Econometrica Supplementary Material

SUPPLEMENT TO "NONPARAMETRIC INSTRUMENTAL VARIABLE ESTIMATION UNDER MONOTONICITY"

(*Econometrica*, Vol. 85, No. 4, July 2017, 1303–1320)

DENIS CHETVERIKOV
Department of Economics, University of California at Los Angeles

DANIEL WILHELM
Department of Economics, University College London

This supplement provides the proofs of the results from the main text.

APPENDIX A: PROOF OF THEOREM 1

FOR ANY $h \in L^1[0,1]$, let $\|h\|_1 := \int_0^1 |h(x)|\,dx$, $\|h\|_{1,t} := \int_{\delta_2}^{1-\delta_2} |h(x)|\,dx$ and define the operator norm by $\|T\|_2 := \sup_{h\in L^2[0,1]:\|h\|_2>0} \|Th\|_2/\|h\|_2$. Note that $\|T\|_2^2 \le \int_0^1\int_0^1 f_{X,W}^2(x, w)\,dx\,dw$, and so under Assumption 2, $\|T\|_2 \le \sqrt{C_T}$. Also, let $\mathcal{M}$ denote the set of all monotone functions in $L^2[0,1]$. To prove Theorem 1 from the main text, we first establish some auxiliary results.

LEMMA A.1—Lower Bound on T: *Let Assumptions* 1 *and* 2 *be satisfied. Then there exists a finite constant* $\bar{C}$ *such that*

$$\|h\|_{2,t} \le \bar{C}\|Th\|_2 \tag{19}$$

for any function $h \in \mathcal{M}$. *Here* $\bar{C}$ *depends only on the constants appearing in Assumptions* 1 *and* 2, *and on* x_1, x_2.

PROOF: We first show that for any $h \in \mathcal{M}$,

$$\|h\|_{2,t} \le C_1\|h\|_{1,t} \tag{20}$$

for $C_1 := (x_2 - x_1)^{1/2}/\min\{x_1 - \delta_2, 1 - \delta_2 - x_2\}$. Indeed, by monotonicity of h,

$$\begin{aligned}\|h\|_{2,t} &= \left(\int_{x_1}^{x_2} h(x)^2\,dx\right)^{1/2} \\ &\le \sqrt{x_2 - x_1}\max\{|h(x_1)|, |h(x_2)|\} \\ &\le \sqrt{x_2 - x_1}\frac{\int_{\delta_2}^{1-\delta_2} |h(x)|\,dx}{\min\{x_1 - \delta_2, 1 - \delta_2 - x_2\}},\end{aligned}$$

so that (20) follows. Therefore, for any increasing continuously differentiable $h \in \mathcal{M}$,

$$\|h\|_{2,t} \le C_1\|h\|_{1,t} \le C_1C_2\|Th\|_1 \le C_1C_2\|Th\|_2,$$

Denis Chetverikov: chetverikov@econ.ucla.edu
Daniel Wilhelm: d.wilhelm@ucl.ac.uk

 DOI: 10.3982/ECTA13639

where the first inequality follows from (20), the second from Lemma A.2 below (which is the main step in the proof of Theorem 1), and the third by Jensen's inequality. Hence, conclusion (19) of Lemma A.1 holds for increasing continuously differentiable $h \in \mathcal{M}$ with $\bar{C} := C_1 C_2$ and C_2 as defined in Lemma A.2.

Next, for any increasing function $h \in \mathcal{M}$, it follows from Lemma I.5 that one can find a sequence of increasing continuously differentiable functions $h_k \in \mathcal{M}$, $k \geq 1$, such that $\|h_k - h\|_2 \to 0$ as $k \to \infty$. Therefore, by the triangle inequality,

$$\begin{aligned}
\|h\|_{2,t} &\leq \|h_k\|_{2,t} + \|h_k - h\|_{2,t} \\
&\leq \bar{C}\|Th_k\|_2 + \|h_k - h\|_{2,t} \\
&\leq \bar{C}\|Th\|_2 + \bar{C}\big\|T(h_k - h)\big\|_2 + \|h_k - h\|_{2,t} \\
&\leq \bar{C}\|Th\|_2 + \bar{C}\|T\|_2\|h_k - h\|_2 + \|h_k - h\|_{2,t} \\
&\leq \bar{C}\|Th\|_2 + \big(\bar{C}\|T\|_2 + 1\big)\big\|(h_k - h)\big\|_2 \\
&\leq \bar{C}\|Th\|_2 + (\bar{C}\sqrt{C_T} + 1)\|h_k - h\|_2,
\end{aligned}$$

where the third line follows from the Cauchy–Schwarz inequality, the fourth from $\|h_k - h\|_{2,t} \leq \|h_k - h\|_2$, and the fifth from Assumption 2(i). Taking the limit as $k \to \infty$ of both the left-hand and the right-hand sides of this chain of inequalities yields conclusion (19) of Lemma A.1 for all increasing $h \in \mathcal{M}$.

Finally, since for any decreasing $h \in \mathcal{M}$, we have that $-h \in \mathcal{M}$ is increasing, $\|-h\|_{2,t} = \|h\|_{2,t}$, and $\|Th\|_2 = \|T(-h)\|_2$, conclusion (19) of Lemma A.1 also holds for all decreasing $h \in \mathcal{M}$, and thus for all $h \in \mathcal{M}$. This completes the proof of the lemma. *Q.E.D.*

LEMMA A.2: *Let Assumptions* 1 *and* 2 *hold*. *Then for any increasing continuously differentiable* $h \in L^1[0,1]$,

$$\|h\|_{1,t} = \int_{\delta_2}^{1-\delta_2} \big|h(x)\big|\,dx \leq C_2\|Th\|_1,$$

where $C_2 := ((c_W c_f/4)\min\{1 - w_2, w_1\}\min\{(C_F - 1)/2, 1\})^{-1}$.

