

# ECON2390 Fall 2024

## Applied Econometrics

### Problem Set 2

Adrien Foutelet

October 2024

## 1 Regression Discontinuity Design

### 1.1

We run OLS to estimate the following model:

$$Y_m = \beta_0 + \beta_1 \text{PANwin}_m + \epsilon_m$$

We get the following estimates (see table 1).

Table 1: Simple OLS regression model

	Y
(mean) PANwin	0.904 (0.06)
Constant	43.63*** (4.58)
Observations	152

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

The positive correlation between variables Y and PANwin that the coefficient suggests does not appear to be significant.

We may be facing endogeneity: it is where the homicide rate is high and the drug trafficking organisations have

strong presence that the party promoting radical anti drug policies is elected (reverse causality); PAN may have prioritized campaigning in dangerous areas where their agenda gave them the best chances to win (selection); where the local culture is more martial, making PAN most popular, results from weaker state capacity which is suitable for criminal organisations to thrive (omitted variable bias). We need to isolate exogenous variations in PANwin to infer causal results.

For these reasons,  $\beta_1$  should be positively biased and  $\beta_0$  should be negatively biased.

## 1.2

The key assumptions necessary for the fuzzy regression discontinuity design to be valid in Dell's analysis are:

- Local monotonicity.

There exists a neighborhood  $G$  around  $X = c$  where any individual in the population has one of the following treatment selection responses:

- $T = t_c \implies D_i(x) = 1_{\{x \geq c\}}$
- $T = t_{at} \implies D_i(x) = 1$
- $T = t_{at} \implies D_i(x) = 0.$

- Continuity:

$\forall t \in \{t_c, t_{at}, t_{nt}\}, E(Y(1) | T = t, X = x), E(Y(0) | T = t, X = x), Pr(T = t | X = x)$  are continuous at  $x = c$ .

- First stage

$$\lim_{x \rightarrow c^-} Pr(D = 1 | X = x) \neq \lim_{x \rightarrow c^+} Pr(D = 1 | X = x)$$

From the text (need to reconcile that):

- All relevant factors besides treatment vary smoothly at the threshold between a PAN victory and loss.
- Absence of selective sorting around the PAN win-loss threshold.

## 1.3

Investigating whether the regression discontinuity design shows evidence of manipulation, we do a McCrary test for the following null hypothesis:

$$H_0 : \text{"The density of the running variable is continuous at the cutoff"}$$

We get the following results (see table 2):

Table 2: McCrary density test

	margin_victory
LogDensityRatio	-0.0763
StdError	24.19
PValue	0.939
Bandwidth_left	0.0135
Bandwidth_right	0.0157

Our high p-value implies that we fail to reject the null hypothesis, meaning that we find no evidence of significant difference in the density of the running variable on either side of the cutoff.

We get the following plot of the estimated density of the running variable on both sides of the cutoff (see figure 1). No visual discontinuity can be found in the density at the cutoff: the density is smooth.

