Online supplement to: Controlling for ability using test scores

Benjamin Williams *

October 31, 2018

Contents

*George Washington University, Monroe Hall 309, 2115 G Street NW, Washington, DC 20052, bdwilliams@gwu.edu

B Proofs

*Proof of Theorem **??**.* I consider a sequence $\{J_n : n \geq 1\}$ and show that $plim_{n\to\infty}\hat{\beta}_{1J_n} = \beta_{10}$. First, $m^*(V_i^*, \beta_1, \mathbf{h})$ is linear in β_1 so

$$\begin{aligned}
&M^*(\hat{\beta}_{1J_n}, \mathbf{h}_0, \tau_{0,J_n}) - M^*(\beta_1, \mathbf{h}_0, \tau_{0,J_n}) \\
&= E(\tau_{0,J_n}(\theta_i)(Y_i - h_{y,0}(\theta_i) - \hat{\beta}'_{1J_n}(X_i - h_{x,0}(\theta_i)))(X_i - h_{x,0}(\theta_i))) \\
&- E(\tau_{0,J_n}(\theta_i)(Y_i - h_{y,0}(\theta_i) - \beta'_1(X_i - h_{x,0}(\theta_i)))(X_i - h_{x,0}(\theta_i))) \\
&= -Q^*_{0,J_n}\left(\hat{\beta}_{1J_n} - \beta_1\right)
\end{aligned}$$

Since, by assumption, equation (**??**) holds for β_{10} with $E(e_i \mid X_i, \theta_i) = 0$, $Y_i - h_{y,0}(\theta_i) - \beta'_{10}(X_i - h_{x,0}(\theta_i)) = e_i$ and

$$\begin{aligned}
M^*(\beta_{10}, \mathbf{h}_0, \tau_{0,J}) &= E(\tau_{0,J}(\theta_i)e_i(X_i - h_{x,0}(\theta_i))) \\
&= E(\tau_{0,J}(\theta_i)E(e_i \mid X_i, \theta_i)(X_i - h_{x,0}(\theta_i))) = 0
\end{aligned}$$

Then, since $Q^*_{0,J}$ is invertible by Assumption **??**(b), $\hat{\beta}_{1J_n} - \beta_{10} = -Q^{*-1}_{0,J_n}M^*(\hat{\beta}_{1J_n}, \mathbf{h}_0, \tau_{0,J_n})$ and hence for any $\epsilon > 0$

$$Pr(|\hat{\beta}_{1J_n} - \beta_{10}| > \epsilon) \leq Pr(||Q^{*-1}_{0,J_n}|| \cdot |M^*(\hat{\beta}_{1J_n}, \mathbf{h}_0, \tau_{0,J_n})| \geq \epsilon) \leq Pr(|M^*(\hat{\beta}_{1J_n}, \mathbf{h}_0, \tau_{0,J_n})| \geq c\epsilon)$$

where the second inequality follows from the bound on $||Q^{*-1}_{0,J}||$ provided by Assumption **??**(b). It will thus be sufficient to show that $Pr(|M^*(\hat{\beta}_{1J_n}, \mathbf{h}_0, \tau_{0,J_n})| \geq \varepsilon) \to 0$ as $n \to \infty$ for all $\varepsilon > 0$.

By the triangle inequality, $|M^*(\hat{\beta}_{1J_n}, \mathbf{h}_0, \tau_{0,J_n})| \leq |\hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w}) - M^*(\hat{\beta}_{1J_n}, \mathbf{h}_0, \tau_{0,J_n})| + |\hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w})|$. I will first show that

$$|\hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w}) - M^*(\hat{\beta}_{1J_n}, \mathbf{h}_0, \tau_{0,J_n})| = o_p(1) \tag{B.1}$$

using the following decomposition

$$\begin{aligned}
&|\hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w}) - M^*(\hat{\beta}_{1J_n}, \mathbf{h}_0, \tau_{0,J_n})| \qquad \text{(B.2)}\\
&\leq |M(\hat{\beta}_{1J_n}, \tilde{\mathbf{g}}, w_{0,J_n}) - M^*(\hat{\beta}_{1J_n}, \mathbf{h}_0, \tau_{0,J_n})| + |M(\hat{\beta}_{1J_n}, \tilde{\mathbf{g}}, \hat{w}) - M(\hat{\beta}_{1J_n}, \tilde{\mathbf{g}}, w_{0,J_n})| \\
&+ |M(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w}) - M(\hat{\beta}_{1J_n}, \tilde{\mathbf{g}}, \hat{w})| + |\hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w}) - M(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w})|
\end{aligned}$$

Define $\xi_i(\beta_1) = Y_i - \beta'_1 X_i$ and $\gamma(\beta_1) = (1, \beta'_1)'$ and let $\hat{\xi}_i = \xi_i(\hat{\beta}_{1J})$ and $\hat{\gamma} = \gamma(\hat{\beta}_{1J})$. Then

$M(\beta_1, \mathbf{g}, w) = E(w(\bar{M}_{iJ})(\xi_i(\beta_1) - \gamma(\beta_1)'\mathbf{g}(\bar{M}_{iJ}))(X_i - g_x(\bar{M}_{iJ})))$ so

$$\begin{aligned}
|M(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w}) - M(\hat{\beta}_{1J_n}, \tilde{\mathbf{g}}, \hat{w})| &\leq |E(\hat{w}(\bar{M}_{iJ_n})\hat{\xi}_i(\hat{g}_x(\bar{M}_{iJ_n}) - \tilde{g}_x(\bar{M}_{iJ_n})))| \\
&\quad + |E(\hat{w}(\bar{M}_{iJ_n})\gamma'(\hat{\mathbf{g}}(\bar{M}_{iJ_n}) - \tilde{\mathbf{g}}(\bar{M}_{iJ_n}))X_i)| \\
&\quad + |E(\hat{w}(\bar{M}_{iJ_n})\hat{\gamma}'(\hat{\mathbf{g}}(\bar{M}_{iJ_n})\hat{g}_x(\bar{M}_{iJ_n}) - \tilde{\mathbf{g}}(\bar{M}_{iJ_n})\tilde{g}_x(\bar{M}_{iJ_n})))| \\
&\leq C\left(1 + \sup_{m\in\hat{\mathcal{M}}} |\hat{\mathbf{g}}(m)| + \sup_{m\in\hat{\mathcal{M}}} |\tilde{\mathbf{g}}(m)|\right) \sup_{m\in\hat{\mathcal{M}}} |\hat{\mathbf{g}}(m) - \tilde{\mathbf{g}}(m)|
\end{aligned} \tag{B.3}$$

where the second inequality follows for some constant $C > 0$ by (a), (c), and (d) of Assumption **??**. Then, using Theorem C.1 and Assumption **??**(d), for any $\varepsilon > 0$,

$$\begin{aligned}
&Pr(|M(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w}) - M(\hat{\beta}_{1J_n}, \tilde{\mathbf{g}}, \hat{w})| \geq \varepsilon) \\
&\leq Pr\left(C\left(1 + \sup_{m\in\mathcal{M}_{\delta_1}} |\hat{\mathbf{g}}(m)| + \sup_{m\in\mathcal{M}_{\delta_1}} |\tilde{\mathbf{g}}(m)|\right) \sup_{m\in\mathcal{M}_{\delta_1}} |\hat{\mathbf{g}}(m) - \tilde{\mathbf{g}}(m)| > \varepsilon\right) \\
&+ \left(1 - Pr(\hat{\mathcal{M}} \subset \mathcal{M}_{\delta_1})\right) \to 0
\end{aligned}$$

Next,

$$\begin{aligned}
&|M(\hat{\beta}_{1J_n}, \tilde{\mathbf{g}}, w_{0,J_n}) - M^*(\hat{\beta}_{1J_n}, \mathbf{h}_0, \tau_{0,J_n})| \\
&\quad \leq E\Big(|w_{0,J_n}(\bar{M}_{iJ_n})| \Big|(\hat{\xi}_i - \hat{\gamma}'\tilde{\mathbf{g}}(\bar{M}_{iJ_n}))(X_i - \tilde{g}_x(\bar{M}_{iJ_n})) \\
&\qquad\qquad - (\hat{\xi}_i - \hat{\gamma}'\tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i)))(X_i - \tilde{g}_x(\bar{p}_{J_n}(\theta_i)))\Big|\Big) \\
&\quad + E(|w_{0,J_n}(\bar{M}_{iJ_n}) - w_{0,J_n}(\bar{p}_{J_n}(\theta_i))||(\hat{\xi}_i - \hat{\gamma}'\tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i)))(X_i - \tilde{g}_x(\bar{p}_{J_n}(\theta_i)))|)
\end{aligned} \tag{B.4}$$

By Assumption **??**(a), the first term in equation (B.4) is bounded by a positive constant times

$$\begin{aligned}
&E(|w_{0,J_n}(\bar{M}_{iJ_n})(|X_i| + |Y_i|)|\tilde{\mathbf{g}}(\bar{M}_{iJ_n}) - \tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i))|) \\
&+ E(|w_{0,J_n}(\bar{M}_{iJ_n})|\tilde{\mathbf{g}}(\bar{M}_{iJ_n})||\tilde{\mathbf{g}}(\bar{M}_{iJ_n}) - \tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i))|) \\
&+ E(|w_{0,J_n}(\bar{M}_{iJ_n})|\tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i))||\tilde{\mathbf{g}}(\bar{M}_{iJ_n}) - \tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i))|)
\end{aligned}$$

It can be shown that each of these three terms can be bounded by an $o(1)$ sequence using essentially the same argument. First, by conditions (a) and (b) of Assumption **??**, $\tilde{\mathbf{g}}$ is continuously differentiable so $|\tilde{\mathbf{g}}(\bar{M}_{iJ_n}) - \tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i))| \leq |D\tilde{\mathbf{g}}(p_i^*)||\eta_i|$ for some p_i^* between $\bar{p}_{J_n}(\theta_i)$ and $\bar{M}_{iJ_n}$. Next, by Assumption **??**(e), $|D\tilde{\mathbf{g}}(m)| \leq \bar{D}(\bar{p}_{J_n}^{-1}(m))$ where $\bar{D}(\cdot)$ is nonincreasing on the interval $(-\infty, q_{\delta_2}(\theta_i)]$ and nondecreasing on $[q_{1-\delta_2}(\theta_i), \infty)$. If $w_{0,J_n}(\bar{M}_{iJ_n}) > 0$ and

$\theta_i \in \Theta_{\delta_2}$ then $p_i^* \in \mathcal{M}_{\delta_2}$. If $w_{0,J_n}(\bar{M}_{iJ_n}) > 0$ and $\theta_i \notin \Theta_{\delta_2}$ then $\bar{D}(\bar{p}_{J_n}^{-1}(p_i^*)) \leq \bar{D}(\theta_i)$. Then

$$
\begin{aligned}
& E(|w_{0,J_n}(\bar{M}_{iJ_n})(|X_i| + |Y_i|)|\tilde{\mathbf{g}}(\bar{M}_{iJ_n}) - \tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i))|) \\
& \leq \int_{-\infty}^{q_{\delta_2}(\theta_i)} B\bar{D}(t)E(|X_i| + |Y_i| \mid \theta_i = t)f_\theta(t)E(|\eta_i| \mid \theta_i = t)dt \\
& + \int_{\Theta_{\delta_2}} B\left(\sup_{m \in \mathcal{M}_{\delta_2}} |D\tilde{\mathbf{g}}(m)|\right) E(|X_i| + |Y_i| \mid \theta_i = t)E(|\eta_i| \mid \theta_i = t)f_\theta(t)dt \\
& + \int_{q_{1-\delta_2}(\theta_i)}^{\infty} B\bar{D}(t)E(|X_i| + |Y_i| \mid \theta_i = t)E(|\eta_i| \mid \theta_i = t)f_\theta(t)dt \\
& \leq \frac{B}{J_n^{1/2}}\left\{E\left(\bar{D}(\theta_i)(|X_i| + |Y_i|)\right) + \left(\sup_{m \in \mathcal{M}_{\delta_2}} |D\tilde{\mathbf{g}}(m)|\right) E(|X_i| + |Y_i|)\right\} = o(1)
\end{aligned}
$$

where the second inequality follows since $\sup_{t\in\mathbb{R}} E(|\eta_i| \mid \theta_i = t) \leq (\sup_{t\in\mathbb{R}} E(\eta_i^2 \mid \theta_i = t))^{1/2} \leq J_n^{-1/2}$ and the final equality follows because $\sup_{m\in\mathcal{M}_{\delta_2}} |D\tilde{\mathbf{g}}(m)| = O(1)$ by Theorem C.1 and $E\left(\bar{D}(\theta_i)(|X_i| + |Y_i|)\right)$ and $E(|X_i| + |Y_i|)$ are both bounded by (c) and (e) of Assumption **??**.

The second term in equation (B.4) is bounded by

$$
\begin{aligned}
& BE(\mathbf{1}(\theta_i \in \Theta_{\delta_2})|\eta_i||(\hat{\xi}_i - \hat{\gamma}'\tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i)))(X_i - \tilde{g}_x(\bar{p}_{J_n}(\theta_i)))|) \\
& \leq \frac{B}{J_n^{1/2}} E(\mathbf{1}(\theta_i \in \Theta_{\delta_2})|(\hat{\xi}_i - \hat{\gamma}'\tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i)))(X_i - \tilde{g}_x(\bar{p}_{J_n}(\theta_i)))|) \\
& \leq \frac{B}{J_n^{1/2}} \sup_{\beta_1 \in \mathcal{B}, m \in \mathcal{M}_{\delta_2}} E(|(\xi_i(\beta_1) - \gamma(\beta_1)'\tilde{\mathbf{g}}(m))(X_i - \tilde{g}_x(m))|) = o(1),
\end{aligned}
$$

where convergence follows by Assumption **??**(c) and Theorem C.1.

Next, the second term in equation (B.2) can be bounded as follows.

$$
\begin{aligned}
& |M(\hat{\beta}_{1J_n}, \tilde{\mathbf{g}}, \hat{w}) - M(\hat{\beta}_{1J_n}, \tilde{\mathbf{g}}, w_{0,J_n})| \\
& \leq E(|\hat{w}(\bar{M}_{iJ_n}) - w_{0,J_n}(\bar{M}_{iJ_n})||(\hat{\xi}_i - \hat{\gamma}'\tilde{\mathbf{g}}(\bar{M}_{iJ_n}))(X_i - \tilde{g}_x(\bar{M}_{iJ_n}))|) \\
& \leq C\left(1 + \left(\sup_{m \in \hat{\mathcal{M}} \cup \mathcal{M}_{\delta_2}} |\hat{\mathbf{g}}(m)|\right)^2\right) \sup_{m \in \hat{\mathcal{M}} \cup \mathcal{M}_{\delta_2}} |\hat{w}(m) - w_{0,J_n}(m)| = o_p(1)
\end{aligned}
$$

where the second inequality follows for some constant $C > 0$ by conditions (a) and (c) of Assumption **??** and the final equality follows by using Theorem C.1 and Assumption **??**(d) as above since $\delta_2 \geq \delta_1$ implies that $\mathcal{M}_{\delta_2} \subseteq \mathcal{M}_{\delta_1}$ and hence $Pr(\hat{\mathcal{M}} \cup \mathcal{M}_{\delta_2} \subset \mathcal{M}_{\delta_1}) = Pr(\hat{\mathcal{M}} \subset \mathcal{M}_{\delta_1}) \to 0$.

Next, $|\hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w}) - M(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w})| = o_p(1)$ by applying Theorem B.2. Let $\Gamma_n =$

$\mathcal{B} \times \{(w, \mathbf{g}) : w(m) = 0 \forall m \notin \mathcal{M}_{\delta_1}, \sup_{m \in \mathcal{M}_{\delta_1}} |\mathbf{g}(m)| < B, \sup_{m \in \mathcal{M}_{\delta_1}} |D\mathbf{g}(m)| < B,$ $\sup_{m \in \mathcal{M}_{\delta_1}} |w(m)| < B, \sup_{m \in \mathcal{M}_{\delta_1}} |Dw(m)| < B\}$. Define the metric $d_n((\beta_1', \mathbf{g}', w'), (\beta_1, \mathbf{g}, w)) = |\beta_1' - \beta_1| + \sup_{m \in \mathcal{M}_{\delta_1}} |\mathbf{g}'(m) - \mathbf{g}(m)| + \sup_{m \in \mathcal{M}_{\delta_1}} |w'(m) - w(m)|$. Both Γ_n and d_n vary with n because $\mathcal{M}_{\delta_1}$ varies with J_n.

The space Γ_n is uniformly totally bounded because Θ_{δ_1} is compact and because of the conditions in Assumption **??**(b) controlling $\{\bar{p}_J : J \geq J_0\}$. Condition (c) in Theorem B.2 is satisfied under conditions (a) and (c) of Assumption **??**. Condition (b) in the theorem follows from Theorem B.1 since the random variable $|w(\bar{M}_{iJ_n})(\xi_i(\beta_1) - \gamma(\beta_1)'\mathbf{g}(\bar{M}_{iJ_n}))(X_i - g_x(\bar{M}_{iJ_n}))|$ is bounded by a random variable that has finite absolute mean when $(\beta_1, \mathbf{g}, w) \in \Gamma_n$ by condition (c) of Assumption **??**. Lastly, $Pr((\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w}) \in \Gamma_n) \leq Pr(\sup_{m \in \mathcal{M}_{\delta_1}} |\hat{\mathbf{g}}(m)| < B) + Pr(\sup_{m \in \mathcal{M}_{\delta_1}} |D\hat{\mathbf{g}}(m)| < B) + Pr(\sup_{m \notin \mathcal{M}_{\delta_1}} |\hat{w}(m)| = 0, |\hat{w}(m)| \leq B, |D\hat{w}(m)| \leq B)$ and each of the first two terms converges to 1 by Theorem C.1 and the third converges to 1 by Assumption **??**(d).

