and Assumption 5.2 and if* $\|\delta_h\| = +\infty$, *then* $\mathcal{S}_v$ *converges in distribution to* $\mathcal{N}\left(-A_v P_g \delta_g, \sigma^2\right)$. *Furthermore,* $\tilde{\mathcal{S}}_v$, $\mathcal{LM}_v$ *and* $\mathcal{J}_v$ *are asymptotically equivalent to* $\mathcal{S}_v$.

Lemma SM2.11 *Suppose Assumption 5.1 holds. Under a sequence* $\{\mathrm{P}_n\}_{n=1}^\infty \in \mathrm{Seq}\,(\beta_0, \gamma^*, \eta^*, \delta_h, \Omega, D_g, D_h, A_v)$ *satisfying* $\sqrt{n}\mathrm{E}_{\mathrm{P}_n}(g_{in}(\beta_{0,\mathrm{P}_n})) \to \delta_g$ *and Assumption 5.2 and if* $\|\delta_h\| = +\infty$, $\hat{\sigma}_j^2 \xrightarrow{p} \sigma^2$, $j = 1, ...4$.

Lemma SM2.12 *Suppose Assumptions 5.1 holds. Under a sequence* $\{\mathrm{P}_n\}_{n=1}^\infty \in \mathrm{Seq}\,(\beta_0, \gamma^*, \eta^*, \delta_h, \Omega, D_g, D_h, A_v)$ *with* $\|\delta_h\|_\infty < \infty$ *and satisfying* $\sqrt{n}\mathrm{E}_{\mathrm{P}_n}(g_{in}(\beta_{0,\mathrm{P}_n})) \to \delta_g$ *with* $\|\delta_g\|_\infty < +\infty$ *and let* $\delta = (\delta'_g, \delta'_h)'$, *then*

1. $\sqrt{n}\mathcal{S}_v \xrightarrow{d} \mathcal{T}(\delta,\Omega,Q_1)$ *and* $\sqrt{n}\mathcal{J}_{v,2}$, $\sqrt{n}\mathcal{J}_{v,3}$, $\sqrt{n}\mathcal{LM}_{v,2}$ *and* $\sqrt{n}\mathcal{LM}_{v,3}$ *are asymptotically equivalent to* $\sqrt{n}\mathcal{S}_v$*;*
2. $\sqrt{n}\tilde{\mathcal{S}}_v \xrightarrow{d} \mathcal{T}(\delta,\Omega,Q_2)$ *and* $\sqrt{n}\mathcal{J}_{v,1}$, $\sqrt{n}\mathcal{J}_{v,4}$, $\sqrt{n}\mathcal{LM}_{v,1}$ *and* $\sqrt{n}\mathcal{LM}_{v,4}$ *are asymptotically equivalent to* $\sqrt{n}\tilde{\mathcal{S}}_v$*;*
3. $n\hat{\sigma}_1^2 \xrightarrow{d} \mathcal{T}(\delta,\Omega,Q_3)$ *and* $n\hat{\sigma}_4^2$ *is asymptotically equivalent to* $n\hat{\sigma}_1^2$*;*
4. $n\hat{\sigma}_2^2 \xrightarrow{d} \mathcal{T}(\delta,\Omega,Q_4)$ *and* $n\hat{\sigma}_3^2$ *is asymptotically equivalent to* $n\hat{\sigma}_2^2$*;*
5. $n\hat{\sigma}_j^2$ *is non-negative with probability approaching one for* $j = 1,2,3,4$*.*

Proofs:

Proof of 1: $\sqrt{n}\mathcal{S}_v$ is considered first. By a second order Taylor expansion,

$$v_a(\hat{p}_i n) = v_a(1) + (\hat{p}_i n - 1) + v_{a,2}(\hat{\sigma}_i)(\hat{p}_i n - 1)^2/2, \tag{5}$$

where $\hat{\sigma}_i = \alpha_i + (1-\alpha_i)\hat{p}_i n$ and $\alpha_i \in (0,1)$. Hence,

$$\begin{aligned}\sqrt{n}\mathcal{S}_v = \sqrt{n}\sum_{i=1}^n \sqrt{n}\left(\hat{p}_i - \frac{1}{n}\right) n\hat{p}_i \left[v_b\left(\frac{\hat{q}_i^{v_c}}{\hat{p}_i^{v_c}}\right) - \sum_{\ell=1}^n v_b\left(\frac{\hat{q}_\ell^{v_c}}{\hat{p}_\ell^{v_c}}\right)\hat{p}_\ell\right] \\ + n\sum_{i=1}^n \frac{v_{a,2}(\hat{\sigma}_i)}{2}(\hat{p}_i n - 1)^2 \hat{p}_i \left[v_b\left(\frac{\hat{q}_i^{v_c}}{\hat{p}_i^{v_c}}\right) - \sum_{\ell=1}^n v_b\left(\frac{\hat{q}_\ell^{v_c}}{\hat{p}_\ell^{v_c}}\right)\hat{p}_\ell\right] = B_{1,n} + B_{2,n}.\end{aligned}$$

Using Lemma SM2.9 $n^{1/2}(\hat{p}_i - 1/n) = n^{-1}\hat{g}_{in}'\sqrt{n}\hat{\lambda}(1+o_p(1)) + O_p(n^{-3/2})$ and $\sum_{i=1}^n \hat{p}_i\hat{g}_{in}' = 0$ we have $B_{1,n} = \sqrt{n}\hat{A}_{v,3}\sqrt{n}\hat{\lambda}(1+o_p(1))] + B_{1r,n}$, where $B_{1r,n}$ is defined below.

Note that by a first-order Taylor expansion

$$v_b\left(\frac{n\hat{q}_i^{v_c}}{n\hat{p}_i^{v_c}}\right) = v_b(1) + v_{b,1}\left(\frac{\hat{\sigma}_{1,i}}{\hat{\sigma}_{2,i}}\right)\frac{(n\hat{q}_i^{v_c} - 1)}{\hat{\sigma}_{2,i}} - \frac{\hat{\sigma}_{1,i}}{\hat{\sigma}_{2,i}^2}v_{b,1}\left(\frac{\hat{\sigma}_{1,i}}{\hat{\sigma}_{2,i}}\right)(n\hat{p}_i^{v_c} - 1),$$

where $\hat{\sigma}_{1,i} = \alpha_{1,i} + (1-\alpha_{1,i})n\hat{q}_i^{v_c}$ and $\alpha_{1,i} \in (0,1)$ and $\hat{\sigma}_{2,i} = \alpha_{2,i} + (1-\alpha_{2,i})n\hat{p}_i^{v_c}$ and $\alpha_{2,2} \in (0,1)$. Thus,

$$\begin{aligned}\sqrt{n}\hat{A}_{v,3} &= \sum\nolimits_{i=1}^n n\hat{p}_i\hat{g}_{in}' v_{b,1}\left(\frac{\hat{\sigma}_{1,i}}{\hat{\sigma}_{2,i}}\right)\sqrt{n}\left(\hat{q}_i^{v_c} - \frac{1}{n}\right)\frac{1}{\hat{\sigma}_{2,i}} \\ &\quad - \sum\nolimits_{i=1}^n n\hat{p}_i\hat{g}_{in}'\frac{\hat{\sigma}_{1,i}}{\hat{\sigma}_{2,i}^2}v_{b,1}\left(\frac{\hat{\sigma}_{1,i}}{\hat{\sigma}_{2,i}}\right)\sqrt{n}\left(\hat{p}_i^{v_c} - \frac{1}{n}\right) \\ &= W_{1,n} - W_{2,n}.\end{aligned}$$

