Econometrica Supplementary Material

SUPPLEMENT TO "NONPARAMETRIC INSTRUMENTAL REGRESSION"

(*Econometrica*, Vol. 79, No. 5, September 2011, 1541–1565)

BY S. DAROLLES, Y. FAN, J. P. FLORENS, AND E. RENAULT

APPENDIX B: VERIFICATION OF ASSUMPTIONS A.3 AND A.4

The objective of this appendix is to give a set of *primitive* conditions that imply the main assumptions of the paper for the kernel estimator. For notational compactness, in this appendix, we suppress the subscripts in $\widehat{f}_W(w)$, $\widehat{f}_Z(z)$, and $\widehat{f}_{Z,W}(z,w)$, and in the corresponding probability density dunctions (p.d.f.s): they are distinguished by their arguments. We also suppress the subscript in h_N. We use C to denote a generic positive constant which may take different values in different places and we adopt the following assumptions.

ASSUMPTION B.1: (i) *The data* (y_n, z_n, w_n), $n = 1, \dots, N$, *define an independent and identically distributed sample of* (Y, Z, W). (ii) *The p.d.f.* $f(z, w)$ *is* d *times continuously differentiable in the interior of* $[0,1]^p \times [0,1]^q$.

ASSUMPTION B.2: *The p.d.f.* $f(z,w)$ *is bounded away from zero on the support* $[0,1]^p \times [0,1]^q$.

ASSUMPTION B.3: *Both multivariate kernels* $K_{Z,h}$ *and* $K_{W,h}$ *are product kernels generated from the univariate generalized kernel function* K_h *satisfying* (i) *the kernel function* $K_h(\cdot,\cdot)$ *is a generalized kernel function of order* l *and* (ii) *for each* $t \in [0,1]$, *the function* $K_h(h\cdot, t)$ *is supported on* $[(t-1)/h, t/h] \cap \mathcal{K}$, *where* $\mathcal{K}$ *is a compact interval not depending on* t *and*

$$\sup_{h>0, t\in[0,1], u\in\mathcal{K}} |K_h(hu, t)| < \infty.$$

ASSUMPTION B.4: *The smoothing parameter satisfies* $h \to 0$ *and* $(Nh^{p+q})^{-1} \times \log N \to 0$.

The independence assumption is a simplifying assumption and could be extended to weakly dependent (stationary mixing) observations. Assumption B.3 is the same as A.5 in Hall and Horowitz (2005). We first provide a result on the uniform convergence of $\widehat{f}(w)$, $\widehat{f}(z)$, and $\widehat{f}(z,w)$ with rates. For density functions with compact support, uniform convergence of kernel density estimators using ordinary kernel functions must be restricted to a proper subset of the compact support. Using generalized kernel functions, we show uniform convergence over the entire support. A similar result is provided in Proposition 2(ii) in Rothe (2010). However, the assumptions in Rothe (2010) differ from our assumptions and no proof is provided.

 DOI: 10.3982/ECTA6539

LEMMA B.1: *Suppose Assumptions* B.1–B.4 *hold*. *Let* $\rho = \min\{l, d\}$. *Then*

(i) $$\sup_{w\in[0,1]^q} |\widehat{f}(w) - f(w)| = O_P\big([(Nh^q)^{-1}\log N]^{1/2} + h^\rho\big) = o_P(1),$$

(ii) $$\sup_{z\in[0,1]^p, w\in[0,1]^q} |\widehat{f}(z,w) - f(z,w)| = O_P\big([(Nh^{p+q})^{-1}\log N]^{1/2} + h^\rho\big) = o_P(1),$$

(iii) $$\sup_{z\in[0,1]^p} |\widehat{f}(z) - f(z)| = O_P\big([(Nh^p)^{-1}\log N]^{1/2} + h^\rho\big) = o_P(1).$$