Thus, I have shown that

$$|M^*(\hat{\beta}_{1J_n}, \mathbf{h}_0, \tau_{0,J_n})| \leq o_p(1) + |\hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w})|$$

But, $|\hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w})| = \inf_{\beta_1 \in \mathcal{B}} |\hat{M}_n(\beta_1, \hat{\mathbf{g}}, \hat{w})|$ and

$$\begin{aligned}
\inf_{\beta_1 \in \mathcal{B}} |\hat{M}_n(\beta_1, \hat{\mathbf{g}}, \hat{w})| &\leq \inf_{\beta_1 \in \mathcal{B}} \left\{ \left| \hat{M}_n(\beta_1, \hat{\mathbf{g}}, \hat{w}) - M^*(\beta_1, \mathbf{h}_0, \tau_{0,J_n}) \right| + M^*(\beta_1, \mathbf{h}_0, \tau_{0,J_n}) \right\} \\
&\leq \inf_{\beta_1 \in \mathcal{B}} \left| \hat{M}_n(\beta_1, \hat{\mathbf{g}}, \hat{w}) - M^*(\beta_1, \mathbf{h}_0, \tau_{0,J_n}) \right| + \inf_{\beta_1 \in \mathcal{B}} |M^*(\beta_1, \mathbf{h}_0, \tau_{0,J_n})| \\
&\leq \left| \hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w}) - M^*(\hat{\beta}_{1J_n}, \mathbf{h}_0, \tau_{0,J_n}) \right| + \inf_{\beta_1 \in \mathcal{B}} |M^*(\beta_1, \mathbf{h}_0, \tau_{0,J_n})| \\
&= o_p(1) + \inf_{\beta_1 \in \mathcal{B}} |M^*(\beta_1, \mathbf{h}_0, \tau_{0,J_n})| = o_p(1),
\end{aligned}$$

where the third inequality follows since $\hat{\beta}_{1J_n} \in \mathcal{B}$, the first equality follows from (B.1), and the second equality follows because $\beta_{10} \in \mathcal{B}$ and $M^*(\beta_{10}, \mathbf{h}_0, \tau_{0,J_n}) = 0$.

Therefore,

$$|M^*(\hat{\beta}_{1J_n}, \mathbf{h}_0, \tau_{0,J_n})| \leq o_p(1) + |\hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w})| = o_p(1)$$

□

Theorem **??** can be proved through a few lemmas. Let $\hat{Z}_n(\mathbf{g}, w) = n^{-1} \sum_{i=1}^n w(\bar{M}_i)(X_i - g_x(\bar{M}_i))(Z_i - \mathbf{g}(\bar{M}_i))'$ and $\hat{Z}_n^*(\mathbf{h}, \tau) = n^{-1} \sum_{i=1}^n \tau(\theta_i)(X_i - h_x(\theta_i))(Z_i - \mathbf{h}(\theta_i))'$. Then $\hat{M}_n(\beta_1, \mathbf{g}, w) =$

$\hat{Z}_n(\mathbf{g}, w)\gamma(\beta_1)'$ where $\gamma(\beta_1) = (1, \beta_1')'$. I can also define $\hat{M}_n^*(\beta_1, \mathbf{h}, \tau) = \hat{Z}_n^*(\mathbf{h}, \tau)\gamma(\beta_1)'$ and

$$\hat{Q}_n(\mathbf{g}, w) := n^{-1} \sum_{i=1}^{n} w(\bar{M}_i)(X_i - g_x(\bar{M}_i))(X_i - g_x(\bar{M}_i))' = \hat{Z}_n(\mathbf{g}, w)A'$$

where A is the $K \times K+1$ matrix $[0_{K\times 1} \quad I_K]$. Then let $\hat{Q}_n := \hat{Q}_n(\hat{\mathbf{g}}, \hat{w})$ and $\hat{Q}_n^*(\mathbf{h}, \tau) := \hat{Z}_n^*(\mathbf{h}, \tau)A'$.

As in the proof of Theorem **??**, I consider a sequence $\{J_n : n \geq 1\}$ and then derive the stated results as $n \to \infty$. I first state the following lemmas, which will then be used to prove Theorem **??**.

Lemma B.1. *Under the assumptions of Theorem* **??**,

$$\text{(a) } \sqrt{n}\left\{\hat{Z}_n(\hat{\mathbf{g}}, \hat{w}) - \hat{Z}_n(\tilde{\mathbf{g}}, w_{0,J_n}) - E\left(\hat{Z}_n(\hat{\mathbf{g}}, \hat{w}) - \hat{Z}_n(\tilde{\mathbf{g}}, w_{0,J_n})\right)\right\} \to_p 0 \quad \text{and}$$
$$\text{(b) } \sqrt{n}\left\{\hat{Z}_n(\tilde{\mathbf{g}}, w_{0,J_n}) - \hat{Z}_n^*(\mathbf{h}_0, \tau_{0,J_n}) - E\left(\hat{Z}_n(\tilde{\mathbf{g}}, w_{0,J_n}) - \hat{Z}_n^*(\mathbf{h}_0, \tau_{0,J_n})\right)\right\} \to_p 0$$

Lemma B.2. *Under the assumptions of Theorem* **??**, $\sqrt{n}(E\left(\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w})\right) - B_{1J_n}) = o_p(1)$ *and* $B_{1J_n} = O(J_n^{-1})$ *where* $B_{1J} = E\left(\tau_{0J}(\theta_i)\eta_i^2 Dh(\theta_i)Dh_x(\theta_i)\right)$.

Lemma B.3. *Under the assumptions of Theorem* **??**, $\hat{Q}_n - Q_{0,J_n}^* = O_p(r_n) + O\left((J_n^{-1}\log(J_n))^{1/2}\right)$ *where* $r_n = h_n^2 + \frac{\log(n)}{\sqrt{nh_n}} + \frac{\log(J_n)^{p/2}}{h_n^{(p-1)}J_n^{p/2}}$.

Proof of Theorem **??**. First,

$$\begin{aligned}
&\sqrt{n}\left(\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) - B_{1J_n}\right) \qquad \text{(B.5)}\\
&= \sqrt{n}\left\{\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) - \hat{M}_n^*(\beta_{10}, \mathbf{h}_0, \tau_{0,J_n}) - E\left(\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) - \hat{M}_n^*(\beta_{10}, \mathbf{h}_0, \tau_{0,J_n})\right)\right\}\\
&+ \sqrt{n}\left(E\left(\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w})\right) - B_{1J_n}\right) + \sqrt{n}\left(\hat{M}_n^*(\beta_{10}, \mathbf{h}_0, \tau_{0,J_n}) - E(\hat{M}_n^*(\beta_{10}, \mathbf{h}_0, \tau_{0,J_n}))\right)
\end{aligned}$$

By Lemma B.1, since $\hat{M}_n(\beta_{10}, \mathbf{g}, w) = \hat{Z}_n(\mathbf{g}, w)\gamma_0'$ and $\hat{M}_n^*(\beta_{10}, \mathbf{h}, \tau) = \hat{Z}_n^*(\mathbf{h}, \tau)\gamma_0'$, the first term is $o_p(1)$. By Lemma B.2 the second term is also $o_p(1)$. Therefore,

$$\begin{aligned}
&\sqrt{n}\left(\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) - B_{1J_n}\right) \qquad \text{(B.6)}\\
&= \sqrt{n}\left(\hat{M}_n^*(\beta_{10}, \mathbf{h}_0, \tau_{0,J_n}) - E(\hat{M}_n^*(\beta_{10}, \mathbf{h}_0, \tau_{0,J_n}))\right) + o_p(1)
\end{aligned}$$

Since condition (b) of Assumption **??** implies that $\sup_n ||V_{1J_n}^{-1/2}|| < \infty$,

$$\begin{aligned}
&\sqrt{n} V_{1J_n}^{-1/2} (\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) - B_{1J_n}) \\
&= \sqrt{n} V_{1J_n}^{-1/2} \left(\hat{M}_n^*(\beta_{10}, \mathbf{h}_0, \tau_{0,J_n}) - E(\hat{M}_n^*(\beta_{10}, \mathbf{h}_0, \tau_{0,J_n})) \right) + o_p(1) \\
&\to_d N(0, I)
\end{aligned} \tag{B.7}$$

where the last line follows from the Lindeberg-Feller central limit theorem for triangular arrays since condition (c) of Assumption **??** implies the Lyapounov condition and condition (b) implies that $\sup_n ||V_{1J_n}^{-1/2}|| < \infty$ where $V_{1J_n} = E(\tau_{0J_n}(\theta_i)^2 e_i^2 (X_i - h_{0x}(\theta_i))(X_i - h_{0x}(\theta_i))') = Var(\sqrt{n}\hat{M}_n^*(\beta_{10}, \mathbf{h}_0, \tau_{0,J_n}))$.

Next, for any β_1, $\hat{M}_n(\beta_1, \hat{\mathbf{g}}, \hat{w}) = \hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) - \hat{Q}_n(\beta_1 - \beta_{10})$. Therefore, $\sqrt{n}\hat{Q}_n(\hat{\beta}_{1J_n} - \beta_{10}) = \sqrt{n}\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) - \sqrt{n}\hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w})$. Rather than assuming that $\hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w}) = 0$, the following argument shows that $\sqrt{n}\hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w}) = o_p(1)$.

The result in (B.7) and condition (b) of Assumption **??** together imply that $\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) - B_{1J_n} = O_p(n^{-1/2})$. Further, since $B_{1J_n} = o(1)$ by Lemma B.2, we have that $\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) = o_p(1)$. The, for each β_1, define $\tilde{M}_n(\beta_1, \hat{\mathbf{g}}, \hat{w}) := \hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) - Q_{0,J_n}^*(\beta_1 - \beta_{10})$. Let $\bar{\beta}_1 = \beta_{10} + Q_{0,J_n}^{*-1}\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w})$ so that $\tilde{M}_n(\bar{\beta}_1, \hat{\mathbf{g}}, \hat{w}) = 0$. Then $\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) = o_p(1)$, so condition (b) of Assumption **??** implies that $\bar{\beta}_1 - \beta_{10} = Q_{0,J_n}^{*-1}\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) \to_p 0$. By condition (a) of Assumption **??**, $\beta_{10} \in int(\mathcal{B})$, so I can assume that $\bar{\beta}_1 \in \mathcal{B}$. Therefore, condition (a) of Assumption **??** also implies that $|\hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w})| = \inf_{\beta_1 \in \mathcal{B}} |\hat{M}_n(\beta_1, \hat{\mathbf{g}}, \hat{w})| \leq |\hat{M}_n(\bar{\beta}_1, \hat{\mathbf{g}}, \hat{w})| + o_p(n^{-1/2})$. So it remains to show that $\sqrt{n}\hat{M}_n(\bar{\beta}_1, \hat{\mathbf{g}}, \hat{w}) = o_p(1)$.

But since $\tilde{M}_n(\bar{\beta}_1, \hat{\mathbf{g}}, \hat{w}) = 0$, $|\hat{M}_n(\bar{\beta}_1, \hat{\mathbf{g}}, \hat{w})| \leq |\tilde{M}_n(\bar{\beta}_1, \hat{\mathbf{g}}, \hat{w})| + |(\hat{Q}_n - Q_{0,J_n}^*)(\bar{\beta}_1 - \beta_{10})| = |(\hat{Q}_n - Q_{0,J_n}^*)(\bar{\beta}_1 - \beta_{10})|$. And $0 = \tilde{M}_n(\bar{\beta}_1, \hat{\mathbf{g}}, \hat{w}) = \hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) - Q_{0,J_n}^*(\bar{\beta}_1 - \beta_{10})$ so that

$$\begin{aligned}
(\hat{Q}_n - Q_{0,J_n}^*)(\bar{\beta}_1 - \beta_{10}) &= (\hat{Q}_n - Q_{0,J_n}^*)(\bar{\beta}_1 - \beta_{10} - B_{J_n}) + (\hat{Q}_n - Q_{0,J}^*)B_{J_n} \\
&= (\hat{Q}_n - Q_{0,J_n}^*)Q_{0,J_n}^{*-1}(\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) - B_{1J_n}) + (\hat{Q}_n - Q_{0,J}^*)B_{J_n}
\end{aligned}$$

I have already shown that $(\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) - B_{1J_n}) = O_p(n^{-1/2})$ so the first term here is $o_p(n^{-1/2})$ by Lemma B.3. Applying both Lemmas B.2 and B.3, the second term is

$$\left(O_p\left(h_n^2 + \frac{\log(n)}{\sqrt{n h_n}} + \frac{\log(J_n)^{p/2}}{h_n^{(p-1)} J_n^{p/2}} \right) + O((J_n^{-1}\log(J_n))^{1/2}) \right) O(J_n^{-1}),$$

which is $o_p(n^{-1/2})$ by conditions (e) and (g) of Assumption **??**. Thus I have shown that $\sqrt{n}\hat{M}_n(\hat{\beta}_{1J_n}, \hat{\mathbf{g}}, \hat{w}) = o_p(1)$ and therefore $\sqrt{n}\hat{Q}_n(\hat{\beta}_{1J_n} - \beta_{10}) = \sqrt{n}\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w}) + o_p(1)$.

Next, since $B_J = Q_{0,J}^{*-1}B_{1J}$, $\sup_n ||V_{1J_n}^{-1/2}|| < \infty$, and, as just shown, $(\hat{Q}_n - Q_{0,J_n}^*)B_{J_n} =$

$o_p(n^{-1/2})$,

$$\begin{aligned}
&\sqrt{n}V_{1J_n}^{-1/2}Q_{0,J_n}^*(\hat{\beta}_{1J_n}-\beta_{10}-B_{J_n})\\
&=\sqrt{n}V_{1J_n}^{-1/2}\hat{Q}_n(\hat{\beta}_{1J_n}-\beta_{10}-B_{J_n})+\sqrt{n}V_{1J_n}^{-1/2}(Q_{0,J_n}^*-\hat{Q}_n)(\hat{\beta}_{1J_n}-\beta_{10}-B_{J_n})\\
&=\sqrt{n}V_{1J_n}^{-1/2}\left(\hat{M}_n(\beta_{10},\hat{\mathbf{g}},\hat{w})-\hat{Q}_nB_{J_n}\right)+\sqrt{n}V_{1J_n}^{-1/2}(Q_{0,J_n}^*-\hat{Q}_n)(\hat{\beta}_{1J_n}-\beta_{10}-B_{J_n})\\
&=\sqrt{n}V_{1J_n}^{-1/2}(\hat{M}_n(\beta_{10},\hat{\mathbf{g}},\hat{w})-B_{1J_n})\\
&+\left(V_{1J_n}^{-1/2}(Q_{0,J_n}^*-\hat{Q}_n)Q_{0,J_n}^{*-1}V_{1J_n}^{1/2}\right)\left(\sqrt{n}V_{1J_n}^{-1/2}Q_{0,J_n}^*(\hat{\beta}_{1J_n}-\beta_{10}-B_{J_n})\right)+o_p(1)
\end{aligned}$$

Then conditions (b) and (c) of Assumption **??** and Lemma B.3 imply that

$$\left(V_{1J_n}^{-1/2}(Q_{0,J_n}^*-\hat{Q}_n)Q_{0,J_n}^{*-1}V_{1J_n}^{1/2}\right)=o_p(1)$$

so that

$$\sqrt{n}V_{1J_n}^{-1/2}Q_{0,J_n}^*(\hat{\beta}_{1J_n}-\beta_{10}-B_{J_n})=\frac{\sqrt{n}V_{1J_n}^{-1/2}(\hat{M}_n(\beta_{10},\hat{\mathbf{g}},\hat{w})-B_{1J_n})}{1-o_p(1)}+o_p(1)\to_d N(0,I)$$

Since $B_{J_n}=O(J_n^{-1})$ it follows from conditions (b) and (c) of Assumption **??** that $\hat{\beta}_{1J_n}-\beta_{10}=O_p(n^{-1/2})+O(J_n^{-1})$.

If $V_{1J_n}\to\bar{V}_1$ and $Q_{0,J_n}^*\to\bar{Q}_0^*$ then $\sqrt{n}(\hat{\beta}_{1J_n}-\beta_{10}-B_{J_n})=\left(Q_{0,J_n}^{*-1}V_{1J_n}^{1/2}\right)\sqrt{n}V_{1_{J_n}}^{-1/2}Q_{0,J_n}^*(\hat{\beta}_{1_{J_n}}-\beta_{10}-B_{J_n})\to_d N(0,\bar{Q}_0^{*-1}\bar{V}_1\bar{Q}_0^{*-1})$. If, in addition, $\sqrt{n}B_{J_n}\to\gamma\bar{B}$ then $\sqrt{n}(\hat{\beta}_{1_{J_n}}-\beta_{10})=\sqrt{n}(\hat{\beta}_{1_{J_n}}-\beta_{10}-B_{J_n})+\sqrt{n}B_{J_n}\to_d N(\gamma\bar{B},\bar{V})$. □

Proof of Lemma B.1. Proof of (a): Let $\bar{\mathcal{M}}^*=\bar{p}_\infty(\Theta_\delta)$ for some $0<\delta<\delta_1$. By conditions (d) and (f) of Assumption **??**, $Pr(\hat{M}\subset\bar{\mathcal{M}}^*)\to 1$.