Now by Lemmata SM2.9 $\sqrt{n}(\hat{p}_i^{v_c} - 1/n) = n^{-1}\hat{g}_{in}'\sqrt{n}\hat{\lambda}(1+o_p(1)) + O_p(n^{-3/2})$ and $\sqrt{n}(\hat{q}_i^{v_c} - 1/n) = n^{-1}\hat{h}_{in}'\sqrt{n}\hat{\eta}(1+o_p(1)) + O_p(n^{-3/2})$ as $\kappa_{v_c} = 1$. Additionally, note that similarly to (4) we have $\sqrt{n}\hat{\lambda} =$

$-P_g\sqrt{n}\hat{g}(\beta_{0,\mathrm{P}_n})+o_p(1)$, and $\sqrt{n}\hat{\eta}=-P_h\sqrt{n}\hat{h}(\gamma^*_{\mathrm{P}_n})+o_p(1)$. Combining these results with the fact that $v_{b,1}=1$ and Lemma SM2.2, we obtain

$$\begin{aligned} W_{1,n} &= \sqrt{n}\hat{\eta}'\frac{1}{n}\sum_{i=1}^{n}\hat{h}_{in}\hat{g}'_{in}v_{b,1}\left(\frac{\hat{\sigma}_{1,i}}{\hat{\sigma}_{2,i}}\right)\frac{1}{\hat{\sigma}_{2,i}}+O_p(n^{-2}) \\ &= -\sqrt{n}\hat{s}(\varphi^*_n)'S'_hP_h\Omega_{hg}+O_p(n^{-2}), \qquad (6) \end{aligned}$$

and

$$\begin{aligned} W_{2,n} &= \sum_{i=1}^{n}n\hat{p}_i\hat{g}'_{in}\frac{\hat{\sigma}_{1,i}}{\hat{\sigma}^2_{2,i}}v_{b,1}\left(\frac{\hat{\sigma}_{1,i}}{\hat{\sigma}_{2,i}}\right)\left[\frac{1}{n}\hat{g}'_{in}\sqrt{n}\hat{\lambda}\left(1+o_p\left(1\right)\right)+O_p(n^{-3/2})\right] \\ &= -\sqrt{n}\hat{s}(\varphi^*_n)'S_gP_g\Omega_g+O_p(n^{-2}), \end{aligned}$$

where $\varphi^*_n=(\beta'_{0,\mathrm{P}_n},\gamma^{*\prime}_{\mathrm{P}_n})'$.

Hence

$$\begin{aligned} \sqrt{n}\hat{A}_{3,v} &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}n\hat{p}_i\hat{g}'_{in}v_b\left(\frac{\hat{q}_i^{v_c}}{\hat{p}_i^{v_c}}\right)=-\sqrt{n}\hat{s}(\varphi^*_n)'S'_hP_h\Omega_{hg} \qquad (7) \\ &\quad +\sqrt{n}\hat{s}(\varphi^*_n)'S'_gP_g\Omega_g+O_p(n^{-2}). \end{aligned}$$

Now note that $B_{1r,n}\equiv n^{-1}\sum_{i=1}^{n}n\hat{p}_i[v_b\left(\hat{q}_i^{v_c}/\hat{p}_i^{v_c}\right)-n^{-1}\sum_{i=1}^{n}v_b\left(\hat{q}_i^{v_c}/\hat{p}_i^{v_c}\right)n\hat{p}_i]O_p(1)=o_p(1)$.

Additionally, by Lemma SM2.9 we have

$$\begin{aligned} B_{2,n} &\equiv n\sum_{i=1}^{n}\frac{v_{a,2}\left(\hat{\sigma}_i\right)}{2}\left[\sqrt{n}\left(\hat{p}_i-\frac{1}{n}\right)\right]^2 n\hat{p}_i\left[v_b\left(\frac{\hat{q}_i^{v_c}}{\hat{p}_i^{v_c}}\right)-\sum_{\ell=1}^{n}v_b\left(\frac{\hat{q}_\ell^{v_c}}{\hat{p}_\ell^{v_c}}\right)\hat{p}_\ell\right] \\ &= n\sum_{i=1}^{n}\frac{v_{a,2}\left(\hat{\sigma}_i\right)}{2}\left[\frac{1}{n}\hat{g}'_{in}\sqrt{n}\hat{\lambda}\left(1+o_p\left(1\right)\right)+O_p(n^{-3/2})\right]^2 n\hat{p}_i\left[v_b\left(\frac{\hat{q}_i^{v_c}}{\hat{p}_i^{v_c}}\right)-\sum_{\ell=1}^{n}v_b\left(\frac{\hat{q}_\ell^{v_c}}{\hat{p}_\ell^{v_c}}\right)\hat{p}_\ell\right]. \end{aligned}$$

By T and the fact that $(a+b)^2\leq 2a^2+2b^2$ we have

$$|B_{2,n}|\leq\frac{2}{n}\sum_{i=1}^{n}[\|\hat{g}_{in}\|^2\left\|\sqrt{n}\hat{\lambda}\right\|^2\left(1+o_p\left(1\right)\right)+O_p(n^{-3})]\left|\frac{v_{a,2}\left(\hat{\sigma}_i\right)}{2}n\hat{p}_i\left[v_b\left(\frac{\hat{q}_i^{v_c}}{\hat{p}_i^{v_c}}\right)-\sum_{\ell=1}^{n}v_b\left(\frac{\hat{q}_\ell^{v_c}}{\hat{p}_\ell^{v_c}}\right)\hat{p}_\ell\right]\right|.$$

Since $\sum_{i=1}^{n}v_b\left(\hat{q}_i^{v_c}/\hat{p}_i^{v_c}\right)\hat{p}_i=v_b\left(1\right)+o_p(1)$ by Lemmata SM2.9 and $\left\|\sqrt{n}\hat{\lambda}\right\|^2=O_p(1)$ by Lemma SM2.8 it follows by Lemma SM2.2 that $|B_{2,n}|=o_p(1)$.

Therefore,

$$\begin{aligned} \sqrt{n}\mathcal{S}_v &= -\sqrt{n}\hat{s}(\varphi^*_n)'S'_hP_h\Omega_{hg}\sqrt{n}\hat{\lambda}+\sqrt{n}\hat{s}(\varphi^*_n)'S'_gP_g\Omega_g\sqrt{n}\hat{\lambda}+O_p(n^{-2}) \\ &= \sqrt{n}\hat{s}(\varphi^*_n)'S'_hP_h\Omega_{hg}P_gS_g\sqrt{n}\hat{s}(\varphi^*_n)-\sqrt{n}\hat{s}(\varphi^*_n)'S'_gP_gS_g\sqrt{n}\hat{s}(\varphi^*_n)+O_p(n^{-2}) \end{aligned}$$

because $\sqrt{n}\hat{\lambda}=-P_g\sqrt{n}\hat{g}(\beta_{0,\mathrm{P}_n})+o_p(1)$ and the fact $P_g\Omega_gP_g=P_g$. Hence, since $\sqrt{n}\hat{s}(\varphi^*_n)\xrightarrow{d}\mathcal{N}(\delta,\Omega)$, $\delta=(\delta'_g,\delta'_h)'$ by Lemma SM2.8 it follows that $\sqrt{n}\mathcal{S}_v\xrightarrow{d}\mathcal{T}\left(\delta,\Omega,Q_1\right)$.