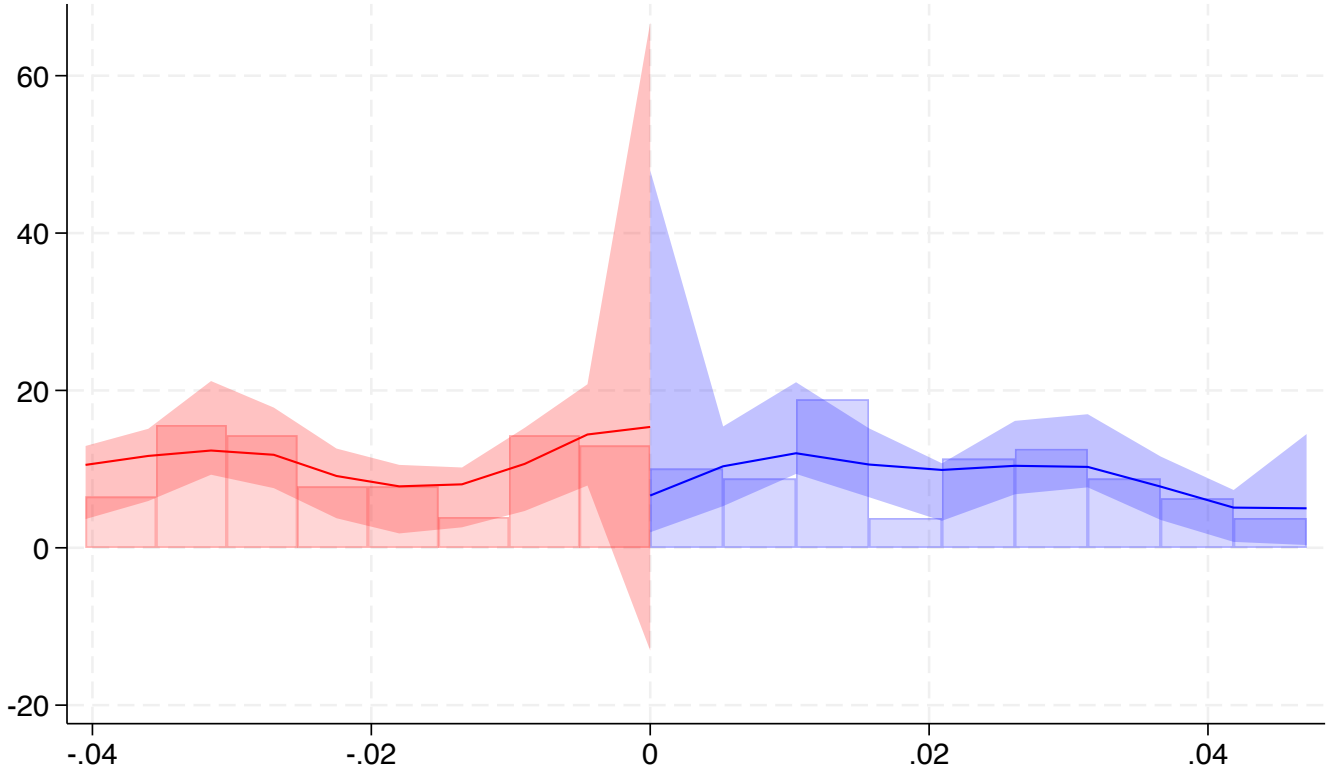


Figure 1: McCrary test result plot

All in all, there is no evidence of manipulation or sorting of the running variable at the cutoff. This supports the validity of the RDD design.

## 1.4

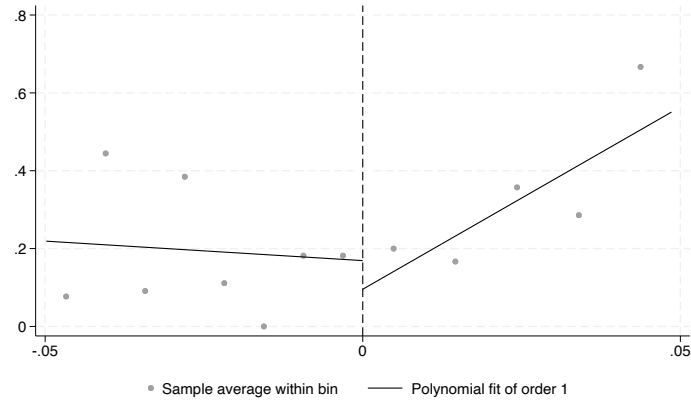


Figure 2: Binned scatter plot of PAN victory margin and homicide rate (number of bins selected via mimicking variance evenly-spaced method using spacings estimators, linear specification)