PROOF: Take any increasing continuously differentiable function $h \in L^1[0,1]$ such that $\|h\|_{1,t} = 1$. Define $M(w) := \mathrm{E}[h(X)|W = w]$ for all $w \in [0,1]$ and note that

$$\begin{aligned}
\|Th\|_1 &= \int_0^1 \big|M(w)f_W(w)\big|\,dw \\
&\geq c_W \int_0^1 \big|M(w)\big|\,dw,
\end{aligned}$$

where the inequality follows from Assumption 2(iii). Therefore, the asserted claim follows if we can show that $\int_0^1 |M(w)|\,dw \geq (c_W C_2)^{-1}$.

To do so, first note that $M(w)$ is increasing. This is because, by integration by parts,

$$M(w) = \int_0^1 h(x)f_{X|W}(x|w)\,dx = h(1) - \int_0^1 Dh(x)F_{X|W}(x|w)\,dx,$$

so that condition (7) of Assumption 1 and $Dh(x) \geq 0$ for all x imply that the function $M(w)$ is increasing.

Next, consider the case in which $h(x) \geq 0$ for all $x \in [0, 1]$. Then $M(w) \geq 0$ for all $w \in [0, 1]$. Therefore,

$$\begin{aligned}
\int_0^1 \left|M(w)\right| dw &\geq \int_{w_2}^1 \left|M(w)\right| dw \\
&\geq (1 - w_2) M(w_2) \\
&= (1 - w_2) \int_0^1 h(x) f_{X|W}(x|w_2)\, dx \\
&\geq (1 - w_2) \int_{\delta_2}^{1-\delta_2} h(x) f_{X|W}(x|w_2)\, dx \\
&\geq (1 - w_2) c_f \int_{\delta_2}^{1-\delta_2} h(x)\, dx \\
&= (1 - w_2) c_f \|h\|_{1,t} \\
&= (1 - w_2) c_f \\
&\geq (c_W C_2)^{-1}
\end{aligned}$$

by Assumption 2(ii). Similarly,

$$\int_0^1 \left|M(w)\right| dw \geq w_1 c_f \geq (c_W C_2)^{-1}$$

when $h(x) \leq 0$ for all $x \in [0, 1]$. Therefore, it remains to consider the case in which there exists $x^* \in (0, 1)$ such that $h(x) \leq 0$ for $x \leq x^*$ and $h(x) \geq 0$ for $x > x^*$. Since $h(x)$ is continuous, $h(x^*) = 0$, and so integration by parts yields

$$\begin{aligned}
M(w) &= \int_0^{x^*} h(x) f_{X|W}(x|w)\, dx + \int_{x^*}^1 h(x) f_{X|W}(x|w)\, dx \\
&= -\int_0^{x^*} Dh(x) F_{X|W}(x|w)\, dx + \int_{x^*}^1 Dh(x)\big(1 - F_{X|W}(x|w)\big)\, dx.
\end{aligned} \tag{21}$$

For $k = 1, 2$, let $A_k := \int_{x^*}^1 Dh(x)(1 - F_{X|W}(x|w_k))\, dx$ and $B_k := \int_0^{x^*} Dh(x) F_{X|W}(x|w_k)\, dx$, so that

$$M(w_k) = A_k - B_k, \quad k = 1, 2.$$

Consider the following five cases separately, depending on where x^* lies relative to δ_2, δ_1, $1 - \delta_1$, and $1 - \delta_2$ (note that we have $0 \leq \delta_2 \leq \delta_1 < 1 - \delta_1 \leq 1 - \delta_2 \leq 1$).

Case I ($\delta_1 < x^* < 1 - \delta_1$): First, we have

$$\begin{aligned}
A_1 + B_2 &= \int_{x^*}^1 Dh(x)\big(1 - F_{X|W}(x|w_1)\big)\, dx + \int_0^{x^*} Dh(x) F_{X|W}(x|w_2)\, dx \\
&= \int_{x^*}^1 h(x) f_{X|W}(x|w_1)\, dx - \int_0^{x^*} h(x) f_{X|W}(x|w_2)\, dx
\end{aligned}$$

$$
\begin{aligned}
&\geq \int_{x^*}^{1-\delta_2} h(x) f_{X|W}(x|w_1)\,dx - \int_{\delta_2}^{x^*} h(x) f_{X|W}(x|w_2)\,dx \\
&\geq c_f \int_{x^*}^{1-\delta_2} h(x)\,dx + c_f \int_{\delta_2}^{x^*} \big|h(x)\big|\,dx \qquad (22)\\
&= c_f \int_{\delta_2}^{1-\delta_2} \big|h(x)\big|\,dx \\
&= c_f \|h\|_{1,t} \\
&= c_f,
\end{aligned}
$$

where the fourth line follows from Assumption 2(ii). Second, by (7) and (8) of Assumption 1,

$$
\begin{aligned}
M(w_1) &= \int_{x^*}^{1} Dh(x)\big(1 - F_{X|W}(x|w_1)\big)\,dx - \int_0^{x^*} Dh(x) F_{X|W}(x|w_1)\,dx \\
&\leq \int_{x^*}^{1} Dh(x)\big(1 - F_{X|W}(x|w_2)\big)\,dx - C_F \int_0^{x^*} Dh(x) F_{X|W}(x|w_2)\,dx \\
&= A_2 - C_F B_2,
\end{aligned}
$$

