Online Supplemental Material to Aucejo et al. (2015)

Esteban M. Aucejo
Department of Economics
London School of Economics and Political Science
e.m.aucejo@lse.ac.uk

Federico A. Bugni
Department of Economics
Duke University
federico.bugni@duke.edu

V. Joseph Hotz
Department of Economics
Duke University, NBER, and IZA
hotz@econ.duke.edu

July 17, 2015

Abstract

This supplement contains the proofs of Theorems 4.1 and 4.2 of Aucejo et al. (2015, Section 4), along with two intermediate results that are required for these proofs. All definitions required for this supplement can be found in Aucejo et al. (2015). Finally, the derivations in this supplement follow closely results in Andrews and Shi (2013) (hereafter, referred to as AS13).

S1 Results on inference

Proof of Theorem 4.1. The proof of this result follows closely the arguments in AS13 (Theorem 2(a)). Notice that Assumption A.2 implies their Assumptions S1-S2, Assumption A.5 implies the manageability of the stochastic processes implied by their Assumption M, and Assumption A.6 implies their Assumption GMS1.

Suppose that Eq. (4.6) does not hold. In this case, we can find a subsequence $\{a_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$ and a sequence $\{(\theta_{a_n}, F_{a_n}) \in \bar{\mathcal{F}}_0\}_{n\geq 1}$ s.t. $P_{F_{a_n}}(\theta_{a_n} \notin CS_{a_n}) > \alpha\ \forall n \in \mathbb{N}$. By the compactness implicit in the definition of $\bar{\mathcal{F}}_0$, we can find a further subsequence $\{b_n\}_{n\geq 1}$ of $\{a_n\}_{n\geq 1}$ s.t. $\{(\theta_{b_n}, F_{b_n}) \in \bar{\mathcal{F}}_0\}_{n\geq 1} \in SubSeq(h_2)$ for some limiting variance-covariance kernel h_2, where $SubSeq(h_2)$ is as in Definition A.4. By this and Assumption A.5, Lemmas S2.1-S2.2 imply that:

$$\left(\begin{array}{c} v_{b_n,F_{b_n}}(\theta_{b_n},\cdot) \\ \hat{h}_{2,b_n,F_{b_n}}(\theta_{b_n},\cdot) \end{array}\right) \xrightarrow{d} \left(\begin{array}{c} v_{h_2}(\cdot) \\ h_2(\cdot) \end{array}\right)$$

as stochastic processes indexed by $(x,\nu) \in \mathcal{G}$. This and Assumptions A.2 and A.6 allow us to establish AS13 (Lemmas A2-A5). In turn, these can be used to contradict $P_{F_{b_n}}(\theta_{b_n} \notin CS_{b_n}) > \alpha\ \forall n \in \mathbb{N}$, thus concluding the proof. □

Proof of Theorem 4.2. The proof of this result follows closely the arguments in AS13 (Theorem 3), with the exception of certain steps. For the sake of completeness, we sketch the main steps of the proof and point out the differences with the one in AS13.

Consider the following derivation:

$$\begin{aligned} P_F(\theta \in CS_n) &= P_F(T_n(\theta) \leq c(\varphi_n(\theta,\cdot), \hat{h}_{2,n}(\theta,\cdot), 1-\alpha+\eta)+\eta) \\ &\leq P_F(T_n(\theta) \leq c(\mathbf{0}, \hat{h}_{2,n}(\theta,\cdot), 1-\alpha+\eta)+\eta) \\ &= P_F(n^{-\chi/2}T_n(\theta) \leq n^{-\chi/2}(c(\mathbf{0}, \hat{h}_{2,n}(\theta,\cdot), 1-\alpha+\eta)+\eta)), \end{aligned}$$

