Supplementary Material for
"The Effects of Foreign Exchange Intervention: Evidence from a Rule-Based Policy in Colombia"

Borradores de ECONOMÍA

Por: Guido M. Kuersteiner
David C. Phillips
Mauricio Villamizar-Villegas

Núm. 965
2016

BANCO DE LA REPÚBLICA
COLOMBIA

tá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Col

Supplementary Material for "The Effects of Foreign Exchange Intervention: Evidence from a Rule-Based Policy in Colombia"*

Guido M. Kuersteiner† David C. Phillips‡ Mauricio Villamizar-Villegas§

Abstract

This Appendix provides a more detailed discussion of the technical results, including proofs of theorems reported in the main paper. For ease of reference notation and definitions are repeated from the main paper.

Key Words: Rule-Based Foreign Exchange Interventions, Portfolio Balance, Central Bank Policy, Regression Discontinuity, Non-linear Impulse Response.

JEL Codes: E58, F31, C22

*We especially thank Juan Jose Echavarria, Hernando Vargas, Diego Rojas and Pamela Cardozo from the Central Bank of Colombia (Banco de la Republica) for answering numerous questions. The views expressed herein are those of the authors and not necessarily those of the Banco de la República nor its Board of Directors. All errors and omissions remain our responsibility.

†University of Maryland. Email: kuersteiner@econ.umd.edu

‡University of Notre Dame. Email: david.phillips.184@nd.edu

§Central Bank of Colombia. Email: mvillavi@banrep.gov.co

1 Model and Assumptions

Let χ_t be a strictly stationary stochastic process defined on a probability space $(\Omega, \mathcal{F}, P)$ and taking values in $\mathbb{R}^d$. Let $\mathcal{F}_l^k = \sigma(\chi_t : l \leq t \leq k)$ be the sigma field generated by $\{\chi_t\}_{t=l}^k$. The strong mixing coefficient α_m is defined as

$$\alpha_m = \sup_{A \in \mathcal{F}_{-\infty}^0, B \in \mathcal{F}_m^\infty} |P(A \cap B) - P(A) P(B)|.$$

The process χ_t is called strongly mixing (Doukhan, 1994) if $\alpha_m \to 0$ as $m \to \infty$.

Let (y_t, X_t) be macroeconomic variables measurable with respect to $\mathcal{F}_a^t$ for some $-\infty < a \leq t$. If χ_t is strongly mixing then also (y_t, X_t) are strongly mixing. We will refer to y_t as the outcome variable and X_t as the running variable. As in Hahn, Todd and Van der Klaauw (2001), consider the sharp regression discontinuity design where the policy D_t is enacted if X_t crosses a threshold c,

$$D_t = 1\{X_t > c\}.$$

The setup is flexible enough to include cases where $y_t = X_t$. This situation may occur when past realizations of the outcome variable are used to trigger the policy D_t. HTV demonstrate under what conditions a regression discontinuity design can be used to identify treatment effects in the potential outcomes framework of Rubin (1974). Here we show how to identify non-linear impulse response functions using regression discontinuity designs. Recall that the outcome variable y_{t+j} can be represented as

$$y_{t+j} = F_{t,j}(0, \chi_t) + \theta_j^D(\chi_t) D_t.$$

Because χ_t is strictly stationary it follows that the marginal and conditional distributions of $F_{t,j}(0, \chi_t)$ and $\theta_j^D(\chi_t)$ depend only on the horizon j but not on t. In order to define the treatment effect of interest we impose the following assumption which is similar to Assumption A1 in HTV.

Assumption 1 *The expectations $E[F_{t,j}(0, \chi_t) | X_t = x]$ and $E[F_{t,j}(1, \chi_t) | X_t = x]$ exist and are continuous in x at c for all $j \geq 1$.*

Assumption 1 is stronger than Assumption A1 in HTV in the sense that we not only require continuity of the conditional mean of $F_{t,j}(0, \chi_t)$ but also of that of $F_{t,j}(1, \chi_t)$. On the other hand, we do not assume that $\theta_j^D(\chi_t)$ is a fixed constant.

The parameter of interest is the expectation of the impulse response $\theta_j^D(\chi_t)$ conditional on $X_t = c$, given by

$$\theta_j(c) = E\left[\theta_j^D(\chi_t) | X_t = c\right].$$

The parameter $\theta_j(c)$ is well defined by Assumption 1 and can be estimated by local linear regression (LLR) as advocated in HTV. LLR goes back to Fan (1992) and was studied in the context of regression discontinuity designs in HTV and Porter (2003). Masry and Fan (1997) establish asymptotic properties as well as bandwidth selection rules for LLR with dependent data. HTV propose to estimate $\theta_j(c)$ with two separate LLR on the subsamples where $D_t = 1$ and $D_t = 0$.

As noted in the proof of Theorem 1 and 1' of Hahn, Todd and Van der Klaauw (1999) (henceforth HTV99), the asymptotic distribution for the combined estimator of $\theta_j(c)$ can be easily obtained in the case of cross-sectional samples because the subsamples where $D_t = 0$ and $D_t = 1$ are independent. This is clearly not the case in the context of time series data considered here. Imbens and Lemieux (2008) note that the combined method of HTV can be represented in a numerically equivalent regression using appropriate dummies and interaction terms. The advantage of their formulation in our context is that it automatically produces joint inference that accounts for the (temporal) dependence in our data. An additional complication that arises in our case is the fact that we are also interested in the joint distribution of estimators of all $\theta_j(c)$ for $j = 0, ..., J$. Neither the RD design nor the inclusion of multiple outcomes does directly correspond to the model considered in Masry and Fan (1997). T The necessary extensions are given here. Thus, let $a = (a_1, ..., a_J)'$, $b = (b_1, ..., b_J)$, $\gamma = (\gamma_1, ..., \gamma_J)'$ and $\theta = (\theta_1(c), ..., \theta_J(c))'$. Extending Imbens and Lemieux (2008) define the estimator $\hat{\theta}$ of the parameter θ as the solution to

$$\left(\hat{a}, \hat{b}, \hat{\gamma}, \hat{\theta}\right) = \underset{a,b,\gamma,\theta}{\arg\min} \sum_{j=1}^{J} \sum_{t=1}^{T-J} \left(y_{t+j} - a_j - b_j (X_t - c) - \theta_j D_t - \gamma_j (X_t - c) D_t\right)^2 K\left(\frac{X_t - c}{h}\right)$$

where $K(.) \geq 0$ is a kernel function and h is a bandwidth parameter, both to be specified in more detail below. Let $\Pi = (a, b, \theta, \gamma)$, be a $J \times 4$ matrix of parameters, $Y_t = (y_{t+1}, ..., y_{t+J})'$ and $Z_t = (1, (X_t - c), D_t, (X_t - c) D_t)'$. Now define the data-matrices $Y = (Y_1, ...Y_{T-J})'$, $Z = (Z_1, ...Z_{T-J})'$ and $W = \operatorname{diag}(K((X_1 - c)/h), ..., K((X_{T-J} - c)/h))$. Then,

$$\operatorname{vec} \hat{\Pi}' = \left(I_J \otimes \left(Z'WZ\right)^{-1} Z'W\right) \operatorname{vec} Y$$

where I_J is the $J \times J$ dimensional identity matrix. The expression for $\hat{\Pi}$ is formally the same as for weighted least squares in a system of seemingly unrelated regressions (SUR) and indicates in particular that $\hat{\theta}_j$ for a particular horizon j can be obtained by an individual weighted least squares regression for that horizon with the weights given by the kernel function. However, for joint inference on θ the joint distribution of these estimators needs to be derived. The following assumptions correspond to assumptions made in Masry and Fan (1997).

Assumption 2 *(i)Let $f(x)$ be the marginal distribution of X_t. Assume that $f(x)$ is continuous and bounded.*
(ii) $|f_l(u,v) - f(u)f(v)| \leq M < \infty$ for all $l > 0$ where $f_l(u,v)$ is the joint density of X_0 and X_l.
(iii) The process χ_t is strong mixing with $\sum_{m=1}^{\infty} m^a \alpha_m^{1-2/\delta} < \infty$ for some $\delta > 2$ and $a > 1 - 2/\delta$.
(iv) The kernel function $K(.)$ is a bounded density function satisfying $u^{4\delta+2}K(u) \to 0$ as $|u| \to \infty$.