PROOF: We provide a proof of (i) only. First we evaluate the bias of $\widehat{f}(w)$. Let $w = (w_1, \dots, w_q)'$. Then

$$\begin{aligned}
E(\widehat{f}(w)) &= \frac{1}{h^q} E[K_{W,h}(w - w_n, w)] \\
&= \frac{1}{h^q}\int_{[0,1]^q} K_{W,h}(w - v, w) f(v)\, dv \\
&= \int_{\prod_{j=1}^q [\frac{w_j - 1}{h}, \frac{w_j}{h}]} K_{W,h}(hv, w) f(w - hv)\, dv \\
&= \int_{\prod_{j=1}^q [\frac{w_j - 1}{h}, \frac{w_j}{h}]} K_{W,h}(hv, w) \\
&\quad \times \Bigg[f(w) + (-h)\sum_{j=1}^q \frac{\partial f(w)}{\partial w_j} v_j + \cdots \\
&\quad + \frac{1}{\rho!}\sum_{j_1=1}^q \cdots \sum_{j_\rho=1}^q \frac{\partial^\rho f(w^*)}{\partial w_{j_1}\cdots \partial w_{j_\rho}} (-h)^\rho v_{j_1}\cdots v_{j_\rho}\Bigg]\, dv,
\end{aligned}$$

where w^* lies between w and $(w - hv)$. Now making use of Assumptions B.2–B.4, we get

$$\sup_{w\in[0,1]^q} \big|E(\widehat{f}(w)) - f(w)\big| \le Ch^\rho \Big[\sup_{h>0, t\in[0,1], u\in\mathcal{K}} |K_h(hu, t)|\Big]^q = O(h^\rho).$$

It remains to show $\sup_{w\in[0,1]^q} |\widehat{f}(w) - E[\widehat{f}(w)]| = O_P([(Nh^q)^{-1}\log N]^{1/2})$. This can be shown by the standard arguments in the proof of uniform consistency of kernel density estimators based on ordinary kernel functions, see, for example, Hansen (2008) and references therein. *Q.E.D.*

The next lemma shows that Assumption A.3 is satisfied under the previous conditions. Actually, this lemma proves a stronger result than is needed for Assumption A.3, because the convergence is proved in the Hilbert–Schmidt norm, which implies convergence for the supremum norm.

LEMMA B.2: *Suppose Assumptions* B.1–B.4 *hold*. *Then consider the following equalities*:

(i) $\|\widehat{T}-T\|_{\mathrm{HS}}^2 = O_P((Nh^{p+q})^{-1}+h^{2\rho})$,

(ii) $\|\widehat{T}^*-T^*\|_{\mathrm{HS}}^2 = O_P((Nh^{p+q})^{-1}+h^{2\rho})$.

Here $\|\cdot\|_{\mathrm{HS}}$ *denotes the Hilbert–Schmidt norm*, *that is*,

$$\begin{aligned}\|\widehat{T}-T\|_{\mathrm{HS}}^2 &= \int_{[0,1]^q}\int_{[0,1]^p}\frac{[\widehat{f}(z|w)-f(z|w)]^2}{f^2(z)}f(z)f(w)\,dz\,dw\\ &= \int_{[0,1]^q}\int_{[0,1]^p}\left[\frac{\widehat{f}(z,w)}{\widehat{f}(w)}-\frac{f(z,w)}{f(w)}\right]^2\frac{f(w)}{f(z)}\,dz\,dw.\end{aligned}$$

PROOF: (i) Let $\int\int\cdot\,dz\,dw = \int_{[0,1]^q}\int_{[0,1]^p}\cdot\,dz\,dw$. Note that