where the first line holds by definition of $\hat{c}_n(\theta, 1-\alpha)$, the second line holds by definition of $\varphi_n(\theta,\cdot)$ and $c(\cdot, \hat{h}_{2,n}(\theta,\cdot), 1-\alpha+\eta)$, combined with Assumptions A.2(b) and A.6, which imply that $\varphi_n(\theta,\cdot) \geq \mathbf{0}$, and in the last line χ is as in Assumption A.2(g). The proof is completed by showing that (a) $P_F(n^{-\chi/2}T_n(\theta) \geq C) \to 1$ for some $C > 0$ and (b) $c(\mathbf{0}, \hat{h}_{2,n}(\theta,\cdot), 1-\alpha+\eta) = O_p(1)$, which imply that $n^{-\chi/2}(c(\mathbf{0}, \hat{h}_{2,n}(\theta,\cdot), 1-\alpha+\eta)+\eta) = o_p(1)$. The proof of (b) is identical to the proof in AS13 (which requires our Assumptions A.2 and A.5). On the other hand, our proof of (a) is slightly different, and so we devote the remainder of this proof to develop this argument.

By definition, $\theta \notin \Theta_S(F)$ implies that $\exists j \leq p$ s.t. $E_F[M_j(Z,\theta,x,\nu)] < 0$ for some $(x,\nu) \in \mathcal{G}$. Under Assumptions A.2(c,e,f) and A.4, we can use the arguments in the proof of Theorem A.3 to define a set $A \subset \mathcal{G}$ with positive Lebesgue measure s.t. $E_F[M_j(Z,\theta,x,\nu)] \leq -\epsilon\ \forall (x,\nu) \in A$. As a consequence,

$$S(E_F[M(Z,\theta,x,\nu)], Var_F[M(Z,\theta,x,\nu)] + \lambda D_F(\theta)) \geq \eta\ \ \forall (x,\nu) \in A$$

for some $\delta > 0$. By Assumptions A.2(a) and A.3, this implies that:

$$\int_{(x,\nu)\in A} S(D_F^{-1/2}(\theta) E_F[M(Z,\theta,x,\nu)], h_{2,F}(\theta,x,\nu) + \lambda I_{p\times p}) d\mu(x,\nu) \geq \eta\mu(A) > 0. \qquad \text{(S1.1)}$$

To complete the proof, consider the following derivation:

$$\begin{aligned} n^{-\chi/2}T_n(\theta) &= n^{-\chi/2}\int_{(x,\nu)\in\mathcal{G}} S(v_{n,F}(\theta,x,\nu)+h_{1,n,F}(\theta,x,\nu),\hat{h}_{2,n,F}(\theta,x,\nu)+\lambda I_{p\times p})d\mu(x,\nu) \\ &= \int_{(x,\nu)\in\mathcal{G}} S(n^{-1/2}v_{n,F}(\theta,x,\nu)+D_F^{-1/2}(\theta)E_F[M(Z,\theta,x,\nu)],\hat{h}_{2,n,F}(\theta,x,\nu)+\lambda I_{p\times p})d\mu(x,\nu) \\ &\geq \int_{(x,\nu)\in A} S(n^{-1/2}v_{n,F}(\theta,x,\nu)+D_F^{-1/2}(\theta)E_F[M(Z,\theta,x,\nu)],\hat{h}_{2,n,F}(\theta,x,\nu)+\lambda I_{p\times p})d\mu(g) \\ &\overset{p}{\to} \int_{(x,\nu)\in A} S(D_F^{-1/2}(\theta)E_F[M(Z,\theta,x,\nu)],h_{2,F}(\theta,x,\nu)+\lambda I_{p\times p})d\mu(x,\nu) \geq \eta\mu(A)>0, \quad \text{(S1.2)} \end{aligned}$$

where the first line holds by definition of $T_n(\theta)$, the second line holds by Assumption A.2(g) and by definition of $h_{1,n,F}(\theta,x,\nu)$, the third line holds by $A\subset\mathcal{G}$ and Assumption A.2(c), the convergence in the fourth line holds by the same argument described in the next paragraph, and the last expression is positive by Eq. (S1.1). By Eq. (S1.2), $P_F(n^{-\chi/2}T_n(\theta)\geq\eta\mu(A)/2)\to 1$ for some $\eta\mu(A)/2>0$, which implies the desired result.