Next, $\hat{Z}_n(\mathbf{g},w)=\sum_{s=1}^4\hat{Z}_{ns}(\mathbf{g},w)$ where $\hat{Z}_{n1}(\mathbf{g},w)=n^{-1}\sum_{i=1}^n w(\bar{M}_{iJ_n})X_iZ_i'$, $\hat{Z}_{n2}(\mathbf{g},w)=-n^{-1}\sum_{i=1}^n w(\bar{M}_{iJ_n})X_i\mathbf{g}(\bar{M}_{iJ_n})'$, $\hat{Z}_{n3}(\mathbf{g},w)=-n^{-1}\sum_{i=1}^n w(\bar{M}_{iJ_n})g_x(\bar{M}_{iJ_n})Z_i'$, and $\hat{Z}_{n4}(\mathbf{g},w)=n^{-1}\sum_{i=1}^n w(\bar{M}_{iJ_n})g_x(\bar{M}_{iJ_n})\mathbf{g}(\bar{M}_{iJ_n})'$.

Then stochastic equicontinuity results of Andrews (1994) can be applied to each of these four terms separately. For positive integers r,s let $\Gamma_{0,r,s}$ be the space of $r\times s$ matrix-valued functions, $\{\mathbf{f}:\mathbf{f}(m)=0\forall m\notin\bar{\mathcal{M}}^*,\sup_{m\in\bar{\mathcal{M}}^*}|\mathbf{f}(m)|<B,\sup_{m\in\bar{\mathcal{M}}^*}|D\mathbf{f}(m)|<B\}$. Then let $\Gamma_1=\{xz'\}\times\Gamma_{0,1,1}$ and let $\rho_1(f^*,f)=\sup_n E\left(|(f^*(\bar{M}_{iJ_n})-f(\bar{M}_{iJ_n}))X_iZ_i|^2\right)^{1/2}$. Then, by Theorems 1-3 of Andrews (1994) and condition **LIx** of Assumption **??** and condition (c) of Assumption **??**, for any sequence $\delta_n\to 0$,

$$\sup_{f,f^*\in\Gamma_1,\rho_1(f^*,f)<\delta_n}||v_{n1}(f^*)-v_{n1}(f))||\to_p 0$$

where $v_{n1}(f) = n^{-1/2} \sum_{i=1}^n f(\bar{M}_{iJ_n}) X_i Z_i' = \sqrt{n}\hat{Z}_{n1}(\mathbf{g}, w)$.

Similarly, let $\Gamma_2 = \{x\} \times \Gamma_{0,K+1,1}$ and let $\rho_2(\mathbf{f}^*, \mathbf{f}) = \sup_n E\left(|X_i(\mathbf{f}^*(\bar{M}_{iJ_n}) - \mathbf{f}(\bar{M}_{iJ_n}))'|^2\right)^{1/2}$. Then, by Theorems 1-3 of Andrews (1994) and condition **LIx** of Assumption **??** and condition (c) of Assumption **??**, for any sequence $\delta_n \to 0$,

$$\sup_{\mathbf{f},\mathbf{f}^* \in \Gamma_2, \rho_2(\mathbf{f}^*,\mathbf{f}) < \delta_n} ||v_{n2}(\mathbf{f}^*) - v_{n2}(\mathbf{f}))|| \to_p 0$$

where $v_{n2}(\mathbf{f}) = n^{-1/2} \sum_{i=1}^n X_i \mathbf{f}(\bar{M}_{iJ_n})'$ and $v_{n2}(w\mathbf{g}) = \sqrt{n}\hat{Z}_{n2}(\mathbf{g}, f)$.

Third, let $\Gamma_3 = \{z\} \times \Gamma_{0,K,1}$ and let $\rho_3(\mathbf{f}^*, \mathbf{f}) = \sup_n E\left(|(\mathbf{f}^*(\bar{M}_{iJ_n}) - \mathbf{f}(\bar{M}_{iJ_n}))Z_i'|^2\right)^{1/2}$. Then, by Theorems 1-3 of Andrews (1994) and condition **LIx** of Assumption **??** and condition (c) of Assumption **??**, for any sequence $\delta_n \to 0$,

$$\sup_{\mathbf{f},\mathbf{f}^* \in \Gamma_3, \rho_3(\mathbf{f}^*,\mathbf{f}) < \delta_n} ||v_{n3}(\mathbf{f}^*) - v_{n3}(\mathbf{f})|| \to_p 0$$

where $v_{n3}(\mathbf{f}) = n^{-1/2} \sum_{i=1}^n \mathbf{f}(\bar{M}_{iJ_n}) Z_i'$ and $v_{n3}(wg_x) = \sqrt{n}\hat{Z}_{n3}(\mathbf{g}, f)$.

Lastly, let $\Gamma_4 = \Gamma_{0,K,K+1}$ and let $\rho_4(\mathbf{f}^*, \mathbf{f}) = \sup_n E\left(||\mathbf{f}^*(\bar{M}_{iJ_n}) - \mathbf{f}(\bar{M}_{iJ_n})||^2\right)^{1/2}$. Then, by Theorems 1-3 of Andrews (1994) and condition **LIx** of Assumption **??** and condition (c) of Assumption **??**, for any sequence $\delta_n \to 0$,

$$\sup_{\mathbf{f},\mathbf{f}^* \in \Gamma_4, \rho_4(\mathbf{f}^*,\mathbf{f}) < \delta_n} ||v_{n4}(\mathbf{f}^*) - v_{n4}(\mathbf{f})|| \to_p 0$$

where $v_{n4}(\mathbf{f}) = n^{-1/2} \sum_{i=1}^n \mathbf{f}(\bar{M}_{iJ_n})$ and $v_{n4}(wg_x\mathbf{g}') = \sqrt{n}\hat{Z}_{n4}(\mathbf{g}, f)$.

Then (a) follows since Theorem C.2 implies (1) that $Pr(\hat{w} \in \Gamma_{0,1,1})$, $Pr(\hat{w}\hat{\mathbf{g}} \in \Gamma_{0,K+1,1})$, $Pr(\hat{w}\hat{g}_x \in \Gamma_{0,K,1})$ and $Pr(\hat{w}\hat{g}_x\hat{\mathbf{g}}' \in \Gamma_{0,K,K+1})$ each converge to 1 and (2) that $\rho_1(\hat{w}, w_{0,J_n}) \to_p 0$, $\rho_2(\hat{w}\hat{\mathbf{g}}, w_{0,J_n}\tilde{\mathbf{g}}) \to_p 0$, $\rho_3(\hat{w}\hat{g}_x, w_{0,J_n}\tilde{\mathbf{g}}) \to_p 0$, and $\rho_4(\hat{w}\hat{g}_x\hat{\mathbf{g}}', w_{0,J_n}\tilde{g}_x\tilde{\mathbf{g}}') \to_p 0$.

Proof of (b): Let

$$\begin{aligned}
\hat{m}_1 &= \frac{1}{n}\sum_{i=1}^n w(\bar{M}_{iJ_n})(X_i - \tilde{g}_x(\bar{M}_{iJ_n}))(Z_i - \tilde{\mathbf{g}}(\bar{M}_{iJ_n}))' \\
&\quad - \frac{1}{n}\sum_{i=1}^n w(\bar{M}_{iJ_n})(X_i - h_{x,0}(\theta_i))(Z_i - \mathbf{h}_0(\theta_i))' \\
&\quad - E\left(w(\bar{M}_{iJ_n})(X_i - \tilde{g}_x(\bar{M}_{iJ_n}))(Z_i - \tilde{\mathbf{g}}(\bar{M}_{iJ_n}))' - w(\bar{M}_{iJ_n})(X_i - h_{x,0}(\theta_i))(Z_i - \mathbf{h}_0(\theta_i))'\right)
\end{aligned}$$

and

$$
\begin{aligned}
\hat{m}_2 = &\frac{1}{n}\sum_{i=1}^{n} w(\bar{M}_{iJ_n})(X_i - h_{x,0}(\theta_i))(Z_i - \mathbf{h}_0(\theta_i))' \\
&- \frac{1}{n}\sum_{i=1}^{n} w(\bar{p}_{J_n}(\theta_i))(X_i - h_{x,0}(\theta_i))(Z_i - \mathbf{h}_0(\theta_i))' \\
&- E\left(w(\bar{M}_{iJ_n})(X_i - h_{x,0}(\theta_i))(Z_i - \mathbf{h}_0(\theta_i))' - w(\bar{p}_{J_n}(\theta_i))(X_i - h_{x,0}(\theta_i))(Z_i - \mathbf{h}_0(\theta_i))'\right)
\end{aligned}
$$

Then (b) follows if $\sqrt{n}\hat{m}_1 = o_p(1)$ and $\sqrt{n}\hat{m}_2 = o_p(1)$.

First, consider $Var(w(\bar{M}_{iJ_n})(\tilde{g}_s(\bar{M}_{iJ_n}) - h_{0,s}(\theta_i))V_i)$ for V_i equal to Y_i, a component of the vector X_i, or a component of the vector $\mathbf{h}_0(\theta_i)$ where $\tilde{g}_s$ and $h_{0,s}$ represent any component of the vectors $\tilde{\mathbf{g}}$ and $\mathbf{h}_0$, respectively. By a Taylor expansion, $\tilde{g}_s(\bar{M}_{iJ_n}) - h_{0,s}(\theta_i) = D\tilde{g}_s(p_i^*)\eta_i$ for some p_i^* between $\bar{p}_{J_n}(\theta_i)$ and $\bar{M}_{iJ_n}$ so

$$
Var(w(\bar{M}_{iJ_n})(\tilde{g}_s(\bar{M}_{iJ_n}) - h_{0,s}(\theta_i))V_i) \leq E(w(\bar{M}_{iJ_n})^2 D\tilde{g}_s(p_i^*)^2\eta_i^2 V_i^2)
$$

Next, $E(w(\bar{M}_{iJ_n})^2 D\tilde{g}_s(p_i^*)^2\eta_i^2 V_i^2) = \int E(w(\bar{M}_{iJ_n})^2 D\tilde{g}_s(p_i^*)^2\eta_i^2 V_i^2 \mid \theta_i = \theta) f_\theta(\theta) d\theta$. For $\theta \in \Theta_{\delta_2}$, both $\bar{p}_{J_n}(\theta_i)$ and $\bar{M}_{iJ_n}$ are in $\mathcal{M}_{\delta_2}$ (unless $w(\bar{M}_{iJ_n}) = 0$) so $p_i^* \in \mathcal{M}_{\delta_2}$ and, therefore,

$$
\begin{aligned}
&\int_{\Theta_{\delta_2}} E(w(\bar{M}_{iJ_n})^2 D\tilde{g}_s(p_i^*)^2\eta_i^2 V_i^2 \mid \theta_i = \theta) f_\theta(\theta) d\theta \\
&\leq B^2 \left(\sup_{m\in\mathcal{M}_{\delta_2}} |D\tilde{g}_s(m)|\right)^2 \int_{\Theta_{\delta_2}} E(V_i^2 \mid \theta_i = t) E(\eta_i^2 \mid \theta_i = t) f_\theta(\theta) d\theta \\
&\leq \frac{1}{J_n} B^2 \left(\sup_{m\in\mathcal{M}_{\delta_2}} |D\tilde{g}_s(m)|\right)^2 \int_{\Theta_{\delta_2}} E(V_i^2) f_\theta(\theta) d\theta
\end{aligned}
$$

where I have used (a) the fact that $E(V_i^2\eta_i^2 \mid \theta_i) = E(V_i^2 \mid \theta_i)E(\eta_i^2 \mid \theta_i)$ for V_i equal to Y_i, a component of the vector X_i, or a component of the vector $\mathbf{h}_0(\theta_i)$, by Assumption **??**, (b) the fact that $\sup_\theta E(\eta_i^2 \mid \theta_i = \theta) \leq J_n^{-1}$, by Lemma C.1, and (c) condition (d) of Assumption **??** which implies that the function $w(m)$ is bounded uniformly by B.

Next, by condition (e) of Assumption **??**, $|D\tilde{g}_s(m)| \leq \bar{D}(\bar{p}_{J_n}^{-1}(m))$ where $\bar{D}(\cdot)$ is nonincreasing on the interval $(-\infty, q_{\delta_2}(\theta_i)]$ and nondecreasing on $[q_{1-\delta_2}(\theta_i), \infty)$. Then

$$
\begin{aligned}
&\int_{\Theta\backslash\Theta_{\delta_2}} E(w(\bar{M}_{iJ_n})^2 D\tilde{g}_s(p_i^*)^2\eta_i^2 V_i^2 \mid \theta_i = \theta) f_\theta(\theta) d\theta \\
&\le B^2 \int_{\Theta\backslash\Theta_{\delta_2}} \bar{D}(t)^2 E(V_i^2 \mid \theta_i = \theta) E(\eta_i^2 \mid \theta_i = \theta) f_\theta(\theta) d\theta \\
&\le \frac{1}{J_n} B^2 E(\bar{D}(\theta_i)^2 V_i^2)
\end{aligned}
$$

so $Var(w(\bar{M}_{iJ_n})(\tilde{g}_s(\bar{M}_{iJ_n}) - h_{0,s}(\theta_i))V_i) = O(J_n^{-1})$ by condition (e) of Assumption **??** and Theorem C.2, and by Chebyschev's inequality,

$$
n^{-1/2} \sum_{i=1}^{n} \left\{ w(\bar{M}_{iJ_n})(\tilde{g}_s(\bar{M}_{iJ_n}) - h_{0,s}(\theta_i))V_i - E(w(\bar{M}_{iJ_n})(\tilde{g}_s(\bar{M}_{iJ_n}) - h_{0,s}(\theta_i))V_i) \right\} \to_p 0
$$

Since $\sqrt{n}\hat{m}_1$ can be expanded into a sum of (a finite, fixed number of) terms of this form, the desired result follows.

Next, using the Taylor series approximation $w(\bar{M}_{iJ_n}) - w(\bar{p}_{J_n}(\theta_i)) = w'_{0,J_n}(p_i^*)\eta_i$,

$$
\begin{aligned}
&Var((w(\bar{M}_{iJ_n}) - w(\bar{p}_{J_n}(\theta_i))(X_{ik} - h_{x_k,0}(\theta_i))(Z_{il} - \mathbf{h}_{0,l}(\theta_i))) \\
&\le B^2 E(\eta_i^2 (X_{ik} - h_{x_k,0}(\theta_i))^2 (Z_{il} - \mathbf{h}_{0,l}(\theta_i))^2) \\
&\le \frac{1}{J_n} E((X_{ik} - h_{x_k,0}(\theta_i))^2 (Z_{il} - \mathbf{h}_{0,l}(\theta_i))^2) = O(J_n^{-1})
\end{aligned}
$$

by condition (f) of Assumption **??** and conditions (c) and (d) of Assumption **??** so that by Chebyschev's inequality, $\sqrt{n}\hat{m}_2 = o_p(1)$.