$$\begin{aligned}&\|\widehat{T}-T\|_{\mathrm{HS}}^2\\ &\quad= \int\int\left[\frac{\widehat{f}(z,w)}{\widehat{f}(w)}-\frac{f(z,w)}{f(w)}\right]^2\frac{f(w)}{f(z)}\,dz\,dw\\ &\quad= \int\int\left[\frac{\widehat{f}(z,w)f(w)-f(z,w)\widehat{f}(w)}{\widehat{f}(w)f(w)}\right]^2\frac{f(w)}{f(z)}\,dz\,dw\\ &\quad\le \frac{1}{\inf_{w\in[0,1]^q}[\widehat{f}(w)]^2}\int\int\left[\frac{\widehat{f}(z,w)f(w)-f(z,w)\widehat{f}(w)}{f(w)}\right]^2\\ &\qquad\times\frac{f(w)}{f(z)}\,dz\,dw\\ &\quad= O_P(1)\int\int\left[\widehat{f}(z,w)-f(z,w)-\frac{f(z,w)[\widehat{f}(w)-f(w)]}{f(w)}\right]^2\\ &\qquad\times\frac{f(w)}{f(z)}\,dz\,dw\\ &\quad= O_P(1)\int\int\frac{[\widehat{f}(z,w)-f(z,w)]^2}{f(z)}f(w)\,dz\,dw\\ &\qquad+O_P(1)\int\int\frac{f^2(z,w)[\widehat{f}(w)-f(w)]^2}{f(w)f(z)}\,dz\,dw\\ &\quad\equiv O_P(1)(A_1+A_2),\end{aligned}$$

where we have used the fact that $\frac{1}{\inf_{w\in[0,1]^q}[\widehat{f}(w)]^2}=O_P(1)$, which is implied by Assumptions B.2, B.4, and Lemma B.1. Now, we show

$$A_1 = O_P((Nh^{p+q})^{-1}+h^{2\rho})\quad\text{and}\quad A_2 = O_P((Nh^q)^{-1}+h^{2\rho}).$$

As a result, we obtain $\|\widehat{T}-T\|_{\mathrm{HS}}^2 = O_P((Nh^{p+q})^{-1}+h^{2\rho})$.

We prove the result for A_1. Note that

$$\begin{aligned}
E(|A_1|) &= \int\int \frac{E[\widehat{f}(z,w)-f(z,w)]^2}{f(z)} f(w)\,dz\,dw \\
&= \int\int \mathrm{Var}(\widehat{f}(z,w))\frac{f(w)}{f(z)}\,dz\,dw \\
&\quad + \int\int \big[E(\widehat{f}(z,w))-f(z,w)\big]^2 \frac{f(w)}{f(z)}\,dz\,dw \\
&= O((Nh^{p+q})^{-1}) + O(h^{2\rho}).
\end{aligned}$$

This follows from standard arguments for evaluating the first term and from the proof of Lemma B.1 for the second term. By Markov inequality, we obtain $A_1 = O_P((Nh^{p+q})^{-1}+h^{2\rho})$. *Q.E.D.*

The next lemma shows that Assumption A.4 is satisfied under the primitive conditions.

LEMMA B.3: *Suppose Assumptions* B.1–B.4 *hold*. *In addition*, *assume* $E(U^2|W=w)$ *is uniformly bounded in* $w\in[0,1]^q$. *Then* $\|\widehat{T}^*\widehat{r}-\widehat{T}^*\widehat{T}\varphi\|^2 = O_P(N^{-1}+h^{2\rho})$.

PROOF: By definition,

$$\begin{aligned}
&(\widehat{T}^*\widehat{r}-\widehat{T}^*\widehat{T}\varphi)(z) \\
&\quad = [\widehat{T}^*(\widehat{r}-\widehat{T}\varphi)](z) \\
&\quad = \int (\widehat{r}-\widehat{T}\varphi)(w)\frac{\widehat{f}(z,w)}{\widehat{f}(z)}\,dw \\
&\quad = \int \left(\widehat{r}(w) - \int \varphi(z')\frac{\widehat{f}(z',w)}{\widehat{f}(w)}\,dz'\right)\frac{\widehat{f}(z,w)}{\widehat{f}(z)}\,dw \\
&\quad = \int \left(\frac{1}{Nh^q}\sum_{n=1}^N y_n K_{W,h}(w-w_n,w) - \int \varphi(z')\widehat{f}(z',w)\,dz'\right) \\
&\qquad \times \frac{\widehat{f}(z,w)}{\widehat{f}(z)\widehat{f}(w)}\,dw \\
&\quad \equiv \int A_N(w)\frac{\widehat{f}(z,w)}{\widehat{f}(z)\widehat{f}(w)}\,dw.
\end{aligned}$$