To conclude the proof, it suffices to justify the convergence in the fourth line of Eq. (S1.2). For a fixed parameter $(\theta,F)\in\mathcal{F}$, Lemmas S2.1-S2.2 (see Section S2 in this supplement) imply that:

$$\left(\begin{array}{c} v_{n,F}(\theta,\cdot) \\ \hat{h}_{2,n,F}(\theta,\cdot) \end{array}\right) \overset{d}{\to} \left(\begin{array}{c} v_{h_{2,F}}(\theta,\cdot) \\ h_{2,F}(\theta,\cdot) \end{array}\right)$$

as stochastic processes indexed by $(x,\nu)\in\mathcal{G}$. In turn, this implies that:

$$\sup_{(x,\nu)\in\mathcal{G}}\left\|\left(\begin{array}{c} n^{-1/2}v_{n,F}(\theta,x,\nu) \\ \hat{h}_{2,n,F}(\theta,x,\nu) \end{array}\right)-\left(\begin{array}{c} \mathbf{0} \\ h_{2,F}(\theta,x,\nu) \end{array}\right)\right\| \overset{p}{\to} 0.$$

The convergence in the fourth line of Eq. (S1.2) is a result of this, the almost sure representation theorem, the bounded convergence theorem, and Assumption A.2(d). □

S2 Auxiliary results

Lemma S2.1. *Assume Assumption A.5 and that $\{(\theta_{k_n},F_{k_n})\in\bar{\mathcal{F}}_0\}_{n\geq1}\in SubSeq(h_2)$ for an arbitrary subsequence $\{k_n\}_{n\geq1}$ of $\{n\}_{n\geq1}$. Then,*

$$v_{k_n,F_{k_n}}(\theta_{k_n},\cdot)\overset{d}{\to}v_{h_2}(\cdot),$$

as stochastic processes indexed by $(x,\nu)\in\mathcal{G}$, where v_{h_2} is a $\mathbb{R}^p$-valued Gaussian process with zero mean and variance-covariance kernel $h_2(\cdot,\cdot)$ on $\mathcal{G}\times\mathcal{G}$.

Proof. This result follows from AS13 (Lemmas A1(a) and E3). We describe the main ideas behind these arguments for the sake of completeness. Throughout this proof, we replace the subsequence $\{k_n\}_{n\geq1}$ by the original sequence $\{n\}_{n\geq1}$ in order to simplify the notation.

Suppose that $\{(\theta_n,F_n)\in\bar{\mathcal{F}}_0\}_{n\geq1}\in SubSeq(h_2)$. By Pollard (1990, Theorem 10.2), the desired result is a consequence of the following conditions:

(1) $(\mathcal{G},\rho)$ is a totally bounded pseudo-metric space, where ρ is the following pseudo-metric:

$$\rho^2((x,\nu),(\tilde{x},\tilde{\nu})) \equiv \lim_{n\to\infty}(Trace(Var_{F_n}[D_{F_n}^{-1/2}(\theta_n)(M(Z,\theta_n,x,\nu)-M(Z,\theta_n,\tilde{x},\tilde{\nu})]),$$

(2) The finite dimensional convergence holds, i.e., $\forall(a,L)\in\mathbb{R}^p/\mathbf{0}\times\mathbb{N}$ and $\forall\{(x_s,\nu_s)\}_{s=1}^L\subset\mathcal{G}$, $\{a'v_{n,F_n}(\theta_n,x_s,\nu_s)\}_{s=1}^L$ converges in distribution to an L-dimensional Gaussian distribution with zero mean and variance covariance matrix with (s_1,s_2) component given by $a'h_2((x_{s_1},\nu_{s_1}),(x_{s_2},\nu_{s_2}))a$.