Consider now $\sqrt{n}\mathcal{J}_{v,2} = -\sqrt{n}\hat{A}_{v,2}\hat{\Omega}_g^{-1}\sqrt{n}\hat{g}(\hat{\beta})$, note that

$$\begin{aligned}\sqrt{n}\mathcal{J}_{v,2} &= n^{-1/2}\sum_{i=1}^{n}(n\hat{p}_i-1)v_b(n\hat{q}_i/(n\hat{p}_i))\hat{g}_{in}'\hat{\Omega}_g^{-1}\sqrt{n}\hat{g}(\hat{\beta}) \\ &\quad -\sqrt{n}\hat{A}_{v,3}\hat{\Omega}_g^{-1}\sqrt{n}\hat{g}(\hat{\beta}).\end{aligned}$$

We prove that the first term converges in probability to zero. Note that by Lemma SM2.9

$$\begin{aligned}n^{-1/2}\sum_{i=1}^{n}(n\hat{p}_i-1)v_b\left(\frac{n\hat{q}_i^{v_c}}{n\hat{p}_i}\right)\hat{g}_{in}' &= \frac{1}{n}\sum_{i=1}^{n}v_b\left(\frac{n\hat{q}_i^{v_c}}{n\hat{p}_i}\right)\hat{g}_{in}\hat{g}_{in}'\sqrt{n}\hat{\lambda}\left(1+o_p\left(1\right)\right) \\ &\quad +\sum_{i=1}^{n}O_p(n^{-3/2})v\left(n\hat{q}_i\right)\hat{g}_{in}'.\end{aligned}$$

Both terms of the rhs converge to zero in probability by Lemma SM2.9, Lemma SM2.2, the facts that $\sqrt{n}\hat{\lambda} = -P_g\sqrt{n}\hat{g}_n(\beta_{0,\mathrm{P}_n}) + o_p(1)$, $\Omega P = 0$ and $\hat{g}(\hat{\beta}) \xrightarrow{p} 0$. Hence

$$n^{-1/2}\sum_{i=1}^{n}(n\hat{p}_i-1)v_b\left(\frac{n\hat{q}_i^{v_c}}{n\hat{p}_i}\right)\hat{g}_{in}' = o_p(1). \tag{8}$$

Combining (7) and (8), one obtains

$$\sqrt{n}\hat{A}_{v,2} = \sqrt{n}\hat{A}_{v,3} + o_p(1). \tag{9}$$

Consequently as

$$\begin{aligned}-\sqrt{n}\Omega_g^{-1}\hat{g}(\hat{\beta}) &= \sqrt{n}\hat{\lambda} + o_p(1) \\ &= -P_g\sqrt{n}\hat{g}(\beta_{0,\mathrm{P}_n}) + o_p(1),\end{aligned} \tag{10}$$

we have $\sqrt{n}\mathcal{J}_{2,v} = \sqrt{n}\mathcal{J}_{3,v} + o_p(1) = \sqrt{n}\mathcal{S}_v + o_p(1)$.

Concerning the Lagrange multiplier test statistics note that for $\sqrt{n}\mathcal{LM}_{j,v}$ $j = 2, 3$ we have $\sqrt{n}\mathcal{LM}_{v,j} - \sqrt{n}\mathcal{J}_{v,j} = n^{1/2}\hat{A}_{v,j}\left[n^{1/2}\hat{\lambda} + \hat{\Omega}_g^{-1}n^{1/2}\hat{g}(\hat{\beta})\right] = n^{1/2}\hat{A}_{v,j}[-\Omega_g^{-1} + \hat{\Omega}_g^{-1} + o_p(1)]\sqrt{n}\hat{g}(\hat{\beta}) = o_p(1)$ using (10) and the facts that $\sqrt{n}\hat{A}_{j,v} = O_p(1)$, $\Omega_g^{-1} - \hat{\Omega}_g^{-1} = o_p(1)$ by Lemma SM2.2 and $\sqrt{n}\hat{g}(\hat{\beta}) = O_p(1)$.

Proof of 2: Let us now consider $\sqrt{n}\tilde{\mathcal{S}}_v$. By (5), it follows that

$$\begin{aligned}\sqrt{n}\tilde{\mathcal{S}}_v &= \sqrt{n}\sum_{i=1}^{n}\sqrt{n}\left(\hat{p}_i - \frac{1}{n}\right)n\hat{p}_i\left[v_b\left(n\hat{q}_i^{v_c}\right) - \sum_{\ell=1}^{n}v_b\left(n\hat{q}_\ell^{v_c}\right)\hat{p}_\ell\right] \\ &\quad + n\sum_{i=1}^{n}\frac{v_{2,a}\left(\hat{\sigma}_i\right)}{2}\left(\hat{p}_i n - 1\right)^2\hat{p}_i\left[v_b\left(n\hat{q}_i^{v_c}\right) - \sum_{\ell=1}^{n}v_b\left(n\hat{q}_\ell^{v_c}\right)\hat{p}_\ell\right] = C_{1,n} + C_{2,n}.\end{aligned}$$

Hence, by Lemma SM2.9,

$$\begin{aligned}C_{1,n} &= n^{-1/2}\sum_{i=1}^{n}n\hat{p}_i\left[v_b\left(\hat{q}_i^{v_c}n\right) - \sum_{\ell=1}^{n}v_b\left(\hat{q}_\ell^{v_c}n\right)\hat{p}_\ell\right]\hat{g}_{in}'\sqrt{n}\hat{\lambda}\left(1+o_p\left(1\right)\right) \\ &\quad + C_{1r,n}.\end{aligned}$$

Note that $n^{-1/2}\sum_{i=1}^{n} n\hat{p}_i[v_b\left(\hat{q}_i^{v_c}n\right)-\sum_{\ell=1}^{n} v_b\left(\hat{q}_\ell^{v_c}n\right)\hat{p}_\ell]\hat{g}_{in}' = n^{-1/2}\sum_{i=1}^{n} n\hat{p}_i v_b\left(\hat{q}_i^{v_c}n\right)\hat{g}_{in}'$ because $\sum_{i=1}^{n}\hat{p}_i\hat{g}_{in}' = 0$. Additionally by a Taylor expansion $v_b\left(n\hat{q}_i^{v_c}\right) = v_b(1) + v_{b,1}\left(\hat{\sigma}_{3,i}\right)\left(n\hat{q}_i^{v_c}-1\right)$, where $\hat{\sigma}_{3,i} = \alpha_{3,i} + (1-\alpha_{3,i})\, n\hat{q}_i^{v_c}$ and $\alpha_{3,i}\in(0,1)$. Thus

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} n\hat{p}_i\hat{g}_{in}' v_b\left(\hat{q}_i^{v_c}n\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} n\hat{p}_i\hat{g}_{in}' v_{b,1}\left(\hat{\sigma}_{3,i}\right)\left(n\hat{q}_i^{v_c}-1\right).$$

Now note that $\sum_{i=1}^{n} n\hat{p}_i\hat{g}_{in}'\sqrt{n}\left(n\hat{q}_i^{v_c}-1/n\right) = O_p(1)$ by Lemma SM2.9, Lemma SM2.2 and the fact that $\sqrt{n}\hat{\eta} = O_p(1)$. Thus using continuity of $v_{b,1}(.)$ at 1, we have

$$\begin{aligned}\sqrt{n}\hat{A}_{4,v} &= n^{-1/2}\sum_{i=1}^{n} n\hat{p}_{in}\hat{g}_{in}' v_b\left(\hat{q}_i^{v_c}n\right) = W_{1,n} + o_p(1) \qquad (11)\\ &= -\sqrt{n}\hat{s}(\varphi_n^*)' S_h' P_h\Omega_{hg} + O_p(n^{-2})\end{aligned}$$

by (6). Also $C_{1r,n} = n^{-1/2}\sum_{i=1}^{n} n\hat{p}_i[v_b\left(\hat{q}_i^{v_c}n\right)-\sum_{\ell=1}^{n} v_b\left(\hat{q}_\ell^{v_c}n\right)\hat{p}_\ell]\hat{g}_i' O_p(n^{-3/2}) = o_p(1/n)$ by Lemmata SM2.9 and SM2.1 and $C_{2,n} = o_p(1)$ using a proof similar to the proof that $B_{2,n} = o_p(1)$. Hence

$$\begin{aligned}\sqrt{n}\tilde{\mathcal{S}}_v &= -\sqrt{n}\hat{s}(\varphi_n^*)' S_h' P_h\Omega_{hg}\sqrt{n}\hat{\lambda} + o_p(1)\\ &= \sqrt{n}\hat{s}(\varphi_n^*)' S_h' P_h\Omega_{hg} P_g S_g\sqrt{n}\hat{s}(\varphi_n^*) + o_p(1).\end{aligned}$$

Thus, since $\sqrt{n}\hat{s}(\varphi_n^*)\xrightarrow{d}\mathcal{N}(\delta,\Omega)$ we have $\sqrt{n}\tilde{\mathcal{S}}_v\xrightarrow{d}\mathcal{T}(\delta,\Omega,Q_2)$.

