Econometrica Supplementary Material

SUPPLEMENT TO "INFERENCE ON CAUSAL EFFECTS IN A GENERALIZED REGRESSION KINK DESIGN": APPENDICES

(*Econometrica*, Vol. 83, No. 6, November 2015, 2453–2483)

BY DAVID CARD, DAVID S. LEE, ZHUAN PEI, AND ANDREA WEBER

APPENDIX A: IDENTIFICATION

A.1. *Proofs of Propositions 1 and 2*

IN ORDER TO PROVE PROPOSITION 1, we first present and prove the following lemmas.

LEMMA 1: *Let $\varphi(x,t) : I \times [c,d] \to R$, where I is a compact subset of $\mathbb{R}^m$. Suppose $\varphi(x,t)$ and its partial derivative, $\varphi_2(x,t)$, are continuous and that φ is integrable with respect to the probability measure α for each t. Then $f(t) = \int \varphi(x,t)\,d\alpha(x)$ is continuously differentiable on $[c,d]$.*

PROOF: By Theorem 5 (p. 97) of Roussas (2004), $f'(t) = \int \varphi_2(x,t)\,d\alpha(x)$ for all $t \in [c,d]$. Let $s_1, s_2 \in [c,d]$:

$$\begin{aligned}\left|f'(s_1) - f'(s_2)\right| &= \left|\int \varphi_2(x,s_1)\,d\alpha(x) - \int \varphi_2(x,s_2)\,d\alpha(x)\right| \\ &\le \int \left|\varphi_2(x,s_1) - \varphi_2(x,s_2)\right|\,d\alpha(x).\end{aligned}$$

The continuity of $\varphi_2(x,t)$ on the compact set $I \times [c,d]$ implies uniform continuity, and therefore we can choose a δ such that $|s_1 - s_2| < \delta$ implies $|\varphi_2(x,s_1) - \varphi_2(x,s_2)| < \frac{\varepsilon}{\alpha(I)}$ for all $x \in I$, which in turn implies that $|f'(s_1) - f'(s_2)| < \varepsilon$. *Q.E.D.*

LEMMA 2: *Under Assumptions* 1(i), 3(ii), *and* 4, *$f(v)$ is continuously differentiable and strictly positive on I_V.*

PROOF: The continuous differentiability of $f(v)$ follows from Assumption 1(i), Assumption 4, and Lemma 1. $f(V) \ge \int_{A_U} f_{V|U=u}(v)\,dF_U(u) > 0$ follows directly from Assumption 3(ii). *Q.E.D.*

In order to prove Proposition 2, we present and prove the following lemmas. Let S be a sub-vector of the random vector $(U, \varepsilon, U_{B'}, U_{V'})$ that at least includes U and ε, and let S^* denote the vector of the random variables in $(U, \varepsilon, U_{B'}, U_{V'})$ but not in S. Let Q be a sub-vector of the random vector $(\varepsilon, U_{B'}, U_{V'})$ that at least includes ε, and let Q^* denote the vector of the random variables in $(U, \varepsilon, U_{B'}, U_{V'})$ but not in Q. Let I_{V^*}, I_S, I_{S^*}, I_Q,

 DOI: 10.3982/ECTA11224

and I_{Q^*} be the smallest closed rectangle that contains the support of V^*, S, S^*, Q, and Q^*, respectively. In the proofs, we only consider the case where $\pi^{ij}(V, U, \varepsilon, U_{V'}, U_{B'}) > 0$ for all $i, j = 0, 1$ since the proofs for the cases where some of the $\pi^{ij} = 0$ are similar in spirit and simpler.

LEMMA 3: (a) *$f_{S^*|S=s}(s^*)$ is continuous on $I_{S^*,S}$*; (b) *$f_{Q^*|Q=q}(q^*)$ is continuous on $I_{Q^*,Q}$.*

PROOF: We prove part (a), and the proof for part (b) is similar. There are three cases: (1) $S^* = U_{V'}$, (2) $S^* = U_{B'}$ and (3) $S^* = (U_{V'}, U_{B'})$. For case (1),

$$
\text{(A.1)} \qquad f_{U_{V'}|U_{B'}=u_{B'},U=u,\varepsilon=e}(u_{V'}) = \frac{f_{U_{V'},U_{B'}|U=u,\varepsilon=e}(u_{V'}, u_{B'})}{f_{U_{B'}|U=u,\varepsilon=e}(u_{B'})} = \frac{\int f_{V,U_{V'},U_{B'}|U=u,\varepsilon=e}(v, u_{V'}, u_{B'})\,dv}{\int\int f_{V,U_{V'},U_{B'}|U=u,\varepsilon=e}(v, u_{V'}, u_{B'})\,dv\,du_{V'}}.
$$

Note that the numerator of (A.1) is exactly $f_{S^*|S=s}(s^*)$ in case (3). Both the numerator and the denominator are continuous as guaranteed by Assumption 4a and Proposition 1 in Section 17.5 of Zorich (2004). Since the denominator is strictly positive, $f_{U_{V'}|U_{B'}=u_{B'},U=u,\varepsilon=e}(u_{V'})$ is continuous. The proof for case (2) is analogous with the roles of $U_{V'}$ and $U_{B'}$ exchanged. *Q.E.D.*

LEMMA 4: (a) *$\frac{\partial}{\partial v} f_{V|S=s}(v)$ is continuous on $I_{V,S}$*; (b) *$\frac{\partial}{\partial v} f_{V|Q=q}(v)$ is continuous on $I_{V,Q}$.*

PROOF: We only prove part (a), and the proof of part (b) is similar. Note that

$$
f_{V|S=s,S^*=s^*}(v) = \frac{f_{V,U_{V'},U_{B'}|U=u,\varepsilon=e}(v, u_{V'}, u_{B'})}{f_{U_{V'},U_{B'}|U=u,\varepsilon=e}(u_{V'}, u_{B'})},
$$

$$
f_{V|S=s}(v) = \int \frac{f_{V,U_{V'},U_{B'}|U=u,\varepsilon=e}(v, u_{V'}, u_{B'})}{f_{U_{V'},U_{B'}|U=u,\varepsilon=e}(u_{V'}, u_{B'})} f_{S^*|S=s}(s^*)\,ds^*,
$$

where the first line follows by Bayes' rule, and we integrate both sides over $f_{S^*|S=s}(s^*)$ to arrive at the second line. Taking derivatives with respect to v on both sides, interchanging differentiation and integration (permitted by Assumption 4a, Lemma 3, and Roussas (2004)), we obtain the result following Lemma 1. *Q.E.D.*

LEMMA 5: (a) *$\frac{\partial}{\partial v^*} f_{V^*|S=s}(v^*)$ is continuous on $I_{V^*,S}$*; (b) *$\frac{\partial}{\partial v^*} f_{V^*|Q=q}(v^*)$ is continuous on $I_{V^*,Q}$.*

PROOF: We only prove part (a), and the proof of part (b) is similar. Note that after applying Bayes' rule and rearranging, we obtain

$$\begin{aligned}
&f_{V^*|S=s,S^*=s^*}(v^*)\\
&\quad= \Pr[U_V=0|S=s,S^*=s^*]f_{V|S=s,S^*=s^*,U_V=0}(v^*)\\
&\qquad+\Pr[U_V\neq 0|S=s,S^*=s^*]f_{V^*|S=s,S^*=s^*,U_V\neq 0}(v^*)\\
&\quad= \Pr[U_V=0|S=s,S^*=s^*]\Pr[U_V=0|V=v^*,S=s,S^*=s^*]\\
&\qquad\times\frac{f_{V|S=s,S^*=s^*}(v^*)}{\Pr[U_V=0|S=s,S^*=s^*]}+\Pr[U_V\neq 0|S=s,S^*=s^*]\\
&\qquad\times\Pr[U_V\neq 0|V=v^*-u_{V'},S=s,S^*=s^*]\\
&\qquad\times\frac{f_{V|S=s,S^*=s^*}(v^*-u_{V'})}{\Pr[U_V\neq 0|S=s,S^*=s^*]}\\
&\quad= \Pr[U_V=0|V=v^*,S=s,S^*=s^*]f_{V|S=s,S^*=s^*}(v^*)\\
&\qquad+\Pr[U_V\neq 0|V=v^*-u_{V'},S=s,S^*=s^*]f_{V|S=s,S^*=s^*}(v^*-u_{V'}).
\end{aligned}$$

Multiplying both sides of the last line by $f_{S^*|S=s}(s^*)$ and integrating over s^*, taking the partial derivative with respect to v^*, and applying Assumptions 4a and 5 and Lemmas 3 and 4, we have the desired result. *Q.E.D.*

LEMMA 6: *(a)* $\frac{\partial}{\partial v^*}\Pr[G_V=i,G_B=j|V^*=v^*,S=s]$ *and* $\frac{\partial}{\partial v^*}\Pr[G_V=i|V^*=v^*,S=s]$ *are continuous on the set* $\{(v^*,s):f_{V^*|S=s}(v^*)>0\}$ *for* $i,j=0,1$; *(b)* $\frac{\partial}{\partial v^*}\Pr[G_V=i,G_B=j|V^*=v^*,Q=q]$ *and* $\frac{\partial}{\partial v^*}\Pr[G_V=i|V^*=v^*,Q=q]$ *are continuous on the set* $\{(v^*,q):f_{V^*|Q=q}(v^*)>0\}$ *for* $i,j=0,1$.