(3) $\{v_{n,F_n}(\theta_n,x,\nu):(x,\nu)\in\mathcal{G}\}_{n\geq 1}$ is stochastically equicontinuous with respect to ρ.

To prove these conditions, AS13 use the Crámer-Wold device. In particular, AS13 (Lemma A1(a)) shows that these conditions hold if, for all $a\in\mathbb{R}^p/\mathbf{0}$, the following three conditions hold:

(a) $(\mathcal{G},\rho_a)$ is a totally bounded pseudo-metric space, where ρ_a is the following pseudo-metric:

$$\rho_a^2((x,\nu),(\tilde{x},\tilde{\nu}))\equiv\lim_{n\to\infty}Var_{F_n}[D_{F_n}^{-1/2}(\theta_n)a'(M(Z,\theta_n,x,\nu)-M(Z,\theta_n,\tilde{x},\tilde{\nu})],\tag{S2.1}$$

(b) The finite dimensional convergence holds, i.e., $\forall L$ and $\forall\{(x_s,\nu_s)\}_{s=1}^L\subset\mathcal{G}$, $\{a'v_{n,F_n}(\theta_n,x_s,\nu_s)\}_{s=1}^L$ converges in distribution to an L-dimensional Gaussian distribution with zero mean and variance covariance matrix with (s_1,s_2) component given by $a'h_2((x_{s_1},\nu_{s_1}),(x_{s_2},\nu_{s_2}))a$. This convergence uniquely determines a Gaussian distribution v_a concentrated on the space of uniformly $\rho_a(\cdot)$-continuous bounded functionals on $\mathcal{G}$, $U_{\rho_a}(\mathcal{G})$,

(c) $a'v_{n,F_n}(\theta_n,\cdot)\overset{d}{\to}v_a$.

To prove conditions (a)-(c), we rely on AS13 (Lemma E3), which extends Pollard (1990, Theorem 10.6, page 53) to triangular array stochastic processes. Fix $a\in\mathbb{R}^p/\mathbf{0}$ and $(x,\nu),(\tilde{x},\tilde{\nu})\in\mathcal{G}$ arbitrarily and define:

$$\begin{aligned}f_{a,n,i}(\omega,x,\nu) &\equiv n^{-1/2}a'D_{F_n}^{-1/2}(\theta_n)(M_n(Z_i,\theta_n,x,\nu)-E_{F_n}[M_n(Z_i,\theta_n,x,\nu)]),\\ \rho_{n,a}^2((x,\nu),(\tilde{x},\tilde{\nu})) &\equiv n\,E_{F_n}[f_{a,n,i}(\omega,x,\nu)-f_{a,n,i}(\omega,\tilde{x},\tilde{\nu})]^2.\end{aligned}\tag{S2.2}$$

By definition, notice that $a'v_{n,F_n}(\theta_n,x,\nu)=\sum_{i=1}^n f_{a,n,i}(\omega,x,\nu)$. AS13 (Lemma E3) show that conditions (a)-(c) hold provided that, $\forall a\in\mathbb{R}^p/\mathbf{0}$, the following results hold:

(i) $\{f_{a,n,i}(\omega,x,\nu):(x,\nu)\in\mathcal{G}\}_{i=1}^n$ is manageable with respect to some envelopes $\{F_{a,n,i}(\omega)\}_{i=1}^n$,

(ii) $\lim_{n\to\infty}E_{F_n}[f_{a,n,i}(\omega,x,\nu)f_{a,n,i}(\omega,\tilde{x},\tilde{\nu})]=a'h_2((x,\nu),(\tilde{x},\tilde{\nu}))a$ for all $(x,\nu),(\tilde{x},\tilde{\nu})\in\mathcal{G}$,