Assumption 3 *(i) The kernel $K(.)$ is bounded with bounded support $[-1,1]$.*
(ii) Assume that $f_l(u,v) \leq M_1$ and $E\left[y_1^2 + y_j^2 | X_0 = u, X_l = v\right] \leq M_2 < \infty$ for all l and u,v in a neighborhood of c.
(iii) Let $\Sigma(x) = \operatorname{Var}(Y_t|X_t = x)$ and assume that $\Sigma(x)$ is positive definite and bounded for all x. For $\delta > 2$ as in Assumption 2, $E\left[|y_1|^\delta |X = u\right] \leq M_3 < \infty$ for all u in a neighborhood of c.
(iv) assume $h_T \to 0$ and $Th_T \to \infty$. (we often used the notation h instead of h_T). Assume that there is a sequence $s_T > 0$ such that $s_T \to \infty$ and $s_T = o\left((Th_T)^{1/2}\right)$ such that $(T/h_T)^{1/2} \alpha_{s_T} \to 0$ as $T \to \infty$.

An additional set of technical assumptions specific to the RD estimator are similar to assumptions made in HTV.

Assumption 4 *Let $m_j(x) = E[y_{t+j}|X_t = x]$, $m_j^+(x_0) = \lim_{x\to x_0^+} E[y_{t+j}|X_t = x]$ and*

$$m_j^-(x_0) = \lim_{x\to x_0^-} E[y_{t+j}|X_t = x].$$

For $x > c$, assume that $m_j^+(x)$ is twice continuously differentiable with uniformly bounded derivatives $m_j^{'+}(x), m_j^{''+}(x)$ on $(c, c+M]$. Similarly, for $x < c$, $m_j^-(x)$ is twice continuously differentiable with uniformly bounded derivatives $m_j^{'-}(x), m_j^{''-}(x)$ on $(c-M, c)$ for some M. Let $\Sigma^+(x_0) = \lim_{x\to x_0^+} \operatorname{Var}(Y_t|X_t = x)$, $\Sigma^-(x_0) = \lim_{x\to x_0^-} \operatorname{Var}(Y_t|X_t = x)$ and assume that $\Sigma^+(x_0)$, $\Sigma^-(x_0)$ positive definite for $x_0 = c$.

2 Results

This section summarizes the results for the identification of the impulse response function and the asymptotic distribution of $\operatorname{vec}\hat{\Pi}'$ and of individual components of this vector. In particular the joint distribution of the impulse response function θ as well as of individual components θ_j is of interest. The latter are useful for optimal bandwidth selection which leads to similar results as in Masry and Fan (1997) and Imbens and Kalyanaraman (2012).

We start with a result on the identification of the impulse response function.

Theorem 5 *Assume that there is a non-random function* $g : \mathbb{R}^d \to \mathbb{R}$ *of* χ_t *such that* $g(\chi_t) = X_t$ *and* c *is a known threshold. Let* $D_t = 1\{X_t > c\}$. *If Assumption 1 holds, it follows that*

$$\theta_j(c) = \lim_{x \to c^+} E[y_{t+j}|X_t = x] - \lim_{x \to c^-} E[y_{t+j}|X_t = x]$$

Assume that there exists at least one continuous path $\varepsilon(\delta) : [0,\infty) \to \mathbb{R}^d$ *such that* $g(\chi_t + \varepsilon(\delta)) - g(\chi_t) = \delta$ *for all* $\delta \geq 0$ *and* $\varepsilon(0) = 0$.*Assume that* $E[F_{t,j}(D_t, \chi_t)|\chi_t = x]$ *is continuous in* x *a.s.,* $|E[F_{t,j}(D_t, \chi_t)|\chi_t]| \leq B(\chi_t)$ *and* $E[B(\chi_t)|X_t = c] < \infty$ *a.s. Let* $\theta_j(\varepsilon, \chi_t) = F_{t,j}(D_t(\chi_t + \varepsilon), \chi_t + \varepsilon) - F_{t,j}(D_t(\chi_t), \chi_t)$. *Then it follows that*

$$\theta_j(c) = \lim_{\delta \downarrow 0} E[\theta_j(\varepsilon(\delta), \chi_t)|X_t = c].$$

Remark 1 *Note that the local conditional independence assumption used in HTV, Theorem 2 is not needed here because we only consider sharp regression discontinuity designs.*

We introduce the following notation needed for the statements of the asymptotic distributions. Let $\mu_{lk} = \int_{-\infty}^{\infty} 1\{u > 0\}^k u^l K(u)\, du$ and define the matrix

$$\Gamma = \begin{bmatrix} \mu_{00} & \mu_{10} & \mu_{01} & \mu_{11} \\ \mu_{10} & \mu_{20} & \mu_{11} & \mu_{21} \\ \mu_{01} & \mu_{11} & \mu_{01} & \mu_{11} \\ \mu_{11} & \mu_{21} & \mu_{11} & \mu_{21} \end{bmatrix}.$$

Let $v_l^+ = \int_0^\infty u^l K^2(u)\, du$ and $v_l^- = \int_{-\infty}^0 u^l K^2(u)\, du$ and define the matrices

$$V^+ = \begin{bmatrix} v_0^+ & v_1^+ \\ v_1^+ & v_2^+ \end{bmatrix},\ V^- = \begin{bmatrix} v_0^- & v_1^- \\ v_1^- & v_2^- \end{bmatrix},$$

and

$$\Omega^+ = \begin{bmatrix} V^+ & V^+ \\ V^+ & V^+ \end{bmatrix},\ \Omega^- = \begin{bmatrix} V^- & 0 \\ 0 & 0 \end{bmatrix}$$

Also define

$$\Lambda_{lk}^- = 1\{k = 0\} \int_{-\infty}^0 u^{l+2} K(u)\, du$$

and

$$\Lambda_{lk}^+ = \int_0^\infty u^{l+2} K(u)\, du$$

and let $\Lambda^- = \left(\Lambda^-_{00}, \Lambda^-_{10}, \Lambda^-_{01}, \Lambda^-_{11}\right)'$ and similarly for $\Lambda^+ = \left(\Lambda^+_{00}, \Lambda^+_{10}, \Lambda^+_{01}, \Lambda^+_{11}\right)'$ as well as $m''^-(c) = \left(m''^-_0(c), ..., m''^-_J(c)\right)'$ and similarly for $m''^+(c)$. With this notation we can now state the first result.

Theorem 6 *Assume that Assumptions 1-4[GK: additional reg conditions as in C4 of MF96 are not assumed because of the RD design - double check that this is not causing problems] hold and that* $h = O\left(T^{-1/5}\right)$. *Let* $H_T = \operatorname{diag}\left(1, h^{-1}, h^{-1}, 1\right)$. *Then,*

$$\sqrt{Th}\left(H_T \operatorname{vec}\left(\hat{\Pi}' - \Pi'\right) - \frac{h^2}{2}\left(m''^-(c) \otimes \Gamma^{-1}\Lambda^- + m''^+(c) \otimes \Gamma^{-1}\Lambda^+\right)\right)$$
$$\to_d N\left(0, f(c)^{-1}\left(\Sigma^+(c) \otimes \Gamma^{-1}\Omega^+\Gamma'^{-1} + \Sigma^-(c) \otimes \Gamma^{-1}\Omega^-\Gamma'^{-1}\right)\right)$$

as $T \to \infty$.

It is interesting to note that the product structure of the asymptotic variance covariance matrix is similar to systems estimators with iid errors even though we have not made any such assumptions. The next result considers the marginal limiting distribution of $\hat{\theta}$. Let b^+ be the third element of $\Gamma^{-1}\Lambda^+$, b^- the third element of $\Gamma^{-1}\Lambda^-$, ω^+ the third diagonal element of $\Gamma^{-1}\Omega^+\Gamma'^{-1}$ and ω^- the third diagonal element of $\Gamma^{-1}\Omega^-\Gamma'^{-1}$.