$\square$

Proof of Lemma B.2. First, by Assumption **??**, $E(Y_i \mid X_i, \theta_i, \mathbf{M}_i) = \beta'_{10}X_i + h_0(\theta_i)$, $E(X_i \mid \theta_i, \mathbf{M}_i) = h_{x,0}(\theta_i)$, and $h_{y,0}(\theta_i) := E(Y_i \mid \theta_i) = \beta'_{10}h_{x,0}(\theta_i) + h_0(\theta_i)$, and therefore,

$$
\begin{aligned}
&E\left(\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w})\right) \\
&= E\left(\hat{w}(\bar{M}_{iJ_n})(X_i - \hat{g}_x(\bar{M}_{iJ_n}))(\beta'_{10}X_i + h_0(\theta_i) - \hat{g}_y(\bar{M}_{iJ_n}) - \beta'_{10}(X_i - \hat{g}_x(\bar{M}_{iJ_n})))\right) \\
&= E\left(\hat{w}(\bar{M}_{iJ_n})(h_{x,0}(\theta_i) - \hat{g}_x(\bar{M}_{iJ_n}))(\mathbf{h}_0(\theta_i) - \hat{\mathbf{g}}(\bar{M}_{iJ_n}))'\right)\gamma_0 \\
&= E\left(\hat{w}(\bar{M}_{iJ_n})\left\{(h_{x,0}(\theta_i) - \tilde{g}_x(\bar{M}_{iJ_n})) + (\tilde{g}_x(\bar{M}_{iJ_n}) - \hat{g}_x(\bar{M}_{iJ_n}))\right\}\right. \\
&\qquad \left. \cdot \left\{(\mathbf{h}_0(\theta_i) - \tilde{\mathbf{g}}(\bar{M}_{iJ_n})) + (\tilde{\mathbf{g}}(\bar{M}_{iJ_n}) - \hat{\mathbf{g}}(\bar{M}_{iJ_n}))\right\}'\right)\gamma_0
\end{aligned}
$$

By condition (d) of Assumption **??**, I can assume that $\hat{\mathcal{M}} \subset \mathcal{M}_{\delta_1}$ since

$$\begin{aligned}
&Pr(|\sqrt{n}E\left(\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w})\right) - B_{1J_n}| \geq \varepsilon) \\
&\leq Pr(|\sqrt{n}E\left(\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w})\right) - B_{1J_n}| \geq \varepsilon, \hat{\mathcal{M}} \subset \mathcal{M}_{\delta_1}) + (1 - Pr(\hat{\mathcal{M}} \subset \mathcal{M}_{\delta_1})) \\
&= Pr(|\sqrt{n}E\left(\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w})\right) - B_{1J_n}| \geq \varepsilon, \hat{\mathcal{M}} \subset \mathcal{M}_{\delta_1}) + o(1)
\end{aligned}$$

Let $\hat{m}_1 = E\left(\hat{w}(\bar{M}_{iJ_n})(h_{x,0}(\theta_i) - \tilde{g}_x(\bar{M}_{iJ_n}))(\mathbf{h}_0(\theta_i) - \tilde{\mathbf{g}}(\bar{M}_{iJ_n}))'\right)\gamma_0$. Then, using a second order taylor expansion of $\tilde{\mathbf{g}}$ and $\tilde{g}_x$,

$$\begin{aligned}
\hat{m}_1 &= E\left(\hat{w}(\bar{M}_{iJ_n})D\tilde{g}_x(\bar{p}_{J_n}(\theta_i))D\tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i))'\eta_i^2\right)\gamma_0 + E\left(\hat{w}(\bar{M}_{iJ_n})D\tilde{g}_x(\bar{p}_{J_n}(\theta_i))D^2\tilde{\mathbf{g}}(p_i^*)'\eta_i^3\right)\gamma_0 \\
&+ E\left(\hat{w}(\bar{M}_{iJ_n})D^2\tilde{g}_x(p_i^{**})D\tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i))'\eta_i^3\right)\gamma_0 + E\left(\hat{w}(\bar{M}_{iJ_n})D^2\tilde{g}_x(p_i^{**})D^2\tilde{\mathbf{g}}(p_i^*)'\eta_i^4\right)\gamma_0 \\
&= E\left(\hat{w}(\bar{M}_{iJ_n})D\tilde{g}_x(\bar{p}_{J_n}(\theta_i))D\tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i))'\eta_i^2\right)\gamma_0 + O_p(J_n^{-3/2})
\end{aligned}$$

where the second equality follows from conditions (d) and (e) of Assumption **??** by applying the same argument used above in the proof of Lemma B.1. In addition,

$$\begin{aligned}
&E\left(\hat{w}(\bar{M}_{iJ_n})D\tilde{g}_x(\bar{p}_{J_n}(\theta_i))D\tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i))'\eta_i^2\right)\gamma_0 \\
&= E\left(w_{0,J_n}(\bar{p}_{J_n}(\theta_i))D\tilde{g}_x(\bar{p}_{J_n}(\theta_i))D\tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i))'\eta_i^2\right)\gamma_0 \\
&+ E\left((w_{0,J_n}(\bar{M}_{iJ_n}) - w_{0,J_n}(\bar{p}_{J_n}(\theta_i)))D\tilde{g}_x(\bar{p}_{J_n}(\theta_i))D\tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i))'\eta_i^2\right)\gamma_0 \\
&+ E\left((\hat{w}(\bar{M}_{iJ_n}) - w_{0,J_n}(\bar{M}_{iJ_n}))D\tilde{g}_x(\bar{p}_{J_n}(\theta_i))D\tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i))'\eta_i^2\right)\gamma_0 \\
&= E\left(w_{0,J_n}(\bar{p}_{J_n}(\theta_i))D\tilde{g}_x(\bar{p}_{J_n}(\theta_i))D\tilde{\mathbf{g}}(\bar{p}_{J_n}(\theta_i))'\eta_i^2\right)\gamma_0 + O_p(J_n^{-3/2}) + o_p(n^{-1/2})
\end{aligned}$$

by (e) and (f) of Assumption **??** and (d) of Assumption **??**. Thus $\sqrt{n}(\hat{m}_1 - B_{1J_n}) = o_p(1)$.

Next, let $\hat{m}_2 = E\left(\hat{w}(\bar{M}_{iJ_n})(\tilde{g}_x(\bar{M}_{iJ_n}) - \hat{g}_x(\bar{M}_{iJ_n}))(\mathbf{h}_0(\theta_i) - \tilde{\mathbf{g}}(\bar{M}_{iJ_n}))'\right)\gamma_0$. Then using a first order Taylor approximation of $\tilde{\mathbf{g}}$,

$$\begin{aligned}
\sqrt{n}|\hat{m}_2| &\leq \sqrt{n}\sup_{m \in \mathcal{M}_{\delta_1}}|\hat{g}_x(m) - \tilde{g}_x(m)|E\left(\hat{w}(\bar{M}_{iJ_n})|D\tilde{\mathbf{g}}(p_i^*)\eta_i|\right)|\gamma_0| \\
&= \sqrt{n}O_p\left((J_n^{-1}\log(J_n))^{1/2}\right)\sup_{m \in \mathcal{M}_{\delta_1}}|\hat{g}_x(m) - \tilde{g}_x(m)||\gamma_0| \\
&= o_p(1)
\end{aligned}$$

where the first equality is due to Lemma C.1 and the second because Theorem C.2 and conditions (e) and (g) of Assumption **??** imply that $\sup_{m \in \mathcal{M}_{\delta_1}}|\hat{g}_x(m) - \tilde{g}_x(m)| = O_p(r_n)$

$= o_p\left((J_n n^{-1}\log(J_n))^{1/2}\right)$ where $r_n = h_n^2 + \frac{\log(n)}{\sqrt{nh_n}} + \frac{\log(J_n)^{p/2}}{h_n^{(p-1)}J_n^{p/2}}$. By essentially the same argument,

$$\sqrt{n}|\hat{m}_3| \le \sqrt{n} O_p\left((J_n^{-1}\log(J_n))^{1/2}\right) \sup_{m\in\mathcal{M}_{\delta_1}} |\hat{g}_x(m) - \tilde{g}_x(m)||\gamma_0| = o_p(1)$$

and $\sqrt{n}|\hat{m}_4| \le \sqrt{n}\sup_{m\in\mathcal{M}_{\delta_1}} |\hat{g}_x(m) - \tilde{g}_x(m)|^2|\gamma_0| = o_p(1)$ where

$$\begin{aligned}
\hat{m}_3 &= E\left(\hat{w}(\bar{M}_{iJ_n})(h_{x,0}(\theta_i) - \tilde{g}_x(\bar{M}_{iJ_n}))(\tilde{\mathbf{g}}(\bar{M}_{iJ_n}) - \hat{\mathbf{g}}(\bar{M}_{iJ_n}))'\right)\gamma_0 \\
\hat{m}_4 &= E\left(\hat{w}(\bar{M}_{iJ_n})(\tilde{g}_x(\bar{M}_{iJ_n}) - \hat{g}_x(\bar{M}_{iJ_n}))(\tilde{\mathbf{g}}(\bar{M}_{iJ_n}) - \hat{\mathbf{g}}(\bar{M}_{iJ_n}))'\right)\gamma_0
\end{aligned}$$

noting that

$$\begin{aligned}
\sqrt{n}\sup_{m\in\mathcal{M}_{\delta_1}} |\hat{g}_x(m) - \tilde{g}_x(m)|^2 &= \left(n^{1/4}\sup_{m\in\mathcal{M}_{\delta_1}} |\hat{g}_x(m) - \tilde{g}_x(m)|\right)^2 \\
&= \left(O\left(\sqrt{n}(J_n^{-1}\log(J_n))^{1/2}\right)\sup_{m\in\mathcal{M}_{\delta_1}} |\hat{g}_x(m) - \tilde{g}_x(m)|\right)^2
\end{aligned}$$

Therefore

$$\begin{aligned}
\sqrt{n}\left(E\left(\hat{M}_n(\beta_{10}, \hat{\mathbf{g}}, \hat{w})\right) - B_{1J_n}\right) &= \sqrt{n}(\hat{m}_1 - B_{1J_n}) + \sqrt{n}\hat{m}_2 + \sqrt{n}\hat{m}_3 + \sqrt{n}\hat{m}_4 \\
&= o_p(1)
\end{aligned}$$

□

Proof of Lemma B.3. Let $A = [0_{K\times 1} \quad I_K]$ and recall that $\hat{Q}_n(\mathbf{g}, w) = \hat{Z}_n(\mathbf{g}, w)A'$ and $\hat{Q}_n^*(\mathbf{h}, \tau) = \hat{Z}_n^*(\mathbf{h}, \tau)A'$. The desired result follows from the following expansion,

$$\begin{aligned}
\hat{Q}_n - Q_{0,J_n}^* = &\left\{\hat{Z}_n(\hat{\mathbf{g}}, \hat{w}) - \hat{Z}_n(\tilde{\mathbf{g}}, w_{0,J_n}) - E\left(\hat{Z}_n(\hat{\mathbf{g}}, \hat{w}) - \hat{Z}_n(\tilde{\mathbf{g}}, w_{0,J_n})\right)\right\} A' + \\
&\left\{\hat{Z}_n(\tilde{\mathbf{g}}, w_{0,J_n}) - \hat{Z}_n^*(\mathbf{h}_0, \tau_{0,J_n}) - E\left(\hat{Z}_n(\tilde{\mathbf{g}}, w_{0,J_n}) - \hat{Z}_n^*(\mathbf{h}_0, \tau_{0,J_n})\right)\right\} A' \\
&+ \hat{Q}_n^*(\mathbf{h}_0, \tau_{0,J_n}) - E\left(\hat{Q}_n^*(\mathbf{h}_0, \tau_{0,J_n})\right) + E\left(\hat{Q}_n(\hat{\mathbf{g}}, \hat{w}) - \hat{Q}_n^*(\mathbf{h}_0, \tau_{0,J_n})\right)
\end{aligned}$$

The first two terms are $o_p(n^{-1/2})$ by Lemma B.1. The third term is $O_p(n^{-1/2})$ by application of the Lindeberg-Feller central limit theorem for triangular arrays since condition (b) of Assumption **??** implies the relevant Lyapounov conditions.

Lastly,

$$E\left(\hat{Q}_n(\hat{\mathbf{g}},\hat{w})-\hat{Q}_n^*(\mathbf{h}_0,\tau_{0,J_n})\right)=E\left(\hat{Q}_n(\hat{\mathbf{g}},\hat{w})-\hat{Q}_n(\tilde{\mathbf{g}},w_{0,J_n})\right)$$
$$+E\left(\hat{Q}_n(\tilde{\mathbf{g}},w_{0,J_n})-\hat{Q}_n^*(\mathbf{h}_0,\tau_{0,J_n})\right)$$

The first term is $O_p(h_n^2+\frac{\log(n)}{\sqrt{nh_n}}+\frac{\log(\tilde{J}_n)^{p/2}}{h_n^{(p-1)}\tilde{J}_n^{p/2}})$ by Theorem C.2 and conditions (c) and (d) of Assumption **??**. The second term is $O\left((J_n^{-1}\log(J_n))^{1/2}\right)$ under conditions (d) and (e) of Assumption **??**, the proof of which is nearly identical to the proof of Lemma B.2. □

B.0.1 Some useful weak laws of large numbers

The following is an extension of Khintchin's WLLN that can be proved using the same methods employed to prove the well-known Kolmogorov-Feller WLLN.

Theorem B.1. *Suppose that for each n, the random variables $V_{1n,},\ldots,V_{nn}$ are i.i.d. Moreover, suppose that there exists an i.i.d. sequence of random variables $V_{1\infty},\ldots,V_{i\infty},\ldots$ such that $Pr(|V_{in}|>|V_{i\infty}|)=0$ and $E|V_{i\infty}|<\infty$. Then $n^{-1}\sum_{i=1}^n(V_{in}-E(V_{in}))\to_p 0$.*

Next, I provide a uniform WLLN. Let $V_{in}, 1\le i\le n$ be a triangular array of random variables where each V_{in} takes values in a (measurable) space $\mathcal{V}_n$ and for each $n\ge 1$ and each $\gamma\in\Gamma_n$, $h(v,\gamma)$ is a measurable function from $\mathcal{V}_n$ to $\mathbb{R}$. The following theorem extends Theorem 3(a) in Andrews (1992) by explicitly allowing for a triangular array and by allowing the parameter space to vary with n. Each parameter space Γ_n is assigned a metric $d_n(\cdot,\cdot)$. Moreover, a uniform version of the totally bounded assumption in Andrews (1992) is required. The family of parameter spaces, $\{\Gamma_n : n\ge 1\}$, is said to be uniformly totally bounded if for all $\varepsilon>0$ there exists an integer K such that each space Γ_n can be covered by no more than K balls of radius ε.

Theorem B.2. *If (a) $\{\Gamma_n : n\ge 1\}$ is a uniformly totally bounded family of parameter spaces, (b) for any sequence $\gamma_n\in\Gamma_n$, $n^{-1}\sum_{i=1}^n(h(V_{in},\gamma_n)-E(h(V_{in},\gamma_n)))\to_p 0$, and (c) $|h(V_{in},\gamma')-h(V_{in},\gamma)|\le|f_1(V_{in})|f_2(d_n(\gamma',\gamma))$ for all $\gamma',\gamma\in\Gamma_n$ almost surely, for a function $f_2(d)$ that converges to 0 as $d\to 0$ and a function f_1 such that $\sup_{n\ge1}n^{-1}\sum_{i=1}^n E(|f_1(V_{in})|)<\infty$ then*

$$\sup_{\gamma\in\Gamma_n}\left|\frac{1}{n}\sum_{i=1}^n h(V_{in},\gamma)-E(h(V_{in},\gamma))\right|\to_p 0$$

The proof of this follows as a variation in the proofs in Andrews (1992).

C Uniform convergence of kernel regression estimators

Consistent estimation of β_1 in the partially linear model in Section **??** requires uniform convergence of estimators of $E(W_i \mid \theta_i)$ for a random variable W_i. In this section, I provide three such results for the kernel regression estimator

$$\hat{g}_w(m) = \frac{\sum_{i=1}^n W_i K\left(\frac{\bar{M}_i - m}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{\bar{M}_i - m}{h_n}\right)}$$

where $\bar{M}_i = \tilde{J}^{-1}\sum_{j=1}^{\tilde{J}} \tilde{M}_{ij}$. Dependence on $\tilde{J}$ is left implicit in the notation for $\bar{M}_i$ for convenience. Let $\bar{p}_{\tilde{J}}(\theta) = E(\bar{M}_i \mid \theta_i = \theta)$. The results in this section will be applicable for the case where $\tilde{J} = J$ and $\tilde{M}_{ij} = M_{ij}$ for each j but also cases where $\tilde{\mathbf{M}}_i := (\tilde{M}_{i1}, \ldots, \tilde{M}_{i\tilde{J}})$ is some subset of the full vector of J items, $\mathbf{M}_i$. A statement of the main results and the sufficient conditions are collected in the first subsection and proofs are all in a separate section below.

C.1 Assumptions and statement of convergence results

Before stating the main uniform convergence results for $\hat{g}_w(m)$ I first state two important results regarding the convergence of $\bar{M}_i$ to $\bar{p}_{\tilde{J}}(\theta_i)$ under the following assumption.

Assumption C.1.

(a) *The binary random variables, $\tilde{M}_{i1}, \ldots, \tilde{M}_{i\tilde{J}}$ are mutually independent conditional on θ_i*

(b) *$\exists J_0$ such that, for each $\tilde{J} \geq J_0$, $\bar{p}_{\tilde{J}}(t)$ is strictly increasing, continuous and differentiable at all $t \in \mathbb{R}$ with derivative $D\bar{p}_{\tilde{J}}(t)$ such that for each $t \in \mathbb{R}$, the family of functions $\{D\bar{p}_{\tilde{J}} : \tilde{J} \geq J_0\}$ is equicontinuous at t. Moreover, for each $t \in \mathbb{R}$, $\inf_{\tilde{J} \geq J_0} D\bar{p}_{\tilde{J}}(t) > 0$.*

(c) *θ has absolutely continuous distribution function F_θ and density f_θ that is continuous and satisfies $0 < f_\theta(t) \leq \bar{f}_\theta$ for all $t \in \Theta :=$ support(θ_i).*

Lemma C.1. *Under Assumption C.1(a), if the sequence of random vectors $\tilde{\mathbf{M}}_i = (\tilde{M}_{i1}, \ldots, \tilde{M}_{i\tilde{J}})$, $i = 1, \ldots, n$ is i.i.d. for each $\tilde{J}$ then*

(a) *for any $\varepsilon > 0$, $Pr(|\bar{M}_i - \bar{p}_{\tilde{J}}(\theta_i)| > \varepsilon) \leq 2\exp(-2\tilde{J}\varepsilon^2)$*

(b) *for any $\varepsilon > 0$, $Pr(\max_{1\leq i\leq n} |\bar{M}_i - \bar{p}_{\tilde{J}}(\theta_i)| > \varepsilon) \leq 2n\exp(-2\tilde{J}\varepsilon^2)$*

(c) *for any $s > 0$, $\sup_{\theta\in\Theta} E(|\bar{M}_i - \bar{p}_{\tilde{J}}(\theta_i)|^s \mid \theta_i = \theta) = O\left(\left(\tilde{J}^{-1}\log\tilde{J}\right)^{s/2}\right)$*

The first two conclusions of this lemma are due to Douglas (2001). Theorem A.2 in Williams (2017) provides a similar result under a more general mixing condition in place of C.1(a). The proof is short but instructive.