(iii) $\limsup_{n\to\infty}\sum_{i=1}^n E_{F_n}[F_{a,n,i}^2]<\infty$,

(iv) $\lim_{n\to\infty}\sum_{i=1}^n E_{F_n}[F_{a,n,i}^2 1[F_{a,n,i}>\varepsilon]]=0$ for all $\varepsilon>0$,

(v) The pseudo-metric ρ_a in Eq. (S2.1) satisfies $\rho_a((x,\nu),(\tilde{x},\tilde{\nu}))\equiv\lim_{n\to\infty}\rho_{n,a}((x,\nu),(\tilde{x},\tilde{\nu}))$ for all $(x,\nu),(\tilde{x},\tilde{\nu})\in\mathcal{G}$ and, for all deterministic sequences $\{(x_n,\nu_n)\in\mathcal{G}\}_{n\geq1}$ and $\{(\tilde{x}_n,\tilde{\nu}_n)\in\mathcal{G}\}_{n\geq1}$, $\rho_a((x_n,\nu_n),(\tilde{x}_n,\tilde{\nu}_n))\to0$ implies that $\rho_{n,a}((x_n,\nu_n),(\tilde{x}_n,\tilde{\nu}_n))\to0$,

The verification of these conditions is similar to that in AS13.

Condition (i). By Assumption A.5, $\{a'M(Z_i,\theta,x,\nu):(x,\nu)\in\mathcal{G}\}_{i=1}^n$ is manageable with respect to the envelopes $\{a'M(Z_i,\theta)\}_{i=1}^n$. By the definitions in Eq. (S2.2) and AS13 (Lemma E1), it then follows that $\{f_{a,n,i}(\omega,x,\nu):(x,\nu)\in\mathcal{G}\}_{i=1}^n$ is manageable with respect to envelopes $\{F_{a,n,i}(\omega)\}_{i=1}^n$ defined as follows:

$$F_{a,n,i}(\omega)\equiv n^{-1/2}a'D_{F_n}^{-1/2}(\theta_n)(M_n(Z_i,\theta_n)+E_{F_n}[M_n(Z_i,\theta_n)]).$$

Condition (ii)-(v). While the definitions of our stochastic processes and envelopes are slightly different from those in AS13, one can still complete this proof by using similar arguments to those in AS13 (Lemma E3). □

Lemma S2.2. *Assume Assumption A.5 and that $\{(\theta_{k_n},F_{k_n})\in\bar{\mathcal{F}}_0\}_{n\geq 1}\in SubSeq(h_2)$ for an arbitrary subsequence $\{k_n\}_{n\geq1}$ of $\{n\}_{n\geq1}$. Then,*

$$\sup_{(x_n,\nu_n),(\tilde{x}_n,\tilde{\nu}_n)\in\mathcal{G}}||\hat{h}_{2,k_n,F_{k_n}}(\theta_{k_n},(x_n,\nu_n),(\tilde{x}_n,\tilde{\nu}_n))-h_2((x_n,\nu_n),(\tilde{x}_n,\tilde{\nu}_n))||\overset{p}{\to}0.$$

Proof. This result follows from AS13 (Lemmas A1(b)). We describe the main ideas behind these arguments for the sake of completeness. Throughout this proof, we replace the subsequence $\{k_n\}_{n\geq1}$ by the original sequence $\{n\}_{n\geq1}$ in order to simplify the notation.

Consider the following derivation:

$$\sup_{(x,\nu),(\tilde{x},\tilde{\nu})\in\mathcal{G}}||\hat{h}_{2,n,F_n}((x,\nu),(\tilde{x},\tilde{\nu}))-h_2((x,\nu),(\tilde{x},\tilde{\nu}))||\leq$$
$$\left\{\begin{array}{c}\sup_{(x,\nu),(\tilde{x},\tilde{\nu})\in\mathcal{G}}||\hat{h}_{2,n,F_n}((x,\nu),(\tilde{x},\tilde{\nu}))-h_{2,F_n}((x,\nu),(\tilde{x},\tilde{\nu}))||\\+\sup_{(x,\nu),(\tilde{x},\tilde{\nu})\in\mathcal{G}}||h_{2,F_n}((x,\nu),(\tilde{x},\tilde{\nu}))-h_2((x,\nu),(\tilde{x},\tilde{\nu}))||\end{array}\right\}.$$