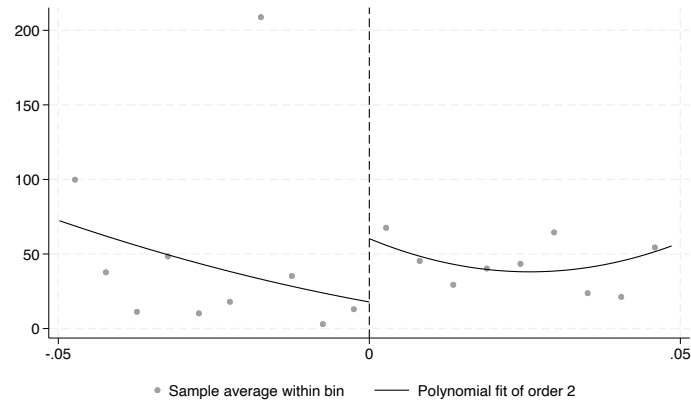


Figure 3: Binned scatter plot of PAN victory margin and homicide rate (number of bins selected via mimicking variance evenly-spaced method using spacings estimators, quadratic specification)

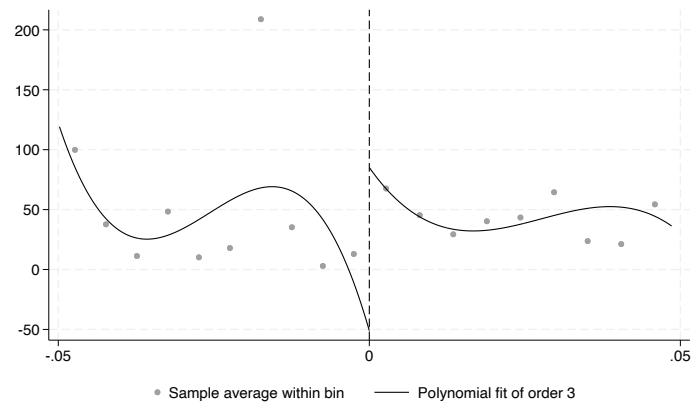


Figure 4: Binned scatter plot of PAN victory margin and homicide rate (number of bins selected via mimicking variance evenly-spaced method using spacings estimators, cubic specification)

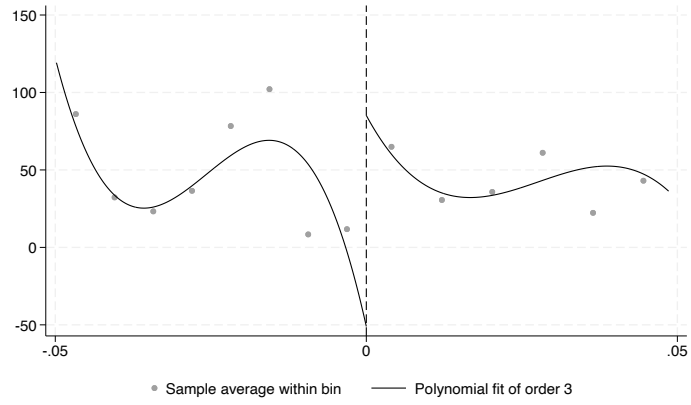


Figure 5: Binned scatter plot of PAN victory margin and homicide rate (number of bins selected via IMSE-optimal evenly-spaced method using spacings estimators, cubic specification)

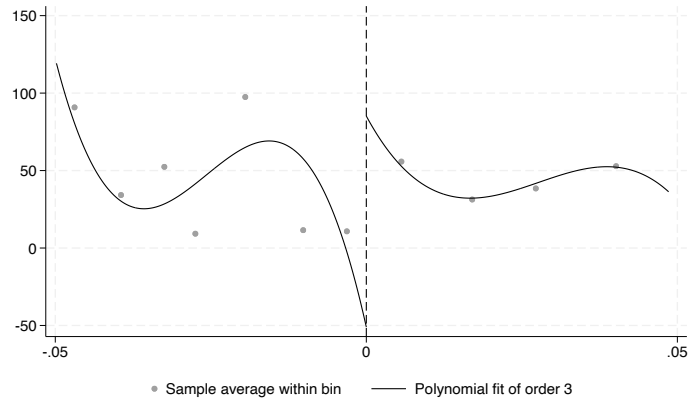


Figure 6: Binned scatter plot of PAN victory margin and homicide rate (number of bins selected via IMSE-optimal quantile-spaced method using spacings estimators, cubic specification)