Proof of Lemma C.1. First, (a) follows from Hoeffding's inequality since

$$\begin{aligned} Pr(|\bar{M}_i - \bar{p}_{\tilde{J}}(\theta_i)| > \varepsilon) &= \int Pr(|\bar{M}_i - \bar{p}_{\tilde{J}}(\theta_i)| > \varepsilon \mid \theta_i = \theta) f_\theta(\theta) d\theta \\ &\leq \int 2\exp(-2\tilde{J}\varepsilon^2) f_\theta(\theta) d\theta \end{aligned}$$

This then implies (b) since

$$\begin{aligned} Pr(\max_{1 \leq i \leq n} |\bar{M}_i - \bar{p}_{\tilde{J}}(\theta_i)| > \varepsilon) &\leq \sum_{i=1}^{n} Pr(|\bar{M}_i - \bar{p}_{\tilde{J}}(\theta_i)| > \varepsilon) \\ &\leq 2n\exp(-2\tilde{J}\varepsilon^2) \end{aligned}$$

Let $\eta_i = \bar{M}_i - \bar{p}_{\tilde{J}}(\theta_i)$ and define a sequence $\rho_{\tilde{J}} = \left(\frac{s}{4}\tilde{J}^{-1}\log(\tilde{J})\right)^{1/2}$. Then

$$\begin{aligned} &\sup_\theta E(|\bar{M}_i - \bar{p}_{\tilde{J}}(\theta_i)|^s \mid \theta_i = \theta) \\ &\leq \sup_\theta E(|\eta_i|^s \mathbf{1}(|\eta_i|^s \leq \rho_{\tilde{J}}^s) \mid \theta_i = \theta) + E(|\eta_i|^s \mathbf{1}(|\eta_i|^s > \rho_{\tilde{J}}^s) \mid \theta_i = \theta) \\ &\leq \rho_{\tilde{J}}^s + Pr(|\eta_i|^s > \rho_{\tilde{J}}^s \mid \theta_i = \theta) \\ &= \rho_{\tilde{J}}^s + Pr(|\eta_i| > \rho_{\tilde{J}} \mid \theta_i = \theta) \\ &\leq \rho_{\tilde{J}}^s + 2\tilde{J}^{-s/2} \end{aligned}$$

where the final line follows from an application of Hoeffding's inequality and (c) then follows from the definition of $\rho_{\tilde{J}}$. □

In addition to Assumption C.1, I will impose additional regularity conditions and assumptions on the rate of convergence of the bandwidth sequence h_n and impose properties for the kernel function, K to derive asymptotic convergence of $\hat{g}_w$. The following conditions are used for the first result.

Assumption C.2.

(a) $W_i \perp\!\!\!\perp \tilde{\mathbf{M}}_i \mid \theta_i$

(b) *The function $h_{w,0}(t) := E(W_i \mid \theta_i = t)$ is continuous for all $t \in \mathbb{R}$ and is differentiable at all $t \in \mathbb{R}$ with derivative $Dh_{w,0}(t)$ that is also continuous at all $t \in \mathbb{R}$.*

(c) $E|e_i|^3 < \infty$ *and for any* $\delta > 0$, $\sup_{\theta \in \Theta_\delta} E(|e_i|^3 \mid \theta_i = \theta) < \infty$ *where* $e_i = W_i - h_{w,0}(\theta_i)$.

(d) K *is nonnegative and twice differentiable with continuous first and second derivatives* K' *and* K''. *All three functions* $K(u)$, $K'(u)$, *and* $K''(u)$ *are bounded by* $\bar{K}\mathbf{1}(|u| \leq 1)$, *and* $K(u) \geq \underline{K}\mathbf{1}(|u| \leq 1/2)$ *for all* $u \in \mathbb{R}$, *for constants* $0 < \underline{K} < \bar{K} < \infty$.

By Assumption C.1(b), for each $\tilde{J}$, the function $\bar{p}_{\tilde{J}}$ has an inverse which is well-defined on its range, which is an interval in $[0,1]$. The inverse $\bar{p}_{\tilde{J}}^{-1}(m)$ can be extended to $[0,1]$ by assigning the values $\inf \Theta$ and $\sup \Theta$ for values of m below and above this interval, respectively. Then define $\tilde{g}_w(m) = h_{w,0}(\bar{p}_{\tilde{J}}^{-1}(m))$.

Also, for a fixed $0 < \delta < 1/2$, let Θ_δ denote the interval $[q_\delta(\theta_i), q_{1-\delta}(\theta_i)]$ and define $\mathcal{M}_\delta = \bar{p}_J(\Theta_\delta) = \{m \in [0,1] : m = \bar{p}_{\tilde{J}}(\theta) \text{ for some } \theta \in \Theta_\delta\}$. Though it is suppressed in the notation, $\mathcal{M}_\delta$ varies with $\tilde{J}$.

Theorem C.1. *Under Assumptions C.1 and C.2, if* $\tilde{J}_n$ *is a sequence such that* $\tilde{J}_n = O(n^r)$ *and* $\tilde{J}_n^{-1} = O(n^{-r})$ *for some* $r > 0$, $h_n \to 0$, $nh_n^3 \to \infty$, *and* $(\tilde{J}_n^{-1}\log(\tilde{J}_n))^{1/2}h_n^{-1} = o(1)$ *then there exists a constant* $0 < B < \infty$ *such that*

(a) $\lim_{J\to\infty} \sup_{m \in \mathcal{M}_\delta} |\tilde{g}_w(m)| \leq B$ *and* $\lim_{J\to\infty} \sup_{m \in \mathcal{M}_\delta} |D\tilde{g}_w(m)| \leq B$

(b) $\lim_{n\to\infty} Pr(\sup_{m \in \mathcal{M}_\delta} \max_{i:|\bar{M}_i - m| \leq h} \sup_{t \in [0,1]} |D\tilde{g}_w(t\bar{p}_{\tilde{J}_n}(\theta_i) + (1-t)m)| > B) = 0$ *and* $\lim_{n\to\infty} Pr(\inf_{m \in \mathcal{M}_\delta} \sup_{t \in [0,1]} |D\bar{p}_{\tilde{J}_n}(\bar{p}_{\tilde{J}_n}^{-1}(t2h_n + (1-t)m))| > B^{-1}) = 0$

(c) $\lim_{n\to\infty} Pr(\sup_{m \in \mathcal{M}} \left|(nh_n)^{-1}\sum_{i=1}^n K\left(\frac{\bar{M}_i - m}{h_n}\right)\right| \leq B^{-1}) = 0$

(d) $\sup_{m \in \mathcal{M}_\delta} |\hat{g}_w(m) - \tilde{g}_w(m)| = O_p(h_n) + O_p(\log(n)(nh_n)^{-1/2})$

(e) $\lim_{n\to\infty} Pr(\sup_{m \in \mathcal{M}_\delta} |\hat{g}_w(m)| \leq B) = 1$

(f) $\lim_{n\to\infty} Pr(\sup_{m \in \mathcal{M}_\delta} |D\hat{g}_w(m)| \leq B) = 1$

The convergence rate in conclusion (d) of Theorem C.1 is not sufficient for $\sqrt{n}-$convergence of semiparametric estimators based on $\hat{g}$ because the convergence rate is not faster than $n^{-1/4}$ when $r \leq 1/2$, which is the case if $\sqrt{n}/\tilde{J}_n \to \gamma > 0$. This is because if $\tilde{J}_n = O(n^r)$ for $r \leq 1/2$ then the restriction $(\tilde{J}_n^{-1}\log(\tilde{J}_n))^{1/2}h_n^{-1} = o(1)$ implies that $h_n^{-1} = O(n^{r/2}) = O(n^{1/4})$ which implies that $O_p(h_n)$ is not $o_p(n^{-1/4})$. Fortunately, this convergence rate can be improved under the following assumption, which implies the conditions of Assumption C.2 but imposes several additional smoothness restrictions.

Assumption C.3.

(a) $W_i \perp\!\!\!\perp \tilde{\mathbf{M}}_i \mid \theta_i$.

(b) *The function* $h_{w,0}(t) := E(W_i \mid \theta_i = t)$ *is continuous for all* $t \in \mathbb{R}$ *and is twice differentiable at all* $t \in \mathbb{R}$ *with first and second derivatives* $Dh_{w,0}(t)$ *and* $D^2 h_{w,0}(t)$ *that are both continuous at all* $t \in \mathbb{R}$.

(c) $\exists J_0$ *such that, for each* $\tilde{J} \geq J_0$, $\bar{p}_{\tilde{J}}(t)$ *is twice differentiable at all* $t \in \mathbb{R}$ *with second derivative* $D^2\bar{p}_{\tilde{J}}(t)$ *such that for each* $t \in \mathbb{R}$, *the family of functions* $\{D^2\bar{p}_{\tilde{J}} : \tilde{J} \geq J_0\}$ *is equicontinuous at* t.

(d) *the density function* f_θ *is differentiable with derivative* $Df_\theta(t)$ *that is continuous at all* $t \in \mathbb{R}$.

(e) *For each* $s \in \mathbb{N}$, $2 \leq s < p$, *the function* $\omega_{s\tilde{J}}(t) = \tilde{J}^{s/2} E(\eta_i^s \mid \theta_i = t)$ *is differentiable with derivative* $D\omega_{s\tilde{J}}(t)$ *such that for each* $t \in \mathbb{R}$, *the family of functions* $\{\omega_{s\tilde{J}} : \tilde{J} \geq J_0\}$ *is equicontinuous at* t *and the family of functions* $\{D\omega_{s\tilde{J}} : \tilde{J} \geq J_0\}$ *is equicontinuous at* t.

(f) $E|e_i|^q < \infty$, $E|W_i|^q < \infty$ *and for any* $\delta > 0$, $\sup_{\theta \in \Theta_\delta} E(|e_i|^q \mid \theta_i = \theta) < \infty$ *and* $\sup_{\theta \in \Theta_\delta} E(|W_i|^q \mid \theta_i = \theta) < \infty$ *for some* $q \geq 3$, *where* $e_i = W_i - h_w(\theta_i)$.

(g) K *is nonnegative and* $p+1$*-times differentiable and, for* $0 \leq s \leq p+1$, $K^{(s)}(u)$ *is continuous for all* $u \in \mathbb{R}$, *where* $K^{(s)}(u) := \frac{d^s}{du^s}K(u)$. *Also, for each* $0 \leq s \leq p+1$, $|K^{(s)}(u)| \leq \bar{K}\mathbf{1}(|u| \leq 1)$ *and* $K(u) \geq \underline{K}\mathbf{1}(|u| \leq 1/2)$ *for all* $u \in \mathbb{R}$, *for constants* $0 < \underline{K} < \bar{K} < \infty$.

Theorem C.2. *Under Assumptions C.1 and C.3, if* $\tilde{J}_n$ *is a sequence such that* $\tilde{J}_n = O(n^r)$ *and* $\tilde{J}_n^{-1} = O(n^{-r})$ *for some* $r > 0$, $h_n \to 0$, $nh_n^3 \to \infty$, $(\tilde{J}_n^{-1}\log(\tilde{J}_n))^{1/2} h_n^{-1} = o(1)$ *then*

$$\sup_{m \in \mathcal{M}_\delta} |\hat{g}_w(m) - \tilde{g}_w(m)| = O_p\left(h_n^2 + \frac{\log(n)}{\sqrt{nh_n}} + \frac{\log(\tilde{J}_n)^{p/2}}{h_n^{(p-1)}\tilde{J}_n^{p/2}}\right).$$

C.2 Proofs

The proof of Theorems C.1 and C.2 both rely on the following lemma. This result is proved below using arguments that are standard in the literature (see, e.g., Hansen (2008)).

Lemma C.2. *Let* $\Delta_n^{Vsa}(m) = (nh_n)^{-1}\sum_{i=1}^n V_i \eta_i^s \kappa\left(\frac{\bar{M}_{iJ}^a \bar{p}_J(\theta_i)^{(1-a)} - m}{h_n}\right)$ *for an i.i.d random vector* $\{V_i\}_{i=1}^n$, *nonnegative integer* s, *and* $a \in \{0,1\}$. *If* $(V_i, \theta_i, \tilde{\mathbf{M}}_i), i = 1, \ldots, n$ *is an i.i.d. random sequence, Assumption C.1 holds and, in addition,*

(a) $V_i \perp\!\!\!\perp \tilde{\mathbf{M}}_i \mid \theta_i$, $E|V_i|^q < \infty$ *for some* $q > 2$, *and for any* $\delta > 0$, $\sup_{\theta \in \Theta_\delta} E(|V_i|^q \mid \theta_i = \theta) < \infty$.

(b) $|\kappa(u)| \leq B\mathbf{1}(|u| \leq 1)$ *and* κ *has a derivative,* κ' *which is continuous and is also bounded by* B.

(c) $\tilde{J}_n$ *is a sequence such that* $\tilde{J}_n = O(n^r)$ *and* $\tilde{J}_n^{-1} = O(n^{-r})$ *for some* $r > 0$, $h_n \to 0$, $h_n^{-1} = O(n^\alpha)$ *for some* $\alpha > 0$ *such that* $q(1-\alpha) > 2$, *and* $(\tilde{J}_n^{-1}\log(\tilde{J}_n))^{1/2} h_n^{-1} = o(1)$.

then

$$\sup_{m\in\mathcal{M}_\delta} |\Delta_n^{Vsa}(m) - E(\Delta_n^{Vsa}(m))| = O_p\left(\log(n)(nh_n)^{-1/2}\left(\tilde{J}_n^{-1}\log(\tilde{J}_n)\right)^{s/2}\right)$$

Moreover, if $\Delta_n^{Vs}(m) = (nh_n)^{-1}\sum_{i=1}^n V_i \int_{\bar{p}_J(\theta_i)}^{\bar{M}_i} (\bar{M}_i - t)^{s-1}\kappa\left(\frac{t-m}{h_n}\right) dt$ *then*

$$\sup_{m\in\mathcal{M}_\delta} |\Delta_n^{Vs}(m) - E(\Delta_n^{Vs}(m))| = O_p\left(\log(n)(nh_n)^{-1/2}\left(\tilde{J}_n^{-1}\log(\tilde{J}_n)\right)^{s/2}\right)$$

I now provide the proofs of the three main uniform convergence results. Where it is not necessary for understanding the notation is simplified by omitting the n subscript on $\tilde{J}_n$ and the $\tilde{J}$ subscript on $\bar{p}_{\tilde{J}}$.

Proof of Theorem C.1. (a) $sup_{m\in\mathcal{M}_\delta}|\tilde{g}_w(m)| \leq \sup_{\theta\in\Theta_\delta}|\mathbf{h}_0(\theta)|$, which is bounded since Θ_δ is compact and h_0 is continuous, by Assumption C.2(a). The function $\tilde{g}_w(m) = \mathbf{h}_0(\bar{p}^{-1}(m))$ is differentiable with $D\tilde{g}_w(m) = D\mathbf{h}_0(\bar{p}^{-1}(m))\frac{1}{D\bar{p}(\bar{p}^{-1}(m))}$ since $D\bar{p} > 0$ by Assumption C.2(b). Then

$$\sup_{m\in\mathcal{M}_\delta} |D\tilde{g}_w(m)| \leq \frac{\sup_{\theta\in\Theta_\delta}|D\mathbf{h}_0(\theta)|}{\inf_{\theta\in\Theta_\delta}|D\bar{p}(\theta)|}$$

By Assumption C.2(a), the function $\mathbf{h}_0$ is continuous and hence bounded on the compact set Θ_δ and by Assumption C.2(b) $\inf_{\theta\in\Theta_\delta} D\bar{p}(\theta)$ is bounded away from 0 as $\tilde{J} \to \infty$.

(b) Let the bound found in the proof of (a) above be $B/2$. If $\sup_{m\in\mathcal{M}_\delta}\max_{i:|\bar{M}_i - m|\leq h_n}\sup_{t\in[0,1]}$ $D\tilde{g}_w(t\bar{p}(\theta_i) + (1-t)m) > B$ then there must be $m^* \in \mathcal{M}_\delta$ such that $\bar{M}_i - m^*| \leq h_n$ and $|D\tilde{g}_w(t\bar{p}(\theta_i) + (1-t)m^*) - D\tilde{g}_w(m^*)| > B/2$. By (a) and (b) of Assumption C.2, this implies that there is a constant $\epsilon > 0$ such that $|m^* - \bar{p}(\theta_i)| > \epsilon$. The result follows by Lemma C.1 and Assumption C.2(c) and Assumption C.2(f) since

$$Pr(\sup_{m\in\mathcal{M}_\delta}\max_{i:|\bar{M}_i - m|\leq h_n}|m - \bar{p}(\theta_i)| > \epsilon) \leq Pr(\max_{1\leq i\leq n}|\bar{M}_i - \bar{p}(\theta_i)| \geq \epsilon - h_n) = o(1)$$

The second part follows from Assumption C.2(b) by a similar argument.

(c) First, define $\hat{f}_1(m) = (nh_n)^{-1}\sum_{i=1}^n K\left(\frac{\bar{M}_i - m}{h_n}\right)$. Then

$$
\begin{aligned}
|\hat{f}_1(m)| &\geq \underline{K}(nh_n)^{-1}\sum_{i=1}^n \mathbf{1}(|\bar{M}_i - m| \leq h_n/2) \\
&\geq \underline{K}(nh_n)^{-1}\sum_{i=1}^n \mathbf{1}(|\bar{p}(\theta_i) - m| \leq h_n/4) - \underline{K}(nh_n)^{-1}\sum_{i=1}^n \mathbf{1}(|\bar{M}_i - \bar{p}(\theta_i)| \leq h_n/4) \\
&\geq \underline{K}(nh_n)^{-1}\sum_{i=1}^n \mathbf{1}(|\bar{p}(\theta_i) - m| \leq h_n/4) - o(1)
\end{aligned}
$$

where the second inequality follows from Assumption C.2(d) and the last line follows from Lemma C.1 and Assumptions C.2(c) and (f) since $\tilde{J}h_n^2 = \tilde{J}\rho_n^2(h_n/\rho_n)^2 \to \infty$.