Theorem 7 *Assume that Assumptions 1-4 hold and that* $h = O\left(T^{-1/5}\right)$. *Then,*

$$\sqrt{Th}\left(\hat{\theta} - \theta - \frac{h^2}{2}\left(m''^-(c)\, b^- + m''^+(c)\, b^+\right)\right) \to_d N\left(0, f(c)^{-1}\left(\Sigma^-(c)\,\omega^- + \Sigma^+(c)\,\omega^+\right)\right)$$

as $T \to \infty$.

Finally, consider the limiting distribution of an individual impulse coefficient $\hat{\theta}_j$ for the response at horizon j. In this case, let $\sigma^2_{+j}(c)$ be the corresponding diagonal element of $\Sigma^+(c)$ and $\sigma^2_{-j}(c)$ the corresponding diagonal element of $\Sigma^-(c)$. We obtain the following result.

Theorem 8 *Assume that Assumptions 1-4 hold and that* $h = O\left(T^{-1/5}\right)$. *Then,*

$$\sqrt{Th}\left(\hat{\theta}_j - \theta_j - \frac{h^2}{2}\left(m''^-_j(c)\, b^- + m''^+_j(c)\, b^+\right)\right) \to_d N\left(0, f(c)^{-1}\left(\sigma^2_{+j}(c)\,\omega^+ + \sigma^2_{-j}(c)\,\omega^-\right)\right)$$

as $T \to \infty$.

The results in Theorem 7 and 8 can be used to obtain optimal bandwidth rules analogous to the ones obtained by Masry and Fan (1997) and Imbens and Kalyanaraman (2012). Since we are often interested in joint estimation of the impulse response function θ it makes sense to consider the average means squared error across all impulse response coefficients. Let $\lambda \in \mathbb{R}^{J+1}$ with $\|\lambda\| = 1$. For example, we consider equal weights $\lambda = 1_J/J$ where 1_a is a vector of length a with all elements set equal to one. A second case of interest is $\lambda = e_j$ where e_j is the j-th unit vector in $\mathbb{R}^{J+1}$. Then, $\lambda'\hat{\theta}$ is the weighted average of the impulse response coefficients with bias

$$\frac{h^2}{2}\lambda'\left(m''^{-}(c)\, b^{-} + m''^{+}(c)\, b^{+}\right)$$

and variance

$$\frac{\lambda'\Sigma^{-}(c)\,\lambda\omega^{-} + \lambda'\Sigma^{+}(c)\,\lambda\omega^{+}}{hTf(c)}.$$

The optimal bandwidth is obtained by minimizing the asymptotic mean squared error with respect to h and given as

$$h_{opt} = \left(\frac{\lambda'\Sigma^{-}(c)\,\lambda\omega^{-} + \lambda'\Sigma^{+}(c)\,\lambda\omega^{+}}{f(c)\left(\lambda'\left(m''^{-}(c)\, b^{-} + m''^{+}(c)\, b^{+}\right)\right)^2}\right)^{1/5} T^{-1/5}.$$

For a given horizon j this formula specializes to

$$h_{opt,j} = \left(\frac{\sigma^2_{+j}(c)\,\omega^{+} + \sigma^2_{-j}(c)\,\omega^{-}}{f(c)\left(m''^{-}_j(c)\, b^{-} + m''^{+}_j(c)\, b^{+}\right)^2}\right)^{1/5} T^{-1/5}$$

Note that we are not assuming $b^{-} = -b^{+}$. However, with a symetric kernel it follows that $b^{-} = -b^{+}$ and $\omega^{+} = \omega^{-}$. The bandwidth formula the futher simplifies to

$$h_{opt,j} = \left(\frac{\omega^{+}}{(b^{-})^2}\right)^{1/5}\left(\frac{\sigma^2_{+j}(c) + \sigma^2_{-j}(c)}{f(c)\left(m''^{-}_j(c) - m''^{+}_j(c)\right)^2}\right)^{1/5} T^{-1/5}$$

which corresponds to the plug in formula of Imbens and Kalyanaraman (2012). For example, for the Bartlett kernel $K(u) = (1 - |u|)\, 1\{|u| \leq 1\}$ it can be shown that $\omega^{+} = \omega^{-} = 24/5$ and $b^{-} = -b^{+} = 1/10$ which leads to $\left(\omega^{+}/(b^{-})^2\right)^{1/5} = 2\,(15)^{1/5} \approx 3.4375$ which is the same as the constant obtained in Imbens and Kalyanaraman (2012).

3 Proofs

Proof of Theorem 5. The proof of the first part closely follows the argument in HTV. First note that

$$E\left[y_{t+j}|X_t = x\right] = E\left[y_{t+j}|g(\chi_t) = x\right].$$

Using the representation for y_{t+j} one obtains $y_{t+j} = F_{t+j}(0, \chi_t) + \theta_j^D(\chi_t) D_t$

$$E[y_{t+j}|g(\chi_t) = x] = E\left[F_{t+j}(0, \chi_t) + \theta_j^D(\chi_t) D_t(\chi_t)|g(\chi_t) = x\right]$$

Since $D_t(\chi_t) = 1$ conditional on $g(\chi_t) = x > c$ and $D_t(\chi_t) = 0$ conditional on $g(\chi_t) = x \leq c$ we obtain for all $x > c$ that

$$E[y_{t+j}|g(\chi_t) = x] = E[F_{t+j}(0, \chi_t)|g(\chi_t) = x] + E\left[\theta_j^D(\chi_t)|g(\chi_t) = x\right]$$

and for all $x \leq c$ that

$$E[y_{t+j}|g(\chi_t) = x] = E[F_{t+j}(0, \chi_t)|g(\chi_t) = x].$$

By Assumption 1 it follows that $\lim_{x\downarrow c} E[F_{t+j}(0, \chi_t)|g(\chi_t) = x] = \lim_{x\uparrow c} E[F_{t+j}(0, \chi_t)|g(\chi_t) = x]$ and $\lim_{x\downarrow c} E\left[\theta_j^D(\chi_t)|g(\chi_t) = x\right] = \lim_{x\uparrow c} E\left[\theta_j^D(\chi_t)|g(\chi_t) = x\right] = E\left[\theta_j^D(\chi_t)|g(\chi_t) = c\right]$. The first result now follows.

For the second part of the Theorem note that $\theta_j(\varepsilon(\delta), \chi_t) = F_{t+j}(D_t(\chi_t + \varepsilon(\delta)), \chi_t + \varepsilon(\delta)) - F_{t+j}(D_t(\chi_t), \chi_t)$ and for $\delta > 0$,

$$\begin{aligned} E[\theta_j(\varepsilon(\delta), \chi_t)|X_t = c] &= E[F_{t+j}(D_t(\chi_t + \varepsilon(\delta)), \chi_t + \varepsilon(\delta)) - F_{t+j}(D_t(\chi_t), \chi_t)|X_t = c] \\ &= E[F_{t+j}(1, \chi_t + \varepsilon(\delta))|X_t = c] - E[F_{t+j}(0, \chi_t)|X_t = c]. \end{aligned}$$

Consider

$$E[F_{t+j}(1, \chi_t + \varepsilon(\delta))|X_t = c] = E[E[F_{t+j}(1, \chi)|\chi = \chi_t + \varepsilon(\delta)]|X_t = c].$$

By the assumption of continuity, $\lim_{\delta\downarrow 0} E[F_{t+j}(1, \chi)|\chi = \chi_t + \varepsilon(\delta)] = E[F_{t+j}(1, \chi)|\chi = \chi_t]$ a.s. and $|E[F_{t+j}(1, \chi)|\chi = \chi_t + \varepsilon(\delta)]| \leq B(\chi_t)$ with $E[|B(\chi_t)||X_t = c] < \infty$. Then it follows from the Lebesque Convergence Theorem (see Royden, 1988, p.91) that

$$\lim_{\delta\downarrow 0} E[F_{t+j}(1, \chi_t + \varepsilon(\delta))|X_t = c] = E[F_{t+j}(1, \chi_t)|X_t = c] \text{ a.s.}$$

and the result follows from the definition of $\theta_j^D(\chi_t)$. ■

Proof of Theorem 7. Note that, similar to Masry and Fan (1997, p.168) the estimator $\operatorname{vec}\hat{\Pi}$ can be written as

$$\begin{aligned} \operatorname{vec}\left(H_T\hat{\Pi}'\right) &= (I \otimes H_T)\operatorname{vec}\hat{\Pi}' = \left(I_J \otimes H_T\left(Z'WZ\right)^{-1} Z'W\right)\operatorname{vec} Y \\ &= \left(I_J \otimes \left(H_T Z'WZH_T\right)^{-1} H_T Z'W\right)\operatorname{vec} Y \end{aligned}$$