PROOF: Again we only prove part (a). First, note that the continuous differentiability of $\Pr[G_V=i,G_B=j|V^*=v^*,S=s]$ and $\Pr[G_V=i|V^*=v^*,S=s]$ is only needed on the set $\{(v^*,s):f_{V^*|S=s}(v^*)>0\}$ for the purpose of proving Proposition 2, because these quantities are always multiplied by $f_{V^*|S=s}(v^*)$ when they appear in subsequent proofs. We consider the two cases of $i=0,1$ separately. For case 1, where $i=0$,

$$\begin{aligned}
&\Pr[G_V=0,G_B=j|V^*=v^*,S=s]\\
&\quad= f_{V^*|S=s,G_V=0,G_B=j}(v^*)\frac{\Pr[G_V=0,G_B=j|S=s]}{f_{V^*|S=s}(v^*)}\\
&\quad= f_{V|S=s,G_V=0,G_B=j}(v^*)\frac{\Pr[G_V=0,G_B=j|S=s]}{f_{V^*|S=s}(v^*)}
\end{aligned}$$

$$= f_{V|S=s,G_V=0,G_B=j}(v^*) \frac{\Pr[G_V = 0, G_B = j|S = s]}{f_{V|S=s}(v^*)} \frac{f_{V|S=s}(v^*)}{f_{V^*|S=s}(v^*)}$$

$$= \Pr[G_V = 0, G_B = j|V = v^*, S = s] \frac{f_{V|S=s}(v^*)}{f_{V^*|S=s}(v^*)}$$

$$= \int \pi^{0j}(v^*, u, \varepsilon, u_{V'}, u_{B'}) f_{S^*|V=v^*,S=s}(s^*)\, ds^* \frac{f_{V|S=s}(v^*)}{f_{V^*|S=s}(v^*)}$$

$$= \int \pi^{0j}(v^*, u, \varepsilon, u_{V'}, u_{B'}) f_{V|S^*=s^*,S=s}(v^*) \frac{f_{S^*|S=s}(s^*)}{f_{V|S=s}(v^*)}\, ds^* \frac{f_{V|S=s}(v^*)}{f_{V^*|S=s}(v^*)}.$$

The partial derivative of the right-hand side w.r.t. v^* in the last line is continuous on $I_{V^*,S}$ by Assumption 5 and Lemmas 3, 4, and 5. For case 2 where $i = 1$,

$$\Pr[G_V = 1, G_B = j|V^* = v^*, S = s]$$

$$= \int \Pr[G_V = 1, G_B = j|V^* = v^*, S = s, S^* = s^*] f_{S^*|V^*=v^*,S=s}(s^*)\, ds^*$$

$$= \int \Pr[G_V = 1, G_B = j|V = v^* - u_{V'}, S = s, S^* = s^*]$$

$$\times \frac{f_{V^*|S=s,S^*=s^*}(v^*) f_{S^*|S=s}(s^*)}{f_{V^*|S=s}(v^*)}\, ds^*$$

$$= \int \pi^{1j}(v^* - u_{V'}, u, \varepsilon, u_{V'}, u_{B'}) \frac{f_{V^*|S=s,S^*=s^*}(v^*) f_{S^*|S=s}(s^*)}{f_{V^*|S=s}(v^*)}\, ds^*.$$

Its partial derivative w.r.t. v^* is continuous on $I_{V^*,S}$ for the same reason as in case 1.

Since $\Pr[G_V = i|V^* = v^*, S = s] = \sum_j \Pr[G_V = i, G_B = j|V^* = v^*, S = s]$, the continuous differentiability with respect to v^* of $\Pr[G_V = i, G_B = j|V^* = v^*, S = s]$ implies that of $\Pr[G_V = i|V^* = v^*, S = s]$. *Q.E.D.*

PROOF OF PROPOSITION 2: For part (a), the proof is the same as for part (a) in Proposition 1, replacing V with V^*, letting the pair (U, ε) serve the role of U and using Lemma 5.

For part (b), we can write

$$E[Y|V^* = v^*] \tag{A.2}$$

$$= \int E[Y|V^* = v^*, U = u, \varepsilon = e]\, dF_{U,\varepsilon|V^*=v^*}(u, e)$$

$$
\begin{aligned}
&= \int \big(E\big[Y|U_V = 0, V^* = v^*, U = u, \varepsilon = e\big] \\
&\quad \times \Pr\big[U_V = 0|V^* = v^*, U = u, \varepsilon = e\big] \\
&\quad + E\big[Y|U_V \neq 0, V^* = v^*, U = u, \varepsilon = e\big] \\
&\quad \times \Pr\big[U_V \neq 0|V^* = v^*, U = u, \varepsilon = e\big]\big)\, dF_{U,\varepsilon|V^*=v^*}(u, e) \\
&= \int \left(z_1 z_2 + \left[\int z_3 z_4\, du_{V'}\right] \cdot [1 - z_2]\right) z_5\, dF_{U,\varepsilon}(u, e),
\end{aligned}
$$

where the second line follows from the law of iterated expectations, and to ease exposition below, we use the notation:

$$
\begin{aligned}
z_1 &\equiv y\big(b(v^*, e), v^*, u\big), \\
z_2 &\equiv \Pr\big[V = V^*|V^* = v^*, U = u, \varepsilon = e\big], \\
z_3 &\equiv y\big(b(v^* - u_{V'}, e), v^* - u_{V'}, u\big), \\
z_4 &\equiv f_{U_{V'}|U_V \neq 0, V^*=v^*, U=u, \varepsilon=e}(u_{V'}), \\
z_5 &\equiv \frac{f_{V^*|U=u,\varepsilon=e}(v^*)}{f_{V^*}(v^*)}.
\end{aligned}
$$

The derivative of $E[Y|V^* = v^*]$ in equation (A.2) with respect to v^* is

$$
\begin{aligned}
\text{(A.3)} \qquad \frac{dE\big[Y|V^* = v^*\big]}{dv^*} &= \int z_1' z_2 z_5\, dF_{U,\varepsilon}(u, e) + \int z_1 \frac{\partial(z_2 z_5)}{\partial v^*}\, dF_{U,\varepsilon}(u, e) \\
&\quad + \int \frac{\partial\left[\left(\int z_3 z_4\, du_{V'}\right)[1 - z_2] z_5\right]}{\partial v^*}\, dF_{U,\varepsilon}(u, e),
\end{aligned}
$$

where z_j' denotes the partial derivative of z_j with respect to v^*, provided that the integrands are continuous.

In a parallel fashion, we can write

$$
\begin{aligned}
&E\big[B^*|V^* = v^*\big] \\
&\quad = \int \Big\{\big[z_6 + z_8(1 - z_7)\big] z_{13} \\
&\qquad + \left[\left(\int z_9 z_{10}\, du_{V'}\right) z_{11} + \left(\int (z_9 + u_{B'}) z_{12}\, du_{V'}\, du_{B'}\right)(1 - z_{11})\right] \\
&\qquad \times (1 - z_{13})\Big\} z_{14}\, dF_{\varepsilon}(e),
\end{aligned}
$$

with

$$
\begin{aligned}
z_6 &\equiv b(v^*, e), \\
z_7 &\equiv \Pr[U_B = 0 | U_V = 0, V^* = v^*, \varepsilon = e], \\
z_8 &\equiv \int u_{B'} f_{U_{B'}|U_V=0, U_B\neq 0, V^*=v^*, \varepsilon=e}(u_{B'})\, du_{B'}, \\
z_9 &\equiv b(v^* - u_{V'}, e), \\
z_{10} &\equiv f_{U_{V'}|U_B=0, U_V\neq 0, V^*=v^*, \varepsilon=e}(u_{V'}), \\
z_{11} &\equiv \Pr[U_B = 0 | U_V \neq 0, V^* = v^*, \varepsilon = e], \\
z_{12} &\equiv f_{U_{V'}, U_{B'}|U_V\neq 0, U_B\neq 0, V^*=v^*, \varepsilon=e}(u_{V'}, u_{B'}), \\
z_{13} &\equiv \Pr[V = V^* | V^* = v^*, \varepsilon = e], \\
z_{14} &\equiv \frac{f_{V^*|\varepsilon=e}(v^*)}{f_{V^*}(v^*)}.
\end{aligned}
$$

And the analogous derivative with respect to v^* is

$$
\begin{aligned}
&\frac{dE[B^* | V^* = v^*]}{dv^*} \\
&\quad = \int z_6' z_{13} z_{14}\, dF_\varepsilon(e) + \int z_6 \frac{\partial}{\partial v^*}(z_{13} z_{14})\, dF_\varepsilon(e) \\
&\qquad + \frac{\partial}{\partial v^*} \int \bigl[z_8(1 - z_7) z_{13} z_{14}\bigr]\, dF_\varepsilon(e) \\
&\qquad + \frac{\partial}{\partial v^*} \int \biggl\{ \int z_9 z_{10}\, du_{V'} \cdot z_{11} \\
&\qquad + \int \int (z_9 + u_{B'}) z_{12}\, du_{V'}\, du_{B'} \cdot (1 - z_{11}) \biggr\} (1 - z_{13} z_{14}) \}\, dF_\varepsilon(e),
\end{aligned}
\tag{A.4}
$$

provided that the integrands are continuous.

The proof of part (b) follows from showing that the partial derivatives of z_2, $\int z_3 z_4\, du_{V'}$, z_5, z_7, z_8, $\int z_9 z_{10}\, du_{V'}$, z_{11}, $\int \int (z_9 + u_{B'}) z_{12}\, du_{V'}\, du_{B'}$, and $z_{13} z_{14}$ with respect to v^* are continuous, and noting that z_1 and z_6 are continuous by Assumptions 1a, 2, and 3a. From this, it follows that there is no discontinuity in all but the first term on the right-hand side of (A.3) and (A.4) at $v^* = 0$ and that the RKD estimand is the ratio of the discontinuities in the first terms of those two equations.

As shown by Lemma 6, z_2 is continuously differentiable in v^*.

z_4 is continuously differentiable in v^* because

$$
\begin{aligned}
& f_{U_{V'}|U_V\neq 0, V^*=v^*, U=u, \varepsilon=e}(u_{V'}) \\
&\quad = \big(\Pr\big[U_V \neq 0 | U_{V'} = u_{V'}, V^* = v^*, U = u, \varepsilon = e\big] \\
&\qquad \times f_{U_{V'}|V^*=v^*, U=u, \varepsilon=e}(u_{V'})\big) \\
&\qquad / \big(\Pr\big[U_V \neq 0 | V^* = v^*, U = u, \varepsilon = e\big]\big) \\
&\quad = \Big(\big(1 - \Pr\big[U_V = 0 | U_{V'} = u_{V'}, V^* = v^*, U = u, \varepsilon = e\big]\big) \\
&\qquad \times \frac{f_{V^*|U_{V'}=u_{V'}, U=u, \varepsilon=e}(v^*) f_{U_{V'}|U=u, \varepsilon=e}(u_{V'})}{f_{V^*|U=u, \varepsilon=e}(v^*)}\Big) \Big/ (1 - z_2)
\end{aligned}
$$

and the derivative of the last line is continuous by Lemmas 3, 5, and 6.