Then, $\mathbf{1}(|\bar{p}(\theta_i) - m| \leq h_n/4) = \mathbf{1}(\bar{p}(\theta_i) \leq m + h_n/4) - \mathbf{1}(\bar{p}(\theta_i) \leq m - h_n/4)$ so

$$
\begin{aligned}
&\inf_{m\in\mathcal{M}_\delta} \underline{K}(nh_n)^{-1}\sum_{i=1}^n \mathbf{1}(|\bar{p}(\theta_i) - m| \leq h_n/4) \\
&\geq \underline{K}(h_n)^{-1}\inf_{m\in\mathcal{M}_\delta} Pr(|\bar{p}(\theta_i) - m| \leq h_n/4) \\
&- 2\underline{K}(h_n)^{-1}\sup_{s\in[0,1]}\left(n^{-1}\sum_{i=1}^n \mathbf{1}(\bar{p}(\theta_i) \leq s) - Pr(\bar{p}(\theta_i) \leq s)\right)
\end{aligned}
$$

The second term is $O_p(h_n^{-1}n^{-1/2})$ by the DKW inequality (see, e.g., p. 268 of Van der Vaart, 2000) applied to $\sup_{s^*\in\mathbb{R}}\left(n^{-1}\sum_{i=1}^n \mathbf{1}(\theta_i \leq s^*) - Pr(\theta_i \leq s^*)\right)$

Finally, for n large enough, either $m + h_n/4 \in \mathcal{M}_\delta$ or $m - h_n/4 \in \mathcal{M}_\delta$, or both, so I will assume wlog that $m + h_n/4 \in \mathcal{M}_\delta$. Then $Pr(\bar{p}(\theta_i) \leq m + h_n/4) = F_\theta(\bar{p}^{-1}(m + h_n/4))$ and $Pr(\bar{p}(\theta_i) \leq m) = F_\theta(\bar{p}^{-1}(m))$, so

$$
\begin{aligned}
&\underline{K}(h_n)^{-1}\inf_{m\in\mathcal{M}_\delta} Pr(|\bar{p}(\theta_i) - m| \leq h_n/4) \\
&\geq \underline{K}(h_n)^{-1}\inf_{m\in\mathcal{M}_\delta} Pr(\bar{p}(\theta_i) \leq m + h_n/4) - Pr(\bar{p}(\theta_i) \leq m) \\
&= \underline{K}(h_n)^{-1}\inf_{m\in\mathcal{M}_\delta}\left(F_\theta(\bar{p}^{-1}(m + h_n/4)) - F_\theta(\bar{p}^{-1}(m))\right) \\
&\geq \frac{\underline{K}\inf_{\theta\in\Theta_\delta} f_\theta(\theta)}{\sup_{m\in\mathcal{M}_\delta} D\bar{p}(\bar{p}^{-1}(m))}
\end{aligned}
$$

which is bounded away from 0 by Assumptions C.2(b) and (e).

(d) First, $W_i = h_{w,0}(\theta_i) + e_i = \tilde{g}_w(\bar{p}(\theta_i)) + e_i = \tilde{g}_w(m) + \tilde{g}_w(\bar{p}(\theta_i)) - \tilde{g}_w(m) + e_i$ where

Assumption C.1(c) implies that $E(e_i \mid \theta_i, \mathbf{M}_i) = 0$. Then

$$
\begin{aligned}
|\hat{g}_w(m) - \tilde{g}_w(m)| &\leq \frac{(nh_n)^{-1}\left|\sum_{i=1}^{n}(W_i - \tilde{g}_w(m))K\left(\frac{\bar{M}_i - m}{h_n}\right)\right|}{|\hat{f}_1(m)|} \\
&\leq |\hat{f}_1(m)|^{-1}\left|(nh_n)^{-1}\sum_{i=1}^{n}\left(\tilde{g}_w(\bar{p}(\theta_i)) - \tilde{g}_w(m)\right)K\left(\frac{\bar{M}_i - m}{h_n}\right)\right. \\
&\qquad\qquad\qquad\qquad\qquad\qquad \left. + (nh_n)^{-1}\sum_{i=1}^{n} e_i K\left(\frac{\bar{M}_i - m}{h_n}\right)\right|
\end{aligned}
$$

Next,

$$
\begin{aligned}
&\left|(nh_n)^{-1}\sum_{i=1}^{n}\left(\tilde{g}_w(\bar{p}(\theta_i)) - \tilde{g}_w(m)\right)K\left(\frac{\bar{M}_i - m}{h_n}\right)\right| \\
&\leq \left(\sup_{m \in \hat{\mathcal{M}}_\delta}|D\tilde{g}_w(m)|\right)(nh_n)^{-1}\sum_{i=1}^{n}\left(|\bar{M}_i - \bar{p}(\theta_i)| + |\bar{M}_i - m|\right)K\left(\frac{\bar{M}_i - m}{h_n}\right) \\
&\leq \left(\sup_{m \in \hat{\mathcal{M}}_\delta}|D\tilde{g}_w(m)|\right)\left\{(nh_n)^{-1}\bar{K}\sum_{i=1}^{n}\mathbf{1}(|\bar{M}_i - \bar{p}(\theta_i)| > \rho_n)\right. \\
&\qquad\qquad\qquad\qquad\qquad \left. + (1 + h_n^{-1}\rho_n)n^{-1}\sum_{i=1}^{n}\mathbf{1}(|\bar{M}_i - m| \leq h_n)\right\}
\end{aligned}
$$

where $\hat{\mathcal{M}}_\delta = \{t\bar{p}(\theta_i) + (1-t)m : m \in \mathcal{M}_\delta, t \in [0,1], |\bar{M}_i - m| < h_n\}$. The probability that the first term in braces is nonzero is bounded by $Pr(\max_{1 \leq i \leq n}|\bar{M}_i - \bar{p}(\theta_i)| > \rho_n)$ which is $o(n^{-1/2})$ by Assumption C.2(c), Assumption C.2(f), and Lemma C.1. Next, $(1 + h_n^{-1}\rho_n) = 1 + o(1)$, $n^{-1}\sum_{i=1}^{n}\mathbf{1}(|\bar{M}_i - m| \leq h_n) = n^{-1}\sum_{i=1}^{n}\mathbf{1}(\bar{M}_i \leq m + h_n) - n^{-1}\sum_{i=1}^{n}\mathbf{1}(\bar{M}_i \geq m - h_n)$, and

$$
n^{-1}\sum_{i=1}^{n}\mathbf{1}(\bar{M}_i \leq m + h_n) \leq n^{-1}\sum_{i=1}^{n}\mathbf{1}(\bar{p}(\theta_i) \leq m + 2h_n) + n^{-1}\sum_{i=1}^{n}\mathbf{1}(|\bar{M}_i - \bar{p}(\theta_i)| > h_n)
$$

The second term is $o_p(n^{-1/2})$, again by Lemma C.1, since $h_n \geq \rho_n$, at least for n sufficiently large. Therefore,

$$
\begin{aligned}
\sup_{m \in \mathcal{M}_\delta} n^{-1}\sum_{i=1}^{n}\mathbf{1}(|\bar{M}_i - m| \leq h_n) &\leq \sup_{m \in \mathcal{M}_\delta}|Pr(\bar{p}(\theta_i) \leq m + 2h_n) - Pr(\bar{p}(\theta_i) \leq m - 2h_n)| \\
&\quad + 2\sup_{s \in [0,1]}\left|n^{-1}\sum_{i=1}^{n}\mathbf{1}(\bar{p}(\theta_i) \leq s) - E(\mathbf{1}(\bar{p}(\theta_i) \leq s)\right| + o_p(n^{-1/2})
\end{aligned}
$$

Here, the first term is bounded by $8h\bar{f}_\theta/(\inf_{m\in\hat{\mathcal{M}}_\delta}\inf_{t\in[0,1]} D\bar{p}(\bar{p}^{-1}(tm+(1-t)2h_n))$ by (b) and (e) of Assumption C.2. The second term is $o_p(n^{-1/2})$ (see proof of (c) above). Next, since $E\left(e_i K\left(\frac{\bar{M}_i-m}{h_n}\right)\right)=0$, it remains to show that

$$|\Delta_n(m)-E(\Delta_n(m))|=O_p(r_n)$$

where $\Delta_n(m)=(nh_n)^{-1}\sum_{i=1}^n e_i K\left(\frac{\bar{M}_i-m}{h_n}\right)$ and $r_n=\log(n)(nh_n)^{-1/2}$. This follows by applying Lemma C.2 with $\kappa=K$, $V_i=e_i$, $s=0$, and $a=1$. Conditions (a)-(c) of the lemma are implied by Assumption C.2.

(e) This follows from parts (a) and (d) of the lemma, which have already been proved.

(f) First,

$$D\hat{g}_w(m)=\frac{D\hat{f}_w(m)}{\hat{f}_1(m)}-\frac{D\hat{f}_1(m)\hat{f}_w(m)}{(\hat{f}_1(m))^2}$$

where $\hat{f}_w(m)=(nh_n)^{-1}\sum_{i=1}^n W_i K\left(\frac{\bar{M}_i-m}{h_n}\right)$. Using $W_i=\tilde{g}_w(m)+\tilde{g}_w(\bar{p}(\theta_i))-\tilde{g}_w(m)+e_i$, and abbreviating $K_i=K\left(\frac{\bar{M}_i-m}{h_n}\right)$ and $K_i'=K'\left(\frac{\bar{M}_i-m}{h_n}\right)$, the same arguments used in the proof of (d) can be used to show that

$$|D\hat{g}_w(m)|\leq(|\hat{f}_1(m)|)^{-2}\left\{O_p(1)\left|(nh_n^2)^{-1}\sum_{i=1}^n e_iK_i'\right|+O_p(1)\left|(nh_n^2)^{-1}\sum_{i=1}^n e_iK_i\right|+O_p(1)\right\}$$

Then, since $E(e_iK_i)=E(e_iK_i')=0$, Lemma C.2 can be applied to conclude that

$$\sup_m\left|(nh_n^2)^{-1}\sum_{i=1}^n e_iK_i\right|=h_n^{-1}\sup_m\left|(nh_n)^{-1}\sum_{i=1}^n e_iK_i\right|=O_p((nh_n^3)^{-1/2})$$
$$\sup_m\left|(nh_n^2)^{-1}\sum_{i=1}^n e_iK_i'\right|=h_n^{-1}\sup_m\left|(nh_n)^{-1}\sum_{i=1}^n e_iK_i'\right|=O_p((nh_n^3)^{-1/2})$$

Conditions (a)-(c) of the lemma are implied Assumption C.2. This then implies by part(c) that $\sup_m|D\hat{g}_w(m)|=O_p(1)+O_p((nh_n^3)^{-1/2})$. The desired result follows since $nh_n^3\to\infty$.

□

Proof of Theorem C.2. First, define $\tilde{e}_i = W_i - \tilde{g}_w(m) = \tilde{g}_w(\bar{p}(\theta_i)) - \tilde{g}_w(m) + e_i$. Then

$$\begin{aligned}|\hat{g}_w(m) - \tilde{g}_w(m)| &\le |\hat{f}_1(m)|^{-1} \left| (nh_n)^{-1} \sum_{i=1}^n \tilde{e}_i K\left(\frac{\bar{M}_i - m}{h}\right) \right| \\ &\le |\hat{f}_1(m)|^{-1} \left| (nh_n)^{-1} \sum_{i=1}^n \tilde{e}_i K\left(\frac{\bar{p}(\theta_i) - m}{h}\right) \right| \\ &+ |\hat{f}_1(m)|^{-1} \left| (nh_n)^{-1} \sum_{i=1}^n \tilde{e}_i \left\{ K\left(\frac{\bar{M}_i - m}{h}\right) - K\left(\frac{\bar{p}(\theta_i) - m}{h}\right) \right\} \right|\end{aligned}$$

where $\hat{f}_1(m) = (nh_n)^{-1} \sum_{i=1}^n K\left(\frac{\bar{M}_i - m}{h_n}\right)$. Since Assumption C.3 implies Assumption C.2, $\sup_{m \in \mathcal{M}_\delta} |\hat{f}_1(m)|^{-1} = O_p(1)$ follows from conclusion (c) of Theorem C.1.

Under condition (f) of Assumption C.3, by a p^{th} order Taylor series expansion, $K(u') - K(u) = \sum_{s=1}^{p-1} \frac{(u'-u)^s}{s!} K^{(s)}(u) + \int_u^{u'} \frac{(u'-t)^{p-1}}{p!} K^{(p)}(t)dt$. Therefore,

$$\begin{aligned}|\hat{g}_w(m) - \tilde{g}_w(m)| &\le |\hat{f}_1(m)|^{-1} \sum_{s=0}^{p-1} (nh_n)^{-1} \left| \sum_{i=1}^n \tilde{e}_i \frac{\eta_i^s}{h_n^s s!} K^{(s)}\left(\frac{\bar{p}(\theta_i) - m}{h}\right) \right| \\ &+ |\hat{f}_1(m)|^{-1} (nh_n)^{-1} \left| \sum_{i=1}^n \tilde{e}_i \int_{\bar{p}(\theta_i)}^{\bar{M}_i} \frac{(\bar{M}_i - t)^{p-1}}{p! h_n^p} K^{(p)}\left(\frac{t-m}{h}\right) dt \right|\end{aligned}$$

Since $\tilde{e}_i = W_i - \tilde{g}_w(m)$, for each $0 \le s < p$,

$$\sup_{m \in \mathcal{M}_\delta} (nh_n)^{-1} \left| \sum_{i=1}^n \tilde{e}_i \frac{\eta_i^s}{h_n^s s!} K^{(s)}\left(\frac{\bar{p}(\theta_i) - m}{h}\right) \right| \tag{C.1}$$

$$\begin{aligned}&\le \sup_{m \in \mathcal{M}_\delta} (nh_n)^{-1} \left| \sum_{i=1}^n \left\{ W_i \frac{\eta_i^s}{h_n^s s!} K^{(s)}\left(\frac{\bar{p}(\theta_i) - m}{h}\right) - E\left(W_i \frac{\eta_i^s}{h_n^s s!} K^{(s)}\left(\frac{\bar{p}(\theta_i) - m}{h}\right)\right) \right\} \right| \\ &+ \left(\sup_{m \in \mathcal{M}_\delta} |\tilde{g}_w(m)| \right) \\ &\times \sup_{m \in \mathcal{M}_\delta} (nh_n)^{-1} \left| \sum_{i=1}^n \left\{ \frac{\eta_i^s}{h_n^s s!} K^{(s)}\left(\frac{\bar{p}(\theta_i) - m}{h}\right) - E\left(\frac{\eta_i^s}{h_n^s s!} K^{(s)}\left(\frac{\bar{p}(\theta_i) - m}{h}\right)\right) \right\} \right| \\ &+ (nh_n)^{-1} \left| \sum_{i=1}^n E\left(\tilde{e}_i \frac{\eta_i^s}{h_n^s s!} K^{(s)}\left(\frac{\bar{p}(\theta_i) - m}{h}\right)\right) \right|\end{aligned} \tag{C.2}$$

By application of Lemma C.2, first with $V_i = W_i$ and second with $V_i = 1$, each of the first two terms is $O_p\left(\log(\tilde{J}_n)^{s/2} h_n^{-s} \tilde{J}_n^{-s/2} \frac{\log(n)}{\sqrt{nh_n}}\right)$. In addition, I show below that

$$\sup_{m \in \mathcal{M}_\delta} (nh_n)^{-1} \left| \sum_{i=1}^{n} E\left(\tilde{e}_i \frac{\eta_i^s}{h_n^s s!} K^{(s)}\left(\frac{\bar{p}(\theta_i) - m}{h}\right)\right)\right| = O\left(h_n^{-(s-2)} \tilde{J}_n^{-s/2}\right) \tag{C.3}$$

Next, since $E(e_i \mid \bar{M}_i, \theta_i) = E(e_i \mid \theta_i) = 0$,

$$\begin{aligned}
&(nh_n)^{-1} \left| \sum_{i=1}^{n} \tilde{e}_i \int_{\bar{p}(\theta_i)}^{\bar{M}_i} \frac{(\bar{M}_i - t)^{p-1}}{p! h_n^p} K^{(p)}\left(\frac{t-m}{h}\right) dt \right| \\
&\leq (nh_n)^{-1} \left| \sum_{i=1}^{n} \left\{ e_i \int_{\bar{p}(\theta_i)}^{\bar{M}_i} \frac{(\bar{M}_i - t)^{p-1}}{p! h_n^p} K^{(p)}\left(\frac{t-m}{h}\right) dt \right.\right. \\
&\qquad\qquad \left.\left. - E\left(e_i \int_{\bar{p}(\theta_i)}^{\bar{M}_i} \frac{(\bar{M}_i - t)^{p-1}}{p! h_n^p} K^{(p)}\left(\frac{t-m}{h}\right) dt\right)\right\}\right| \\
&+ (nh_n)^{-1} \left| \sum_{i=1}^{n} \left(\tilde{g}_w(\bar{p}(\theta_i)) - \tilde{g}_w(m)\right) \int_{\bar{p}(\theta_i)}^{\bar{M}_i} \frac{(\bar{M}_i - t)^{p-1}}{p! h_n^p} K^{(p)}\left(\frac{t-m}{h}\right) dt \right|
\end{aligned}$$