Consider now $\sqrt{n}\mathcal{J}_{v,1} = -n^{-1/2}\sum_{i=1}^{n} v_b\left(n\hat{q}_i^{v_c}\right)\hat{g}_i'\hat{\Omega}_g^{-1} n^{1/2}\hat{g}(\hat{\beta})$. Note that $\sqrt{n}\hat{A}_{v,1} = n^{-1/2}\sum_{i=1}^{n} v\left(n\hat{q}_i\right)\hat{g}_{in}' = -n^{-1/2}\sum_{i=1}^{n}(n\hat{p}_i-1)v_b\left(n\hat{q}_i^{v_c}\right)\hat{g}_{in}' + \sqrt{n}\hat{A}_{v,4}$. The second term of the rhs has the asymptotic representation given in (11). Similarly to the proof of (8) the first term converges in probability to zero and therefore

$$\sqrt{n}\hat{A}_{v,1} = \sqrt{n}\hat{A}_{v,4} + o_p(1). \qquad (12)$$

Thus, using equation (10), and the fact that $P_g\sqrt{n}\hat{g}(\beta_{0,\mathrm{P}_n}) = O_p(1)$ it follows that $\sqrt{n}\mathcal{J}_{1,v} = \sqrt{n}\mathcal{J}_{4,v} + o_p(1) = \sqrt{n}\tilde{\mathcal{S}}_v + o_p(1)$. The proof that $\sqrt{n}\mathcal{LM}_{v,j} - \sqrt{n}\mathcal{J}_{v,j} = o_p(1)$, $j = 1,4$ is similar to the case $j = 2,3$ and therefore omitted.

Proof of 3: Let us first consider $n\hat{\sigma}_1^2 = n\hat{A}_{1,v}\hat{P}_g\hat{A}_{1,v}'$. Note that by (12) and (11), $n\hat{\sigma}_1^2 = n\hat{s}(\varphi_n^*)' K_h P_g K_h'\hat{s}(\varphi_n^*) + o_p(1)$. Hence, it follows that $n\hat{\sigma}_1^2\xrightarrow{d}\mathcal{T}(\delta,\Omega,Q_3)$ because of the fact that $\sqrt{n}\hat{s}(\varphi_n^*)\xrightarrow{d}\mathcal{N}(\delta,\Omega)$. Additionally, using (12) and (11) we have $n\hat{\sigma}_4^2 = n\hat{\sigma}_1^2 + o_p(1)$.

Proof of 4: Now consider $n\hat{\sigma}_2^2 = n\hat{A}_{2,v}\hat{P}_g\hat{A}_{2,v}'$. By (9) and (7) we have

$$n\hat{\sigma}_2^2 = n\hat{s}(\varphi_n^*)'(K_h - K_g)P_g(K_h - K_g)'\hat{s}(\varphi_n^*) + o_p(1).$$

Hence $n\hat{\sigma}_1^2\xrightarrow{d}\mathcal{T}(\delta,\Omega,Q_4)$ as $\sqrt{n}\hat{s}(\varphi_n^*)\xrightarrow{d}\mathcal{N}(\delta,\Omega)$. Additionally, also by (9) and (7) we have $n\hat{\sigma}_2^2 = n\hat{\sigma}_3^2 + o_p(1)$.

Proof of 5: Because P_g is a positive semidefinite matrix, it follows that $n\hat{\sigma}_j^2 \geq 0$, $j = 1,2,3,4$ wpa1. ■

Proof of Theorem 5.1: (1) follows from Lemmata SM2.10 and SM2.11 with $\delta_g = 0$, while (2) follows from Lemma SM2.12 with $\delta = S_h' \delta_h$. ■

Proof of Theorem 5.2: The proof of this theorem is similar to the proof of Theorem 4.1 of Shi (2015). First let $\hat{\Omega}_n = \hat{\Omega}$, $\hat{Q}_{n,i} = \hat{Q}_i$ and $\hat{Q}_{n,j} = \hat{Q}_j$, $cv_{ij,n}^* = cv^*(1-\tau, \hat{\Omega}_n, \hat{Q}_{n,i}, \hat{Q}_{n,j})$. We take a sequence $\{\mathrm{P}_n \in \mathcal{P}_0\}$ and a subsequence $\{b_n\}$ of $\{n\}$ such that

$$\lim_{n\to\infty} \sup \sup_{\mathrm{P}\in P_0} \Pr\left(|W_n^s(i,j)| > cv_n^*\right) = \lim_{n\to\infty} \Pr_{\mathrm{P}_{b_n}}\left(\left|W_{b_n}^s(i,j)\right| > cv_{ij,b_n}^*\right).$$

Such sequences and subsequences always exist. Assumption 5.1 and condition (c) of Definition 5.1 imply that elements in the arrays $\beta_{o,\mathrm{P}}$, γ_{P}^*, η_{P}^*, $\Omega_{\mathrm{P}}(\varphi^*)$, $D_{\mathrm{P},g}(\beta^*)$, $D_{\mathrm{P},h}(\gamma^*)$, $A_{\mathrm{P},v}(\mu^*)$ are uniformly bounded over $\mathrm{P}\in\mathcal{P}$. Thus, there exists a subsequence $\{a_n\}$ of $\{b_n\}$ and some $(\beta_0, \gamma^*, \delta_h, \Omega, D_g, D_h, A_v)$ such that $\left(\beta_{0,\mathrm{P}_{a_n}}, \gamma_{\mathrm{P}_{a_n}}^*, \sqrt{n}\mathcal{E}_{\mathrm{P}_{a_n},h}, \Omega_{\mathrm{P}_{a_n}}(\varphi^*), D_{\mathrm{P}_{a_n},g}(\beta^*), D_{\mathrm{P}_{a_n},h}(\gamma^*), A_{\mathrm{P}_{a_n},v}(\mu^*)\right) \to (\beta_0, \gamma^*, \delta_h, \Omega, D_g, D_h, A_v)$. It suffices to show that $\lim_{n\to\infty} \Pr_{\mathrm{P}_{a_n}}\left(\left|W_{a_n}^s(i,j)\right| > cv_{ij,a_n}^*\right) \leq \tau$.

Note now that

$$W_{a_n}^s(i,j) = \frac{T_{a_n}(i) - tr(\hat{Q}_{a_n,i}\hat{\Omega}_{a_n})/\sqrt{a_n}}{\sqrt{V_{a_n}(j) + c_{ij}\cdot tr(\hat{Q}_{a_n,j}\hat{\Omega}_{a_n})/a_n}}, i = 1,2, j = 3,4.$$

By Theorem 5.1 (1) if $\|\delta_h\|_\infty = +\infty$, we have $T_{a_n}(i) \xrightarrow{d} \mathcal{N}(0,\sigma^2)$ and $V_{a_n}(j) \xrightarrow{p} \sigma^2$. Additionally, $\hat{Q}_{a_n,i} \xrightarrow{p} Q_i$, $\hat{Q}_{a_n,j} \xrightarrow{p} Q_j$, $\hat{\Omega}_{a_n} \xrightarrow{p} \Omega$ by Lemma SM2.2 and the Slutsky theorems. Therefore $W_{a_n}^s(i,j) \xrightarrow{d} \mathcal{N}(0,1)$ and

$$\begin{aligned}\lim_{n\to\infty} \Pr_{\mathrm{P}_{a_n}}\left(\left|W_{a_n}^s(i,j)\right| > cv_{ij,a_n}^*\right) &\leq \lim_{n\to\infty} \Pr_{\mathrm{P}_{a_n}}\left(\left|W_{a_n}^s(i,j)\right| > z_{1-\tau/2}\right) \\ &= \tau\end{aligned}$$

because $cv_{ij,a_n}^* \geq z_{1-\tau/2}$.