Similar to the proof of Lemma B.2, we can show by using Lemma B.1 that, uniformly in $z \in [0,1]^p$, the equality

$$(\widehat{T}^*\widehat{r} - \widehat{T}^*\widehat{T}\varphi)(z) = \int A_N(w)\frac{f(z,w)}{f(z)f(w)}\,dw + o_P\biggl(\int A_N(w)\frac{f(z,w)}{f(z)f(w)}\,dw\biggr)$$

holds. Thus, it suffices to show that $\|\int A_N(w)\frac{f(z,w)}{f(z)f(w)}\,dw\|^2 = O_P(N^{-1} + h^{2\rho})$. Writing

$$\begin{aligned} A_N(w) &= \frac{1}{Nh^q}\sum_{n=1}^N U_n K_{W,h}(w - w_n, w) \\ &\quad + \frac{1}{Nh^q}\sum_{n=1}^N \biggl[\varphi(z_n) - \frac{1}{h^p}\int \varphi(z')K_{Z,h}(z' - z_n, z')\,dz'\biggr] \\ &\quad \times K_{W,h}(w - w_n, w) \\ &\equiv A_{N1}(w) + A_{N2}(w), \end{aligned}$$

we obtain

$$\begin{aligned} &E\biggl[\biggl\|\int A_N(w)\frac{f(z,w)}{f(z)f(w)}\,dw\biggr\|^2\biggr] \\ &\quad \le 2E\biggl[\biggl\|\int A_{N1}(w)\frac{f(z,w)}{f(z)f(w)}\,dw\biggr\|^2 + \biggl\|\int A_{N2}(w)\frac{f(z,w)}{f(z)f(w)}\,dw\biggr\|^2\biggr] \\ &\quad = 2\int\int\int E[A_{N1}(w)A_{N1}(w')]\frac{f(z,w)f(z,w')}{f(z)f(w)f(w')}\,dw\,dw'\,dz \\ &\qquad + 2\int\int\int E[A_{N2}(w)A_{N2}(w')]\frac{f(z,w)f(z,w')}{f(z)f(w)f(w')}\,dw\,dw'\,dz \\ &\quad = 2B_{N1} + 2B_{N2}. \end{aligned}$$

Below, we show that $B_{N1} = O(N^{-1})$ and $B_{N2} = O(N^{-1} + h^{2\rho})$. First consider the term

$$\begin{aligned} B_{N1} = \frac{1}{Nh^{2q}}&\int\int\int E[U_n^2 K_{W,h}(w - w_n, w)K_{W,h}(w' - w_n, w')] \\ &\times \frac{f(z,w)f(z,w')}{f(z)f(w)f(w')}\,dw\,dw'\,dz \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{N} \int E\bigg[\int_{-w_n/h}^{(1-w_n)/h} \int_{-w_n/h}^{(1-w_n)/h} U_n^2 K_{W,h}(hw, w_n + hw) \\
&\quad \times K_{W,h}(hw', w_n + hw') \\
&\quad \times \frac{f(z, w_n + hw) f(z, w_n + hw')}{f(z) f(w_n + hw) f(w_n + hw')} \, dw \, dw'\bigg] dz \\
&\le CN^{-1} \Big[\sup_{h>0, t\in[0,1], u\in\mathcal{K}} |K_h(hu, t)|\Big]^{2q} \\
&= O(N^{-1})
\end{aligned}
$$

under the conditions of Lemma B.3.

Now letting $B(z_n) = \varphi(z_n) - \frac{1}{h^p} \int \varphi(z') K_{Z,h}(z' - z_n, z') \, dz'$, we get