We break up the integral $\int z_3 z_4 \, du_{V'}$ into two pieces

$$
\begin{aligned}
\int z_3 z_4 \, du_{V'} &= \int_{c_{U_{V'}}}^{v^*} z_3 z_4 \, du_{V'} + \int_{v^*}^{d_{U_{V'}}} z_3 z_4 \, du_{V'} \\
&= \int_{c_{U_{V'}}}^{v^*} y\big(b^+(v^* - u_{V'}, e), v^* - u_{V'}, u\big) z_4 \, du_{V'} \\
&\quad + \int_{v^*}^{d_{U_{V'}}} y\big(b^-(v^* - u_{V'}, e), v^* - u_{V'}, u\big) z_4 \, du_{V'},
\end{aligned}
$$

where $c_{U_{V'}}$ and $d_{U_{V'}}$ are the lower and upper end point of the support of $U_{V'}$, $b^+(v, e) = b(v, e)$ for $v \geq 0$ and all e, and $b^-(v, e) = b(v, e)$ for $v \leq 0$ and all e. Denote $y(b^\pm(v^* - u_{V'}, e), v^* - u_{V'}, u)$ by $z_3^\pm$. Note that z_3^+ and z_3^- are continuously differentiable in v^* on $[c_{U_{V'}}, v^*]$ and $[v^*, d_{U_{V'}}]$ respectively by Assumptions 1a, 2, and 3a, where $c_{U_{V'}}$ and $d_{U_{V'}}$ are the lower and upper endpoints of the support $I_{U_V'}$. Since z_4 is also continuously differentiable as shown above, we can apply the Newton–Leibniz formula, which yields

$$
\begin{aligned}
& \frac{\partial}{\partial v^*}\left(\int_{c_{U_{V'}}}^{v^*} z_3^+ z_4 \, du_{V'} + \int_{v^*}^{d_{U_{V'}}} z_3^- z_4 \, du_{V'}\right) \\
&\quad = \int_{c_{U_{V'}}}^{v^*} \frac{\partial}{\partial v^*}\big(z_3^+ z_4\big) \, du_{V'} + \int_{v^*}^{d_{U_{V'}}} \frac{\partial}{\partial v^*}\big(z_3^- z_4\big) \, du_{V'} \\
&\qquad + z_3^+ z_4|_{u_{V'}=v^*} - z_3^- z_4|_{u_{V'}=v^*}.
\end{aligned}
$$

By Assumptions 1a and 3a, z_3 is continuous, and it follows that $z_3^+ z_4|_{u_{V'}=v^*} - z_3^- z_4|_{u_{V'}=v^*} = 0$. Since $\int_{c_{U_{V'}}}^{v^*} \frac{\partial}{\partial v^*}(z_3^+ z_4)\, du_{V'}$ and $\int_{v^*}^{d_{U_{V'}}} \frac{\partial}{\partial v^*}(z_3^- z_4)\, du_{V'}$ are continuous, $\int z_3 z_4\, du_{V'}$ is continuously differentiable in v^*.

z_5 is continuously differentiable in v^* by Lemma 5 and Assumption 3a—note that the continuous differentiability of $f_{V^*}(v^*) = \int f_{V^*|U=u,\varepsilon=e}\, dF_{U,\varepsilon}(u, e)$ is a part of Corollary 1, and it follows directly from Lemma 5; z_6 is continuously differentiable by Assumption 3a, and z_7 is continuously differentiable in v^* because

$$\begin{aligned}
&\Pr\left[U_B = 0 | U_V = 0, V^* = v^*, \varepsilon = e\right] \\
&\quad = \frac{\Pr\left[U_V = 0, U_B = 0 | V^* = v^*, \varepsilon = e\right]}{\Pr\left[U_V = 0 | V = v^*, \varepsilon = e\right]},
\end{aligned}$$

where the derivative of the right-hand side is continuous in v^* by Lemma 6.

z_8 is continuously differentiable in v^* because

$$\begin{aligned}
&\int u_{B'} f_{U_{B'}|U_V=0,U_B\neq 0,V^*=v^*,\varepsilon=e}(u_{B'})\, du_{B'} \\
&\quad = \int u_{B'} \Pr\left[U_V = 0 | U_{B'} = u_{B'}, V^* = v^*, \varepsilon = e\right] \\
&\qquad \times \frac{f_{U_{B'}|V^*=v^*,\varepsilon=e}(u_{B'})}{\Pr\left[U_V = 0 | V^* = v^*, \varepsilon = e\right]}\, du_{B'} \\
&\quad = \int u_{B'} \Pr\left[U_V = 0 | U_{B'} = u_{B'}, V^* = v^*, \varepsilon = e\right] \\
&\qquad \times \frac{f_{V^*|U_{B'}=u_{B'},\varepsilon=e}(v^*) f_{U_{B'}|\varepsilon=e}(u_{B'})}{f_{V^*|\varepsilon=e}(v^*) \Pr\left[U_V = 0 | V = v^*, \varepsilon = e\right]}\, du_{B'},
\end{aligned}$$

where the continuous differentiability of the last line in v^* is implied by Lemmas 3, 5, and 6. By a similar application of Bayes' rule, we can show that z_{10} is continuously differentiable in v^*. Consequently, $\int z_9 z_{10}\, du_{V'}$ is continuously differentiable in v^* by applying the same argument used for $\int z_3 z_4\, du_{V'}$.

The quantity z_{11} is continuously differentiable in v^* because of Lemma 6 and

$$z_{11} = \frac{\Pr\left[U_B = 0, U_V \neq 0 | V^* = v^*, \varepsilon = e\right]}{\Pr\left[U_V \neq 0 | V^* = v^*, \varepsilon = e\right]}.$$

z_{12} can be expressed as

$$\begin{aligned}
&f_{U_{V'},U_{B'}|U_V\neq 0,U_B\neq 0,V^*=v^*,\varepsilon=e}(u_{V'},u_{B'})\\
&\quad=\frac{\Pr\big[U_V\neq 0, U_B\neq 0|U_{V'}=u_{V'},U_{B'}=u_{B'},V^*=v^*,\varepsilon=e\big]}{\Pr\big[U_V\neq 0,U_B\neq 0|V^*=v^*,\varepsilon=e\big]}\\
&\qquad\times f_{U_{V'},U_{B'}|V^*=v^*,\varepsilon=e}(u_{V'},u_{B'})\\
&\quad=\frac{\Pr\big[U_V\neq 0, U_B\neq 0|U_{V'}=u_{V'},U_{B'}=u_{B'},V^*=v^*,\varepsilon=e\big]}{\Pr\big[U_V\neq 0,U_B\neq 0|V^*=v^*,\varepsilon=e\big]}\\
&\qquad\times\frac{f_{V^*|U_{V'}=u_{V'},U_{B'}=u_{B'},\varepsilon=e}(v^*)}{f_{V^*|\varepsilon=e}(v^*)}f_{U_{V'},U_{B'}|\varepsilon=e}(u_{V'},u_{B'}),
\end{aligned}$$

and z_{12} is continuously differentiable by Lemmas 3, 5, and 6. It follows that $\int\int(z_9+u_{B'})z_{12}\,du_{V'}\,du_{B'}$ is continuously differentiable by the same argument as that for $\int z_3z_4\,du_{V'}$. Finally, z_{13} is continuously differentiable in v^* by Lemma 6 and z_{14} by Lemma 5 and Assumption 3a.

As a result of the smoothness of the above terms along with Theorem 5 on p. 97 of Roussas (2004), we can write

$$\begin{aligned}
\text{(A.5)}\qquad &\lim_{v_0\to 0^+}\frac{dE\big[Y|V^*=v^*\big]}{dv^*}\bigg|_{v^*=v_0}-\lim_{v_0\to 0^-}\frac{dE\big[Y|V^*=v^*\big]}{dv^*}\bigg|_{v^*=v_0}\\
&\quad=\lim_{v_0\to 0^+}\int z_1'z_2z_5\,dF_{U,\varepsilon}(u,e)-\lim_{v_0\to 0^-}\int z_1'z_2z_5\,dF_{U,\varepsilon}(u,e)\\
&\quad=\int\Big(\lim_{v_0\to 0^+}z_1'-\lim_{v_0\to 0^-}z_1'\Big)z_2z_5|_{v^*=v_0}\,dF_{U,\varepsilon}(u,e)\\
&\quad=\int y_1\big(b(0,e),0,u\big)\big(b_1^+(e)-b_1^-(e)\big)z_2z_5|_{v^*=v_0}\,dF_{U,\varepsilon}(u,e).
\end{aligned}$$

The interchange of limit and integration is allowed by the dominated convergence theorem since $z_1'z_2z_5$ is continuous over a compact rectangle. The last line follows from Assumptions 1a and 3a.

Similarly, we can write

$$\begin{aligned}
\text{(A.6)}\qquad &\lim_{v_0\to 0^+}\frac{dE\big[B^*|V^*=v^*\big]}{dv^*}\bigg|_{v^*=v_0}-\lim_{v_0\to 0^-}\frac{dE\big[B^*|V^*=v^*\big]}{dv^*}\bigg|_{v^*=v_0}\\
&\quad=\lim_{v_0\to 0^+}\int z_6'z_{13}z_{14}\,dF_\varepsilon(e)-\lim_{v_0\to 0^-}\int z_6'z_{13}z_{14}\,dF_\varepsilon(e)
\end{aligned}$$

$$= \int \Big(\lim_{v_0 \to 0^+} z_6' - \lim_{v_0 \to 0^-} z_6' \Big) z_{13} z_{14}|_{v^*=v_0} \, dF_\varepsilon(e)$$

$$= \int \big(b_1^+(e) - b_1^-(e) \big) z_{13} z_{14}|_{v^*=v_0} \, dF_\varepsilon(e).$$

Finally, consider the term $z_2 z_5|_{v^*=v_0}$. First, a similar argument as in (6) leads to

$$z_2 = \Pr\big[V = V^* | V = v^*, U = u, \varepsilon = e\big] \frac{f_{V|U=u,\varepsilon=e}\big(v^*\big)}{f_{V^*|U=u,\varepsilon=e}\big(v^*\big)}.$$

After applying Bayes' rule and rearranging, we have

$$\begin{aligned} z_2 z_5|_{v^*=v_0} &= \Pr\big[V = V^* | V = 0, U = u, \varepsilon = e\big] \\ &\quad \times \frac{f_{V|U=u,\varepsilon=e}(0)}{f_{V^*|U=u,\varepsilon=e}(0)} \frac{f_{V^*|U=u,\varepsilon=e}(0)}{f_{V^*}(0)} \\ &= \Pr\big[V = V^* | V = 0, U = u, \varepsilon = e\big] \frac{f_{V|U=u,\varepsilon=e}(0)}{f_V(0)} \frac{f_V(0)}{f_{V^*}(0)}. \end{aligned}$$

Similarly, we can derive

$$z_{13} z_{14}|_{v^*=v_0} = \Pr\big[V = V^* | V = 0, \varepsilon = e\big] \frac{f_{V|\varepsilon=e}(0)}{f_V(0)} \frac{f_V(0)}{f_{V^*}(0)}.$$