so that, together with $M(w_2) = A_2 - B_2$, we obtain

$$
M(w_2) - M(w_1) \geq (C_F - 1) B_2. \tag{23}
$$

Similarly, by (7) and (9) of Assumption 1,

$$
\begin{aligned}
M(w_2) &= \int_{x^*}^{1} Dh(x)\big(1 - F_{X|W}(x|w_2)\big)\,dx - \int_0^{x^*} Dh(x) F_{X|W}(x|w_2)\,dx \\
&\geq C_F \int_{x^*}^{1} Dh(x)\big(1 - F_{X|W}(x|w_1)\big)\,dx - \int_0^{x^*} Dh(x) F_{X|W}(x|w_1)\,dx \\
&= C_F A_1 - B_1,
\end{aligned}
$$

so that, together with $M(w_1) = A_1 - B_1$, we obtain

$$
M(w_2) - M(w_1) \geq (C_F - 1) A_1. \tag{24}
$$

In conclusion, equations (22), (23), and (24) yield

$$
M(w_2) - M(w_1) \geq (C_F - 1)(A_1 + B_2)/2 \geq (C_F - 1) c_f / 2. \tag{25}
$$

Consider the case $M(w_1) \geq 0$ and $M(w_2) \geq 0$. Then $M(w_2) \geq M(w_2) - M(w_1)$ and thus

$$
\begin{aligned}
\int_0^1 \big|M(w)\big|\,dw &\geq \int_{w_2}^1 \big|M(w)\big|\,dw \\
&\geq (1 - w_2) M(w_2) \\
&\geq (1 - w_2)(C_F - 1) c_f / 2 \\
&\geq (c_W C_2)^{-1}.
\end{aligned} \tag{26}
$$

Similarly,

$$\begin{aligned}\int_0^1 \left|M(w)\right| dw &\geq \int_0^{w_1} \left|M(w)\right| dw \\ &\geq w_1 \left|M(w_1)\right| \\ &\geq w_1 (C_F - 1) c_f / 2 \\ &\geq (c_W C_2)^{-1}\end{aligned} \tag{27}$$

when $M(w_1) \leq 0$ and $M(w_2) \leq 0$.

Finally, consider the case $M(w_1) \leq 0$ and $M(w_2) \geq 0$. If $M(w_2) \geq |M(w_1)|$, then $M(w_2) \geq (M(w_2) - M(w_1))/2$ and the same argument as in (26) shows that

$$\int_0^1 \left|M(w)\right| dw \geq (1 - w_2)(C_F - 1) c_f / 4 \geq (c_W C_2)^{-1}.$$

If $|M(w_1)| \geq M(w_2)$, then $|M(w_1)| \geq (M(w_2) - M(w_1))/2$ and we obtain

$$\int_0^1 \left|M(w)\right| dw \geq \int_0^{w_1} \left|M(w)\right| dw \geq w_1 (C_F - 1) c_f / 4 \geq (c_W C_2)^{-1}.$$

This completes the proof of Case I.

Case II ($1 - \delta_1 \leq x^* \leq 1 - \delta_2$): Note that since

$$\|h\|_{1,t} = \int_{\delta_2}^{x^*} \left|h(x)\right| dx + \int_{x^*}^{1-\delta_2} h(x)\, dx = 1,$$

it follows that either $\int_{\delta_2}^{x^*} |h(x)|\, dx \geq 1/2$ or $\int_{x^*}^{1-\delta_2} h(x)\, dx \geq 1/2$. We first consider the case $\int_{\delta_2}^{x^*} |h(x)|\, dx \geq 1/2$. Suppose that $M(w_1) \geq -c_f/4$. As in Case I, we have $M(w_2) \geq C_F A_1 - B_1$. Together with $M(w_1) = A_1 - B_1$, this inequality yields

$$\begin{aligned}M(w_2) - M(w_1) &= M(w_2) - C_F M(w_1) + C_F M(w_1) - M(w_1) \\ &\geq (C_F - 1) B_1 + (C_F - 1) M(w_1) \\ &= (C_F - 1)\left(\int_0^{x^*} Dh(x) F_{X|W}(x|w_1)\, dx + M(w_1)\right) \\ &= (C_F - 1)\left(\int_0^{x^*} \left|h(x)\right| f_{X|W}(x|w_1)\, dx + M(w_1)\right) \\ &\geq (C_F - 1)\left(\int_{\delta_2}^{x^*} \left|h(x)\right| f_{X|W}(x|w_1)\, dx - \frac{c_f}{4}\right) \\ &\geq (C_F - 1)\left(c_f \int_{\delta_2}^{x^*} \left|h(x)\right| dx - \frac{c_f}{4}\right) \\ &\geq \frac{(C_F - 1) c_f}{4}.\end{aligned}$$

We then proceed as in Case I using this inequality to replace (25) to show that $\int_0^1 |M(w)|\,dw \geq (c_W C_2)^{-1}$. On the other hand, when $M(w_1) < -c_f/4$ we bound $\int_0^1 |M(w)|\,dw$ as in (27), and the proof of Case II with $\int_{\delta_2}^{x^*} |h(x)|\,dx \geq 1/2$ is complete.