Suppose now that $\|\delta_h\|_\infty < +\infty$ in this case

$$W_{a_n}^s(i,j) \xrightarrow{d} \mathcal{W}^s(S_h'\delta_h, i, j) = \frac{\mathcal{T}(S_h'\delta_h, \Omega, Q_i) - tr(Q_i\Omega)}{\sqrt{\mathcal{T}(S_h'\delta_h, \Omega, Q_j) + c_{ij}\cdot tr(Q_j\Omega)}}$$

by Lemma SM2.2 and Theorem 5.1 (2).

Note that $cv_{ij,a_n}^* \geq cv(1-\tau, \hat{\Omega}_{a_n}, \hat{Q}_{a_n,i}, \hat{Q}_{a_n,j})$. Hence

$$\Pr_{\mathrm{P}_{a_n}}\left(\left|W_{a_n}^s(i,j)\right| > cv_{ij,a_n}^*\right) \leq \Pr_{\mathrm{P}_{a_n}}\left(\left|W_{a_n}^s(i,j)\right| > cv(1-\tau, \hat{\Omega}_{a_n}, \hat{Q}_{a_n,i}, \hat{Q}_{a_n,j})\right).$$

Note that by inspection $\mathcal{W}^s(S_h'\delta_h, i, j)$ is continuous in $(\delta_h, \Omega, Q_i, Q_j)$; the continuity of the Cholesky decomposition follows from Lemma 2.1.6 p. 295 of Schatzman (2002). Additionally, $\mathcal{W}^s(S_h'\delta_h, i, j)$ is a continuous random variable and consequently $cv(1-\tau, \delta_h, \Omega, Q_i, Q_j)$ is continuous in $(\delta_h, \Omega, Q_i, Q_j)$.

Since $[0, c_h]$ is a compact set it follows by the maximum theorem that $cv(1-\tau, \Omega, Q_i, Q_j)$ is continuous in Ω, Q_i, Q_j. Now by Lemma SM2.2 $\hat{\Omega}_{a_n} \xrightarrow{p} \Omega$, $\hat{Q}_{a_n,i} \xrightarrow{p} Q_i$ and $\hat{Q}_{a_n,j} \xrightarrow{p} Q_j$ and consequently $cv(1-\tau, \hat{\Omega}_{a_n}, \hat{Q}_{a_n,i}, \hat{Q}_{a_n,j}) \xrightarrow{p} cv_{ij}(1-\tau, [0, c_h])$. Therefore

$$\lim_{n\to\infty} \Pr_{\mathrm{P}_{a_n}} \left(\left| W^s_{a_n}(i,j) \right| > cv(1-\tau, \hat{\Omega}_{a_n}, \hat{Q}_{a_n,i}, \hat{Q}_{a_n,j}) \right) = \Pr\left(\left| \mathcal{W}^s \left(S'_h \delta_h, i, j \right) \right| > cv_{ij}(1-\tau, [0, c_h]) \right).$$

Now by Assumption 5.2 $cv_{ij}(1-\tau, [0, c_h]) = cv_{ij}(1-\tau, [0, +\infty))$, consequently for any $\delta_h \in [0, +\infty)$

$$\begin{aligned} \Pr\left(\left| \mathcal{W}^s \left(S'_h \delta_h, i, j \right) \right| > cv(1-\tau, [0, c_h]) \right) &\leq \Pr\left(\left| \mathcal{W}^s \left(S'_h \delta_h, i, j \right) \right| > cv\left(1-\tau, S'_h \delta_h, \Omega, Q_i, Q_j, c_{ij}\right) \right) \\ &= \tau. \end{aligned}$$

■

Proof of Theorem 5.3: Using the same arguments as in the proof of Theorem 5.1 $cv(1-\tau, \Omega, Q_i, Q_j)$ is continuous in Ω, Q_i, Q_j and therefore $cv(1-\tau, \hat{\Omega}, \hat{Q}_i, \hat{Q}_j) \xrightarrow{p} cv(1-\tau, \Omega, Q_i, Q_j)$ because $\hat{\Omega} \xrightarrow{p} \Omega$, $\hat{Q}_i \xrightarrow{p} Q_i$ and $\hat{Q}_{,j} \xrightarrow{p} Q_j$ for $\|\delta_h\|_\infty \leq +\infty$ by Lemma SM2.2. consequently $cv^*(1-\tau, \hat{\Omega}, \hat{Q}_i, \hat{Q}_j) \xrightarrow{p} cv^*_{ij}$.

Now let us consider first the case $\|\delta_h\|_\infty < +\infty$. In this case $W^s_n(i,j) \xrightarrow{d} \mathcal{W}(\delta, \Omega, Q_i, Q_j, c_{ij})$ by Lemmata SM2.2 and SM2.12, It follows that $\lim_{n\to\infty} \Pr_{\mathrm{P}_n} \left(\left| W^s_n(i,j) \right| > cv^*(1-\tau, \hat{\Omega}, \hat{Q}_i, \hat{Q}_j) \right) = 1 - F_{|\mathcal{W}_{ij}|}(cv^*_{ij})$.

Consider now the case $\|\delta_h\|_\infty = +\infty$. In this case $W^s_n(i,j) \xrightarrow{d} \mathcal{N}\left(-A_v P_g \delta_g/\sigma, 1\right)$ by Lemmata SM2.2 and SM2.12. It follows that $\lim_{n\to\infty} \Pr_{\mathrm{P}_n} \left(\left| W^s_n(i,j) \right| > cv^*(1-\tau, \hat{\Omega}, \hat{Q}_i, \hat{Q}_j) \right) = \Pr\left[|x| > cv^*_{ij} \right]$, where $x \sim \mathcal{N}\left(-A_v P_g \delta_g/\sigma, 1\right)$. Now note that $\Pr\left[|x| > cv^*_{ij} \right] = \Phi(-cv^*_{ij} - A_v P_g \delta_g/\sigma) + \Phi(-cv^*_{ij} + A_v P_g \delta_g/\sigma)$. ■

SM2.2 The limit behavior of $W^*_{i,j}(\delta_h)$

In this subsection we investigate the behavior of the random variable $W^*_{i,j}(\delta_h) = \mathcal{T}\left(S'_h \delta_h, \Omega, Q_i\right) / \mathcal{T}\left(S'_h \delta_h, \Omega, Q_j\right)^{1/2}$, $i = 1, 2$, $j = 3, 4$, as some of the elements of δ_h approach infinity. These random variables are defined in section 5.1 of the paper, page 17. We start by presenting two useful Lemmata that allow us to analyze this limit. The proofs of the Lemmata are presented at the end of this subsection.

Lemma SM2.13 *The random variables $W^*_{i,j}(\delta_h) = \mathcal{T}\left(S'_h \delta_h, \Omega, Q_i\right) / \mathcal{T}\left(S'_h \delta_h, \Omega, Q_j\right)^{1/2}$, $i = 1, 2$, $j = 3, 4$ have the following representation:*

$$W^*_{i,j}(\delta_h) = \frac{s_0(\delta_h)\, x_0 + z' C'_\Omega Q_i C_\Omega z}{\left[s^2_0(\delta_h) + 2 s_j(\delta_h)\, x_j + z' C'_\Omega Q_j C_\Omega z \right]^{1/2}}, \tag{13}$$

where $z \sim \mathcal{N}(0, I_m)$, $x_0 \sim \mathcal{N}(0,1)$, $s^2_0(\delta_h) = \delta'_h L_0 \delta_h \geq 0$, $L_0 \equiv P_h \Omega_{hg} P_g \Omega'_{hg} P_h$, $s^2_3(\delta_h) = \delta'_h L_1 \delta_h \geq 0$, where $L_1 \equiv L_0 \Omega_h L_0$, $s^2_4(\delta_h) = s^2_0(\delta_h) - s^2_3(\delta_h) \geq 0$ and $x_j \sim \mathcal{N}(0,1)$.