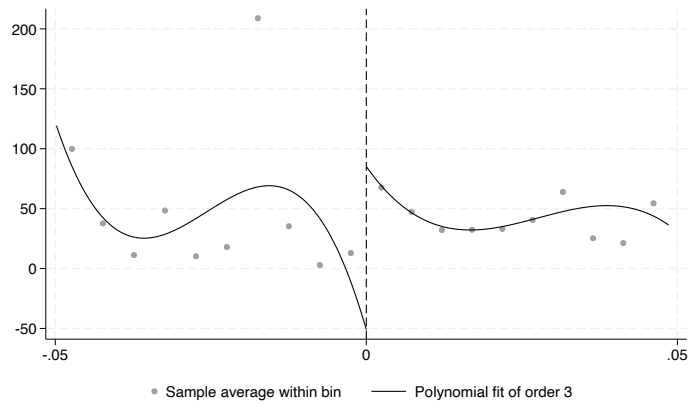


Figure 7: Binned scatter plot of PAN victory margin and homicide rate (number of bins normatively set at 10, cubic specification)

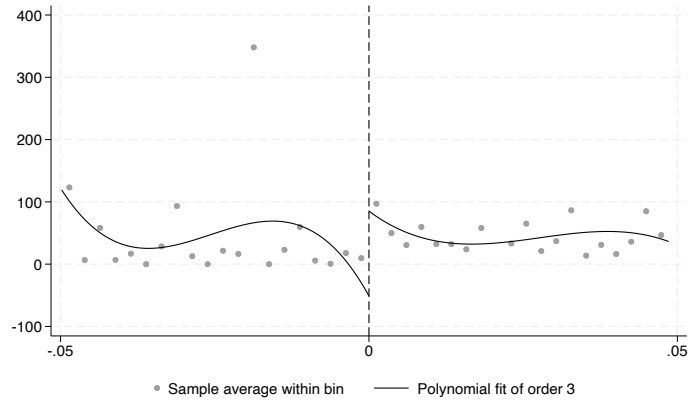


Figure 8: Binned scatter plot of PAN victory margin and homicide rate (number of bins bormatively set at 20, cubic specification)

The discontinuity at the cutoff is robust to the three polynomial specifications.

The smaller the number of bins, the larger the discontinuity.

## 1.5

We get the following estimates (see table 3).

## 1.6

We get the following estimates (see table 4).

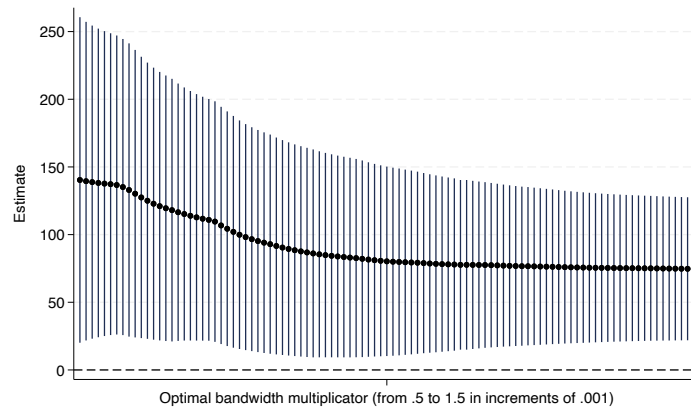


Figure 9: RDD model robustness

Table 3: Non parametric and parametric RDD models

	Para tri lin	Para tri quad	Non-para lin	Non-para quad
RD_Estimate	80.25*	93.93		
	(2.25)	(1.92)		
(mean) spread			979.5	2697.6
			(0.36)	(0.39)
(mean) PANwin			69.99**	81.79*
			(2.91)	(2.57)
(mean) PANwin=1 $\times$ (mean) spread			-5770.9	-12951.7
			(-1.50)	(-1.31)
(mean) spread squared				260133.0
				(0.54)
(mean) PANwin=1 $\times$ (mean) spread squared				138247.6
				(0.21)
Constant			12.95	14.37
			(0.78)	(0.66)
Observations	152	152	38	54

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table 4: RDD model robustness

	Para tri lin	Para tri lin 5%	Para tri lin 10%	Para tri lin 20%
RD_Estimate	80.25*	80.25*	98.36*	48.95
	(2.25)	(2.25)	(2.14)	(1.00)
Observations	152	152	151	144

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

## 1.7

John Marshall's concern about compensating differentials in Politician Characteristic Regression Discontinuity (PCRD) designs revolves around the idea that these designs fail to isolate the pure effect of a politician's characteristic (e.g., gender, party affiliation) because other characteristics, known as compensating differentials, influence both the election outcome and downstream effects.