The RHS is a sum of two terms. By $\{(\theta_n,F_n)\in\bar{\mathcal{F}}_0\}_{n\geq1}\in SubSeq(h_2)$, the second term converges to zero. Hence, it suffices to show that the first term is $o_p(1)$.

For any $s_1,s_2=1,\ldots,p$, the (s_1,s_2)-component of $\hat{h}_{2,n,F_n}((x,\nu),(\tilde{x},\tilde{\nu}))$ is given by:

$$\begin{aligned}&\hat{h}_{2,n,F_n}((x,\nu),(\tilde{x},\tilde{\nu}))_{(s_1,s_2)}\\&=n^{-1}\sigma_{s_1}^{-1}(\theta_n)\sigma_{s_2}^{-1}(\theta_n)\sum_{i=1}^n(M_{s_1}(Z_i,\theta_n,x,\nu)-\bar{M}_{n,s_1}(\theta_n,x,\nu))(M_{s_2}(Z_i,\theta_n,\tilde{x},\tilde{\nu})-\bar{M}_{n,s_2}(\theta_n,\tilde{x},\tilde{\nu}))\\&=n^{-1}\sum_{i=1}^n f_{n,i,s_1,s_2}^{mm}(\omega,(x,\nu),(\tilde{x},\tilde{\nu}))-\left(n^{-1}\sum_{i=1}^n f_{n,i,s_1}^{m}(\omega,x,\nu)\right)\left(n^{-1}\sum_{i=1}^n f_{n,i,s_2}^{m}(\omega,\tilde{x},\tilde{\nu})\right).\end{aligned}$$

where we have relied on the i.i.d. assumption implicit in $(\theta_n,F_n)\in\bar{\mathcal{F}}_0$ and the following definitions:

$$\begin{aligned}f_{n,i,s}^{m}(\omega,x,\nu)&\equiv M_s(Z_i,\theta_n,x,\nu)-E_{F_n}[M_s(Z_i,\theta_n,x,\nu)],\\f_{n,i,s,\check{s}}^{mm}(\omega,(x,\nu),(\tilde{x},\tilde{\nu}))&\equiv f_{n,i,s}^{m}(\omega,x,\nu)\times f_{n,i,\check{s}}^{m}(\omega,\tilde{x},\tilde{\nu}).\end{aligned}$$

Notice that, by definition, $E_{F_n}[f_{n,i,s}^m(\omega,x,\nu)]=E_{F_n}[f_{n,i,\check{s}}^m(\omega,\tilde{x},\tilde{\nu})]=0$ and $E_{F_n}[f_{n,i,s,\check{s}}^{mm}(\omega,(x,\nu),(\tilde{x},\tilde{\nu}))]=$

$h_{2,F_n}((x,\nu),(\tilde{x},\tilde{\nu}))_{(s,\check{s})}$. Based on this argument, the desired result follows from proving that $\forall s,\check{s}=1,\ldots,p$,

$$\sup_{(x,\nu)\in\mathcal{G}}\left\|n^{-1}\sum_{i=1}^{n}f^{m}_{n,i,s}(\omega,x,\nu)-E_{F_n}[f^{m}_{n,i,s}(\omega,x,\nu)]\right\|\overset{p}{\to}0,$$
$$\sup_{(x,\nu),(\tilde{x},\tilde{\nu})\in\mathcal{G}}\left\|n^{-1}\sum_{i=1}^{n}f^{mm}_{n,i,s,\check{s}}(\omega,(x,\nu),(\tilde{x},\tilde{\nu}))-E_{F_n}[f^{mm}_{n,i,s,\check{s}}(\omega,(x,\nu),(\tilde{x},\tilde{\nu}))]\right\|\overset{p}{\to}0.$$