We can see from Lemma SM2.13 that $W^*_{i,j}(\delta_h)$ only depends on δ_h via the quadratic forms $\delta'_h L_0 \delta_h$ and $\delta'_h L_1 \delta_h$. Additionally, it is apparent from equation (13) that if $s^2_0(\delta_h) = \delta'_h L_0 \delta_h \to \infty$, then

$W^*_{i,j}(\delta_h) \to \mathcal{N}(0,1)$ (note that $0 \leq s_j(\delta_h)/s_0^2(\delta_h) \leq s_0^{-1}(\delta_h)$, $j = 3,4$). However, given that the matrix L_0 is positive semidefinite (because P_g is positive semidefinite), the quadratic form $\delta_h' L_0 \delta_h$ does not necessarily diverge as any of the elements of δ_h approach infinity. In fact the following lemma shows that this limit is path dependent. We use the following notation. Let $N(A)$ denote the null space of a matrix A and let $\|\cdot\|$ be the Euclidean norm of $\cdot$. Denote U a matrix of eigenvectors of the matrix L_0 chosen such that they are orthogonal to each other. Let also U_a be the submatrix of U that contains the eigenvectors corresponding to the positive eigenvalues of L_0 and U_b be the submatrix of U that contains the eigenvectors corresponding to the zero eigenvectors of L_0.

Lemma SM2.14 *If $rank(L_0) = r > 0$ and $\delta_h(t_a, t_b) = U_a t_a + U_b t_b$ where $t_a \in \mathbb{R}^r$ and $t_b \in \mathbb{R}^{m_h - r}$, then we have $\|\delta_h(t_a, t_b)\| = \|t_a\| + \|t_b\|$, and the following results hold for $i = 1,2$, $j = 3,4$:*

1. $W^*_{i,j}(\delta_h(t_a, t_b)) = W^*_{i,j}(U_a t_a)$;
2. $\lim_{\|t_b\| \to \infty} W^*_{i,j}(\delta_h(t_a, t_b)) = W^*_{i,j}(U_a t_a)$, *if* $\|t_a\| < \infty$;
3. $\lim_{\|t_a\| \to \infty} W^*_{i,j}(\delta_h(t_a, t_b)) = x_0$ *either if* $\|t_b\| < \infty$ *or if* $\|t_b\| \to \infty$, *where* $x_0 \sim \mathcal{N}(0,1)$.

In Lemma SM2.14 we consider paths of the form $\delta_h(t_a, t_b) = U_a t_a + U_b t_a$ because the eigenvectors are chosen such that they are orthogonal to each other and therefore they form a basis of $\mathbb{R}^{m_h}$. Since $\|\delta_h(t_a, t_b)\| = \|t_a\| + \|t_b\|$, it follows that $\|\delta_h(t_a, t_b)\|$ goes to infinity in the following cases: $\|t_b\| \to \infty$ and $\|t_a\| < \infty$; $\|t_a\| \to \infty$ and $\|t_b\| < \infty$; and $\|t_a\| \to \infty$ and $\|t_b\| \to \infty$. Lemma SM2.14 shows that, for fixed t_a satisfying $\|t_a\| < \infty$, the distribution of $W^*_{i,j}(\delta_h(t_a, t_b))$ does not depend on the value of t_b and consequently this distribution is the same whether $\|t_b\| \to \infty$ or if $\|t_b\| < \infty$. On the other hand, when $\|t_a\| \to \infty$, $W^*_{i,j}(\delta_h(t_a, t_b))$ converges always to the standard normal distribution.

We prove now the above Lemmata.

Proof of Lemma SM2.13: First note that for any matrix Q: $\mathcal{T}(\delta_h' S_h, \Omega, Q) = (\delta_h' S_h + z' C_\Omega') Q (S_h' \delta_h + C_\Omega z) = \delta_h' S_h Q S_h' \delta_h + \delta_h' S_h Q C_\Omega z + z' C_\Omega' Q S_h' \delta_h + z' C_\Omega' Q C_\Omega z$. We prove the result by showing that:

(a) for $i = 1,2: \mathcal{T}(\delta_h' S_h, \Omega, Q_i) = s_0(\delta_h) x_0 + z' C_\Omega' Q_i C_\Omega z$, $s_0^2(\delta_h) = \delta_h' L \delta_h$ and $x_0 \sim \mathcal{N}(0,1)$.

(b) for $j = 3,4: \mathcal{T}(\delta_h' S_h, \Omega, Q_j) = s_0^2(\delta_h) + 2 s_j(\delta_h) x_j + z' C_\Omega' Q_j C_\Omega z$, $s_3^2(\delta_h) = \delta_h' L \Omega_h L \delta_h$, $s_4^2(\delta_h) = s_0^2(\delta_h) - s_3^2(\delta_h) \geq 0$ and $x_j \sim \mathcal{N}(0,1)$.

We start by proving (a). Let us consider $\mathcal{T}(\delta_h' S_h, \Omega, Q_i)$. Note that for $Q = Q_1$ and $Q = Q_2$ we have $\delta_h' S_h Q S_h' \delta_h = 0$, $z' C_\Omega' Q S_h' \delta_h = 0$ and $\delta_h' S_h Q C_\Omega z = \delta_h' P_h \Omega_{hg} P_g S_g C_\Omega z$ because $S_g S_h' = 0$ and $S_h S_h' = I_{m_h}$. Therefore for $i = 1,2$: $\mathcal{T}(\delta, \Omega, Q_i) = \delta_h' P_h \Omega_{hg} P_g S_g C_\Omega z + z' C_\Omega' Q_i C_\Omega z$. Let $s_0^2(\delta_h) = \text{var}(\delta_h' P_h \Omega_{hg} P_g S_g C_\Omega z)$. Note that $s_0^2(\delta_h) = \delta_h' P_h \Omega_{hg} P_g S_g \Omega S_g' P_g \Omega_{hg}' P_h \delta_h = \delta_h' L_0 \delta_h \geq 0$ and let $x_0 \sim \mathcal{N}(0,1)$. Hence for $i = 1,2$ we have $\mathcal{T}(\delta_h' S_h, \Omega, Q_i) = s_0(\delta_h) x_0 + z' C_\Omega' Q_i C_\Omega z$. To see this note that if $s_0^2(\delta_h) = 0$, $\delta_h' P_h \Omega_{hg} P_g S_g C_\Omega z = 0$ and $s_0(\delta_h) x_0 = 0$ and if $s_0(\delta_h) > 0$, then we can define $x_0 = \delta_h' P_h \Omega_{hg} P_g S_g C_\Omega z / s_0(\delta_h)$ and since $z \sim \mathcal{N}(0, I_m)$ it follows that $x_0 \sim \mathcal{N}(0,1)$.

We prove now (b). Note that for $Q = Q_3$ and $Q = Q_4$ we have $\delta_h' S_h Q S_h' \delta_h = \delta_h' L_0 \delta_h = s_0^2(\delta_h)$ because $S_g S_h' = 0$ and $S_h S_h' = I_{m_h}$. Additionally, by symmetry we have for $Q = Q_3$ and $Q = Q_4$: $\delta_h' S_h Q C_\Omega z + z' C_\Omega' Q S_h' \delta_h = 2\delta_h' S_h Q C_\Omega z$.