Next, we consider the case $\int_{x^*}^{1-\delta_2} h(x)\,dx \geq 1/2$. As above, we have $M(w_2) \geq C_F A_1 - B_1$ and $M(w_1) = A_1 - B_1$. Hence,

$$\begin{aligned}
M(w_2) - M(w_1) &\geq (C_F - 1)A_1 \\
&= (C_F - 1)\int_{x^*}^1 Dh(x)\big(1 - F_{X|W}(x|w_1)\big)\,dx \\
&= (C_F - 1)\int_{x^*}^1 h(x) f_{X|W}(x|w_1)\,dx \\
&\geq (C_F - 1)\int_{x^*}^{1-\delta_2} h(x) f_{X|W}(x|w_1)\,dx \\
&\geq (C_F - 1)c_f\int_{x^*}^{1-\delta_2} h(x)\,dx \\
&\geq \frac{(C_F - 1)c_f}{2}.
\end{aligned}$$

We then again proceed as in Case I to show that $\int_0^1 |M(w)|\,dw \geq (c_W C_2)^{-1}$. The proof of Case II with $\int_{x^*}^{1-\delta_2} h(x)\,dx \geq 1/2$ is complete.

Case III ($1 - \delta_2 < x^*$): Suppose $M(w_1) \geq -c_f/2$. As in Case I, we have $M(w_2) \geq C_F A_1 - B_1$. Together with $M(w_1) = A_1 - B_1$, this inequality yields

$$\begin{aligned}
M(w_2) - M(w_1) &= M(w_2) - C_F M(w_1) + C_F M(w_1) - M(w_1) \\
&\geq (C_F - 1)B_1 + (C_F - 1)M(w_1) \\
&= (C_F - 1)\bigg(\int_0^{x^*} Dh(x) F_{X|W}(x|w_1)\,dx + M(w_1)\bigg) \\
&= (C_F - 1)\bigg(\int_0^{x^*} \big|h(x)\big| f_{X|W}(x|w_1)\,dx + M(w_1)\bigg) \\
&\geq (C_F - 1)\bigg(\int_{\delta_2}^{1-\delta_2} \big|h(x)\big| f_{X|W}(x|w_1)\,dx - \frac{c_f}{2}\bigg) \\
&\geq (C_F - 1)\bigg(c_f\int_{\delta_2}^{1-\delta_2} \big|h(x)\big|\,dx - \frac{c_f}{2}\bigg) \\
&= \frac{(C_F - 1)c_f}{2}.
\end{aligned}$$

We then proceed as in Case I to show that $\int_0^1 |M(w)|\,dw \geq (c_W C_2)^{-1}$. On the other hand, when $M(w_1) < -c_f/2$ we bound $\int_0^1 |M(w)|\,dw$ as in (27), and the proof of Case III is complete.

Case IV ($\delta_2 \leq x^* \leq \delta_1$): Similarly to Case II, we first consider the case $\int_{x^*}^{1-\delta_2} h(x)\,dx \geq 1/2$. Suppose first that $M(w_2) \leq c_f/4$. As in Case I, we have $M(w_1) \leq A_2 - C_F B_2$, so that together with $M(w_2) = A_2 - B_2$,

$$\begin{aligned} M(w_2) - M(w_1) &= M(w_2) - C_F M(w_2) + C_F M(w_2) - M(w_1) \\ &\geq (1 - C_F) M(w_2) + (C_F - 1) A_2 \\ &= (C_F - 1)\left(\int_{x^*}^{1} Dh(x)\big(1 - F_{X|W}(x|w_2)\big)\,dx - M(w_2)\right) \\ &= (C_F - 1)\left(\int_{x^*}^{1} h(x) f_{X|W}(x|w_2)\,dx - M(w_2)\right) \\ &\geq (C_F - 1)\left(\int_{x^*}^{1-\delta_2} h(x) f_{X|W}(x|w_2)\,dx - M(w_2)\right) \\ &\geq (C_F - 1)\left(c_f \int_{x^*}^{1-\delta_2} h(x)\,dx - \frac{c_f}{4}\right) \\ &\geq \frac{(C_F - 1)c_f}{4}, \end{aligned}$$

and we proceed as in Case I to show that $\int_0^1 |M(w)|\,dw \geq (c_W C_2)^{-1}$. On the other hand, when $M(w_2) > c_f/4$, we bound $\int_0^1 |M(w)|\,dw$ as in (26), and the proof of Case IV with $\int_{x^*}^{1-\delta_2} h(x)\,dx \geq 1/2$ is complete.

Next, consider the case $\int_{\delta_2}^{x^*} |h(x)|\,dx \geq 1/2$. As above, we have $M(w_1) \leq A_2 - C_F B_2$ and $M(w_2) = A_2 - B_2$. Hence,

$$\begin{aligned} M(w_2) - M(w_1) &\geq (C_F - 1) B_2 \\ &= (C_F - 1) \int_0^{x^*} Dh(x) F_{X|W}(x|w_2)\,dx \\ &= (C_F - 1) \int_0^{x^*} \big|h(x)\big| f_{X|W}(x|w_2)\,dx \\ &\geq (C_F - 1) \int_{\delta_2}^{x^*} \big|h(x)\big| f_{X|W}(x|w_2)\,dx \\ &\geq (C_F - 1) c_f \int_{\delta_2}^{x^*} \big|h(x)\big|\,dx \geq \frac{(C_F - 1)c_f}{2}. \end{aligned}$$

We then again proceed as in Case I to show that $\int_0^1 |M(w)|\,dw \geq (c_W C_2)^{-1}$. The proof of Case IV with $\int_{\delta_2}^{x^*} |h(x)|\,dx \geq 1/2$ is complete.