First consider the matrix

$$\begin{aligned}
& (Th)^{-1} H_T Z' W Z H_T \\
= \; & (Th)^{-1} \sum_{t=2}^{T-J} \begin{bmatrix} 1 & \frac{X_t-c}{h} & D_t & \frac{X_t-c}{h} D_t \\ \frac{X_t-c}{h} & \left(\frac{X_t-c}{h}\right)^2 & \left(\frac{X_t-c}{h}\right) D_t & \left(\frac{X_t-c}{h}\right)^2 D_t \\ D_t & \left(\frac{X_t-c}{h}\right) D_t & D_t & \frac{X_t-c}{h} D_t \\ \frac{X_t-c}{h} D_t & \left(\frac{X_t-c}{h}\right)^2 D_t & \frac{X_t-c}{h} D_t & \left(\frac{X_t-c}{h}\right)^2 D_t \end{bmatrix} K\left(\frac{X_t-c}{h}\right).
\end{aligned}$$

Using results from HTV and Masry and Fan (1997, Theorem 1) one finds that

$$h^{-1} E\left[\left(\frac{X_t-c}{h}\right)^l K\left(\frac{X_t-c}{h}\right)\right] = f(c) \int_{-\infty}^{\infty} u^l K(u)\, du + o_h(1) \;\text{ for } l = 0, 1, 2. \tag{1}$$

where $o_h(1) \to 0$ as $h \to 0$. Similarly,

$$h^{-1} E\left[D_t \left(\frac{X_t-c}{h}\right)^l K\left(\frac{X_t-c}{h}\right)\right] = f(c) \int_{0}^{\infty} u^l K(u)\, du + o_h(1) \;\text{ for } l = 0, 1, 2. \tag{2}$$

Let $\mu_{lk} = \int_{-\infty}^{\infty} 1\{u > 0\}^k u^l K(u)\, du$, which implies that for $h \to 0$ such that $hT \to \infty$ that

$$(Th)^{-1} E\left[H_T Z' W Z H_T\right] = f(c)\, \Gamma.$$

It follows from Masry and Fan (1997, Theorem 1) that

$$Th \operatorname{Var}\left[(Th)^{-1} \sum_{t=1}^{T-J} \left(\frac{X_t-c}{h}\right)^l K\left(\frac{X_t-c}{h}\right)\right] \to f(c)\, \vartheta_{2l} \text{ for } l = 0, 1, 2. \tag{3}$$

where $\vartheta_l = \int_{-\infty}^{\infty} u^l K^2(u)\, du$. For

$$\xi_t^l = D_t \left(\frac{X_t-c}{h}\right)^l K\left(\frac{X_t-c}{h}\right)$$

it remains to consider

$$Th \operatorname{Var}\left[(Th)^{-1} \sum_{t=1}^{T-J} \xi_t^l\right] = (Th)^{-1} \sum_{t=1}^{T-J} \operatorname{Var}\left(\xi_t^l\right) + 2h^{-1} \sum_{j=1}^{T-J} \left(1 - \frac{j}{T}\right) \operatorname{Cov}\left(\xi_1^l, \xi_{1+j}^l\right) \tag{4}$$

where by (2) it follows that

$$\begin{aligned} h^{-1}\operatorname{Var}\left(\xi_t^l\right) &= h^{-1}E\left[D_t\left(\frac{X_t-c}{h}\right)^{2l}K^2\left(\frac{X_t-c}{h}\right)\right] \\ &\quad -h\left(E\left[h^{-1}D_t\left(\frac{X_t-c}{h}\right)^{l}K\left(\frac{X_t-c}{h}\right)\right]\right)^2 \\ &= h^{-1}E\left[D_t\left(\frac{X_t-c}{h}\right)^{2l}K^2\left(\frac{X_t-c}{h}\right)\right]+o_h(1) \\ &= f(c)\int_0^\infty u^{2l}K^2(u)\,du+o_h(1)\,. \end{aligned} \tag{5}$$

For the covariance term in (4) follow Masry and Fan (1997) and choose d_T such that $d_T\to\infty$ and $d_T h\to 0$. Then,

$$h^{-1}\sum_{j=1}^{T-J}\left|\operatorname{Cov}\left(\xi_1^l,\xi_{1+j}^l\right)\right| = h^{-1}\sum_{j=1}^{d_T}\left|\operatorname{Cov}\left(\xi_1^l,\xi_{1+j}^l\right)\right| + h^{-1}\sum_{j=d_T+1}^{T-J}\left|\operatorname{Cov}\left(\xi_1^l,\xi_{1+j}^l\right)\right| \tag{6}$$

where

$$\begin{aligned} & h^{-2}\left|E\left[\xi_1^l\xi_{1+j}^l\right]\right| \\ = \;& h^{-2}\left|\int_{-\infty}^{\infty}1\{v>c\}\,1\{u>c\}\left(\frac{v-c}{h}\right)^l\left(\frac{u-c}{h}\right)^l K\left(\frac{v-c}{h}\right)K\left(\frac{u-c}{h}\right)f_j(u,v)\,dudv\right| \end{aligned}$$

and therefore

$$\begin{aligned} & h^{-2}\left|E\left[D_1D_j\left(\frac{X_0-c}{h}\right)^l\left(\frac{X_j-c}{h}\right)^l K\left(\frac{X_0-c}{h}\right)K\left(\frac{X_j-c}{h}\right)\right]\right| \\ \leq\;& h^{-2}\int_{-\infty}^{\infty}\left|\left(\frac{v-c}{h}\right)^l\left(\frac{u-c}{h}\right)^l K\left(\frac{v-c}{h}\right)K\left(\frac{u-c}{h}\right)\right|\left|f_j(u,v)-f(u)f(v)\right|dudv \\ & +\left|\int_{-\infty}^{\infty}1\{v>c\}\left(\frac{v-c}{h}\right)^l h^{-1}K\left(\frac{v-c}{h}\right)f(v)\,dv\right|^2 \\ \leq\;& M\left(\int_{-\infty}^{\infty}\left|\left(\frac{v-c}{h}\right)^l h^{-1}K\left(\frac{v-c}{h}\right)\right|dv\right)^2+\left|f(c)\int_0^\infty u^lK(u)\,du+o_h(1)\right|^2 \end{aligned}$$

where the first inequality follows from Assumption 2(ii) and the second inequality follows from (2). By a change of variables we have

$$\int_{-\infty}^{\infty}\left|\left(\frac{v-c}{h}\right)^{l}h^{-1}K\left(\frac{v-c}{h}\right)\right|dv=\int_{-\infty}^{\infty}\left|u^{l}K\left(u\right)\right|dv=\int_{0}^{\infty}u^{l}K\left(u\right)dv+\int_{0}^{\infty}u^{l}K\left(-u\right)dv\leq M$$

by Assumption 2(iv). Also note

$$\left|\operatorname{Cov}\left(\xi_{1}^{l},\xi_{1+j}^{l}\right)\right|\leq\left|E\left[\xi_{1}^{l}\xi_{1+j}^{l}\right]\right|+\left|E\left[\xi_{1}^{l}\right]\right|\left|E\left[\xi_{1+j}^{l}\right]\right|$$

where

$$h^{-2}\left|E\left[\xi_{1}^{l}\right]\right|\left|E\left[\xi_{1+j}^{l}\right]\right|=\left(f\left(c\right)\int_{0}^{\infty}u^{l}K\left(u\right)du\right)^{2}+o_{h}\left(1\right)$$

by (2). This implies that

$$h^{-1}\sum_{j=1}^{d_T}\left|\operatorname{Cov}\left(\xi_{1}^{l},\xi_{1+j}^{l}\right)\right|\leq hd_{T}M=o\left(1\right). \tag{7}$$