Because $\frac{f_V(0)}{f_{V^*}(0)}$ can be pulled out of the integral in both (A.5) and (A.6), we have the result

$$\frac{\lim_{v_0 \to 0^+} \frac{dE[Y|V^*=v^*]}{dv^*}|_{v^*=v_0} - \lim_{v_0 \to 0^-} \frac{dE[Y|V^*=v^*]}{dv^*}|_{v^*=v_0}}{\lim_{v_0 \to 0^+} \frac{dE[B^*|V^*=v^*]}{dv^*}|_{v^*=v_0} - \lim_{v_0 \to 0^-} \frac{dE[B^*|V^*=v^*]}{dv^*}|_{v^*=v_0}}$$

$$= \int y_1\big(b(0,e), 0, u\big) \varphi(u,e) \, dF_{U,\varepsilon}(u,e),$$

where

$$\varphi(u,e) = \frac{\Pr[U_V = 0 | V = 0, U = u, \varepsilon = e] \big(b_1^+(e) - b_1^-(e)\big) \frac{f_{V|U=u,\varepsilon=e}(0)}{f_V(0)}}{\int \Pr[U_V = 0 | V = 0, \varepsilon = \omega] \big(b_1^+(\omega) - b_1^-(\omega)\big) \frac{f_{V|\varepsilon=\omega}(0)}{f_V(0)} \, dF_\varepsilon(\omega)}.$$

Note that Assumptions 3a and 6 guarantee nonnegative, finite weights and that $\int \varphi(u, e)\, dF_{U,\varepsilon}(u, e) = 1$. *Q.E.D.*

A.2. *Identification in the Presence of Both Slope and Level Changes—Remark 3*

In Remark 3, we consider the identification of the treatment effect when there is both a level change and a slope change at the threshold $V = 0$. To ease exposition, define $\lim_{v_0\to 0^+} b'(v_0) = b'(0^+)$, $\lim_{v_0\to 0^-} b'(v_0) = b'(0^-)$, $\lim_{v_0\to 0^+} b(v_0) = b(0^+)$, and $\lim_{v_0\to 0^-} b(v_0) = b(0^-)$. We study the case where $b'(0^+) \neq b'(0^-)$ and $b(0^+) \neq b(0^-)$, but $b(\cdot)$ is still a smooth function on $I_V/\{0\}$. Similarly to the derivation in the proof of Proposition 1, we can show that the RK estimand identifies the following parameter:

$$\frac{\lim_{v_0\to 0^+} \left.\frac{dE[Y|V=v]}{dv}\right|_{v=v_0} - \lim_{v_0\to 0^-} \left.\frac{dE[Y|V=v]}{dv}\right|_{v=v_0}}{\lim_{v_0\to 0^+} \left.\frac{db(v)}{dv}\right|_{v=v_0} - \lim_{v_0\to 0^-} \left.\frac{db(v)}{dv}\right|_{v=v_0}}$$

$$= \left(b'(0^+) \int y_1\big(b(0^+), 0, u\big) \frac{f_{V|U=u}(0)}{f_V(0)}\, dF_U(u) - b'(0^-) \int y_1\big(b(0^-), 0, u\big) \frac{f_{V|U=u}(0)}{f_V(0)}\, dF_U(u)\right)$$
$$\Big/ \big(b'(0^+) - b'(0^-)\big)$$
$$+ \left(\int \left\{\big[y_2\big(b(0^+), 0, u\big) - y_2\big(b(0^-), 0, u\big)\big] \frac{f_{V|U=u}(0)}{f_V(0)} + \big[y\big(b(0^+), 0, u\big) - y\big(b(0^-), 0, u\big)\big] \frac{\partial}{\partial v} \frac{f_{V|U=u}(0)}{f_V(0)}\right\} dF_U(u)\right)$$
$$\Big/ \big(b'(0^+) - b'(0^-)\big),$$

which is, in general, not readily interpretable as a weighted average of the causal effect.

On the other hand, we can show that the RD estimand identifies a weighted average of the causal effect of interest:

$$\frac{\lim_{v_0\to 0^+} E[Y|V=v_0] - \lim_{v_0\to 0^-} E[Y|V=v_0]}{\lim_{v_0\to 0^+} b(v_0) - \lim_{v_0\to 0^-} b(v_0)}$$

$$= \Big(\lim_{v_0\to 0^+} E\big[y\big(b(v_0), v_0, U\big)|V=v_0\big]$$

$$
\begin{aligned}
&- \lim_{v_0 \to 0^-} E\big[y\big(b(v_0), v_0, U\big)|V = v_0\big]\Big) \\
&\Big/\Big(\lim_{v_0 \to 0^+} b(v_0) - \lim_{v_0 \to 0^-} b(v_0)\Big) \\
&= E\Big[\frac{y\big(b(0^+), 0, U\big) - y\big(b(0^-), 0, U\big)}{b\big(0^+\big) - b\big(0^-\big)}\Big|V = 0\Big] \\
&= E\big[y_1(\tilde{b}, 0, U)|V = 0\big],
\end{aligned}
$$

where $\tilde{b}$ is between $b(0^+)$ and $b(0^-)$ and the last line follows from the mean value theorem.

Similarly, in the fuzzy framework of Section 2.2.2, it can be shown that the RK estimand no longer identifies the causal effect of interest if we allow a discontinuity in $b(\cdot, e)$ for some e at the threshold. However, the RD estimand still identifies a weighted average of the causal effect y_1. To see this, let $\lim_{v_0^* \to 0^+} b(v_0^*, e) \equiv b(0,^+ e)$, $\lim_{v_0^* \to 0^-} b(v_0^*, e) \equiv b(0^-, e)$ and modify Assumption 3a and Assumption 6 by replacing $b_1^{\pm}(e)$ with $b(0^{\pm}, e)$; using notations from the proof of Proposition 2, we have

$$
\begin{aligned}
&\frac{\lim_{v_0 \to 0^+} E\big[Y|V^* = v_0\big] - \lim_{v_0 \to 0^-} E\big[Y|V^* = v_0\big]}{\lim_{v_0 \to 0^+} E\big[B^*|V^* = v_0\big] - \lim_{v_0 \to 0^-} E\big[B^*|V^* = v_0\big]} \\
&= \frac{\lim_{v_0 \to 0^+} \int z_1 z_2 z_5 \, dF_{U,\varepsilon}(u, e) - \lim_{v_0 \to 0^-} \int z_1 z_2 z_5 \, dF_{U,\varepsilon}(u, e)}{\lim_{v_0 \to 0^+} \int z_6 z_{13} z_{14} \, dF_{U,\varepsilon}(u, e) - \lim_{v_0 \to 0^-} \int z_6 z_{13} z_{14} \, dF_{U,\varepsilon}(u, e)} \\
&= \int \big[y\big(b\big(0^+, e\big), 0, u\big) - y\big(b\big(0^-, e\big), 0, u\big)\big] \\
&\quad \times \Pr[U_V = 0|V = 0, U = u, \varepsilon = e] \frac{f_{V|U=u,\varepsilon=e}(0)}{f_V(0)} \, dF_{U,\varepsilon}(u, e) \\
&\quad \Big/\Big(\int \big[b\big(0^+, e\big) - b\big(0^-, e\big)\big] \Pr[U_V = 0|V = 0, \varepsilon = e] \\
&\quad \times \frac{f_{V|\varepsilon=e}(0)}{f_V(0)} \, dF_\varepsilon(e)\Big) \\
&= \int \frac{y\big(b\big(0^+, e\big), 0, u\big) - y\big(b\big(0^-, e\big), 0, u\big)}{b\big(0^+, e\big) - b\big(0^-, e\big)} \big[b\big(0^+, e\big) - b\big(0^-, e\big)\big] \\
&\quad \times \Pr[U_V = 0|V = 0, U = u, \varepsilon = e] \frac{f_{V|U=u,\varepsilon=e}(0)}{f_V(0)} \, dF_{U,\varepsilon}(u, e)
\end{aligned}
$$

$$
\Big/\bigg(\int \big[b\big(0^+, e\big) - b\big(0^-, e\big)\big] \Pr[U_V = 0 | V = 0, \varepsilon = e] \times \frac{f_{V|\varepsilon=e}(0)}{f_V(0)}\, dF_\varepsilon(e)\bigg)
$$

$$
= \int y\big(\tilde{b}(e), 0, u\big)\psi(e, u)\, dF_{U,\varepsilon}(u, e),
$$

where $\tilde{b}(e)$ is a value between $b(0^+, e)$ and $b(0^-, e)$ for each e and

$$
\psi(e, u) = \big[b\big(0^+, e\big) - b\big(0^-, e\big)\big] \Pr[U_V = 0 | V = 0, U = u, \varepsilon = e] \times \frac{f_{V|U=u,\varepsilon=e}(0)}{f_V(0)} \Big/\bigg(\int \big[b\big(0^+, e\big) - b\big(0^-, e\big)\big] \Pr[U_V = 0 | V = 0, \varepsilon = e] \times \frac{f_{V|\varepsilon=e}(0)}{f_V(0)}\, dF_\varepsilon(e)\bigg).
$$

A.3. *Applying RKD When the Treatment Variable Is Binary—Remark 6*

We provide details on the RK identification result stated in Remark 6. The identifying assumptions are the following:

ASSUMPTION 1c—Regularity: (i) *The support of U and η are bounded*: *they are subsets of the arbitrarily large compact set $I_U \subset \mathbb{R}^m$ and $I_\eta = [c_\eta, d_\eta] \subset \mathbb{R}$, respectively*. (ii) *$y(t, v, u)$ is continuous on $I_{V,U}$ for $t = 0, 1$*. (iii) *$t(b, v, n)$ is continuously differentiable on $I_{b(V),V,\eta}$ and is strictly increasing in n for all $b, v \in I_{b(V),V}$*.

By Assumption 1c and the implicit function theorem, we can define the continuously differentiable function $\tilde{\eta} : I_{b(V)} \times I_V \to \mathbb{R}$ such that $t(b, v, \tilde{\eta}(b, v)) = 0$. Let $\tilde{\eta}(b(V), V)$ be the image of $I_{b(V),V}$ under the mapping $\tilde{\eta}$.

ASSUMPTION 2c—Smooth Effect of V: *$y_2(t, v, u)$ is continuous on $I_{V,U}$ for each $t = 0, 1$*.

ASSUMPTION 3c—First-Stage and Nonnegligible Population at the Kink: (i) *$b(\cdot)$ is a known function*, *everywhere continuous and continuously differentiable on $I_V \setminus \{0\}$*, *but* $\lim_{v\to 0^+} b'(v) \neq \lim_{v\to 0^-} b'(v)$. (ii) *The set $A_U \equiv \{u : f_{V,\eta|U=u}(v, n) > 0\ \forall (v, n) \in I_{V,\tilde{\eta}(b(V),V)}\}$ has a positive measure under U*: $\int_{A_U} dF_U(u) > 0$. (iii) $t_1(b_0, 0, n^0) \neq 0$.

ASSUMPTION 4c—Smooth Density: *The conditional density $f_{V,\eta|U=u}(v, n)$ and its partial derivative w.r.t. v, $\frac{\partial f_{V,\eta|U=u}(v,n)}{\partial v}$, are continuous on $I_{V,\eta,U}$*.