Case V ($x^* < \delta_2$): Similarly to Case III, suppose first that $M(w_2) \leq c_f/2$. As in Case I, we have $M(w_1) \leq A_2 - C_F B_2$, so that together with $M(w_2) = A_2 - B_2$,

$$\begin{aligned} M(w_2) - M(w_1) &= M(w_2) - C_F M(w_2) + C_F M(w_2) - M(w_1) \\ &\geq (1 - C_F) M(w_2) + (C_F - 1) A_2 \end{aligned}$$

$$
\begin{aligned}
&= (C_F - 1)\left(\int_{x^*}^1 Dh(x)\big(1 - F_{X|W}(x|w_2)\big)\,dx - M(w_2)\right)\\
&= (C_F - 1)\left(\int_{x^*}^1 h(x) f_{X|W}(x|w_2)\,dx - M(w_2)\right)\\
&\geq (C_F - 1)\left(\int_{\delta_2}^{1-\delta_2} h(x) f_{X|W}(x|w_2)\,dx - M(w_2)\right)\\
&\geq (C_F - 1)\left(c_f\int_{\delta_2}^{1-\delta_2} h(x)\,dx - \frac{c_f}{2}\right) = \frac{(C_F-1)c_f}{2},
\end{aligned}
$$

and we proceed as in Case I to show that $\int_0^1 |M(w)|\,dw \geq (c_W C_2)^{-1}$. On the other hand, when $M(w_2) > c_f/2$, we bound $\int_0^1 |M(w)|\,dw$ as in (26), and the proof of Case V is complete. The lemma is proven. *Q.E.D.*

LEMMA A.3: *Let Assumptions* 1 *and* 2 *be satisfied. Consider any function* $h \in L^2[0,1]$. *If there exist* $h' \in L^2[0,1]$ *and* $\alpha \in (0,1)$ *such that* $h + h' \in \mathcal{M}$ *and* $\|h'\|_{2,t} + \bar{C}\|T\|_2\|h'\|_2 \leq \alpha\|h\|_{2,t}$, *then*

$$
\|h\|_{2,t} \leq \frac{\bar{C}}{1-\alpha}\|Th\|_2 \tag{28}
$$

for the constant $\bar{C}$ *defined in Lemma* A.1.

PROOF: Define

$$
\widetilde{h}(x) := \frac{h(x) + h'(x)}{\|h\|_{2,t} - \|h'\|_{2,t}}, \quad x \in [0,1].
$$

By assumption, $\|h'\|_{2,t} < \|h\|_{2,t}$, and so the triangle inequality yields

$$
\|\widetilde{h}\|_{2,t} \geq \frac{\|h\|_{2,t} - \|h'\|_{2,t}}{\|h\|_{2,t} - \|h'\|_{2,t}} = 1.
$$

Therefore, since $\widetilde{h} \in \mathcal{M}$, Lemma A.1 gives

$$
\|T\widetilde{h}\|_2 \geq \|\widetilde{h}\|_{2,t}/\bar{C} \geq 1/\bar{C}.
$$

Hence, applying the triangle inequality once again yields

$$
\begin{aligned}
\|Th\|_2 &\geq \big(\|h\|_{2,t} - \|h'\|_{2,t}\big)\|T\widetilde{h}\|_2 - \|Th'\|_2\\
&\geq \big(\|h\|_{2,t} - \|h'\|_{2,t}\big)\|T\widetilde{h}\|_2 - \|T\|_2\|h'\|_2\\
&\geq \frac{\|h\|_{2,t} - \|h'\|_{2,t}}{\bar{C}} - \|T\|_2\|h'\|_2\\
&= \frac{\|h\|_{2,t}}{\bar{C}}\left(1 - \frac{\|h'\|_{2,t} + \bar{C}\|T\|_2\|h'\|_2}{\|h\|_{2,t}}\right).
\end{aligned}
$$

Since the expression in the last parentheses is bounded from below by $1-\alpha$ by assumption, we obtain the inequality

$$\|Th\|_2 \ge \frac{1-\alpha}{\bar{C}}\|h\|_{2,t},$$

which is equivalent to (28). *Q.E.D.*

PROOF OF THEOREM 1: Note that since $\tau(a') \le \tau(a'')$ whenever $a' \le a''$ the claim for $a \le 0$, follows from $\tau(a) \le \tau(0) \le \bar{C}$, where the second inequality holds by Lemma A.1. Therefore, assume that $a > 0$. Fix any $\alpha \in (0,1)$. Take any function $h \in \mathcal{H}(a)$ such that $\|h\|_{2,t} = 1$. Set $h'(x) = ax$ for all $x \in [0,1]$. Note that the function $x \mapsto h(x) + ax$ is increasing and so belongs to the class $\mathcal{M}$. Also, $\|h'\|_{2,t} \le \|h'\|_2 \le a/\sqrt{3}$. Thus, the bound (28) in Lemma A.3 applies whenever $(1+\bar{C}\|T\|_2)a/\sqrt{3} \le \alpha$. Therefore, for all a satisfying the inequality

$$a \le \frac{\sqrt{3}\alpha}{1+\bar{C}\|T\|_2},$$

we have $\tau(a) \le \bar{C}/(1-\alpha)$. This completes the proof of the theorem. *Q.E.D.*

APPENDIX B: PROOF OF THEOREM 2

In this section, we use C to denote a strictly positive constant, whose value may change from place to place. Also, we use $\mathrm{E}_n[\cdot]$ to denote the average over index $i = 1, \dots, n$; for example, $\mathrm{E}_n[X_i] = n^{-1}\sum_{i=1}^n X_i$. Before proving Theorem 2, we prove the following two lemmas.