Another application of Lemma C.2, this time with $V_i = e_i$, implies that the first term is $O_p\left(\log(\tilde{J})^{p/2} h_n^{-p} \tilde{J}^{-p/2} \frac{\log(n)}{\sqrt{nh_n}}\right)$. Lastly, I will show that

$$\begin{aligned}
&\sup_{m \in \mathcal{M}_\delta} (nh_n)^{-1} \left| \sum_{i=1}^{n} \left(\tilde{g}_w(\bar{p}(\theta_i)) - \tilde{g}_w(m)\right) \int_{\bar{p}(\theta_i)}^{\bar{M}_i} \frac{(\bar{M}_i - t)^{p-1}}{p! h_n^p} K^{(p)}\left(\frac{t-m}{h_n}\right) dt \right| \\
&= O_p\left(\log(\tilde{J}_n)^{p/2} h_n^{-p} \tilde{J}_n^{-p/2} h_n\right)
\end{aligned} \tag{C.4}$$

Then, since $h_n^{-s} \tilde{J}_n^{-s/2} = o(\log(\tilde{J}_n)^{s/2} h_n^{-s} \tilde{J}_n^{-s/2})$ and $\log(\tilde{J}_n)^{s/2} h_n^{-s} \tilde{J}_n^{-s/2} = O(1)$ for any $s \geq 0$,

$$\begin{aligned}
\sup_{m \in \mathcal{M}_\delta} |\hat{g}_w(m) - \tilde{g}_w(m)| &= O_p(1) O_p\left(\log(\tilde{J}_n)^{p/2} h_n^{-(p-1)} \tilde{J}_n^{-p/2} + \log(\tilde{J}_n)^{p/2} h_n^{-p} \tilde{J}_n^{-p/2} \frac{\log(n)}{\sqrt{nh_n}}\right. \\
&\qquad \left. + \sum_{s=0}^{p-1} \left\{ h_n^{-(s-2)} \tilde{J}_n^{-s/2} + \log(\tilde{J}_n)^{s/2} h_n^{-s} \tilde{J}_n^{-s/2} \frac{\log(n)}{\sqrt{nh_n}} \right\}\right) \\
&= O_p\left(h_n^2 + \frac{\log(n)}{\sqrt{nh_n}} + \log(\tilde{J}_n)^{p/2} h_n^{-(p-1)} \tilde{J}_n^{-p/2}\right)
\end{aligned}$$

Thus, it remains to prove (C.3) and (C.4). First, $E\left(e_i \eta_i^s K^{(s)}\left(\frac{\bar{p}(\theta_i)-m}{h_n}\right)\right) = 0$ so

$$
\begin{aligned}
&E\left(\tilde{e}_i \eta_i^s K^{(s)}\left(\frac{\bar{p}(\theta_i)-m}{h_n}\right)\right) = E\left((\tilde{g}_w(\bar{p}(\theta_i)) - \tilde{g}_w(m))\eta_i^s K^{(s)}\left(\frac{\bar{p}(\theta_i)-m}{h_n}\right)\right) \\
&= \tilde{J}_n^{-s/2} \int (\tilde{g}_w(\bar{p}(\theta)) - \tilde{g}_w(m))\tilde{\omega}_{s\tilde{J}}(\bar{p}(\theta)) K^{(s)}\left(\frac{\bar{p}(\theta)-m}{h_n}\right) \tilde{f}_{\theta_i}(\bar{p}(\theta)) d\theta \\
&= \tilde{J}_n^{-s/2} h_n \int (\tilde{g}_w(m+uh_n) - \tilde{g}_w(m))\tilde{\omega}_{s\tilde{J}}(m+uh_n) K^{(s)}(u)\, \tilde{f}_{\theta_i}(m+uh_n) d\theta
\end{aligned}
$$

where $\omega_{s\tilde{J}}(\theta) = \tilde{J}^{s/2} E(\eta_i^s \mid \theta_i = \theta)$, $\tilde{\omega}_{s\tilde{J}}(m) = \omega_{s\tilde{J}}(\bar{p}^{-1}(m))$, and $\tilde{f}_{\theta_i}(m) = \frac{f_\theta(\bar{p}^{-1}(m))}{D\bar{p}(\bar{p}^{-1}(m)}$ and the last line follows from the substitution $u = (\bar{p}(\theta) - m)/h_n$.

Next, use three Taylor approximations: $\tilde{g}_w(m^*) - \tilde{g}_w(m) = D\tilde{g}_w(m)(m^* - m) + \frac{1}{2}D^2\tilde{g}_w(m_a)(m^* - m)^2$, $\omega_{s\tilde{J}}(\bar{p}^{-1}(m^*)) = \omega_{s\tilde{J}}(\bar{p}^{-1}(m)) + D\tilde{\omega}_{s\tilde{J}}(m_b)(m^* - m)$, and $\tilde{f}_{\theta_i}(m^*) = \tilde{f}_{\theta_i}(m) + D\tilde{f}_{\theta_i}(m_c)(m^* - m)$ where m_a, m_b and m_c are all between m and m^*. Take n sufficiently large so that m and $m^* = m + uh_n$ are both contained in $\mathcal{M}_{\delta/2}$. Then, by the previous equation

$$
\begin{aligned}
&\tilde{J}_n^{s/2} \sup_{m \in \mathcal{M}_\delta} \left| E\left(\tilde{e}_i \eta_i^s K^{(s)}\left(\frac{\bar{p}(\theta_i)-m}{h}\right)\right)\right| \\
&= h_n^2 \sup_{m \in \mathcal{M}_\delta} \left| D\tilde{g}_w(m)\tilde{\omega}_{s\tilde{J}}(m) \frac{\tilde{f}_{\theta_i}(m)}{D\bar{p}(\bar{p}^{-1}(m))}\right| \int u K^{(s)}(u)\, du + O(h_n^3)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\sup_{m \in \mathcal{M}_\delta} (nh_n)^{-1} \left| \sum_{i=1}^n E\left(\tilde{e}_i \frac{\eta_i^s}{h_n^s s!} K^{(s)}\left(\frac{\bar{p}(\theta_i)-m}{h}\right)\right)\right| \\
&\leq \tilde{J}_n^{-s/2} \frac{1}{h_n^{s+1} s!} O(h_n^3) = O\left(h_n^{-(s-2)} \tilde{J}_n^{-s/2}\right)
\end{aligned}
$$

To prove (C.3), first observe that I can take $|\eta_i| \leq \delta_n := \left(c_0 \tilde{J}_n^{-1} \log(\tilde{J}_n)\right)^{1/2}$ for each i, where $2rc_0 > 1$, by Lemma C.1. Also, take n sufficiently large so that $\delta_n \leq h_n$. Then

$$
\begin{aligned}
&\sup_{m \in \mathcal{M}_\delta} (nh_n)^{-1} \left| \sum_{i=1}^n (\tilde{g}_w(\bar{p}(\theta_i)) - \tilde{g}_w(m)) \int_{\bar{p}(\theta_i)}^{\bar{M}_i} \frac{(\bar{M}_i - t)^{p-1}}{p! h_n^p} K^{(p)}\left(\frac{t-m}{h_n}\right) dt\right| \\
&\leq \left(\sup_{m \in \mathcal{M}_{\delta/2}} |D\tilde{g}_w(m)|\right) \sup_{m \in \mathcal{M}_\delta} \frac{\bar{K}\delta_n^p}{p! h_n^p} n^{-1} \sum_{i=1}^n \mathbf{1}(|\bar{p}(\theta_i) - m| \leq h_n) \\
&= O_p\left(\frac{\delta_n^p}{h_n^{p-1}}\right)
\end{aligned}
$$

The final line follows because, as argued in the proof of Theorem C.1 using the DKW inequality (see, e.g., p. 268 of Van der Vaart, 2000) and the fact that $Pr(|\bar{p}(\theta_i) - m| \leq h_n) = O(h_n)$,

$$n^{-1} \sum_{i=1}^{n} \mathbf{1}(|\bar{p}(\theta_i) - m| \leq h_n) = O(h_n) + O_p(n^{-1/2}) = O_p(h_n)$$

$\square$

Now a proof of Lemma C.2 is provided.

Proof of Lemma C.2. Let $r_n := \log(n)(nh_n)^{-1/2} \left(\tilde{J}_n^{-1} \log(\tilde{J}_n)\right)^{s/2}$. Define b_n such that $b_n^q = n \log(n)$ and let $\bar{V}_{in} = V_i \mathbf{1}(|V_i| \leq b_n)$ and $\bar{\eta}_{in} = \eta_i \mathbf{1}(|\eta_i| \leq \rho_n)$ where $\rho_n = (r^{-1} \tilde{J}_n^{-1} \log(\tilde{J}_n))^{1/2}$ for r given by condition (c) of the lemma. Let

$$\begin{aligned}
\Delta_n^{Vsa,r}(m) &= (nh_n)^{-1} \sum_{i=1}^{n} V_i \eta_i^s \kappa \left(\frac{\bar{M}_i^a \bar{p}(\theta_i)^{(1-a)} - m}{h_n} \right) \\
&\quad - (nh_n)^{-1} \sum_{i=1}^{n} \bar{V}_{in} \bar{\eta}_{in}^s \kappa \left(\frac{\bar{M}_i^a \bar{p}(\theta_i)^{(1-a)} - m}{h_n} \right) \\
&= (nh_n)^{-1} \sum_{i=1}^{n} V_i \mathbf{1}(|V_i| > b_n) \eta_i^s \kappa \left(\frac{\bar{M}_i^a \bar{p}(\theta_i)^{(1-a)} - m}{h_n} \right) \\
&\quad + (nh_n)^{-1} \sum_{i=1}^{n} \bar{V}_{in} \eta_i^s \mathbf{1}(|\eta_i| > \rho_n) \kappa \left(\frac{\bar{M}_i^a \bar{p}(\theta_i)^{(1-a)} - m}{h_n} \right) \\
&= \Delta_n^{Vsa,r1}(m) + \Delta_n^{Vsa,r2}(m)
\end{aligned}$$

Then $|\Delta_n^{Vsa}(m) - E(\Delta_n^{Vsa}(m))| \leq |\bar{\Delta}_n^{Vsa}(m)| + |\Delta_n^{Vsa,r}(m)| + |E(\Delta_n^{Vsa,r}(m))|$ where

$$\bar{\Delta}_n^{Vsa}(m) = (nh_n)^{-1} \sum_{i=1}^{n} \left(\bar{V}_{in} \bar{\eta}_{in}^s \kappa \left(\frac{\bar{M}_i^a \bar{p}(\theta_i)^{(1-a)} - m}{h_n} \right) - E \left(\bar{V}_{in} \bar{\eta}_{in}^s \kappa \left(\frac{\bar{M}_i^a \bar{p}(\theta_i)^{(1-a)} - m}{h_n} \right) \right) \right)$$

First, for any $t > 0$,

$$Pr(\sup_{m \in \mathcal{M}} |\Delta_n^{Vsa,r}(m)| > tr_n) \leq Pr(\max_{1 \leq i \leq n} |V_i| > b_n) + Pr(\max_{1 \leq i \leq n} |\eta_i| > \rho_n)$$

Then $Pr(\max_{1 \leq i \leq n} |V_i| > b_n) \leq nPr(|V_i| > b_n) \leq \frac{E(|V_i|^q)}{\log(n)} \to 0$, where the last inequality follows from Markov's inequality since $\frac{n}{b_n^q} = \frac{1}{\log(n)}$, and the limit holds by condition (a). And $Pr(\max_{1 \leq i \leq n} |\eta_i| > \rho_n) = o(n^{-1})$ by Lemma C.1 and condition (c).

Second, by condition (b)

$$\begin{aligned}\sup_{m\in\mathcal{M}_\delta}|E(\Delta_n^{Vsa,r1}(m))| &\le h_n^{-1}\sup_{m\in\mathcal{M}_\delta}E\left(|V_i|\mathbf{1}(|V_i|>b_n)|\eta_i|^s\mathbf{1}(|\bar{M}_i^a\bar{p}(\theta_i)^{(1-a)}-m|\le h_n|)\right)\\ &\le h_n^{-1}\sup_{m\in\mathcal{M}_\delta}E\left(|V_i|\mathbf{1}(|V_i|>b_n)|\eta_i|^s\mathbf{1}(|\bar{p}(\theta_i)-m|\le 2h_n|)\right) \qquad (C.5)\\ &+\left\{h_n^{-1}E\left(|V_i|\mathbf{1}(|V_i|>b_n)|\eta_i|^s\mathbf{1}(|\bar{M}_i-\bar{p}(\theta_i)|>h_n|)\right)\right\}^a\end{aligned}$$

For n sufficiently large, the first term satisfies

$$\begin{aligned}&h_n^{-1}\sup_{m\in\mathcal{M}_\delta}E\left(|V_i|\mathbf{1}(|V_i|>b_n)|\eta_i|^s\mathbf{1}(|\bar{p}(\theta_i)-m|\le 2h_n|)\right)\\ &\le h_n^{-1}\left(\sup_{\theta\in\Theta_{\delta/2}}|E(|\eta_i|^s\mid\theta_i=\theta)|\right)E\left(|V_i|\mathbf{1}(|V_i|>b_n)\mathbf{1}(|\bar{p}(\theta_i)-m|\le 2h_n)\right)\\ &\le h_n^{-1}\frac{1}{b_n^{q-1}}\left(\sup_{\theta\in\Theta_{\delta/2}}|E(|\eta_i|^s\mid\theta_i=\theta)|\right)E\left(|V_i|^q\mathbf{1}(|\bar{p}(\theta_i)-m|\le 2h_n)\right)\\ &\le h_n^{-1}\frac{1}{b_n^{q-1}}\left(\sup_{\theta\in\Theta_{\delta/2}}E(|\eta_i|^s\mid\theta_i=\theta)\right)\left(\sup_{\theta\in\Theta_{\delta/2}}E(|V_i|^q\mid\theta_i=\theta)\right)Pr(|\bar{p}(\theta_i)-m|\le 2h_n)\end{aligned}$$

where the first inequality is by the conditional independence between V_i and $\bar{M}_i$ conditional on θ_i under condition (a), the second is because $|V_i|>b_n$ implies that $|V_i|\le b_n^{q-1}|V_i|^q$, and the third is valid under condition (a). This term is $O_p(b_n^{-(q-1)}J_n^{-s/2})$ since $Pr(|\bar{p}(\theta_i)-m|\le 2h_n)\le Pr\left(|\theta_i-\bar{p}^{-1}(m)|\le 2h_n/\inf_{\theta\in\Theta_{\delta/2}}D\bar{p}(\theta)\right)\le 4\sup_{\theta\in\Theta}f_\theta(\theta)h_n/\inf_{\theta\in\Theta_{\delta/2}}D\bar{p}(\theta)$ and because
$\sup_{\theta\in\Theta_{\delta/2}}E(|\eta_i|^s\mid\theta_i=\theta)=O\left(\left(\tilde{J}_n^{-1}\log\tilde{J}_n\right)^{s/2}\right)$ by Lemma C.1. Lastly, it is easy to verify that $q>2$ implies that $O_p\left(b_n^{-(q-1)}\left(\tilde{J}^{-1}\log\tilde{J}\right)^{s/2}\right)=o_p(r_n)$ because $b_n^{q-1}>n^{(q-1)/q}>n^{1/2}$.

For any $t>0$, the second term in (C.5), for sufficiently large n, satisfies

$$\begin{aligned}&h_n^{-1}E\left(|V_i|\mathbf{1}(|V_i|>b_n)|\eta_i|^s\mathbf{1}(|\bar{M}_i-\bar{p}(\theta_i)|>h_n|)\right)\\ &\le h_n^{-1}E\left(|V_i|\mathbf{1}(|V_i|>b_n)|\eta_i|^s\mathbf{1}(|\bar{M}_i-\bar{p}(\theta_i)|>t\rho_n|)\right)\\ &\le h_n^{-1}E\left(|V_i|^2\right)^{1/2}Pr(|\bar{M}_i-\bar{p}(\theta_i)|>t\rho_n|)^{1/2}\\ &\le 2h_n^{-1}E\left(|V_i|^2\right)^{1/2}\exp(-J_nt^2\rho_n^2)\end{aligned}$$

where the first inequality follows from condition (c) in the statement of the lemma, the second follows from the Cauchy-Schwarz inequality and the fact that $|\eta_i|\le 1$, the third follows from Hoeffding's inequality. This term is $o_p(r_n)$ if $t^2>\frac{1+\alpha+sr}{2}$ because $E\left(|V_i|^2\right)<\infty$

by condition (a), and condition (d) of Assumption C.1 and condition (c) imply that $h_n^{-1}r_n^{-1} = O(n^{\frac{1}{2}(1+\alpha+sr)}\log(n)^{-1})$ for some $\alpha > 0$ and $\exp(-2J_n t^2\rho_n^2) \leq J_n^{-a^{-1}t^2} = o(n^{-t^2})$.