For the second term in (6) use the mixing inequality in Hall and Heyde (1980, Corollary A2), also given in Masry and Fan (1997, p.171) whereby

$$h^{-2}\left|\operatorname{Cov}\left(\xi_{1}^{l},\xi_{1+j}^{l}\right)\right|\leq 8\alpha_{j}^{1-2/\delta}E\left[\left|h^{-1}\xi_{1}^{l}\right|^{\delta}\right]^{2/\delta}. \tag{8}$$

Since

$$\begin{aligned}
E\left[\left|h^{-1}\xi_{1}^{l}\right|^{\delta}\right] &= E\left[\left|D_{1}\left(\frac{X_{0}-c}{h}\right)^{l}h^{-1}K\left(\frac{X_{0}-c}{h}\right)\right|^{\delta}\right] \qquad (9)\\
&= \int_{-\infty}^{\infty}1\left(u>c\right)\left|\left(\frac{u-c}{h}\right)^{l}h^{-1}K\left(\frac{u-c}{h}\right)\right|^{\delta}f\left(u\right)du\\
&\leq \sup_{u}f\left(u\right)\int_{-\infty}^{\infty}h^{-\delta}\left|\left(\frac{u-c}{h}\right)^{l}K\left(\frac{u-c}{h}\right)\right|^{\delta}du\\
&= \sup_{u}f\left(u\right)h^{-\delta+1}\int_{-\infty}^{\infty}\left|u^{l}K\left(u\right)\right|^{\delta}du\\
&\leq \sup_{u}f\left(u\right)\sup_{u}K\left(u\right)^{\delta-1}h^{-\delta+1}\int_{-\infty}^{\infty}\left|u\right|^{\delta l}K\left(u\right)du\\
&\leq Mh^{-\delta+1}
\end{aligned}$$

where the first inequality follows from Assumption 2(i), the second inequality follows from Assumption 2(iv) and the last inequality similarly follows from Assumption 2(iv). Substituting (9) in (8) leads to, using the same argument as in Mazry and Fan (1997, p.172), i.e. taking $h^{1-2/\delta}d_t^a = 1$ such that $hd_t^\alpha \to 0$ holds,

$$h^{-1}\sum_{j=d_T+1}^{T-J}\left|\text{Cov}\left(\xi_1^l,\xi_{1+j}^l\right)\right| \leq h8M^{2/\delta}h^{-2+2/\delta}\sum_{j=d_T+1}^{T-J}\alpha_j^{1-2/\delta} \leq 8M^{2/\delta}h^{-1+2/\delta}d_t^{-a}\sum_{j=d_T+1}^{T-J}l^a\alpha_j^{1-2/\delta} = o(1). \tag{10}$$

Together, (6), (7) and (10) implies that

$$Th\,\text{Var}\left[(Th)^{-1}\sum_{t=2}^{T-J}\xi_t^l\right] = f(c)\int_0^\infty u^l K^2(u)\,du + o_h(1). \tag{11}$$

By (3) and (11) it follows that

$$(Th)^{-1}H_T Z'WZH_T \to_p \Gamma$$

and that

$$\begin{aligned}\sqrt{Th}\,\text{vec}\left(H_T\left(\hat{\Pi}'-\Pi'\right)\right) &= \left(I_J\otimes Th\left(H_TZ'WZH_T\right)^{-1}(Th)^{-1/2}H_TZ'W\right)\text{vec}\left(Y-Z\Pi'\right)\\ &= \left(I_J\otimes\left(f(c)\Gamma\right)^{-1}\right)\left(I_J\otimes(Th)^{-1/2}H_TZ'W\right)\text{vec}\left(Y-Z\Pi'\right)+o_p(1).\end{aligned}$$

The term $\left(I_J\otimes T^{-1/2}H_TZ'W\right)\text{vec}\,Y$ is analyzed next. Note that

$$\left(I\otimes(Th)^{-1/2}H_TZ'W\right)\text{vec}\left(Y-Z\Pi'\right) = (Th)^{-1/2}\sum_{t=1}^{T-J}\begin{bmatrix} H_TZ_t\left(y_t-Z_t'\pi_0\right)K\left(\frac{X_t-c}{h}\right)\\ \vdots\\ H_TZ_t\left(y_{t+J}-Z_t'\pi_J\right)K\left(\frac{X_t-c}{h}\right)\end{bmatrix}$$

where π_j is the j-th column of Π'. As in HTV and Assumption 4 let $m_j(x) = E\left[y_{t+j}|X_t=x\right]$, $m_j^+(x_0) = \lim_{x\to x_0^+}E\left[y_{t+j}|X_t=x\right]$ and $m_j^-(x_0) = \lim_{x\to x_0^-}E\left[y_{t+j}|X_t=x\right]$. It follows that $E\left(y_{t+j}|X_t\right) = m_j^-(X_t)(1-D_t)+m_j^+(X_t)D_t = m_j^-(X_t)+\left(m_j^+(X_t)-m_j^-(X_t)\right)D_t$. This implies that

$$E\left[y_{t+j}|X_t\right]-Z_t'\pi_0 = m_j^-(X_t)-a_j-b_j(X_t-c)+\left(m_j^+(X_t)-m_j^-(X_t)-\theta_j-\gamma_j(X_t-c)\right)D_t.$$

Following HTV, define

$$\begin{aligned}\zeta_j(x) &= m_j^-(x) - a_j - b_j(x-c) + \left(m_j^+(x) - m_j^-(x) - \theta_j - \gamma_j(x-c)\right) 1(x>c) \\ &\quad -1/2 m''^-(x-c)^2 - 1/2\left(m_j''^+(x) - m_j''^-(x)\right)(x-c)^2 1(x>c).\end{aligned}$$

It follows that

$$\begin{aligned}E[y_t - z_t\pi_0] &= E[E[y_{t+j}|X_t] - z_t\pi_0] \\ &= E\left[\left(m_j^-(X_t) + \left(m_j^+(X_t) - m_j^-(X_t)\right)D_t - z_t\pi_0\right)\right] \\ &= E\left[1/2\left(m''^-(x-c)^2 + \left(m_j''^+(x) - m_j''^-(x)\right)\right)(x-c)^2 1(x>c) + \zeta_j(x)\right]\end{aligned}$$

where $\sup_{x\in[c-h,c+h]}\left|\zeta_j(x)\right| = o\left(h^2\right)$. We then have

$$\begin{aligned}& E\left[(Th)^{-1}\sum_{t=2T}^{T-J} H_T Z_t (y_t - z_t\pi_0) K\left(\frac{X_t-c}{h}\right)\right] \\ =\; & \frac{1}{2}E\left[(Th)^{-1}\sum_{t=2T}^{T-J} H_T Z_t m''^-(X_t)(X_t-c)^2 K\left(\frac{X_t-c}{h}\right)\right] \\ & +\frac{1}{2}E\left[(Th)^{-1}\sum_{t=2T}^{T-J} H_T Z_t\left(\left(m_j''^+(X_t) - m_j''^-(X_t)\right)1(X_t>c)\right)(X_t-c)^2 K\left(\frac{X_t-c}{h}\right)\right] \\ & +\frac{1}{2}E\left[(Th)^{-1}\sum_{t=2T}^{T-J} H_T Z_t \zeta_j(x) K\left(\frac{X_t-c}{h}\right)\right]\end{aligned}$$

where the elements in $H_T Z_t$ are $((X_t-c)/h)^l D_t^k$ for $k,l\in\{0,1\}$. It follows by similar arguments as in HTV that

$$\begin{aligned}& h^{-1}E\left[\left(\frac{X_t-c}{h}\right)^l D_t^k\left\{1/2m''^-(X_t) + 1/2\left(m_j''^+(X_t) - m_j''^-(X_t)\right)1(X_t>c)\right\}(X_t-c)^2 K\left(\frac{X_t-c}{h}\right)\right] \\ =\; & 1/2\left(m_j''^-(c)\Lambda_{lk}^- + m_j''^+(c)\Lambda_{lk}^+\right)f(c)h^2 + o_h\left(h^2\right).\end{aligned}$$