PROPOSITION 3: *Under Assumptions 1c–4c*:
(a) $\Pr(U \le u|V = v)$ *is continuously differentiable in* v *at* $v = 0$ $\forall u \in I_U$.
(b)

$$\frac{\lim_{v_0\to 0^+} \left.\frac{dE[Y|V=v]}{dv}\right|_{v=v_0} - \lim_{v_0\to 0^-} \left.\frac{dE[Y|V=v]}{dv}\right|_{v=v_0}}{\lim_{v_0\to 0^+} \left.\frac{dE[T|V=v]}{dv}\right|_{v=v_0} - \lim_{v_0\to 0^-} \left.\frac{dE[T|V=v]}{dv}\right|_{v=v_0}} = \int \big[y(1,0,u) - y(0,0,u)\big] \frac{f_{V,\eta|U=u}(0, n^0)}{f_{V,\eta}(0, n^0)}\, dF_U(u).$$

PROOF: The proof of (a) is analogous to that of Proposition 1(a).
For part (b), note that

$$\begin{aligned}
&\frac{d}{dv}E[T|V=v] \\
&= \frac{d}{dv}E[1_{[T^*\ge 0]}|V=v] = \frac{d}{dv}\int_{\tilde\eta(b(v),v)}^{d_\eta} f_{\eta|V=v}(n)\, dn \\
&= \int_{\tilde\eta(b(v),v)}^{d_\eta} \frac{\partial}{\partial v}\big[f_{\eta|V=v}(n)\big]\, dn \\
&\quad - \big[\tilde\eta_1\big(b(v),v\big)b'(v) + \tilde\eta_2\big(b(v),v\big)\big] f_{\eta|V=v}\big(\tilde\eta\big(b(v),v\big)\big),
\end{aligned}$$

where $\tilde\eta_k$ denotes the partial derivative of $\tilde\eta$ with respect to its kth argument. The second line follows from Assumption 1c, and the interchange of differentiation and integration in the third line is permitted by Assumption 4c. It follows that the denominator can be expressed as

$$\begin{aligned}
&\lim_{v_0\to 0^+} \left.\frac{dE[T|V=v]}{dv}\right|_{v=v_0} - \lim_{v_0\to 0^-} \left.\frac{dE[T|V=v]}{dv}\right|_{v=v_0} \\
&= -\Big[\lim_{v_0\to 0^+} b'(v_0) - \lim_{v_0\to 0^-} b'(v_0)\Big] \tilde\eta_1(b_0, 0) f_{\eta|V=0}\big(\tilde\eta(b_0, 0)\big) \\
&= -\Big[\lim_{v_0\to 0^+} b'(v_0) - \lim_{v_0\to 0^-} b'(v_0)\Big] \tilde\eta_1(b_0, 0) \frac{f_{V,\eta}(0, n^0)}{f_V(0)}.
\end{aligned}$$

Similarly, by Assumptions 1c, 2c, and 4c,

$$\begin{aligned}
&\frac{d}{dv}E[Y|V=v] \\
&= \frac{d}{dv}E\big[y(T,V,U)|V=v\big]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{d}{dv}\int\bigg\{\int_{\tilde{\eta}(b(v),v)}^{d_\eta} y(1,v,u) f_{\eta|V=v,U=u}(n)\,dn \\
&\quad + \int_{c_\eta}^{\tilde{\eta}(b(v),v)} y(0,v,u) f_{\eta|V=v,U=u}(n)\,dn\bigg\}\,dF_{U|V=v}(u) \\
&= \int\bigg\{\int_{\tilde{\eta}(b(v),v)}^{d_\eta} \frac{\partial}{\partial v}\big[y(1,v,u) f_{\eta|V=v,U=u}(n)\big]\,dn \\
&\quad + \int_{c_\eta}^{\tilde{\eta}(b(v),v)} \frac{\partial}{\partial v}\big[y(0,v,u) f_{\eta|V=v,U=u}(n)\big]\,dn\bigg\}\,dF_{U|V=v}(u) \\
&\quad - \int\big[y(1,v,u) - y(0,v,u)\big] f_{\eta|V=v,U=u}\big(\tilde{\eta}\big(b(v),v\big)\big) \\
&\quad \times \big[\tilde{\eta}_1\big(b(v),v\big)b'(v) + \tilde{\eta}_2\big(b(v),v\big)\big]\,dF_{U|V=v}(u),
\end{aligned}
$$

and it follows that the numerator is

$$
\begin{aligned}
&\lim_{v_0\to 0^+} \frac{dE[Y|V=v]}{dv}\bigg|_{v=v_0} - \lim_{v_0\to 0^-} \frac{dE[Y|V=v]}{dv}\bigg|_{v=v_0} \\
&= -\Big[\lim_{v_0\to 0^+} b'(v_0) - \lim_{v_0\to 0^-} b'(v_0)\Big]\tilde{\eta}_1(b_0,0) \\
&\quad \times \int\big[y(1,0,u) - y(0,0,u)\big] f_{\eta|V=0,U=u}\big(n^0\big)\,dF_{U|V=0}(u) \\
&= -\Big[\lim_{v_0\to 0^+} b'(v_0) - \lim_{v_0\to 0^-} b'(v_0)\Big]\tilde{\eta}_1(b_0,0) \\
&\quad \times \int\big[y(1,0,u) - y(0,0,u)\big] \frac{f_{V,\eta|U=u}\big(0,n^0\big)}{f_V(0)}\,dF_U(u).
\end{aligned}
$$

Assumption 3c(iii) and the implicit function theorem imply that $\tilde{\eta}_1(b_0,0) \neq 0$, and therefore,

$$
\text{(A.7)}\qquad
\begin{aligned}
&\frac{\displaystyle\lim_{v_0\to 0^+} \frac{dE[Y|V=v]}{dv}\bigg|_{v=v_0} - \lim_{v_0\to 0^-} \frac{dE[Y|V=v]}{dv}\bigg|_{v=v_0}}{\displaystyle\lim_{v_0\to 0^+} \frac{dE[T|V=v]}{dv}\bigg|_{v=v_0} - \lim_{v_0\to 0^-} \frac{dE[T|V=v]}{dv}\bigg|_{v=v_0}} \\
&= \int\big[y(1,0,u) - y(0,0,u)\big] \frac{f_{V,\eta|U=u}\big(0,n^0\big)}{f_{V,\eta}\big(0,n^0\big)}\,dF_U(u)
\end{aligned}
$$

by Assumption 3c. *Q.E.D.*

When the benefit variable b directly affects the outcome, that is, $Y = y(T, B, V, U)$, the fuzzy RKD estimand no longer identifies the causal effect of T on Y; rather, the effect of T on Y is confounded by the direct effect of B on Y. If Assumptions 1c–4c are modified accordingly, it can be shown that the RK estimand identifies the parameter

$$\underbrace{\int \big[y(1,0,u) - y(0,0,u)\big] \frac{f_{V,\eta|U=u}\big(0,n^0\big)}{f_{V,\eta}\big(0,n^0\big)}\, dF_U(u)}_{\text{(i)}} - \underbrace{\frac{E\big[y_2\big(T, b(V), V, U\big)|V=0\big]}{\tilde{\eta}_1(b_0,0) f_{\eta|V=0}(n_0)}}_{\text{(ii)}},$$

where term (i) is the same as the RHS of equation (A.7) and term (ii) is the component that depends on the direct impact of B on Y. To the extent that the researcher can determine the sign of (ii), which involves signing $E[y_2(T, b(V), V, U)|V = 0]$ and $\tilde{\eta}_1(b_0, 0)$, she can bound the treatment effect (i) with the RKD estimand. For example, when η represents a student's ability in the empirical example in Remark 6, we may assert that $\tilde{\eta}_1(b_0, 0) < 0$ because the expected return from college attendance increases with the amount of financial aid. The conditional expectation of the direct impact of B on Y, $E[y_2(T, b(V), V, U)|V = 0]$, may be positive because a more generous aid package allows a student more time to focus on her study. If these arguments were true, then the RKD estimand would serve as an upper bound on the economic returns to college attendance.

As stated in Remark 6, we can also allow the relationship between B and V to be fuzzy as in Section 2.2.2: $B = b(V, \varepsilon)$. In addition, we allow measurement error in V, U_V, which has a point mass at 0, and we only observe $V^* = V + U_V$. We do not need to consider the measurement error in B since the observed value of B does not appear in the RK estimand. We abstract away from potential measurement error in T and leave it for future research. The modified set of identifying assumptions are:

ASSUMPTION 1d—Regularity: *In addition to the conditions in Assumption 1c, the support of ε is bounded: it is a subset of the arbitrarily large compact set $I_\varepsilon \subset \mathbb{R}^k$.*

ASSUMPTION 3d—First-Stage and Nonnegligible Population at the Kink: *$b(v, e)$ is continuous on $I_{V,\varepsilon}$ and $b_1(v, e)$ is continuous on $(I_V \setminus \{0\}) \times I_\varepsilon$. Let $b_1^+(e) \equiv \lim_{v\to 0^+} b_1(v, e)$, $b_1^-(e) \equiv \lim_{v\to 0^-} b_1(v, e)$, $A_\varepsilon \equiv \{e : f_{V|\varepsilon=e}(0) > 0\}$, and $n^0(e) \equiv \tilde{\eta}(b(0, e), 0)$; then $\int \{|b_1^+(e) - b_1^-(e)||\tilde{\eta}_1(b(0, e), 0)| \Pr[U_V = 0|V = 0, \varepsilon = e, \eta = n^0(e)] f_{V,\eta|\varepsilon=e}(0, n^0(e))\, dF_\varepsilon(e) > 0$.*

ASSUMPTION 4d—Smooth Density: *The conditional density*

$$f_{V,\eta,U_{V'}|U=u,\varepsilon=e}(v, n, u_{V'})$$

and its partial derivative w.r.t. v,

$$\frac{\partial f_{V,\eta,U_{V'}|U=u,\varepsilon=e}(v, n, u_{V'})}{\partial v},$$

are continuous on $I_{V,\eta,U_{V'},U,\varepsilon}$.

ASSUMPTION 5d—Smooth Probability of No Error in V and B: *As a function of the realized values of* V, U, ε, η, *and* $U_{V'}$, *the conditional probability of* $U_V = 0$, *denoted by* $\pi(v, u, e, u_{V'}, u_{B'})$, *and its partial derivative w.r.t.* v *are continuous on* $I_{V,U,\varepsilon,\eta,U_{V'}}$.