By Lemma C.1, condition (d) of Assumption C.1 and condition (c), and applying the same argument based on Hoeffding's inequality,

$$\begin{aligned}\sup_{m\in\mathcal{M}_\delta} |E(\Delta_n^{Vsa,r2}(m))| &\leq h_n^{-1} \sup_{m\in\mathcal{M}_\delta} E\left(|\bar{V}_{in}|\eta_i^s\mathbf{1}(|\eta_i| > \rho_n)|\mathbf{1}(|\bar{M}_i^a\bar{p}(\theta_i)^{(1-a)} - m| \leq h_n|)\right)\\ &\leq h_n^{-1}b_n Pr(|\eta_i| > \rho_n) = o(r_n)\end{aligned}$$

Third, since $|\bar{V}_{in}\bar{\eta}_{in}^s K\left(\frac{\bar{M}_i^a\bar{p}(\theta_i)^{(1-a)}-m}{h_n}\right)| \leq b_n\rho_n^s B$, I can apply Bernstein's inequality:

$$\begin{aligned}Pr(|\bar{\Delta}_n^{Vsa}(m)| > tr_n) &\leq \exp\left(-\frac{(tr_n nh_n)^2}{2nVar\left(\bar{V}_{in}\bar{\eta}_{in}^s K\left(\frac{\bar{M}_i^a\bar{p}(\theta_i)^{(1-a)}-m}{h_n}\right)\right) + \frac{4}{3}tBb_n\rho_n^s r_n nh_n}\right)\\ &\leq \exp\left(-\frac{t^2\log(n)}{c_1 + c_2 tb_n\log(n)(nh_n)^{-1/2}}\right) \qquad \text{(C.6)}\end{aligned}$$

where the second inequality follows for some positive constants c_1, c_2 because $(r_n nh_n)^2 = O\left(\log(n)nh_n\left(\tilde{J}_n^{-1}\log(\tilde{J}_n)\right)^s\right)$ and

$$\begin{aligned}&Var\left(\bar{V}_{in}\bar{\eta}_{in}^s K\left(\frac{\bar{M}_i^a\bar{p}(\theta_i)^{(1-a)} - m}{h_n}\right)\right)\\ &\leq E\left(\bar{V}_{in}^2\bar{\eta}_{in}^{2s}\mathbf{1}(|\bar{p}(\theta_i) - m| \leq 2h_n)\right) + E\left(\bar{V}_{in}^2\bar{\eta}_{in}^{2s}\mathbf{1}(|\bar{M}_i - \bar{p}(\theta_i)| \geq h_n)\right)\\ &\leq \rho_n^{2s}\left(\sup_{\theta\in\Theta_{\delta/2}} E(V_i^2 \mid \theta_i = \theta)\right) Pr(|\bar{p}(\theta_i) - m| \leq 2h_n) + b_n^2 Pr\left(|\bar{M}_i - \bar{p}(\theta_i)| \geq h_n\right)\\ &= O\left(\left(\tilde{J}_n^{-1}\log(\tilde{J}_n)\right)^s h_n\right)\end{aligned}$$

where the last line follows because, using the same argument based on Hoeffding's inequality $b_n^2 Pr\left(|\bar{M}_i - \bar{p}(\theta_i)| \geq h_n\right) = o(n^{-C})$ for any $C > 0$.

Next, partition $\mathcal{M}_\delta$ into $N \leq \frac{1}{r_n h_n}$ intervals of width $r_n h_n$ centered at $\{m_j\}_{j=1}^N$. Since $\left|K\left(\frac{\bar{M}_i^a\bar{p}(\theta_i)^{(1-a)}-m_j}{h_n}\right) - K\left(\frac{\bar{M}_i^a\bar{p}(\theta_i)^{(1-a)}-m}{h_n}\right)\right| \leq \frac{|m_j-m|}{h_n}\mathbf{1}(|\bar{M}_i^a\bar{p}(\theta_i)^{(1-a)} - m_j| \leq 2h_n)$, for n sufficiently large, following an argument due to Hansen (2008),

$$Pr(\sup_{m\in\mathcal{M}_\delta} |\bar{\Delta}_n^{Vsa}(m)| > 3tr_n) \leq NPr(|\bar{\Delta}_n^{Vsa}(m)| > tr_n) + NPr(|\bar{\Delta}_n^{*Vsa}(m)| > tr_n)$$

provided that $E|\bar{\Delta}_n^{*Vsa}(m)|$ is bounded, where

$$\bar{\Delta}_n^{*Vsa}(m) = (nh_n)^{-1}\sum_{i=1}^n \left(\bar{V}_{in}\bar{\eta}_{in}^s \mathbf{1}(|\bar{M}_i^a \bar{p}(\theta_i)^{(1-a)} - m| \leq h_n)\right.$$
$$\left. -E\left(\bar{V}_{in}\bar{\eta}_{in}^s \mathbf{1}(|\bar{M}_i^a \bar{p}(\theta_i)^{(1-a)} - m| \leq h_n)\right)\right)$$

The same arguments used above can be repeated to show that $E|\bar{\Delta}_n^{*Vsa}(m)|$ is bounded uniformly over $m \in \mathcal{M}_\delta$ and that the bound on $Pr(|\bar{\Delta}_n^{Vsa}(m)| > tr_n)$ derived in equation (C.6) applies to $NPr(|\bar{\Delta}_n^{*Vsa}(m)| > tr_n)$ as well. Therefore,for t large enough

$$Pr(\sup_{m\in\mathcal{M}_\delta} |\bar{\Delta}_n^{*Vsa}(m)| > 3tr_n) \leq 2N\exp\left(-\frac{t^2\log(n)}{c_1 + c_2 t b_n \log(n)(nh_n)^{-1/2}}\right) \to 0$$

where convergence follows because $n^{q^{-1}}(nh_n)^{-1/2} = O(n^{-(\frac{1}{2}-\frac{1}{q}-\frac{\alpha}{2})})$, which implies that $b_n\log(n)(nh_n)^{-1/2} = o(1)$ if $\frac{1}{2}-\frac{1}{q}-\frac{\alpha}{2} > 0$. The latter follows from the restriction in condition (c) that $q(1-\alpha) > 2$.

The result for $\Delta_n^{Vs}(m)$ follows by essentially the same argument. Let

$$\begin{aligned}
\Delta_n^{Vs,r}(m) &= \Delta_n^{Vs}(m) - (nh_n)^{-1}\sum_{i=1}^n \bar{V}_{in}\mathbf{1}(|\eta_i| \leq \rho_n)\int_{\bar{p}(\theta_i)}^{\bar{M}_i}(\bar{M}_i - t)^{s-1}\kappa\left(\frac{t-m}{h_n}\right)dt \\
&= (nh_n)^{-1}\sum_{i=1}^n V_i\mathbf{1}(|V_i| > b_n)\int_{\bar{p}(\theta_i)}^{\bar{M}_i}(\bar{M}_i - t)^{s-1}\kappa\left(\frac{t-m}{h_n}\right)dt \\
&+ (nh_n)^{-1}\sum_{i=1}^n \bar{V}_{in}\mathbf{1}(|\eta_i| > \rho_n)\int_{\bar{p}(\theta_i)}^{\bar{M}_i}(\bar{M}_i - t)^{s-1}\kappa\left(\frac{t-m}{h_n}\right)dt \\
&= \Delta_n^{Vs,r1}(m) + \Delta_n^{Vs,r2}(m)
\end{aligned}$$

Then $|\Delta_n^{Vs}(m) - E(\Delta_n^{Vs}(m))| \leq |\bar{\Delta}_n^{Vs}(m)| + |\Delta_n^{Vs,r}(m)| + |E(\Delta_n^{Vs,r}(m))|$ where

$$\bar{\Delta}_n^{Vs}(m) = (nh_n)^{-1}\sum_{i=1}^n \left(\bar{V}_{in}\mathbf{1}(|\eta_i| \leq \rho_n)\int_{\bar{p}(\theta_i)}^{\bar{M}_i}(\bar{M}_i - t)^{s-1}\kappa\left(\frac{t-m}{h_n}\right)dt\right.$$
$$\left. -E\left(\bar{V}_{in}\mathbf{1}(|\eta_i| \leq \rho_n)\int_{\bar{p}(\theta_i)}^{\bar{M}_i}(\bar{M}_i - t)^{s-1}\kappa\left(\frac{t-m}{h_n}\right)dt\right)\right)$$

First, for any $t > 0$,

$$Pr(\sup_{m\in\mathcal{M}_\delta} |\Delta_n^{Vs,r}(m)| > tr_n) \leq Pr(\max_{1\leq i\leq n}|V_i| > b_n) + Pr(\max_{1\leq i\leq n}|\eta_i| > \rho_n) \to 0$$

Second, by condition (b),

$$\left|\int_{\bar{p}(\theta_i)}^{\bar{M}_i} (\bar{M}_i - t)^{s-1} \kappa\left(\frac{t-m}{h_n}\right) dt\right| \leq |\eta_i|^{s-1} B \int_{\bar{p}(\theta_i)}^{\bar{M}_i} \mathbf{1}(|t-m| \leq h_n) dt$$

which implies that

$$\begin{aligned}\sup_{m \in \mathcal{M}_\delta} |E(\Delta_n^{Vsa,r1}(m))| \leq h_n^{-1} &\sup_{m \in \mathcal{M}_\delta} E\left(|V_i| \mathbf{1}(|V_i| > b_n) |\eta_i|^s \mathbf{1}(|\bar{p}(\theta_i) - m| \leq 2h_n|)\right) \\ &+ h_n^{-1} E\left(|V_i| \mathbf{1}(|V_i| > b_n) |\eta_i|^{s-1} \mathbf{1}(|\bar{M}_i - \bar{p}(\theta_i)| > h_n|)\right)\end{aligned}$$

Both terms are $o(r_n)$, as argued above. And by Lemma C.1 and conditions (b) and (c)

$$\sup_{m \in \mathcal{M}_\delta} |E(\Delta_n^{Vsa,r2}(m))| \leq h_n^{-1} b_n Pr(|\eta_i| > \rho_n) = o(r_n)$$

Third,

$$\left|\bar{V}_{in} \mathbf{1}(|\eta_i| \leq \rho_n) \int_{\bar{p}(\theta_i)}^{\bar{M}_i} (\bar{M}_i - t)^{s-1} \kappa\left(\frac{t-m}{h_n}\right) dt\right| \leq b_n \rho_n^s B$$

and

$$\begin{aligned}&Var\left(\bar{V}_{in} \mathbf{1}(|\eta_i| \leq \rho_n) \int_{\bar{p}(\theta_i)}^{\bar{M}_i} (\bar{M}_i - t)^{s-1} \kappa\left(\frac{t-m}{h_n}\right) dt\right) \\ &\leq E\left(\bar{V}_{in}^2 \bar{\eta}_{in}^{2s} B^2 \mathbf{1}(|\bar{p}(\theta_i) - m| \leq 2h_n)\right) + E\left(\bar{V}_{in}^2 \bar{\eta}_{in}^{2s} B^2 \mathbf{1}(|\bar{M}_i - \bar{p}(\theta_i)| \geq h_n)\right) \\ &= O\left(\left(\tilde{J}_n^{-1} \log(\tilde{J}_n)\right)^s h_n\right)\end{aligned}$$

so Bernstein's inequality can be applied as above to obtain

$$Pr(|\bar{\Delta}_n^{Vs}(m)| > tr_n) \leq \exp\left(-\frac{t^2 \log(n)}{c_1 + c_2 t b_n \log(n)(nh_n)^{-1/2}}\right)$$

The desired result follows by partitioning $\mathcal{M}_\delta$ into $N \leq \frac{1}{r_n h_n}$ intervals of width $r_n h_n$ centered at $\{m_j\}_{j=1}^N$, as above, and combining results since, for n large enough that $r_n < 1$,

$$\begin{aligned}
&\left|\int_{\bar{p}(\theta_i)}^{\bar{M}_i} (\bar{M}_i - t)^{s-1} \kappa\left(\frac{t-m}{h_n}\right) dt - \int_{\bar{p}(\theta_i)}^{\bar{M}_i} (\bar{M}_i - t)^{s-1} \kappa\left(\frac{t-m_j}{h_n}\right) dt\right| \\
&\leq \int_{\bar{p}(\theta_i)}^{\bar{M}_i} |\bar{M}_i - t|^{s-1} \left|\kappa\left(\frac{t-m}{h_n}\right) - \kappa\left(\frac{t-m_j}{h_n}\right)\right| dt \\
&\leq \int_{\bar{p}(\theta_i)}^{\bar{M}_i} |\bar{M}_i - t|^{s-1} \frac{|m-m_j|}{h_n} \mathbf{1}(|t-m_j| \leq 2h_n) dt
\end{aligned}$$

□

References

Andrews, D. W. (1992): "Generic uniform convergence," *Econometric theory*, 8, 241–257.

——— (1994): "Empirical process methods in econometrics," *Handbook of econometrics*, 4, 2247–2294.

Douglas, J. (2001): "Asymptotic identifiability of nonparametric item response models," *Psychometrika*, 66, 531–540.

Van der Vaart, A. W. (2000): *Asymptotic statistics*, vol. 3, Cambridge university press.

Table B.1. Monte Carlo results for the partially linear regression model

n	J	model 1, a=1						model 1, a=2						model 1, a=4					
		OLS		IRT		PLR		OLS		IRT		PLR		OLS		IRT		PLR	
		bias	sd	bias	sd	bias	sd	bias	sd	bias	sd	bias	sd	bias	sd	bias	sd	bias	sd
	50	0.05	0.08	0.05	0.08	0.03	0.09	-0.02	0.10	0.09	0.10	0.06	0.10	-0.13	0.11	0.13	0.10	0.11	0.13
1000	100	0.03	0.08	0.04	0.08	0.02	0.09	-0.06	0.10	0.07	0.09	0.03	0.10	-0.19	0.12	0.09	0.11	0.06	0.13
	500	0.01	0.08	0.01	0.08	0.00	0.09	-0.10	0.10	0.03	0.09	0.01	0.10	-0.26	0.12	0.03	0.10	0.01	0.13
	50	0.05	0.06	0.05	0.06	0.03	0.06	-0.02	0.07	0.09	0.07	0.06	0.07	-0.13	0.08	0.12	0.07	0.11	0.09
2000	100	0.03	0.05	0.04	0.05	0.02	0.06	-0.06	0.07	0.07	0.07	0.03	0.07	-0.21	0.08	0.10	0.07	0.06	0.09
	500	0.02	0.06	0.01	0.06	0.00	0.06	-0.10	0.07	0.02	0.06	0.01	0.07	-0.26	0.08	0.03	0.07	0.01	0.10
		model 2, a=1						model 2, a=2						model 2, a=4					
		OLS		IRT		PLR		OLS		IRT		PLR		OLS		IRT		PLR	
		bias	sd	bias	sd	bias	sd	bias	sd	bias	sd	bias	sd	bias	sd	bias	sd	bias	sd
	50	0.19	0.07	0.07	0.07	0.01	0.08	0.12	0.08	0.02	0.08	0.02	0.09	-0.03	0.09	-0.09	0.08	0.02	0.09
1000	100	0.19	0.07	0.04	0.07	0.00	0.08	0.11	0.08	-0.02	0.08	0.01	0.09	-0.04	0.09	-0.16	0.09	0.01	0.09
	500	0.18	0.07	0.02	0.07	0.00	0.08	0.09	0.08	-0.09	0.08	0.00	0.08	-0.08	0.10	-0.25	0.09	0.00	0.09
	50	0.19	0.05	0.07	0.05	0.01	0.06	0.11	0.06	0.02	0.06	0.02	0.06	-0.02	0.07	-0.09	0.07	0.03	0.06
2000	100	0.18	0.05	0.05	0.05	0.00	0.06	0.09	0.06	-0.03	0.06	0.01	0.06	-0.06	0.07	-0.16	0.06	0.01	0.06
	500	0.18	0.05	0.02	0.05	0.00	0.06	0.09	0.06	-0.08	0.06	0.00	0.06	-0.09	0.07	-0.25	0.07	0.00	0.07
		model 3, a=1						model 3, a=2						model 3, a=4					
		OLS		IRT		PLR		OLS		IRT		PLR		OLS		IRT		PLR	
		bias	sd	bias	sd	bias	sd	bias	sd	bias	sd	bias	sd	bias	sd	bias	sd	bias	sd
	50	0.10	0.08	0.22	0.09	0.09	0.10	0.29	0.11	0.57	0.10	0.22	0.13	0.48	0.13	0.83	0.11	0.49	0.19
1000	100	0.08	0.08	0.22	0.08	0.05	0.09	0.26	0.11	0.59	0.10	0.14	0.14	0.47	0.13	0.84	0.11	0.37	0.20
	500	0.06	0.08	0.20	0.09	0.01	0.09	0.21	0.11	0.57	0.10	0.04	0.13	0.41	0.13	0.81	0.11	0.14	0.23
	50	0.11	0.06	0.22	0.06	0.09	0.07	0.32	0.08	0.57	0.08	0.23	0.09	0.53	0.09	0.82	0.08	0.52	0.13
2000	100	0.08	0.06	0.23	0.06	0.05	0.07	0.27	0.08	0.58	0.07	0.14	0.09	0.47	0.09	0.84	0.07	0.38	0.14
	500	0.06	0.06	0.21	0.06	0.01	0.07	0.21	0.08	0.57	0.07	0.03	0.09	0.44	0.09	0.80	0.08	0.15	0.16

Notes: This table reports results of the Monte Carlo exercise described in Section 3.3. All entries are expressed as a fraction of the true parameter value. This table reports results for the coefficient on the observed regressor. The IRT scores were obtained using the known values for the item response parameters rather than estimated values.