For example, by a change of variables, one obtains

$$\begin{aligned}
& h^{-1}E\left[\left(\frac{X_t-c}{h}\right)^l D_t^k m''^{-}(X_t)\left(1-1\left(X_t>c\right)\right)\left(X_t-c\right)^2 K\left(\frac{X_t-c}{h}\right)\right] \\
= & h^2\int u^{2+l} m''^{-}(uh+c)\,1(uh>0)^k\,1(uh\le 0)\,K(u)\,f(uh+c)\,du \\
= & h^2 m''^{-}(c)\,f(c)\,1\{k=0\}\int_{-\infty}^{0} u^{2+l}K(u)\,du+o_h\left(h^2\right)=m_j''^{-}(c)\,f(c)\,\Lambda_{lk}^{-}h^2+o\left(h^2\right).
\end{aligned}$$

Similarly,

$$\begin{aligned}
& h^{-1}E\left[\left(\frac{X_t-c}{h}\right)^l D_t^k m''^{+}(X_t)\,1\left(X_t>c\right)\left(X_t-c\right)^2 K\left(\frac{X_t-c}{h}\right)\right] \\
= & h^2\int u^{2+l} m''^{+}(uh+c)\,1(uh>0)\,K(u)\,f(uh+c)\,du \\
= & h^2 m''^{+}(c)\,f(c)\int_{0}^{\infty} u^{2+l}K(u)\,du+o_h\left(h^2\right)=m_j''^{+}(c)\,f(c)\,\Lambda_{lk}^{+}h^2+o\left(h^2\right).
\end{aligned}$$

The asymptotic bias of the score can now be written as

$$E\left[(Th)^{-1}\sum_{t=2T}^{T-J}H_T Z_t\left(y_{t+j}-z_t\pi_0\right)K\left(\frac{X_t-c}{h}\right)\right]=\frac{1}{2}\left(m_j''^{-}(c)\,\Lambda^{-}+m_j''^{+}(c)\,\Lambda^{+}\right)f(c)\,h^2+o_h\left(h^2\right).$$

Furthermore,

$$E\left[\left(I\otimes(Th)^{-1}H_T Z'W\right)\operatorname{vec}\left(Y-Z\Pi'\right)\right]=\frac{1}{2}\left(m''^{-}(c)\otimes\Lambda^{-}+m''^{+}(c)\otimes\Lambda^{+}\right)f(c)\,h^2+o_h\left(h^2\right)$$

where $m''^{-}(c)=\left(m_0''^{-}(c),....,m_J''^{-}(c)\right)'$ and similarly for $m''^{+}(c)$.

Let $m_j(x)=E\left[y_{t+J}|X_t=x\right]$. Next consider, as Masry and Fan (1997, p. 169),

$$Q_T=(Th)^{-1/2}\sum_{t=1}^{T-J}\begin{bmatrix} H_T Z_t\left(y_t-m_0(X_t)\right)K\left(\frac{X_t-c}{h}\right) \\ \vdots \\ H_T Z_t\left(y_{t+J}-m_J(X_t)\right)K\left(\frac{X_t-c}{h}\right)\end{bmatrix}$$

where the sum is over typical elements of the form

$$\xi_{t,lk}^{j}=\left(\frac{X_t-c}{h}\right)^l D_t^k\left(y_{t+j}-m_j(X_t)\right)h^{-1}K\left(\frac{X_t-c}{h}\right)\text{ for } l,k\in\{0,1\}$$

Fix a vector $\lambda \in \mathbb{R}^{4J+4}$ with $\|\lambda\| = 1$. Then,

$$\lambda' Q_T = (T/h)^{-1/2} \sum_{t=1}^{T-J} \sum_{j=0}^{J} \sum_{l,k=0}^{1} \lambda_{lk}^j \xi_{t,lk}^j \equiv (T/h)^{-1/2} \sum_{t=1}^{T-J} \xi_t$$

where $\xi_t = \sum_{j=0}^{J} \sum_{l,k=0}^{1} \lambda_{lk}^j \xi_{t,lk}^j$. Let $\lambda_{lk} = \left(\lambda_{lk}^0, ..., \lambda_{lk}^J\right)$, $\lambda^j = \left(\lambda_{00}^j, \lambda_{10}^j, \lambda_{01}^j, \lambda_{11}^j\right)'$, $\lambda = \left(\lambda^{0\prime}, ..., \lambda^{J\prime}\right)'$ and $m(X_t) = (m_0(X_t), ..., m_J(X_t))'$. First consider

$$\begin{aligned} h \operatorname{Var}(\xi_t) &= h \operatorname{Var}\left(\sum_{l,k=0}^{1} \lambda_{lk}' (Y_t - m(X_t)) \left(\frac{X_t - c}{h}\right)^l D_t^k h^{-1} K\left(\frac{X_t - c}{h}\right)\right) \\ &= h \operatorname{Var}\left(\sum_{l=0}^{1} \left(\frac{X_t - c}{h}\right)^l h^{-1} K\left(\frac{X_t - c}{h}\right) \left(\lambda_{l0}' (Y_t - m(X_t)) (1 - D_t) + (\lambda_{l0} + \lambda_{l1})' (Y_t - m(X_t)) D_t\right)\right) \end{aligned}$$

and noting that the two components associated with $(1 - D_t)$ and D_t are orthogonal leads to

$$\begin{aligned} h \operatorname{Var}(\xi_t) &= \sum_{l_1,l_2=0}^{1} E\left[\left(\lambda_{l_10}' \Sigma(X_t) \lambda_{l_20} (1 - D_t)\right) \left(\frac{X_t - c}{h}\right)^{l_1+l_2} h^{-1} K^2\left(\frac{X_t - c}{h}\right)\right] \\ &\quad + \sum_{l_1,l_2=0}^{1} E\left[(\lambda_{l_10} + \lambda_{l_11})' \Sigma(X_t) (\lambda_{l_20} + \lambda_{l_21}) D_t \left(\frac{X_t - c}{h}\right)^{l_1+l_2} h^{-1} K^2\left(\frac{X_t - c}{h}\right)\right] \\ &= f(c) \sum_{l_1,l_2=0}^{1} \left(\lambda_{l_10}' \Sigma^-(c) \lambda_{l_20} \int_{-\infty}^{0} u^{l_1+l_2} K^2(u)\, du\right) \\ &\quad + f(c) \sum_{l_1,l_2=0}^{1} (\lambda_{l_10} + \lambda_{l_11})' \Sigma^+(c) (\lambda_{l_20} + \lambda_{l_21}) \int_{0}^{\infty} u^{l_1+l_2} K^2(u)\, du + o_h(1). \end{aligned}$$

It also follows by the same arguments as in (7) and (10) that

$$h \sum_{j=1}^{T-J} \left|\operatorname{Cov}\left(\xi_1, \xi_{1+j}\right)\right| = o(1).$$

It then follows

$$Th \operatorname{Var}\left(\lambda' Q_T\right) \to f(c) \sum_{l_1,l_2=0}^{1} \left(\lambda_{l_10}' \Sigma^-(c) \lambda_{l_20} v_{l_1+l_2}^- + (\lambda_{l_10} + \lambda_{l_11})' \Sigma^+(c) (\lambda_{l_20} + \lambda_{l_21}) v_{l_1+l_2}^+\right).$$

By uusing the definitions of Ω^- and Ω^+ the result can be expressed in matrix form as

$$Th\,\mathrm{Var}\left(Q_T\right) \to f\left(c\right)\left(\Sigma^+\left(c\right)\otimes\Omega^+ + \Sigma^-\left(c\right)\otimes\Omega^-\right).$$