To complete this task we rely on AS13 (Lemma E2), which extends Pollard (1990, Theorem 8.2) to triangular array stochastic processes. This result requires that, for arbitrary $s,\check{s}=1,\ldots,p$, we verify certain conditions on the following triangular array of processes:

(i) $\{\{f^{m}_{n,i,s}(\omega,x,\nu)\ :\ (x,\nu)\in\mathcal{G}\}_{i=1}^{n}\}_{n\geq1}$,

(ii) $\{\{f^{mm}_{n,i,s,\check{s}}(\omega,(x,\nu),(\tilde{x},\tilde{\nu}))\ :\ (x,\nu),(\tilde{x},\tilde{\nu})\in\mathcal{G}\}_{i=1}^{n}\}_{n\geq1}$.

Conditions for (i). By Assumption A.5, $\{M(Z_i,\theta,x,\nu):(x,\nu)\in\mathcal{G}\}_{i=1}^{n}$ is manageable with respect to the envelopes $\{M(Z_i,\theta)\}_{i=1}^{n}$. From this, it follows that $\{M_s(Z_i,\theta,x,\nu):(x,\nu)\in\mathcal{G}\}_{i=1}^{n}$ is manageable with respect to the envelopes $\{M_s(Z_i,\theta)\}_{i=1}^{n}$. By AS13 (Lemma E1), it then follows that $\{f^{m}_{n,i,s}(\omega,x,\nu):(x,\nu)\in\mathcal{G}\}_{i=1}^{n}$ is manageable with respect to envelopes $\{F_{n,i,s}(\omega)\}_{i=1}^{n}$ defined as follows:

$$F_{n,i,s}(\omega)\equiv\sigma_s^{-1}(\theta_n)(M_s(Z_i,\theta_n)+E_{F_n}[M_s(Z_i,\theta_n)]).\tag{S2.3}$$

To complete the argument, it suffices to show that $n^{-1}\sum_{i=1}^{n}E_{F_n}[F_{n,i,s}^{1+\eta}]\leq\check{K}$ for some $\check{K}<\infty$, $\eta>0$, and all $n\in\mathbb{N}$. For this purpose, consider the following derivation for $\eta=1+\delta$ with $\delta>0$ as in Definition A.1:

$$E_{F_n}[F_{n,i,s}^{2+\delta}]=E_{F_n}[(\sigma_s^{-1}(\theta_n)(M_s(Z_i,\theta_n)+E_{F_n}[M_s(Z_i,\theta_n)]))^{2+\delta}]\leq2^{2+\delta}E_{F_n}[|\sigma_s^{-1}(\theta_n)M_s(Z_i,\theta_n)|^{2+\delta}],$$

where the equality holds by Eq. (S2.3), the inequality holds by the convexity of $x^{2+\delta}$. The desired result then follows immediately from $(\theta_n,F_n)\in\bar{\mathcal{F}}_0$, as this implies that $F_{n,i,s}^{2+\delta}$ is i.i.d. and that $E_{F_n}[|\sigma_j^{-1}(\theta_n)M_{n,j}(Z,\theta_n)|^{2+\delta}]<K$ for all $j=1,\ldots,p$ and $n\in\mathbb{N}$.