Now note that $\delta_h' S_h Q_3 C_\Omega z = \delta_h' P_h \Omega_{hg} P_g \Omega_{hg}' P_h S_h C_\Omega z$ and consequently $s_3^2(\delta_h) = \text{var}\left(\delta_h' P_h \Omega_{hg} P_g \Omega_{hg}' P_h S_h C_\Omega z\right) = \delta_h' L_0 \Omega_h L_0 \delta_h \geq 0$.

Let us consider now $\delta_h' S_h Q_4 C_\Omega z$. Note that $\delta_h' S_h Q_4 C_\Omega z = \delta_h' P_h \Omega_{hg} P_g (K_h - K_g)' C_\Omega z$, because $S_g S_h' = 0$ and and $S_h S_h' = I_{m_h}$. Let $s_4^2(\delta_h) = \text{var}(\delta_h' P_h \Omega_{hg} P_g (K_h - K_g)' C_\Omega z)$. After some lengthy, but simple algebra we can show that $s_4^2(\delta_h) = s_0^2(\delta_h) - s_3^2(\delta_h)$. Since $s_4^2(\delta_h) \geq 0$ it follows that $s_0^2(\delta_h) \geq s_3^2(\delta_h)$. Consequently $\mathcal{T}(\delta_h' S_h, \Omega, Q_j) = s_0^2(\delta_h) + 2 s_j(\delta_h) x_j + z' C_\Omega' Q_j C_\Omega z$ for $j = 3, 4$ if $s_j^2(\delta_h) \geq 0$, $s_0(\delta_h) \geq 0$ using the same arguments adopted in the proof of (a). ■

Proof of Lemma SM2.14: Since L_0 is a real symmetric matrix it can be factorized as $L_0 = U \Lambda U'$, where $UU' = U'U = I_{m_h}$ and Λ is a diagonal matrix whose entries are the eigenvalues of L_0. Note that $L_0 = U \Lambda U' = U_a \Lambda_* U_a'$, where Λ_* is a $(r \times r)$ matrix with only the r positive eigenvalues in the diagonal. Now notice that $U_b t_b \in N(L_0)$ as $AU_b t_b = U_a \Lambda_* U_a' U_b t_b = 0$, because $U_a' U_b = 0$. Also $(U_a t_a)' U_b t_a = t_a U_a' U_b t_b = 0$, as $U_a' U_b = 0$. Hence $\|\delta_h(t_a, t_b)\| = \|t_a\| + \|t_b\|$.

To prove 1 note that since $U_b t_b \in N(L_0)$, it follows that $U_b t_b \in N(L_1)$ and $U_b t_b \in N(L_0 - L_1)$ because L_0, L_1 and $L_0 - L_1$are positive semidefinite by Lemma SM2.13 (see Abadir and Magnus, 2005, solution of Exercise, 8.41, p. 227). Consequently $s_0^2(\delta_h(t_a, t_b)) = t_a U_a' L_0 U_a t_a$, $s_3^2(\delta_h(t_a, t_b)) = t_a U_a' L_1 U_a t_a$ and $s_4^2(\delta_h(t_a, t_b)) = t_a U_a' (L_0 - L_1) U_a t_a$ and the result follows from Lemma SM2.13. The result 2 is a consequence of 1 because $W_{i,j}^*(U_a t_a)$ does not depend on t_b. To prove 3 we only need to prove that $\lim_{\|t_a\| \to \infty} s_0^2(\delta_h(t_a, t_b)) = +\infty$ as in this case $\lim_{\|t_a\| \to \infty} s_j(\delta_h(t_a, t_b)) / s_0^2(\delta_h(t_a, t_b)) = 0$ due to the fact that $0 \leq s_j(\delta_h(t_a, t_b)) / s_0^2(\delta_h(t_a, t_b)) \leq s_0^{-1}(\delta_h(t_a, t_b))$, $j = 3, 4$ for $s_0(\delta_h(t_a, t_b)) > 0$ by SM2.13. Now since $U_a' U_a = I_r$ we have $s_0^2(\delta_h(t_a, t_b)) = t_a U_a' L_0 U_a t_a = t_a U_a' U_a \Lambda_* U_a' U_a t_a = t_a \Lambda_* t_a \geq c \|t_a\|$, where c is a positive constant. Hence $s_0^2(\delta_h(t_a, t_b)) \to \infty$ as $\|t_a\| \to \infty$. ■

SM3 Monte Carlo study (additional results)

Tables SM1 and SM2 present the empirical sizes and powers for the tests based on the non-nested test statistics computed with the exponential tilting (ET) estimator for the sample sizes 200 and 400. The nominal level for all tests reported is 0.05. We report the results for $\mathcal{S}_\alpha$, $\tilde{\mathcal{S}}_\alpha$, $\mathcal{LM}_\alpha$ and $\mathcal{J}_\alpha$ for $\alpha \to 0$, $\alpha = 1, 1.5, 2, 3$ and for the Ramalho and Smith (2002) statistics based on ET. We use the notation $\mathcal{S}_{\text{RS}}$, $\tilde{\mathcal{S}}_{\text{RS}}$, $\mathcal{LM}_{\text{RS}}$ and $\mathcal{J}_{\text{RS}}$ for the Ramalho and Smith (2002) statistics. We also present in the Tables SM1 and SM2 the results for the tests for overidentifying restrictions based on the likelihood ratio statistic computed with ET [Kitamura and Stutzer (1997) and Imbens *et al.* (1998)], which is labelled $\mathcal{LR}_{\text{ET}}$. ET is calculated using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm.

Table SM1: Rejection frequencies of the tests under the null.

n		200						400					
Statistic		original			Shi-type			original			Shi-type		
Estimator	$\rho_{u_g,w_{h2}}$	0.0	0.2	0.4	0.0	0.2	0.4	0	0.2	0.4	0	0.2	0.4
	$\mathcal{LR}_{\text{ET}}$	5.8	5.6	6.2	–	–	–	5.6	5.3	5.7	–	–	–
	$\mathcal{S}_0$	18.0	15.6	14.5	1.8	1.7	1.9	16.3	13.6	12.5	1.6	1.4	1.7
	$\tilde{\mathcal{S}}_0$	5.5	6.1	8.6	4.6	6.4	7.9	4.9	4.9	7.3	4.0	4.9	6.2
	$\mathcal{LM}_0$	7.8	6.4	7.8	4.5	4.2	4.4	6.4	5.8	6.6	4.3	3.3	3.5
	$\mathcal{J}_0$	5.7	4.8	6.0	3.5	3.1	3.0	5.5	5.1	5.7	3.8	2.8	2.8
	$\mathcal{S}_1$	18.2	16.0	15.8	2.0	1.8	1.8	16.4	13.9	14.2	1.6	1.3	1.3
	$\tilde{\mathcal{S}}_1$	6.1	5.9	8.2	4.4	5.8	7.1	5.4	5.1	6.5	3.9	4.6	5.4
	$\mathcal{LM}_1$	7.5	6.4	7.8	4.7	4.3	4.2	6.3	5.7	6.5	4.3	3.3	3.4
	$\mathcal{J}_1$	5.4	4.8	6.2	3.9	3.3	2.7	5.4	4.9	5.7	3.8	2.8	2.7
	$\mathcal{S}_{1.5}$	18.2	16.3	16.0	2.2	2.0	1.9	16.4	14.1	14.5	1.8	1.4	1.5
	$\tilde{\mathcal{S}}_{1.5}$	6.4	6.0	8.0	4.2	5.5	6.7	5.6	5.2	6.9	3.8	4.0	5.2
	$\mathcal{LM}_{1.5}$	7.3	6.5	7.5	4.7	4.5	4.2	6.2	5.6	5.8	4.2	3.4	3.3
ET	$\mathcal{J}_{1.5}$	5.3	5.1	5.6	3.8	3.4	2.8	5.4	4.8	5.1	3.8	3.0	2.7
	$\mathcal{S}_2$	18.2	16.3	15.9	2.4	2.3	2.4	16.4	14.1	14.3	1.9	1.7	1.7
	$\tilde{\mathcal{S}}_2$	6.7	6.2	7.9	4.2	5.2	6.1	5.8	5.2	7.2	3.8	3.6	4.9
	$\mathcal{LM}_2$	7.1	6.5	6.9	4.7	4.7	4.3	6.2	5.5	5.9	4.2	3.5	3.6
	$\mathcal{J}_2$	5.3	4.8	5.2	3.8	3.6	3.0	5.3	4.7	5.0	3.8	2.9	3.0
	$\mathcal{S}_3$	18.2	16.3	15.3	3.4	3.2	3.5	16.4	13.9	13.7	2.3	2.1	2.5
	$\tilde{\mathcal{S}}_3$	6.9	6.8	8.4	4.1	4.5	4.9	6.3	5.6	7.6	3.7	3.2	4.9
	$\mathcal{LM}_3$	7.0	6.5	7.1	4.8	5.0	5.0	6.0	5.4	6.4	4.2	3.7	4.2
	$\mathcal{J}_3$	5.2	4.8	5.5	4.0	3.9	3.7	5.2	4.5	5.5	3.9	3.2	3.6
	$\mathcal{S}_{\text{RS}}$	16.2	15.0	15.1	2.1	1.0	1.4	15.9	13.6	12.0	1.6	1.1	0.8
	$\tilde{\mathcal{S}}_{\text{RS}}$	5.7	5.3	7.4	3.4	5.8	7.6	5.5	5.4	7.5	3.5	5.2	7.1
	$\mathcal{LM}_{\text{RS}}$	7.1	6.9	7.4	4.1	4.3	3.9	6.4	6.6	5.6	3.7	4.1	2.9
	$\mathcal{J}_{\text{RS}}$	5.2	5.1	5.5	3.2	3.0	2.4	5.6	5.6	4.8	3.4	3.5	2.4