LEMMA B.1—Asymptotic Equivalence of Constrained and Unconstrained Estimators: *Let Assumptions* 2 *and* 4–8 *be satisfied*. *In addition*, *assume that g is continuously differentiable and $Dg(x) \ge c_g$ for all $x \in [0,1]$ and some constant $c_g > 0$*. *If we have $\tau_n^2\xi_n^2 \log n/n \to 0$, $\sup_{x\in[0,1]}\|Dp(x)\|(\tau_n(K/n)^{1/2} + K^{-s}) \to 0$, and $\sup_{x\in[0,1]}|Dg(x) - Dg_n(x)| \to 0$ as $n \to \infty$, then*

$$\mathrm{P}\big(\widehat{g}^c(x) = \widehat{g}^u(x) \quad \textit{for all } x \in [0,1]\big) \to 1 \textit{ as } n \to \infty. \tag{29}$$

PROOF: Observe that if $D\widehat{g}^u(x) \ge 0$ for all $x \in [0,1]$, then $\widehat{g}^c$ coincides with $\widehat{g}^u$, so that to prove (29), it suffices to show that

$$\mathrm{P}\big(D\widehat{g}^u(x) \ge 0 \quad \text{for all } x \in [0,1]\big) \to 1 \text{ as } n \to \infty. \tag{30}$$

In turn, (30) follows if

$$\sup_{x\in[0,1]}\big|D\widehat{g}^u(x) - Dg(x)\big| = o_p(1) \tag{31}$$

since $Dg(x) \ge c_g$ for all $x \in [0,1]$ and some $c_g > 0$.

To prove (31), define a function $\widehat{m} \in L^2[0,1]$ by

$$\widehat{m}(w) = q(w)'\mathrm{E}_n\big[q(W_i)Y_i\big], \quad w \in [0,1], \tag{32}$$

and an operator $\widehat{T}: L^2[0,1] \to L^2[0,1]$ by

$$(\widehat{T}h)(w) = q(w)'\mathrm{E}_n\big[q(W_i)p(X_i)'\big]\mathrm{E}\big[p(U)h(U)\big], \quad w \in [0,1], h \in L^2[0,1].$$

Throughout the proof, we assume that the events

$$\big\|\mathrm{E}_n\big[q(W_i)p(X_i)'\big] - \mathrm{E}\big[q(W)p(X)'\big]\big\| \le C\big(\xi_n^2 \log n/n\big)^{1/2}, \tag{33}$$

$$\big\|\mathrm{E}_n\big[q(W_i)q(W_i)'\big] - \mathrm{E}\big[q(W)q(W)'\big]\big\| \le C\big(\xi_n^2 \log n/n\big)^{1/2}, \tag{34}$$

$$\big\|\mathrm{E}_n\big[q(W_i)g_n(X_i)\big] - \mathrm{E}\big[q(W)g_n(X)\big]\big\| \le C\big(J/(\alpha n)\big)^{1/2}, \tag{35}$$

$$\|\widehat{m} - m\|_2 \le C\big(\big(J/(\alpha n)\big)^{1/2} + \tau_n^{-1}J^{-s}\big) \tag{36}$$

hold for some sufficiently large constant $0 < C < \infty$. It follows from Markov's inequality and Lemmas B.2 and I.6 that all four events hold jointly with probability at least $1 - \alpha - n^{-1}$ since the constant C is large enough.

Next, we derive a bound on $\|\widehat{g}^u - g_n\|_2$. By the definition of τ_n,

$$\begin{aligned}\big\|\widehat{g}^u - g_n\big\|_2 &\le \tau_n \big\|T\big(\widehat{g}^u - g_n\big)\big\|_2 \\ &\le \tau_n \big\|T\big(\widehat{g}^u - g\big)\big\|_2 + \tau_n \big\|T(g - g_n)\big\|_2 \\ &\le \tau_n \big\|T\big(\widehat{g}^u - g\big)\big\|_2 + C_g K^{-s},\end{aligned}$$

where the second inequality follows from the triangle inequality, and the third inequality from Assumption 6(iii). Next, since $m = Tg$,

$$\big\|T\big(\widehat{g}^u - g\big)\big\|_2 \le \big\|(T - T_n)\widehat{g}^u\big\|_2 + \big\|(T_n - \widehat{T})\widehat{g}^u\big\|_2 + \big\|\widehat{T}\widehat{g}^u - \widehat{m}\big\|_2 + \|\widehat{m} - m\|_2 \tag{37}$$

by the triangle inequality. The bound on $\|\widehat{m} - m\|_2$ is given in (36). Also, since $\|\widehat{g}^u\|_2 \le C_b$ by construction,

$$\big\|(T - T_n)\widehat{g}^u\big\|_2 \le C_b C_a \tau_n^{-1} K^{-s}$$

by Assumption 8(ii). In addition, by the triangle inequality,

$$\begin{aligned}\big\|(T_n - \widehat{T})\widehat{g}^u\big\|_2 &\le \big\|(T_n - \widehat{T})\big(\widehat{g}^u - g_n\big)\big\|_2 + \big\|(T_n - \widehat{T})g_n\big\|_2 \\ &\le \|T_n - \widehat{T}\|_2 \big\|\widehat{g}^u - g_n\big\|_2 + \big\|(T_n - \widehat{T})g_n\big\|_2.\end{aligned}$$

Moreover,

$$\|T_n - \widehat{T}\|_2 = \big\|\mathrm{E}_n\big[q(W_i)p(X_i)'\big] - \mathrm{E}\big[q(W)p(X)'\big]\big\| \le C\big(\xi_n^2 \log n/n\big)^{1/2}$$

by (33), and

$$\big\|(T_n - \widehat{T})g_n\big\|_2 = \big\|\mathrm{E}_n\big[q(W_i)g_n(X_i)\big] - \mathrm{E}\big[q(W)g_n(X)\big]\big\| \le C\big(J/(\alpha n)\big)^{1/2}$$

by (35).