ASSUMPTION 6d—Monotonicity: (i) *Either* $b_1^+(e) \geq b_1^-(e)$ *for all* e *or* $b_1^+(e) \leq b_1^-(e)$ *for all* e. (ii) $t_1(b(0, e), 0, n^0(e)) \geq 0$ *for all* e *or* $t_1(b(0, e), 0, n^0(e)) \leq 0$ *for all* e.

PROPOSITION 4: *Under Assumptions* 1d, 2, 3d–6d:

(a) $\Pr(U \leq u, \varepsilon = e, \eta = n|V^* = v^*)$ *is continuously differentiable in* v^* *at* $v^* = 0$ $\forall (u, e, n) \in I_{U,\varepsilon,\eta}$.

(b)

$$\frac{\lim_{v_0 \to 0^+} \left.\frac{dE[Y|V = v]}{dv}\right|_{v=v_0} - \lim_{v_0 \to 0^-} \left.\frac{dE[Y|V = v]}{dv}\right|_{v=v_0}}{\lim_{v_0 \to 0^+} \left.\frac{dE[T|V = v]}{dv}\right|_{v=v_0} - \lim_{v_0 \to 0^-} \left.\frac{dE[T|V = v]}{dv}\right|_{v=v_0}} = \int \big[y(1, 0, u) - y(0, 0, u)\big]\tilde{\varphi}(u, e)\, dF_{U,\varepsilon}(u, e),$$

where $\tilde{\varphi}(u, e) \equiv \frac{[b_1^+(e)-b_1^-(e)]\tilde{\eta}_1(b(0,e),0)\Pr[U_V=0|V=0,U=u,\varepsilon=e,\eta=n^0(e)]f_{V,\eta|U=u,\varepsilon=e}(0,n^0(e))}{\int\{[b_1^+(e)-b_1^-(e)]\tilde{\eta}_1(b(0,e),0)\Pr[U_V=0|V=0,\varepsilon=e,\eta=n^0(e)]f_{V,\eta|\varepsilon=e}(0,n^0(e))\, dF_\varepsilon(e)}$.

PROOF: The proof is similar to that of Proposition 2 and is omitted. *Q.E.D.*

APPENDIX B: ESTIMATION

B.1. *Two-Sample RKD*

As suggested by a referee, the triplet (Y, B, V) may not be jointly observed from a single data source. Instead, the vectors (Y_i, V_i) for $i = 1, \ldots, n_1$ are observed in data set 1 and (B_j, V_j) for $j = 1, \ldots, n_2$ are observed in data set 2.

Because of the requirement of a zero point mass in the U_V distribution in Assumption 3a, an RKD typically calls for administrative data as opposed to surveys based on a complex sampling design. Therefore, we assume that (Y_i, V_i) and (B_j, V_j) are independent i.i.d. samples as per Inoue and Solon (2010). The variances of the first-stage and reduced-form kink estimators, $\hat{\tau}_B = \hat{\kappa}_1^+ - \hat{\kappa}_1^-$ and $\hat{\tau}_Y = \hat{\beta}_1^+ - \hat{\beta}_1^-$, can be calculated by using the sharp RKD variance estimator, and the covariance between $\hat{\tau}_B$ and $\hat{\tau}_Y$ is zero by the independence assumption. It follows that the variance of the fuzzy RKD estimator $\frac{\hat{\tau}_Y}{\hat{\tau}_B}$ can be calculated by an application of the delta method. The robust confidence intervals in Calonico, Cattaneo, and Titiunik (2014c) can be constructed analogously by setting the covariances between the first-stage and reduced-form estimators to zero.

B.2. *Optimal Bandwidth in Fuzzy RKD*

In this section, we propose bandwidth selectors that minimize the asymptotic MSE of the fuzzy RD/RKD estimators, building on that in Imbens and Kalyanaraman (2012) (henceforth, IK bandwidth) and Calonico, Cattaneo, and Titiunik (2014c) (henceforth, CCT bandwidth). First we introduce notation similar to Calonico, Cattaneo, and Titiunik (2014c). Define $\mu_{\cdot+}^{(\nu)}$ and $\mu_{\cdot-}^{(\nu)}$ as the νth right and left derivatives of the conditional expectation of a random variable (Y or B) with respect to V at $V = 0$; let $\tau_{Y,\nu} \equiv \mu_{Y+}^{(\nu)} - \mu_{Y-}^{(\nu)}$ and $\tau_{B,\nu} \equiv \mu_{B+}^{(\nu)} - \mu_{B-}^{(\nu)}$. In addition, let σ_{Y+}^2, σ_{Y-}^2, σ_{B+}^2, σ_{B-}^2, σ_{YB+}, and σ_{YB-} be the conditional variances of Y and B and their conditional covariance on two sides of the threshold. Finally, let

$$\begin{aligned}\tilde{\varsigma}_{\nu,p,s}(h) = {}& \frac{1}{\tau_{B,\nu}}\big[\big(\hat{\mu}_{Y+}^{(\nu)}(h) - (-1)^s \hat{\mu}_{Y-}^{(\nu)}(h)\big) - \big(\mu_{Y+}^{(\nu)} - (-1)^s \mu_{Y-}^{(\nu)}\big)\big] \\ & + \frac{\tau_{Y,\nu}}{\tau_{B,\nu}^2}\big[\big(\hat{\mu}_{B+}^{(\nu)}(h) - (-1)^s \hat{\mu}_{B-}^{(\nu)}(h)\big) - \big(\mu_{B+}^{(\nu)} - (-1)^s \mu_{B-}^{(\nu)}\big)\big],\end{aligned}$$

where $\hat{\mu}_{\cdot+}^{(\nu)}$ and $\hat{\mu}_{\cdot-}^{(\nu)}$ are the pth order local polynomial estimator of $\mu_{\cdot+}^{(\nu)}$ and $\mu_{\cdot-}^{(\nu)}$, respectively.

Next we propose the lemma that generalizes Lemma 2 of Calonico, Cattaneo, and Titiunik (2014c) and serves as the fuzzy analog of its Lemma 1:

LEMMA 7: *Assume that Assumptions* 1–3 *in* Calonico, Cattaneo, and Titiunik (2014c) *are satisfied with* $S \geq p + 1$ *and* $\nu \leq p$. *If* $h \to 0$ *and* $nh \to \infty$, *then*

$$\begin{aligned}\mathrm{MSE}_{\nu,p,s} &= E\big[\big(\tilde{\varsigma}_{\nu,p,s}(h)\big)^2 | \{V_i\}_{i=1}^n\big] \\ &= h^{2(p+1-\nu)}\big[B_{F,\nu,p,p+1,s}^2 + o_p(1)\big] + \frac{1}{nh_n^{1+2\nu}}\big[V_{F,\nu,p} + o_p(1)\big],\end{aligned}$$

where

$$B_{F,\nu,p,r,s} = \left(\frac{1}{\tau_{B,\nu}} \frac{\mu_{Y+}^{(r)} - (-1)^{\nu+r+s}\mu_{Y-}^{(r)}}{r!} - \frac{\tau_{Y,\nu}}{\tau_{B,\nu}^2} \frac{\mu_{B+}^{(r)} - (-1)^{\nu+r+s}\mu_{B-}^{(r)}}{r!}\right)$$
$$\times \nu! e_\nu' \Gamma_p^{-1} \vartheta_{p,r},$$
$$V_{F,\nu,p} = \left(\frac{1}{\tau_{B,\nu}^2} \frac{\sigma_{Y-}^2 + \sigma_{Y+}^2}{f} - \frac{2\tau_{Y,\nu}}{\tau_{B,\nu}^3} \frac{\sigma_{YB-} + \sigma_{YB+}}{f} + \frac{\tau_{Y,\nu}^2}{\tau_{B,\nu}^4} \frac{\sigma_{B-}^2 + \sigma_{B+}^2}{f}\right)$$
$$\times \nu!^2 e_\nu' \Gamma_p^{-1} \Psi_p \Gamma_p^{-1} e_\nu,$$

with e_ν, Γ_p, Ψ_p, and $\vartheta_{p,r}$ as defined in Calonico, Cattaneo, and Titiunik (2014c). If, in addition, $B_{F,\nu,p,r,s} \neq 0$, then the asymptotic MSE-optimal bandwidth is $h_{\mathrm{MSE},F,\nu,p} = C_{F,\nu,p,s}^{1/(2p+3)} n^{-1/(2p+3)}$, where $C_{F,\nu,p,s} = \frac{(2\nu+1)V_{F,\nu,p}}{2(p+1-\nu)B_{F,\nu,p,p+1,s}^2}$.

PROOF: The proof of Lemma 7 is analogous to that of Lemma A2 of Calonico, Cattaneo, and Titiunik (2014c). *Q.E.D.*

Note that Lemma 2 of Calonico, Cattaneo, and Titiunik (2014c) is a special case of Lemma 7 above with $s = 0$. As in the sharp case, the bias of the fuzzy RD estimator depends on the difference *or sum* of the derivative estimator from the first stage and the outcome equation. Whether it is a difference or sum depends on the order of the derivative estimated as well as the order of the estimating polynomial. Based on Lemma 7, we propose procedures to compute the CCT and IK bandwidths adapted to the fuzzy RD/RKD designs in the two following subsections.

B.2.1. *Fuzzy Bandwidth Based on the CCT Procedure*

Define the local variance estimator

$$\hat{V}_{F,\nu,p}(h) = \frac{1}{\tilde{\tau}_{B,\nu}^2} \hat{V}_{YY,\nu,p}(h) - \frac{2\tilde{\tau}_{Y,\nu}}{\tilde{\tau}_{B,\nu}^3} \hat{V}_{YB,\nu,p}(h) + \frac{\tilde{\tau}_{Y,\nu}^2}{\tilde{\tau}_{B,\nu}^4} \hat{V}_{BB,\nu,p}(h),$$

where

$$\hat{V}_{R_1R_2,\nu,p}(h) = \hat{\mathcal{V}}_{R_1R_2+,\nu,p}(h) + \hat{\mathcal{V}}_{R_1R_2-,\nu,p}(h)$$
$$= \nu!^2 e_\nu' \Gamma_{+,p}^{-1}(h) \hat{\Psi}_{R_1R_2+,p}(h) \Gamma_{+,p}^{-1}(h) e_\nu / nh^{2\nu}$$
$$+ \nu!^2 e_\nu' \Gamma_{-,p}^{-1}(h) \hat{\Psi}_{R_1R_2-,p}(h) \Gamma_{-,p}^{-1}(h) e_\nu / nh^{2\nu},$$

with R_1 and R_2 serving as place holders for Y and B, and the quantities e_ν, $\Gamma_{+,p}(h)$, and $\hat{\Psi}_{R_1R_2+,p}(h)$ as defined in Calonico, Cattaneo, and Titiunik

(2014c). The constants Γ_p, $\vartheta_{p,q}$, $\mathcal{B}_{\nu,p}$, and $C_{\nu,p}(K)$ also follow the same definitions in Calonico, Cattaneo, and Titiunik (2014c).