For example, if voters are biased against women, female candidates in close elections may need to compensate by having higher levels of competence or qualifications than their male opponents. The characteristic of interest (e.g., gender) correlates with compensating differentials that offset its impact (e.g., competence).

This means that PCRD designs compare candidates who are not equivalent on other key traits, leading to biased estimates of the effect of the characteristic of interest, since the effect measured is a compound treatment effect of the characteristic of interest and the compensation differentials.

Marshall recommends to attempt to measure compensating differentials to bound or reinterpret the results.

Dell leverages close elections to argue that PAN mayoral victories causally increase drug-related violence. The compensating differentials from winning PAN candidates may be their strong-willed disruptive and uncompromising characters (prone to get drug dealers' backs up), their expected lower level of corruption (prone to increase violence from the mobsters), their stronger connections with federal authorities (prone to make anti crime policies even stronger locally), or their lack of knowledge in sociology of deviance (prone to make them both attractively radical and acting without tact). There would then be selection of these traits for the mayors in the treatment group.

## 1.8

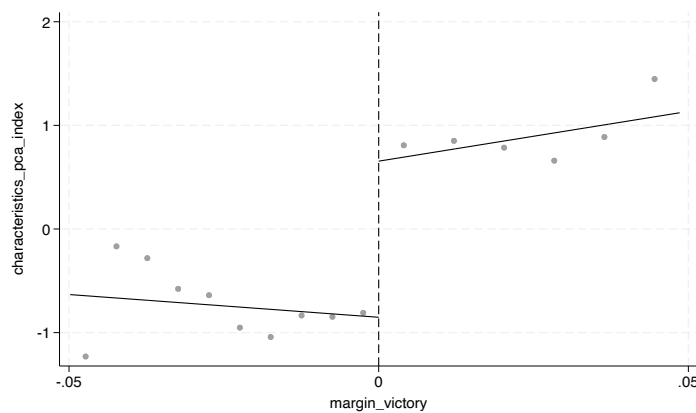


Figure 11: RDD model robustness