Further, by Assumption 2(iii), all eigenvalues of $\mathrm{E}[q(W)q(W)']$ are bounded from below by c_w and from above by C_w, and so it follows from (34) that for large n, all eigenvalues of $Q_n := \mathrm{E}_n[q(W_i)q(W_i)']$ are bounded below from zero and from above. Therefore,

$$\begin{aligned}\|\widehat{T}\widehat{g}^u - \widehat{m}\|_2 &= \|\mathrm{E}_n[q(W_i)(p(X_i)'\widehat{\beta}^u - Y_i)]\| \\ &\le C\|\mathrm{E}_n[(Y_i - p(X_i)'\widehat{\beta}^u)q(W_i)']Q_n^{-1}\mathrm{E}_n[q(W_i)(Y_i - p(X_i)'\widehat{\beta}^u)]\|^{1/2} \\ &\le C\|\mathrm{E}_n[(Y_i - p(X_i)'\beta_n)q(W_i)']Q_n^{-1}\mathrm{E}_n[q(W_i)(Y_i - p(X_i)'\beta_n)]\|^{1/2} \\ &\le C\|\mathrm{E}_n[q(W_i)(p(X_i)'\beta_n - Y_i)]\|\end{aligned}$$

by optimality of $\widehat{\beta}^u$ (note that β_n is feasible in the optimization problem (13) for n large enough since $\|g\|_2 < C_b$ and $g_n(\cdot) = p(\cdot)'\beta_n$ satisfies $\|g - g_n\|_2 \to 0$ as $n \to \infty$). Moreover,

$$\begin{aligned}\|\mathrm{E}_n[q(W_i)(p(X_i)'\beta_n - Y_i)]\| \le \|(\widehat{T} - T_n)g_n\|_2 + \|(T_n - T)g_n\|_2 \\ + \|T(g_n - g)\|_2 + \|m - \widehat{m}\|_2\end{aligned}$$

by the triangle inequality. The terms $\|(\widehat{T} - T_n)g_n\|_2$ and $\|m - \widehat{m}\|_2$ have been bounded above. Also, by Assumptions 8(ii) and 6(iii),

$$\|(T_n - T)g_n\|_2 \le C\tau_n^{-1}K^{-s}, \qquad \|T(g - g_n)\|_2 \le C_g\tau_n^{-1}K^{-s}.$$

Combining the inequalities above shows that the inequality

$$\|\widehat{g}^u - g_n\|_2 \le C\big(\tau_n(J/(\alpha n))^{1/2} + K^{-s} + \tau_n(\xi_n^2 \log n/n)^{1/2}\|\widehat{g}^u - g_n\|_2\big) \tag{38}$$

holds with probability at least $1 - \alpha - n^{-c}$. Since $\tau_n^2\xi_n^2 \log n/n \to 0$, it follows that with the same probability,

$$\|\widehat{\beta}^u - \beta_n\| = \|\widehat{g}^u - g_n\|_2 \le C\big(\tau_n(J/(\alpha n))^{1/2} + K^{-s}\big),$$

and so by the triangle inequality,

$$\begin{aligned}|D\widehat{g}^u(x) - Dg(x)| &\le |D\widehat{g}^u(x) - Dg_n(x)| + |Dg_n(x) - Dg(x)| \\ &\le C \sup_{x\in[0,1]} \|Dp(x)\|\big(\tau_n(K/(\alpha n))^{1/2} + K^{-s}\big) + o(1)\end{aligned}$$

uniformly over $x \in [0, 1]$ since $J \le C_J K$ by Assumption 5. Since by the conditions of the lemma, $\sup_{x\in[0,1]} \|Dp(x)\|(\tau_n(K/n)^{1/2} + K^{-s}) \to 0$, (31) follows by taking $\alpha = \alpha_n \to 0$ slowly enough. This completes the proof of the lemma. *Q.E.D.*

LEMMA B.2: *Suppose that Assumptions* 2, 4, *and* 7 *hold*. *Then* $\|\widehat{m} - m\|_2 \le C((J/(\alpha n))^{1/2} + \tau_n^{-1}J^{-s})$ *with probability at least* $1 - \alpha$ *where* $\widehat{m}$ *is defined in* (32).

PROOF: Using the triangle inequality and an elementary inequality $(a+b)^2 \le 2a^2 + 2b^2$ for all $a, b \ge 0$,

$$\begin{aligned}&\|\mathrm{E}_n[q(W_i)Y_i] - E[q(W)g(X)]\|^2 \\ &\quad \le 2\|\mathrm{E}_n[q(W_i)\varepsilon_i]\|^2 + 2\|\mathrm{E}_n[q(W_i)g(X_i)] - \mathrm{E}[q(W)g(X)]\|^2.\end{aligned}$$

To bound the first term on the right-hand side of this inequality, we have

$$E\big[\big\|\mathrm{E}_n\big[q(W_i)\varepsilon_i\big]\big\|^2\big] = n^{-1}\mathrm{E}\big[\big\|q(W)\varepsilon\big\|^2\big] \le (C_B/n)\mathrm{E}\big[\big\|q(W)\big\|^2\big] \le CJ/n,$$

where the first and the second inequalities follow from Assumptions 4 and 2, respectively. Similarly,

$$\begin{aligned}\mathrm{E}\big[\big\|\mathrm{E}_n\big[q(W_i)g(X_i)\big] - \mathrm{E}\big[q(W)g(X)\big]\big\|^2\big] &\le n^{-1}\mathrm{E}\big[\big\|q(W)g(X)\big\|^2\big]\\ &\le (C_B/n)\mathrm{E}\big[\big\|q(W)\big\|^2\big]\\ &\le CJ/n\end{aligned}$$

by Assumption 4. Therefore, denoting $\bar{m}_n(w) := q(w)'\mathrm{E}[q(W)g(X)]$ for all $w \in [0, 1]$, we obtain