Step 0: Use the CCT bandwidth (optimal in the MSE sense for estimating $\tau_{Y,\nu}$) to obtain preliminary estimates $\tilde{\tau}_{Y,\nu}$ and $\tilde{\tau}_{B,\nu}$.

Step 1: υ and c

1. $\upsilon = Const_K \cdot \min\{S_V, IQR_V/1.349\} \cdot n^{-1/5}$ where $Const_K = (\frac{8\sqrt{\pi}\int K(u)^2\,du}{3(\int u^2 K(u)\,du)^2})^{1/5}$; S_V^2 and IQR_V denote the sample variance and interquartile range of V. The selection of υ_n, which is based on Silverman's rule of thumb, is the same as in Calonico, Cattaneo, and Titiunik (2014c). Use υ to compute the variance estimator $\hat{V}_{F,q+1,q+1}(\hat{\upsilon})$, $\hat{V}_{F,p+1,q}(\hat{\upsilon})$, and $\hat{V}_{F,\nu,p}(\hat{\upsilon})$.

2. Run global polynomials of order $q+2$ separately for B and Y on each side of the threshold. Obtain estimators of the $(q+2)$th derivatives on both sides of the threshold $e'_{q+2}\hat{\gamma}_{Y\pm,q+2}$ and $e'_{q+2}\hat{\gamma}_{B\pm,q+2}$, and use them to calculate the bandwidth c: $c = \check{C}^{1/(2q+5)}_{F,q+1,q+1,\nu+q} n^{-1/(2q+5)}$,

$$\begin{aligned}\check{C}_{F,q+1,q+1,\nu+q} = {} & (2q+3)n\upsilon_n^{2q+3}\hat{V}_{F,q+1,q+1}(\hat{\upsilon}) \\ & \Big/\Bigg(2\mathcal{B}^2_{q+1,q+1}\Bigg\{\frac{1}{\tilde{\tau}_{B,\nu}}\big[e'_{q+2}\hat{\gamma}_{Y+,q+2} - (-1)^{\nu+q}e'_{q+2}\hat{\gamma}_{Y-,q+2}\big] \\ & - \frac{\tilde{\tau}_{Y,\nu}}{\tilde{\tau}^2_{B,\nu}}\big[e'_{q+2}\hat{\gamma}_{B+,q+2} - (-1)^{\nu+q}e'_{q+2}\hat{\gamma}_{B-,q+2}\big]\Bigg\}^2\Bigg).\end{aligned}$$

Step 2: h_q

Perform local regressions with bandwidth c to estimate the $(q+1)$th derivatives on both sides of the threshold and calculate bandwidth h_q: $\hat{h}_q = \hat{C}^{1/(2q+3)}_{F,p+1,q,\nu+q+1} n^{-1/(2q+3)}$,

$$\begin{aligned}&\hat{C}_{F,p+1,q,\nu+q+1} \\ &\quad = (2p+3)n\upsilon_n^{2p+3}\hat{V}_{F,p+1,q}(\hat{\upsilon}) \\ &\qquad \Big/\Bigg(2(q-p) \\ &\qquad \times \mathcal{B}^2_{p+1,q}\Bigg\{\frac{1}{\tilde{\tau}_{B,\nu}}\big[e'_{q+1}\hat{\beta}_{Y+,q+1}(\hat{c}) - (-1)^{\nu+q+1}e'_{q+1}\hat{\beta}_{Y-,q+1}(\hat{c})\big] \\ &\qquad - \frac{\tilde{\tau}_{Y,\nu}}{\tilde{\tau}^2_{B,\nu}}\big[e'_{q+1}\hat{\beta}_{B+,q+1}(\hat{c}) - (-1)^{\nu+q+1}e'_{q+1}\hat{\beta}_{B-,q+1}(\hat{c})\big]\Bigg\}^2\Bigg).\end{aligned}$$

Step 3: h

Perform local regression with bandwidth h_q to estimate the bias in the fuzzy RD/RKD estimator $\hat{\tau}_{F,\nu,p}$ and calculate the resulting main bandwidth h:

$$\hat{h} = \hat{C}^{1/(2p+3)}_{F,\nu,p,\nu+p+1} n^{-1/(2p+3)},$$

$$\begin{aligned}
&\hat{C}_{F,\nu,p,\nu+p+1} \\
&\quad = (2\nu+1) n v_n^{2\nu+1} \hat{V}_{F,\nu,p}(\hat{v}) \\
&\qquad \Big/ \Big(2(p+1-\nu) \\
&\qquad \times \mathcal{B}^2_{\nu,p} \Big\{ \frac{1}{\tilde{\tau}_{B,\nu}} \big[e'_{p+1} \hat{\beta}_{Y+,q}(\hat{h}_q) - (-1)^{\nu+p+1} e'_{p+1} \hat{\beta}_{Y-,q}(\hat{h}_q) \big] \\
&\qquad - \frac{\tilde{\tau}_{Y,\nu}}{\tilde{\tau}^2_{B,\nu}} \big[e'_{p+1} \hat{\beta}_{B+,q}(\hat{h}_q) - (-1)^{\nu+p+1} e'_{p+1} \hat{\beta}_{B-,q}(\hat{h}_q) \big] \Big\}^2 \Big).
\end{aligned}$$

Similarly to Calonico, Cattaneo, and Titiunik (2014c), we have the following consistency result for the fuzzy CCT bandwidth selectors proposed above.

PROPOSITION 5—Consistency of the CCT Bandwidth Selectors: *Let $\nu \le p < q$. Suppose Assumptions* 1–3 *in Calonico, Cattaneo, and Titiunik* (2014c) *hold with $S \ge q+2$ and that*

$$\begin{aligned}
&\frac{1}{\tilde{\tau}_{B,\nu}} \big[e'_{q+2} \hat{\gamma}_{Y+,q+2} - (-1)^{\nu+q} e'_{q+2} \hat{\gamma}_{Y-,q+2} \big] \\
&\quad - \frac{\tilde{\tau}_{Y,\nu}}{\tilde{\tau}^2_{B,\nu}} \big[e'_{q+2} \hat{\gamma}_{B+,q+2} - (-1)^{\nu+q} e'_{q+2} \hat{\gamma}_{B-,q+2} \big] \xrightarrow{p} c \neq 0.
\end{aligned}$$

Step 1. *If $B_{F,p+1,q,q+1,\nu+p+1} \neq 0$, then*

$$\frac{\hat{h}_q}{h_{\mathrm{MSE},F,p+1,q,\nu+p+1}} \xrightarrow{p} 1 \quad \textit{and} \quad \frac{\mathrm{MSE}_{F,p+1,q,\nu+p+1}(\hat{h}_q)}{\mathrm{MSE}_{p+1,q,\nu+p+1}(h_{\mathrm{MSE},p+1,q,\nu+p+1})} \xrightarrow{p} 1.$$

Step 2. *If $B_{F,\nu,p,p+1,0} \neq 0$, then*

$$\frac{\hat{h}}{h_{\mathrm{MSE},F,\nu,p,0}} \xrightarrow{p} 1 \quad \textit{and} \quad \frac{\mathrm{MSE}_{F,\nu,p,0}(\hat{h})}{\mathrm{MSE}_{F,\nu,p,0}(h_{\mathrm{MSE},F,\nu,p,0})} \xrightarrow{p} 1.$$

PROOF: Because the CCT bandwidth optimal for estimating $\tau_{Y,\nu}$ shrinks at the rate of $n^{-1/(2p+3)}$, the preliminary estimators, $\tilde{\tau}_{Y,\nu}$ and $\tilde{\tau}_{B,\nu}$, are consistent. The rest of the proof follows the arguments in the proof of Theorem A4 in Calonico, Cattaneo, and Titiunik (2014c). *Q.E.D.*

The optimal fuzzy RD bandwidth was proposed in Imbens and Kalyanaraman (2012). We suggest an extension to be used in the fuzzy RKD case ($\nu = 1$)

and state the bandwidth selectors for a generic ν. Calonico, Cattaneo, and Titiunik (2014c), Calonico, Cattaneo, and Titiunik (2014b), and Calonico, Cattaneo, and Titiunik (2014a) adapted the IK bandwidth selection procedure to h_q so that it can be used to bias-correct the RD estimator. Building upon these studies, we propose a further extension of the bandwidth selector for h_q to a general fuzzy design with a discontinuity in the νth derivative.

B.2.2. *Fuzzy Bandwidth Based on the IK Procedure*

Step 1: Use the sharp IK bandwidth (optimal in the MSE sense for estimating $\tau_{Y,\nu}$) to obtain preliminary estimates $\tilde{\tau}_{Y,\nu}$ and $\tilde{\tau}_{B,\nu}$.

Step 2: $\boldsymbol{v}$

1. $\hat{v} = 1.84 \cdot S_V \cdot n^{-1/5}$.

2. Use $\hat{h}_1$ to estimate $\hat{\sigma}^2_{Y\pm}(\hat{v})$, $\hat{\sigma}^2_{B\pm}(\hat{v})$, $\hat{\sigma}_{YB\pm}(\hat{v})$, and $\hat{f}(\hat{v})$ as specified in Imbens and Kalyanaraman (2012) (note that Imbens and Kalyanaraman (2012) used W to denote the treatment variable and use h_1 to denote this preliminary bandwidth).

Step 3: h_q

Run global regressions:

$$Y = \delta^Y \cdot 1_{[V\geq 0]} \cdot V^\nu + \alpha_0^Y + \alpha_1^Y V + \cdots + \alpha_{q+2}^Y V^{q+2} + \varepsilon^Y,$$
$$B = \delta^B \cdot 1_{[V\geq 0]} \cdot V^\nu + \alpha_0^B + \alpha_1^B V + \cdots + \alpha_{q+2}^B V^{q+2} + \varepsilon^B,$$

and use $\hat{\alpha}^Y_{q+2}$ and $\hat{\alpha}^B_{q+2}$ to construct

- $\hat{h}_{Y-,q+1} = (C_{q+1,q+1}(K_U)\frac{\hat{\sigma}^2_{Y-}(\hat{v})/\hat{f}(\hat{v})}{n_-(\hat{\alpha}^Y_{q+2})^2})^{1/(2q+5)}$,
- $\hat{h}_{Y+,q+1} = (C_{q+1,q+1}(K_U)\frac{\hat{\sigma}^2_{Y+}(\hat{v})/\hat{f}(\hat{v})}{n_+(\hat{\alpha}^Y_{q+2})^2})^{1/(2q+5)}$,
- $\hat{h}_{B-,q+1} = (C_{q+1,q+1}(K_U)\frac{\hat{\sigma}^2_{B-}(\hat{v})/\hat{f}(\hat{v})}{n_-(\hat{\alpha}^B_{q+2})^2})^{1/(2q+5)}$,
- $\hat{h}_{B+,q+1} = (C_{q+1,q+1}(K_U)\frac{\hat{\sigma}^2_{B+}(\hat{v})/\hat{f}(\hat{v})}{n_+(\hat{\alpha}^B_{q+2})^2})^{1/(2q+5)}$.