By the same arguments as in the proof of Theorem 3 of Masry and Fan (1997) it can be shown that

$$\sqrt{Th}\left(\lambda' Q_T\right) \to_d N\left(0, f\left(c\right)\lambda'\left(\Sigma^+\left(c\right)\otimes\Omega^+ + \Sigma^-\left(c\right)\otimes\Omega^-\right)\lambda\right)$$

which in turn implies by the Cramer-Wold theorem and the continuous mapping theorem that

$$\sqrt{Th}\left(\left(I\otimes f\left(c\right)^{-1}\Gamma^{-1}\right)Q_T\right) \to_d N\left(0, f\left(c\right)^{-1}\left(\Sigma^+\left(c\right)\otimes\Omega^+ + \Sigma^-\left(c\right)\otimes\Omega^-\right)\right).$$

Finally, deduce that

$$\sqrt{Th}\left(H_T\,\mathrm{vec}\left(\hat{\Pi}' - \Pi'\right) - \frac{1}{2}\left(m''^-\left(c\right)\otimes\Lambda^- + m''^+\left(c\right)\otimes\Lambda^+\right)f\left(c\right)\right)$$
$$\to_d N\left(0, f\left(c\right)^{-1}\left(\Sigma^+\left(c\right)\otimes\Omega^+ + \Sigma^-\left(c\right)\otimes\Omega^-\right)\right).$$

■

Proof of Theorem 7. The result follows directly from Theorem 7, the fact that $\Pi = (a, b, \theta, \gamma)$ and that the limting distribution in Theorem 7 is for the parameter vector $\mathrm{vec}\left(\Pi'\right)$. ■

Proof of Theorem 8. The result follows directly from Theorem 7. ■

References

DOUKHAN, P. (1994): *Mixing Properties and Examples*. Springer Verlag, New York.

FAN, J. (1992): "Design Adaptive Nonparametric Regression," *Journal of the American Statistical Association*, 87, 998–1004.

HAHN, J., P. TODD, AND W. V. DER KLAAUW (1999): "Evaluating the Effect of an Antidiscrimination Law Using a Regression-Discontinuity Design," Working Paper 7131, NBER Working Paper 7131.

HAHN, J., P. TODD, AND W. VAN DER KLAAUW (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69(1), 201–09.

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

IMBENS, G. W., AND T. LEMIEUX (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142, 615–635.

MASRY, E., AND J. FAN (1997): "Local Polynomial Estimation of Regression Functions for Mixing Processes," *Scandinavian Journal of Statistics*, 24, 165–179.

PORTER, J. (2003): "Estimation in the Regression Discontinuity Model," .

ROYDEN, H. (1988): *Real Analysis*. Macmillan Publishing Company.

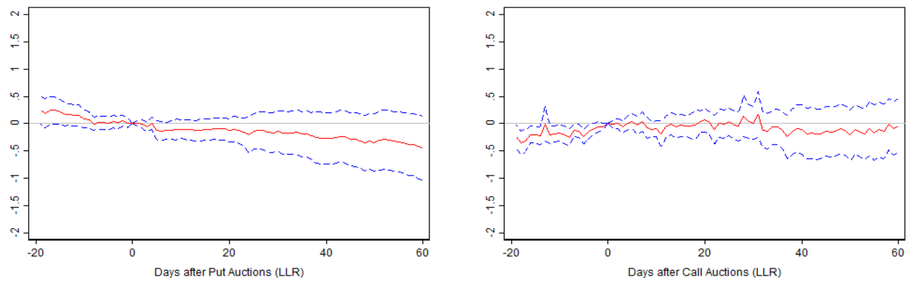
RUBIN, D. B. (1974): "Estimating Causal Effects of Treatments in Randomized and Non-Randomized Studies," *Journal of Educational Psychology*, pp. 688–701.

APPENDIX FIGURES

Figure A-1: Effects on Domestic Interest Rates

(a) Puts - Full Sample

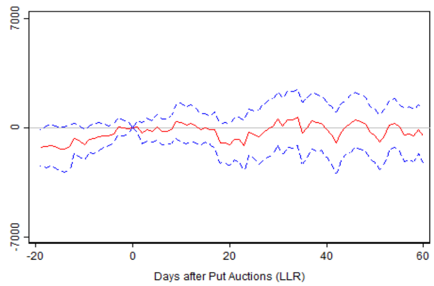
(b) Calls - Full Sample



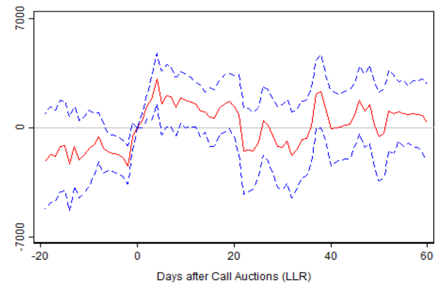
The dependent variable is the interbank interest rate, Colombia's analogue of the federal funds rate in the US, relative to its level at the time of the auction. The solid curve presents a series of regression discontinuity estimates implemented using local linear regression of the dependent variable on the running variable using daily data. Dashed curves display 95% confidence intervals of the estimates.

Figure A-2: 60 day IRFs of Sterilization Mechanisms

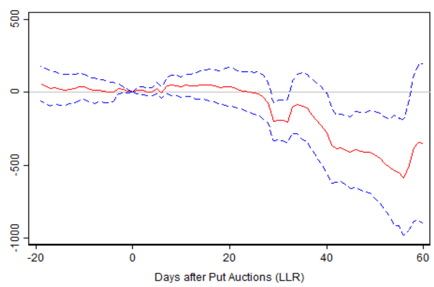
(a) REPOS Full Sample - Put



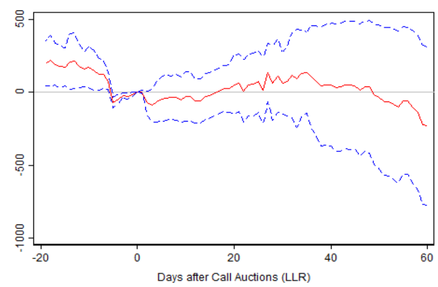
(b) REPOS Full Sample - Call



(c) Col Treasuries Full Sample - Put



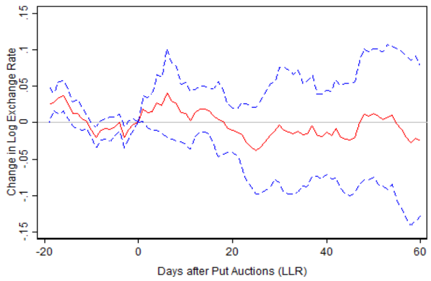
(d) Col Treasuries Full Sample - Call



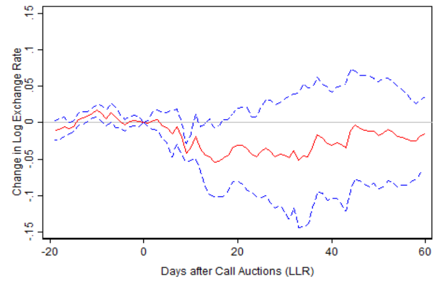
Panes (a) and (b) show the effects of repurchasing agreements, while Panes (c) and (d) show the effects of domestic sovereign bonds. All variables measured in billion (10^9) pesos.