$$\mathrm{E}\big[\|\widehat{m} - \bar{m}_n\|_2^2\big] \le CJ/n,$$

and so by Markov's inequality, $\|\widehat{m} - \bar{m}_n\|_2 \le C(J/(\alpha n))^{1/2}$ with probability at least $1 - \alpha$. Further, using $\gamma_n \in \mathbb{R}^J$ from Assumption 7, so that $m_n(w) = q(w)'\gamma_n$ for all $w \in [0, 1]$, and denoting $r_n(w) := m(w) - m_n(w)$ for all $w \in [0, 1]$, we obtain

$$\begin{aligned}\bar{m}_n(w) &= q(w)'\int_0^1\int_0^1 q(t)g(x)f_{X,W}(x,t)\,dx\,dt\\ &= q(w)'\int_0^1 q(t)m(t)\,dt\\ &= q(w)'\int_0^1 q(t)\big(q(t)'\gamma_n + r_n(t)\big)\,dt\\ &= q(w)'\gamma_n + q(w)'\int_0^1 q(t)r_n(t)\,dt\\ &= m(w) - r_n(w) + q(w)'\int_0^1 q(t)r_n(t)\,dt.\end{aligned}$$

Hence, by the triangle inequality,

$$\|\bar{m}_n - m\|_2 \le \|r_n\|_2 + \left\|\int_0^1 q(t)r_n(t)\,dt\right\| \le 2\|r_n\|_2 \le 2C_m\tau_n^{-1}J^{-s}$$

by Bessel's inequality and Assumption 7. Applying the triangle inequality one more time, we obtain

$$\|\widehat{m} - m\|_2 \le \|\widehat{m} - \bar{m}_n\| + \|\bar{m}_n - m\|_2 \le C\big(\big(J/(\alpha n)\big)^{1/2} + \tau_n^{-1}J^{-s}\big)$$

with probability at least $1 - \alpha$. This completes the proof of the lemma. *Q.E.D.*

PROOF OF THEOREM 2: Consider the event that inequalities (33)–(36) hold for some sufficiently large constant C. As in the proof of Lemma B.1, this event occurs with probability at least $1 - \alpha - n^{-1}$. Also, applying the same arguments as those in the proof of

Lemma B.1, starting from (37), with $\widehat{g}^c$ replacing $\widehat{g}^u$ and using the bound

$$\left\|(T_n-\widehat{T})\widehat{g}^c\right\|_2 \le \|T_n-\widehat{T}\|_2\left\|\widehat{g}^c\right\|_2 \le C_b\|T_n-\widehat{T}\|_2$$

instead of the bound for $\|(T_n-\widehat{T})\widehat{g}^u\|_2$ used in the proof of Lemma B.1, it follows that on this event,

$$\left\|T\left(\widehat{g}^c-g\right)\right\|_2 \le C\left(\left(K/(\alpha n)\right)^{1/2}+\left(\xi_n^2\log n/n\right)^{1/2}+\tau_n^{-1}K^{-s}\right),$$

and so, by Assumption 6(iii),

$$\left\|T\left(\widehat{g}^c-g_n\right)\right\|_2 \le C\left(\left(K/(\alpha n)\right)^{1/2}+\left(\xi_n^2\log n/n\right)^{1/2}+\tau_n^{-1}K^{-s}\right). \tag{39}$$

Further,

$$\left\|\widehat{g}^c-g_n\right\|_{2,t} \le \delta+\tau_{n,t}\left(\frac{\|Dg_n\|_\infty}{\delta}\right)\left\|T\left(\widehat{g}^c-g_n\right)\right\|_2$$

since $\widehat{g}^c$ is increasing (indeed, if $\|\widehat{g}^c-g_n\|_{2,t}\le\delta$, the bound is trivial; otherwise, apply the definition of $\tau_{n,t}$ to the function $(\widehat{g}^c-g_n)/\|\widehat{g}^c-g_n\|_{2,t}$ and use the inequality $\tau_{n,t}(\|Dg_n\|_\infty/\|\widehat{g}^c-g_n\|_{2,t})\le\tau_{n,t}(\|Dg_n\|_\infty/\delta)$). Finally, by the triangle inequality,

$$\left\|\widehat{g}^c-g\right\|_{2,t} \le \left\|\widehat{g}^c-g_n\right\|_{2,t}+\|g_n-g\|_{2,t} \le \left\|\widehat{g}^c-g_n\right\|_{2,t}+C_gK^{-s}.$$

Combining these inequalities gives the asserted claim (15).

To prove (16), observe that combining (39) and Assumption 6(iii) and applying the triangle inequality shows that with probability at least $1-\alpha-n^{-1}$,

$$\left\|T\left(\widehat{g}^c-g\right)\right\|_2 \le C\left(\left(K/(\alpha n)\right)^{1/2}+\left(\xi_n^2\log n/n\right)^{1/2}+\tau_n^{-1}K^{-s}\right),$$

which, by the same argument as that used to prove (15), gives

$$\left\|\widehat{g}^c-g\right\|_{2,t} \le C\left\{\delta+\tau_{n,t}\left(\frac{\|Dg\|_\infty}{\delta}\right)\left(\frac{K}{\alpha n}+\frac{\xi_n^2\log n}{n}\right)^{1/2}+K^{-s}\right\}. \tag{40}$$

The asserted claim (16) now follows by applying (15) with $\delta=0$ and (40) with $\delta=\|Dg\|_\infty/c_\tau$ and using Theorem 1 to bound $\tau(c_\tau)$. This completes the proof of the theorem. *Q.E.D.*

Co-editor Elie Tamer handled this manuscript.

Manuscript received 14 July, 2015; final version accepted 11 April, 2017; available online 18 April, 2017.