Table SM2: Rejection frequencies of the tests under the alternative.

n		200						400					
Statistic		original			Shi-type			original			Shi-type		
Estimator	ω	3	4	5	3	4	5	3	4	5	3	4	5
	$\mathcal{LR}_{\text{ET}}$	35.3	53.1	66.3	–	–	–	34.6	55.6	73.9	–	–	–
	$\mathcal{S}_0$	43.7	57.0	62.7	11.9	19.9	27.2	40.7	55.9	66.8	10.7	19.8	30.6
	$\tilde{\mathcal{S}}_0$	10.3	19.6	33.6	5.1	10.4	19.2	12.5	26.2	42.9	5.7	14.2	27.6
	$\mathcal{LM}_0$	43.6	61.2	73.3	34.5	52.4	65.3	43.8	63.2	79.2	33.4	53.9	72.8
	$\mathcal{J}_0$	37.9	55.5	68.0	27.7	44.3	57.4	40.6	60.3	77.0	29.8	49.7	68.6
	$\mathcal{S}_1$	48.9	62.1	67.6	12.7	20.9	28.7	47.7	64.3	75.4	11.1	21.4	33.3
	$\tilde{\mathcal{S}}_1$	18.4	32.1	49.0	10.9	20.8	35.8	20.1	35.8	54.5	11.1	24.0	40.5
	$\mathcal{LM}_1$	40.2	58.9	71.9	30.4	49.2	62.9	40.4	60.6	77.5	29.0	50.1	70.1
	$\mathcal{J}_1$	34.6	53.0	66.2	23.3	40.9	54.2	38.0	57.7	75.2	25.7	45.4	65.6
	$\mathcal{S}_{1.5}$	50.1	63.9	69.4	14.0	22.9	30.7	49.2	66.5	77.9	12.4	23.4	36.0
	$\tilde{\mathcal{S}}_{1.5}$	23.1	37.3	54.5	15.3	26.3	42.6	23.8	40.1	59.1	14.6	28.8	46.4
	$\mathcal{LM}_{1.5}$	37.2	56.4	70.2	27.3	46.0	60.4	36.6	56.6	75.5	25.6	45.7	67.3
ET	$\mathcal{J}_{1.5}$	31.6	49.9	64.0	20.1	37.8	52.0	34.2	53.6	72.6	22.9	41.1	62.4
	$\mathcal{S}_2$	50.7	64.5	70.2	16.2	25.6	34.0	50.4	67.9	79.5	13.9	26.6	39.5
	$\tilde{\mathcal{S}}_2$	26.3	41.7	58.0	18.5	30.9	47.8	26.5	43.1	61.9	17.4	32.6	50.5
	$\mathcal{LM}_2$	33.7	52.6	67.7	24.2	42.4	57.5	32.4	51.9	71.4	22.0	41.0	61.6
	$\mathcal{J}_2$	28.5	46.2	61.3	17.7	33.4	48.8	29.9	48.8	67.7	19.1	36.1	56.6
	$\mathcal{S}_3$	50.2	65.3	71.0	21.2	32.0	40.6	49.4	67.8	79.7	18.8	33.1	47.1
	$\tilde{\mathcal{S}}_3$	28.2	45.1	62.0	21.2	36.1	53.6	25.7	42.2	59.5	18.7	33.8	51.9
	$\mathcal{LM}_3$	29.1	46.1	61.3	21.2	35.3	51.1	27.1	43.8	60.0	20.3	34.0	51.3
	$\mathcal{J}_3$	23.9	39.0	53.8	15.5	27.3	42.2	24.3	40.1	56.2	17.5	30.3	46.4
	$\mathcal{S}_{\text{RS}}$	52.6	67.5	74.5	13.9	29.2	43.7	50.0	71.5	84.8	11.4	28.1	45.8
	$\tilde{\mathcal{S}}_{\text{RS}}$	10.7	29.7	48.3	7.7	22.9	39.6	5.1	13.4	29.9	1.9	7.1	21.0
	$\mathcal{LM}_{\text{RS}}$	37.7	51.2	62.5	27.5	41.6	54.6	40.7	59.4	68.3	29.2	47.1	58.2
	$\mathcal{J}_{\text{RS}}$	32.1	44.8	55.5	20.0	34.1	46.4	38.2	56.5	64.8	25.5	42.3	52.9

References

Abadir, K. M., and J. R. Magnus (2005): *Matrix Algebra.* Vol. 1. Cambridge University Press.

Andrews, D. W. K. (1992): "Generic Uniform Convergence," *Econometric Theory*, 8, 241–257.

Davidson, J.E.H. (1994): *Stochastic Limit Theory.* Oxford: Oxford University Press.

Imbens, G., R. Spady, and P. Johnson (1998): "Information Theoretic Approaches to Inference in Moment Condition Models," *Econometrica*, 66, 333-357.

Kitamura, Y., and M. Stutzer (1997): "An Information-Theoretic Alternative to Generalized Method of Moments Estimation," *Econometrica*, 65, 861-874.

Newey, W.K., and D.L. McFadden (1994): "Large Sample Estimation and Hypothesis Testing," in Engle, R.F., and D.L. McFadden, eds., *Handbook of Econometrics*, Vol. 4, 2111-2245. New York: North Holland.

Newey, W.K., and R.J. Smith (2004): "Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators," *Econometrica*, 219-255.

Ramalho, J.J.S. and R.J. Smith (2002): "Generalized Empirical Likelihood Non-nested Tests," *Journal of Econometrics*, 107, 99-125.

Ramalho, J.J.S., and R.J. Smith (2004): "Goodness of Fit Tests for Moment Conditions Models," *working paper*, Universidade de Évora.

Rudin, W. (1976): *Principles of Mathematical Analysis*, third edition, McGraw-Hill.

Serfling, R. (1980): *Approximation Theorems of Mathematical Statistics*, Wiley, 1980.

Schatzman, M. (2002): *Numerical analysis: a mathematical introduction.* Oxford University Press.

Shi, X. (2015a): "A nondegenerate Vuong test," *Quantitative Economics*, vol. 6(1), 85-121, 03.