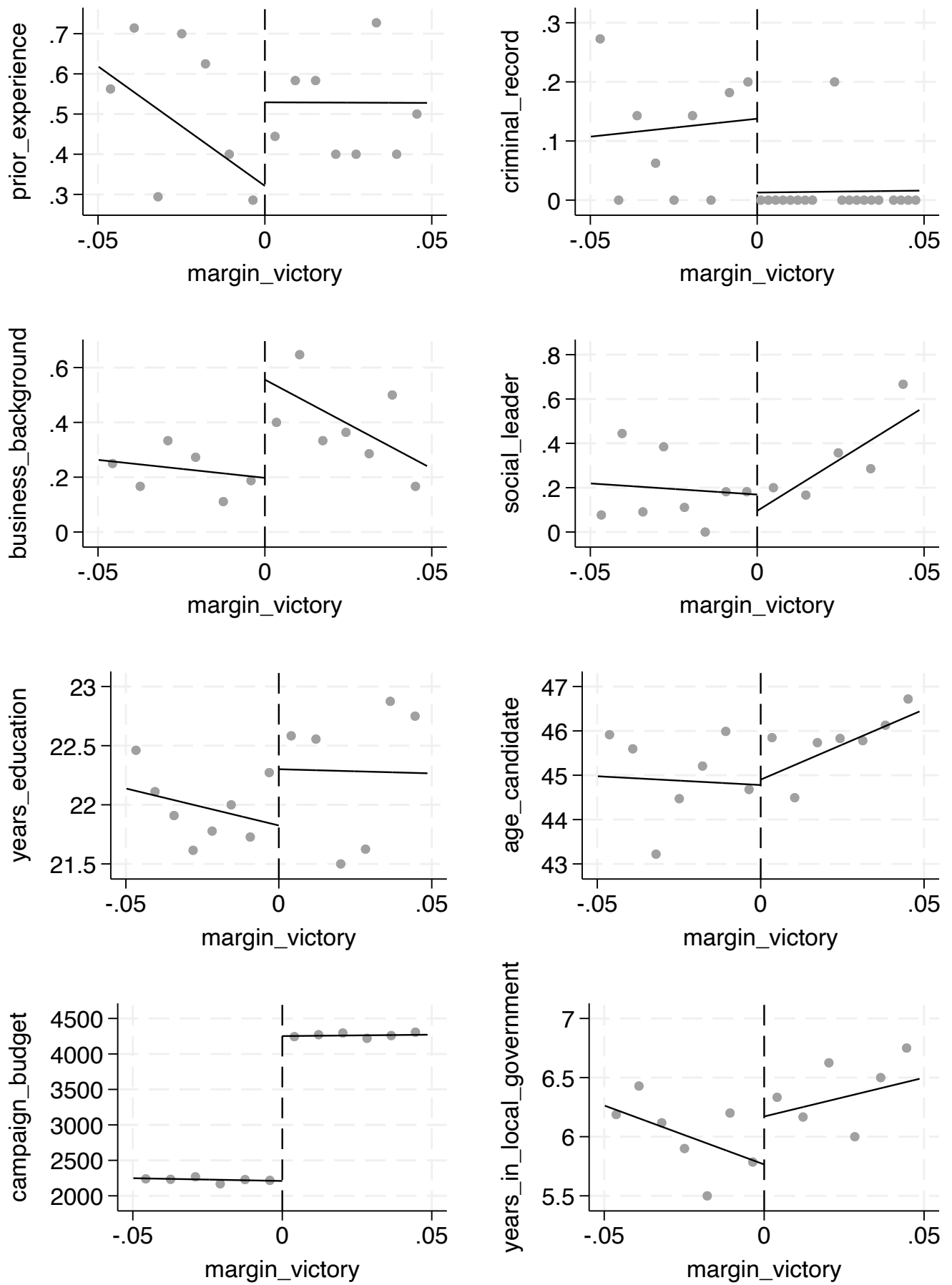


Figure 10: RDD model robustness

Table 5: Candidate characteristic RDD analysis

	pca_index	prior_experience	criminal_record	business_background	social_leader	years_education	age_candidate	campaign_budget	years_in_local_government
RD_Estimate	1.934** (2.84)	0.179 (0.60)	-0.303 (-1.26)	0.0150 (0.05)	0.0274 (0.11)	0.525 (0.60)	1.629 (0.98)	2038.0*** (21.42)	0.673 (0.93)
Observations	152	152	152	152	152	152	152	152	152

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

We get the following estimates (see table 5).

Covariate adjustment in a RDD is not sufficient to address the problem of compensating differentials, because these differentials introduce biases that are inherently linked to both observable and unobservable characteristics. Even if observable covariates balanced or controlled for, unobservable characteristics (e.g., competence) could vary discontinuously at the threshold, biasing the treatment effect. Also, covariates may include variables affected by the treatment itself (e.g., campaign funding postulated by voter - donors - and party organization expectations of winning). Including these covariates introduces post-treatment bias (conditioning on factors influenced by the treatment), contaminating causal estimates.

All in all, covariate adjustment can reduce noise and account for observable imbalances, but it cannot fully address biases introduced by compensating differentials, especially when unobservable factors are at play.