Conditions for (ii). By our previous verification, $\{f^{m}_{n,i,s}(\omega,(x,\nu))\ :\ (x,\nu)\ \in\ \mathcal{G}\}_{i=1}^{n}$ is manageable with respect to envelopes $\{F_{n,i,s}(\omega)\}_{i=1}^{n}$ with $F_{n,i,s}(\omega)$ as in Eq. (S2.3) for $s=1,\ldots,p$. From this, $f^{mm}_{n,i,s,\check{s}}(\omega,(x,\nu),(\tilde{x},\tilde{\nu}))\equiv f^{m}_{n,i,s}(\omega,x,\nu)f^{m}_{n,i,\check{s}}(\omega,\tilde{x},\tilde{\nu})$, and the arguments in the proof of AS13 (Lemma A1(b)), it then follows that $\{f^{mm}_{n,i,s,\check{s}}(\omega,(x,\nu),(\tilde{x},\tilde{\nu})):(x,\nu),(\tilde{x},\tilde{\nu})\in\mathcal{G}\}_{i=1}^{n}$ is manageable with respect to envelopes $\{F_{n,i,s,\check{s}}(\omega)\}_{i=1}^{n}$ defined by:

$$F_{n,i,s,\check{s}}(\omega)\equiv\sigma_s^{-1}(\theta_n)\sigma_{\check{s}}^{-1}(\theta_n)(M_s(Z_i,\theta_n)+E_{F_n}[M_s(Z_i,\theta_n)])(M_{\check{s}}(Z_i,\theta_n)+E_{F_n}[M_{\check{s}}(Z_i,\theta_n)]).\tag{S2.4}$$

To complete the argument, it suffices to show that $n^{-1}\sum_{i=1}^{n}E_{F_n}[F_{n,i,s,\check{s}}^{2+\delta/2}]\leq\check{K}$ for some $\check{K}<\infty$, $\eta>0$, and all $n\in\mathbb{N}$. For this purpose, consider the following derivation for $\eta=1+\delta/2$ with $\delta>0$ as in Definition A.1:

$$\begin{aligned}E_{F_n}[F_{n,i,s,\check{s}}^{2+\delta/2}]&=E_{F_n}[(\sigma_s^{-1}(\theta_n)\sigma_{\check{s}}^{-1}(\theta_n)(M_s(Z_i,\theta_n)+E_{F_n}[M_s(Z_i,\theta_n)])(M_{\check{s}}(Z_i,\theta_n)+E_{F_n}[M_{\check{s}}(Z_i,\theta_n)]))^{2+\delta/2}]\\&\leq4^{2+\delta}E_{F_n}[|\sigma_s^{-1}(\theta_n)M_s(Z_i,\theta_n)||\sigma_{\check{s}}^{-1}(\theta_n)M_{\check{s}}(Z_i,\theta_n)|^{2+\delta/2}]\\&\leq4^{2+\delta}\{E_{F_n}[|\sigma_s^{-1}(\theta_n)M_s(Z_i,\theta_n)|^{2+\delta}]\}^{(2+\delta/2)/(2+\delta)}\{E_{F_n}[|\sigma_{\check{s}}^{-1}(\theta_n)M_{\check{s}}(Z_i,\theta_n)|^{2+\delta}]\}^{(2+\delta/2)/(2+\delta)},\end{aligned}$$

where the first line holds by Eq. (S2.4), the second line holds by the convexity of $x^{2+\delta/2}$, and the third line follows from Hölder's inequality. The desired result then follows immediately from $(\theta_n,F_n)\in\bar{\mathcal{F}}_0$, as this implies that $F_{n,i,s,\check{s}}^{2+\delta/2}$ is i.i.d. and that $E_{F_n}[|\sigma_j^{-1}(\theta_n)M_{n,j}(Z,\theta_n)|^{2+\delta}] < K$ for all $j = 1, \ldots, p$ and $n \in \mathbb{N}$. $\square$

References

Andrews, D. W. K. and X. Shi (2013): "Inference Based on Conditional Moment Inequalities," *Econometrica*, 81, 609–666.

Aucejo, E. M., F. A. Bugni, and V. J. Hotz (2015): "Identification and Inference on Regressions with Missing Covariate Data," Mimeo: London School of Economics and Political Science and Duke University.

Pollard, D. (1990): *Empirical Processes: Theory and Applications*, NSF-CBMS Regional Conference Series in Probability and Statistics, Volume 2.