Perform $(q+1)$th order local regressions of Y and B on each side of the threshold with the uniform kernel K_U and bandwidths $\hat{h}_{Y\pm,q+1}$ and $\hat{h}_{B\pm,q+1}$. Using in the resulting estimators $\hat{\beta}_{Y\pm,q+1}(\hat{h}_{Y\pm,q+1})$ and $\hat{\beta}_{B\pm,q+1}(\hat{h}_{B\pm,q+1})$, we obtain $\hat{h}_q = \hat{C}^{1/(2q+3)}_{F,p+1,q,\nu+q+1} n^{-1/(2q+3)}$,

$$\hat{C}_{F,p+1,q,\nu+q+1}$$
$$= C_{p+1,q}(K) \cdot \Bigg(\frac{1}{\hat{f}(\hat{v})}\Bigg\{\frac{1}{\tilde{\tau}^2_{B,\nu}}\big[\hat{\sigma}^2_{Y+}(\hat{v}) + \hat{\sigma}^2_{Y-}(\hat{v})\big]$$

$$- \frac{2\tilde{\tau}_{Y,\nu}}{\tilde{\tau}^3_{B,\nu}}\big[\hat{\sigma}^2_{YB-}(\hat{v}) + \hat{\sigma}^2_{YB+}(\hat{v})\big] + \frac{\tilde{\tau}^2_{Y,\nu}}{\tilde{\tau}^4_{B,\nu}}\big[\hat{\sigma}^2_{B-}(\hat{v}) + \hat{\sigma}^2_{B+}(\hat{v})\big]\bigg\}\bigg)$$

$$\Big/(D_1 - D_2)^2,$$

where

$$D_1 = \frac{1}{\tilde{\tau}_{B,\nu}}\big[e'_{q+1}\hat{\beta}_{Y+,q+1}(\hat{h}_{Y+,q+1}) - (-1)^{\nu+q+1}e'_{q+1}\hat{\beta}_{Y-,q+1}(\hat{h}_{Y-,q+1})\big]$$

and

$$D_2 = \frac{\tilde{\tau}_{Y,\nu}}{\tilde{\tau}^2_{B,\nu}}\big[e'_{q+1}\hat{\beta}_{B+,q+1}(\hat{h}_{B+,q+1}) - (-1)^{\nu+q+1}e'_{q+1}\hat{\beta}_{B-,q+1}(\hat{h}_{B-,q+1})\big].$$

Step 4: h
Run global regressions:

$$Y = \delta^Y \cdot 1_{[V\geq 0]} \cdot V^\nu + \gamma^Y_0 + \gamma^Y_1 V + \cdots + \gamma^Y_{q+1}V^{q+1} + \varepsilon^Y,$$
$$B = \delta^B \cdot 1_{[V\geq 0]} \cdot V^\nu + \gamma^B_0 + \gamma^B_1 V + \cdots + \gamma^B_{q+1}V^{q+1} + \varepsilon^B,$$

and use $\hat{\gamma}^Y_{q+1}$ and $\hat{\gamma}^B_{q+1}$ to construct

- $\hat{h}_{Y-,q} = (C_{p+1,q}(K_U)\frac{\hat{\sigma}^2_{Y-}(\hat{v})/\hat{f}(\hat{v})}{n_-(\hat{\gamma}^Y_{q+1})^2})^{1/(2q+3)}$,
- $\hat{h}_{Y+,q} = (C_{p+1,q}(K_U)\frac{\hat{\sigma}^2_{Y+}(\hat{v})/\hat{f}(\hat{v})}{n_+(\hat{\gamma}^Y_{q+1})^2})^{1/(2q+3)}$,
- $\hat{h}_{B-,q} = (C_{p+1,q}(K_U)\frac{\hat{\sigma}^2_{B-}(\hat{v})/\hat{f}(\hat{v})}{n_-(\hat{\gamma}^B_{q+1})^2})^{1/(2q+3)}$,
- $\hat{h}_{B+,q} = (C_{p+1,q}(K_U)\frac{\hat{\sigma}^2_{B+}(\hat{v})/\hat{f}(\hat{v})}{n_+(\hat{\gamma}^B_{q+1})^2})^{1/(2q+3)}$.

Perform qth order local regressions of Y and B on each side of the threshold with bandwidths $\hat{h}_{Y\pm,q}$and $\hat{h}_{B\pm,q}$ and obtain local regression estimators $\hat{\beta}_{Y\pm,q}(\hat{h}_{Y\pm,q})$ and $\hat{\beta}_{B\pm,q}(\hat{h}_{B\pm,q})$. Plugging them in, we have an estimate of the main bandwidth h: $\hat{h} = \hat{C}^{1/(2p+3)}_{F,\nu,p,\nu+p+1}n^{-1/(2p+3)}$,

$$\hat{C}_{F,\nu,p,\nu+p+1}$$
$$= C_{\nu,p}(K)\bigg(\frac{1}{\hat{f}(\hat{v})}\bigg\{\frac{1}{\tilde{\tau}^2_{B,\nu}}\big[\hat{\sigma}^2_{Y+}(\hat{v}) + \hat{\sigma}^2_{Y-}(\hat{v})\big]$$
$$- \frac{2\tilde{\tau}_{Y,\nu}}{\tilde{\tau}^3_{B,\nu}}\big[\hat{\sigma}^2_{YB-}(\hat{v}) + \hat{\sigma}^2_{YB+}(\hat{v})\big] + \frac{\tilde{\tau}^2_{Y,\nu}}{\tilde{\tau}^4_{B,\nu}}\big[\hat{\sigma}^2_{B-}(\hat{v}) + \hat{\sigma}^2_{B+}(\hat{v})\big]\bigg\}\bigg)$$

$$\Big/ \Bigg\{ \frac{1}{\tilde{\tau}_{B,\nu}} \big[e'_{p+1} \hat{\beta}_{Y+,q}(\hat{h}_{Y+,q}) - (-1)^{\nu+p+1} e'_{p+1} \hat{\beta}_{Y-,q}(\hat{h}_{Y-,q}) \big]$$
$$- \frac{\tilde{\tau}_{Y,\nu}}{\tilde{\tau}^2_{B,\nu}} \big[e'_{p+1} \hat{\beta}_{B+,q}(\hat{h}_{B+,q}) - (-1)^{\nu+p+1} e'_{p+1} \hat{\beta}_{B-,q}(\hat{h}_{B-,q}) \big] \Bigg\}^2.$$

We have a similar consistency result for the IK bandwidth selectors below.

PROPOSITION 6—Consistency of the IK Bandwidth Selectors: *Let* $\nu \leq p < q$. *Suppose Assumptions* 1–3 *in Calonico, Cattaneo, and Titiunik* (2014c) *hold with* $S \geq q+2$ *and that* α^Y_{q+2}, α^B_{q+2}, γ^Y_{q+1}, *and* γ^B_{q+1} *are nonzero*. *Selector for* h_q: *If* $B_{F,p+1,q,q+1,\nu+p+1} \neq 0$, then

$$\frac{\hat{h}_q}{h_{\text{MSE},F,p+1,q,\nu+p+1}} \xrightarrow{p} 1 \quad \textit{and} \quad \frac{\text{MSE}_{F,p+1,q,\nu+p+1}(\hat{h}_q)}{\text{MSE}_{p+1,q,\nu+p+1}(h_{\text{MSE},p+1,q,\nu+p+1})} \xrightarrow{p} 1.$$

Selector for h. *If* $B_{F,\nu,p,p+1,0} \neq 0$, then

$$\frac{\hat{h}}{h_{\text{MSE},F,\nu,p,0}} \xrightarrow{p} 1 \quad \textit{and} \quad \frac{\text{MSE}_{F,\nu,p,0}(\hat{h})}{\text{MSE}_{F,\nu,p,0}(h_{\text{MSE},F,\nu,p,0})} \xrightarrow{p} 1.$$

PROOF: Because the IK bandwidth optimal for estimating $\tau_{Y,\nu}$ shrinks at the rate of $n^{-1/(2p+3)}$, the preliminary estimators, $\tilde{\tau}_{Y,\nu}$ and $\tilde{\tau}_{B,\nu}$, are consistent. The density, variance, and covariance estimators are consistent as argued in Imbens and Kalyanaraman (2012). Since the higher derivative estimators also converge to their population counterparts, $\hat{C}_{F,p+1,q,\nu+q+1} \xrightarrow{p} C_{F,p+1,q,\nu+q+1}$ and $\hat{C}_{F,\nu,p,\nu+p+1} \xrightarrow{p} C_{F,\nu,p,\nu+p+1}$, and the results of the proposition follow. *Q.E.D.*

REFERENCES

CALONICO, S., M. D. CATTANEO, AND R. TITIUNIK (2014a): "rdrobust: An R Package for Robust Inference in Regression-Discontinuity Designs," Technical Report, University of Michigan. [22]

——— (2014b): "Robust Data-Driven Inference in the Regression-Discontinuity Design," *Stata Journal*, 14, 909–946. [22]

——— (2014c): "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 82, 2295–2326. [18-22,24]

IMBENS, G. W., AND K. KALYANARAMAN (2012): "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *Review of Economic Studies*, 79, 933–959. [18,21,22,24]

INOUE, A., AND G. SOLON (2010): "Two-Sample Instrumental Variables Estimators," *The Review of Economics and Statistics*, 92, 557–561. [18]

ROUSSAS, G. G. (2004): *An Introduction to Measure-Theoretic Probability*. Amsterdam: Academic Press. [1,2,9]

ZORICH, V. A. (2004): *Mathematical Analysis II*. Berlin: Springer. [2]

Dept. of Economics, UC Berkeley, 549 Evans Hall #3880, Berkeley, CA 94720-3880, U.S.A., NBER, and IZA; card@berkeley.edu,

Dept. of Economics, Princeton University, 3 Nassau Hall, Princeton, NJ 08544, U.S.A. and NBER; davidlee@princeton.edu,

Dept. of Policy Analysis and Management, Cornell University, 134 MVR Hall, Ithaca, NY 14853, U.S.A.; zhuan.pei@cornell.edu,

and

Dept. of Economics, University of Mannheim, L 7, 3-5, D-68131 Mannheim, Germany and IZA; a.weber@uni-mannheim.de.

Manuscript received November, 2012; final revision received July, 2015.