## 1.9

To bound the effect of a PAN mayor on the homicide rate, we implement:

$$\hat{\tau}_{PCRD}^{corr} = \hat{\tau}_{PCRD} - \sum_k \hat{\gamma}_k \hat{\delta}_k$$

With:

- index  $i$  for candidates,
- index  $d$  for district,
- index  $k$  for covariates,
- $\hat{\gamma}_k$  an estimate of the LATE of each compensating differential  $Z_{idk}$ ,
- $\delta_k = E[Z_{idk} \mid \Delta_d = 0, X_{id} = 1, X_{jd} = 0] - E[Z_{idk} \mid \Delta_d = 0, X_{id} = 0, X_{jd} = 1]$ ,
- $\Delta_d = V_{id} - V_{0d}$ ,
- $V_{id}$  and  $V_{0d}$  the vote shares of the most popular candidates of types  $X_{id} = 1$  and  $X_{id} = 0$  respectively,
- $X_{id}$  a binary characteristic that helps candidate  $i$  win votes.

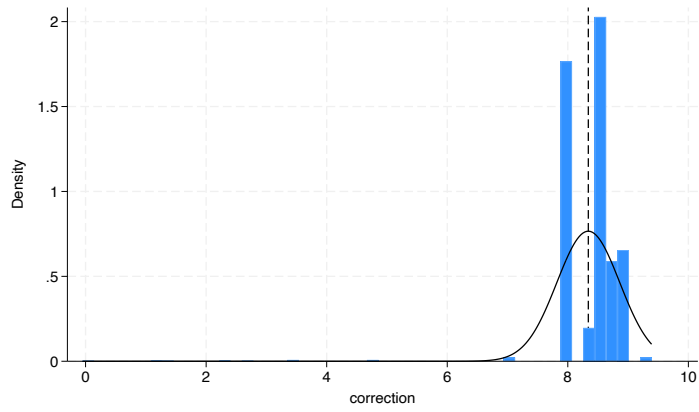


Figure 12: Distribution of the computed correction term

Recalling that  $\hat{\tau}_{PCRD} = 80.25^* (2.25)$ , the sign and the significance level of our estimate are robust to the (observable) effect of compensating differentials.

## 2

### 2.1

Table 6: IV

	2SLS	Wald
D	7.633*** (46.28)	7.633*** (46.02)
x1	0.124* (2.25)	
x2	0.0708 (1.44)	
x1x2	0.282*** (3.84)	
Constant	1.039*** (10.49)	1.208*** (13.17)
Observations	10000	10000

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

The 2SLS and residualized Wald estimators are equal.

## 2.2

A researcher observing  $(Y_i, D_i, Z_i, X_i)$  would identify the conditional-on- $X_i$  LATESs  $E[Y_i(1) - Y_i(0) \mid X_i, G = CP]$  by 2SLS on observations such that  $(X_i = Z_i)$ . This is equivalent to estimating the structural model by OLS on this restricted dataset.

Table 7: IV

	2SLS	Wald	2SLS X=(0,0)	2SLS X=(1,0)	2SLS X=(0,1)	2SLS X=(1,1)
D	7.633*** (46.28)	7.633*** (46.02)	5.993*** (59.52)	7.925*** (88.25)	7.023*** (74.59)	8.930*** (103.58)
x1	0.124* (2.25)					
x2	0.0708 (1.44)					
x1x2	0.282*** (3.84)					
Constant	1.039*** (10.49)	1.208*** (13.17)	-0.00624 (-0.10)	0.0198 (0.48)	-0.00958 (-0.13)	0.125 (1.61)
Observations	10000	10000	417	833	417	833

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

## 2.3

Table 8: Weights and convex weighted average of the conditional LATEs

Beta 2SLS	7.633
Weight00	0.546
Weight10	1.516
Weight01	0.569
Weight11	1.369
LATEsCvxWghtedAvg	7.877

We get  $\hat{\beta}_{2SLS} = 7.8765399$ , which is close to our initial 2SLS estimation.

The group that gets more weight is  $X = (1, 0)$ .

## References

- [1] Dell, Melissa. “Trafficking Networks and the Mexican Drug War.” *American Economic Review* 105, no. 6 (2015): 1738-1779.
- [2] Marshall, J. (2024), Can Close Election Regression Discontinuity Designs Identify Effects of Winning Politician Characteristics?. *American Journal of Political Science*, 68: 494-510.