$$
\begin{aligned}
&B_{N2} \\
&= \frac{1}{(Nh_N^q)^2} \sum\sum_{n\neq n'} \int\int\int E[B(z_n) K_{W,h}(w - w_n, w)] \\
&\quad \times E[B(z_{n'}) K_{W,h}(w' - w_{n'}, w')] \frac{f(z, w) f(z, w')}{f(z) f(w) f(w')} \, dw \, dw' \, dz \\
&\quad + \frac{1}{Nh^{2q}} \int\int\int E\big[[B(z_n)]^2 K_{W,h}(w - w_n, w) K_{W,h}(w' - w_n, w')\big] \\
&\quad \times \frac{f(z, w) f(z, w')}{f(z) f(w) f(w')} \, dw \, dw' \, dz \\
&= \frac{1}{(Nh^q)^2} \sum\sum_{n\neq n'} \int \Big[\int E[B(z_n) K_{W,h}(w - w_n, w)] \frac{f(z, w)}{f(w)} \, dw\Big]^2 \\
&\quad \times \frac{1}{f(z)} \, dz \\
&\quad + \frac{1}{Nh^{2q}} \int\int\int E\Big[[B(z_n)]^2 K_{W,h}(w - w_n, w) \\
&\quad \times K_{W,h}(w' - w_{n'}, w')\Big] \\
&\quad \times \frac{f(z, w) f(z, w')}{f(z) f(w) f(w')} \, dw \, dw' \, dz \\
&= \frac{1}{(Nh^q)^2} \sum\sum_{n\neq n'} \int \Big[\int E[B(z_n) K_{W,h}(w - w_n, w)] \frac{f(z, w)}{f(w)} \, dw\Big]^2 \\
&\quad \times \frac{1}{f(z)} \, dz + O(N^{-1}),
\end{aligned}
$$

where the second term on the right hand side of the second to last equation can be shown to be $O(N^{-1})$ by change of variables and by using Assumptions B.1–B.4. Let B_{N21} denote the first term, that is, let

$$B_{N21} = \frac{1}{(Nh^q)^2} \sum\sum_{n \neq n'} \int \left[\int E[B(z_n) K_{W,h}(w - w_n, w)] \frac{f(z,w)}{f(w)} \, dw \right]^2 \times \frac{1}{f(z)} \, dz.$$

Then it suffices to show that $B_{N21} = O(h^{2\rho})$. Similar to the proof of Lemma B.1, we note that, uniformly in $z \in [0,1]^p$, we get

$$\begin{aligned} &\frac{1}{h^q} \int E[B(z_n) K_{W,h}(w - w_n, w)] \frac{f(z,w)}{f(w)} \, dw \\ &= \frac{1}{h^q} \int E[\varphi(z_n) K_{W,h}(w - w_n, w)] \frac{f(z,w)}{f(w)} \, dw \\ &\quad - \int \int \varphi(z') E\left[\frac{1}{h^{p+q}} K_{Z,h}(z' - z_n, z') K_{W,h}(w - w_n, w) \right] dz' \\ &\quad \times \frac{f(z,w)}{f(w)} \, dw \\ &= O(h^{2\rho}). \end{aligned}$$

As a result, we obtain $B_{N1} + B_{N2} = O(N^{-1} + h^{2\rho})$ or $E[\|\widehat{T}^* \widehat{r} - \widehat{T}^* \widehat{T} \varphi\|^2] = O(N^{-1} + h^{2\rho})$. By Markov inequality, we obtain the result in Lemma B.3.

Q.E.D.

REFERENCES

HANSEN, B. E. (2008): "Uniform Convergence Rates for Kernel Estimation With Dependent Data," *Econometric Theory*, 24, 726–748. [2]

ROTHE, C. (2010): "Nonparametric Estimation of Distributional Policy Effects," *Journal of Econometrics*, 155, 56–70. [1]

DRM—Université Paris Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France and Lyxor Asset Management; serge.darolles@dauphine.fr,

Dept. of Economics, Vanderbilt University, VU Station B 351819, 2301 Vanderbilt Place, Nashville, TN 37235-1819, U.S.A.; yanqin.fan@vanderbilt.edu,

Toulouse School of Economics, Université Toulouse 1 Capitole, Manufacture des Tabacs, 21 Allée de Brienne, 31000 Toulouse, France; florens@cict.fr,

and

Dept. of Economics, P.O. Box B, Brown University, Providence, RI 02912, U.S.A. and CENTER, Tilburg; Eric_Renault@brown.edu.

Manuscript received June, 2002; final revision received December, 2010.