Figure A-3: 60 Day Exchange Rate Response

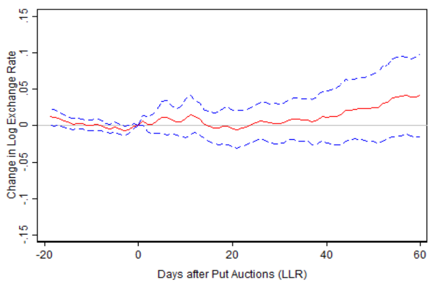
(a) CIP Does Not Hold - Put



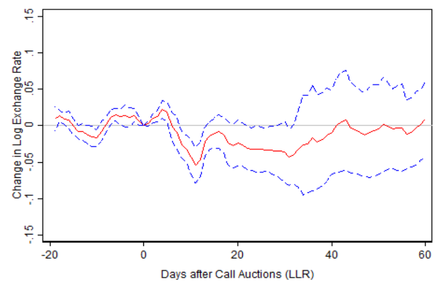
(b) CIP Does Not Hold - Call



(c) CIP Holds - Put



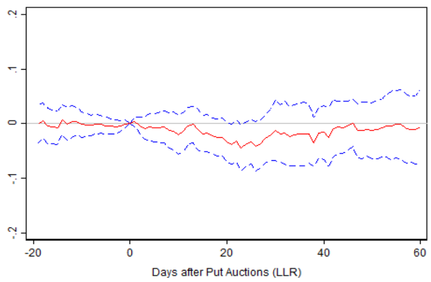
(d) CIP Holds - Call



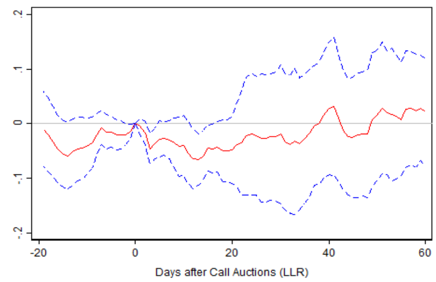
The dependent variable is the log daily average spot rate relative to the log average spot rate the day before the auction. The solid curve presents a series of regression discontinuity estimates implemented using local linear regression on daily data. Dashed curves display 95% confidence intervals of the estimates. We define the time CIP holds as July 1, 2003 to June 30, 2008 and the time CIP does not hold as all dates before and after this time period.

Figure A-4: Ratio of Dollars-Denominated to Peso-Denominated Assets, Top Five Banks, All But Forwards

(a) Full Sample - Put



(b) Full Sample - Call

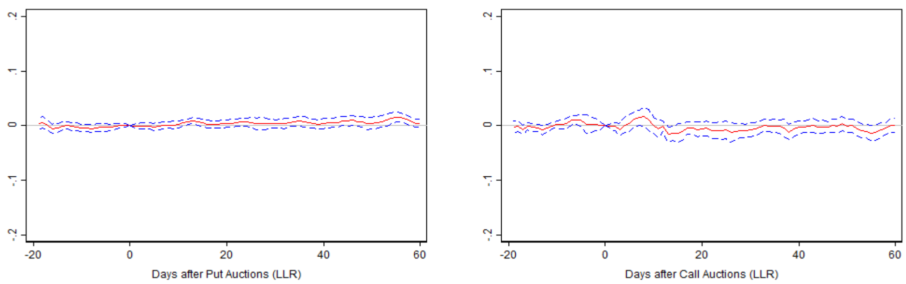


The dependent variable is the ratio of net dollar denominated assets to peso-denominated government bonds held by Colombia's 5 largest banks, relative to its value at the time of the auction. This variable does not include derivatives, i.e. forward contracts. The solid curve presents a series of regression discontinuity estimates implemented using local linear regression on daily data. Dashed curves display 95% confidence intervals of the estimates.

Figure A-5: Ratio of Dollars-Denominated to Peso-Denominated Assets, Top Five Banks, Including Forwards

(a) Full Sample - Put

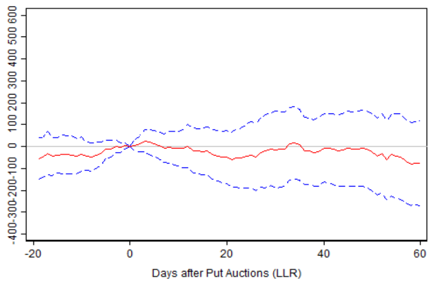
(b) Full Sample - Call



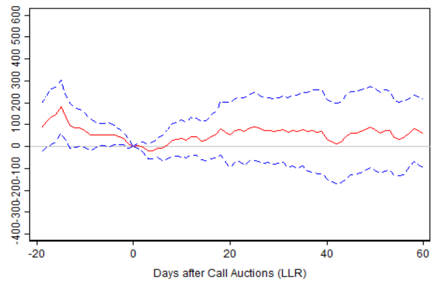
The dependent variable is the ratio of net dollar denominated assets to peso-denominated government bonds held by Colombia's 5 largest banks, relative to its value at the time of the auction. This variable includes derivatives, i.e. forward contracts. The solid curve presents a series of regression discontinuity estimates implemented using local linear regression on daily data. Dashed curves display 95% confidence intervals of the estimates.

Figure A-6: Effects on Net Capital Flows - Domestic Investors

(a) Puts: Net Flows - Full Sample



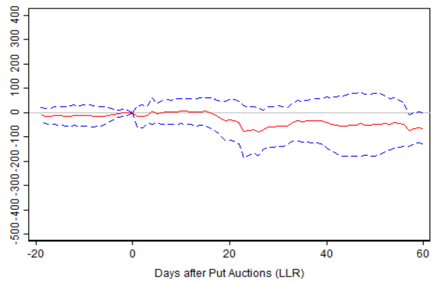
(b) Calls: Net Flows - Full Sample



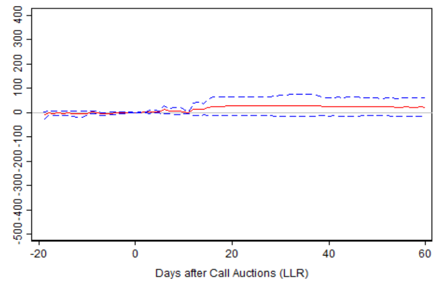
The dependent variable is the net capital inflows by domestic investors in the form of loans since the auction, in millions of dollars. The solid curve presents a series of regression discontinuity estimates implemented using local linear regression on daily data. Dashed curves display 95% confidence intervals of the estimates. We define the time CIP holds as July 1, 2003 to June 30, 2008 and the time CIP does not hold as all dates before and after this time period.

Figure A-7: Effects on Net Capital Flows - Foreign Investors

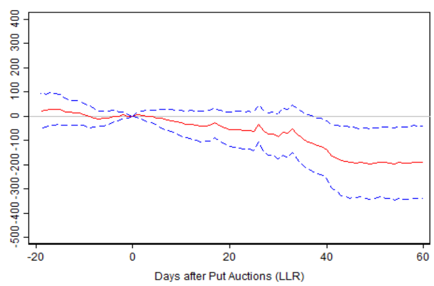
(a) Puts: Net flows - CIP Does Not Hold



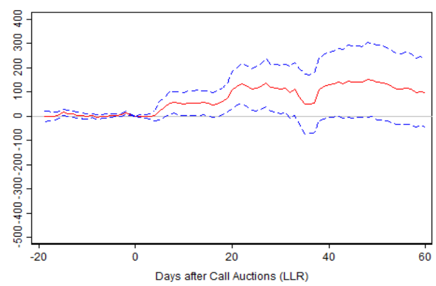
(b) Calls: Net Flows - CIP Does Not Hold



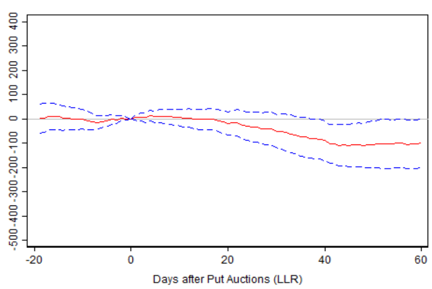
(c) Puts: Net flows - CIP Holds



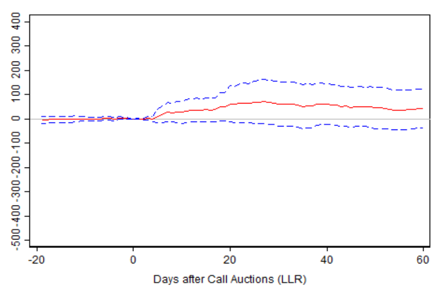
(d) Calls: Net Flows - CIP Holds



(e) Puts: Net Flows - Full Sample



(f) Calls: Net Flows - Full Sample



The dependent variable is cumulative capital inflows by foreign investors in the form of loans since the auction, in millions of dollars. The solid curve presents a series of regression discontinuity estimates implemented using local linear regression on daily data. Dashed curves display 95% confidence intervals of the estimates. We define the time CIP holds as July 1, 2003 to June 30, 2008 and the time CIP does not hold as all dates before and after this time period.


BANCO DE LA REPÚBLICA
COLOMBIA