Breaking the Code: Can a New Penal Procedure Affect Public Safety?[†]

Matias D. Cattaneo[‡] Carlos Díaz[§] Rocío Titiunik[¶]

February 14, 2022

* * * PRELIMINARY DRAFT: PLEASE DO NOT CIRCULATE * * *

Abstract

We explore potential unintended consequences of a new code of penal procedure on criminal behavior. On November 2017, Uruguay switched from an inquisitorial system (where a single judge leads investigation and decides the appropriate punishment for a particular crime) to an adversarial system (judges serve as referees to ensure the correct procedure, whereas investigation is led by prosecutors). Using a regression-discontinuity design for the most common offenses reported in Montevideo, this study finds a local and significant increase in the total number of reports filed when the new adversarial code entered into force. Results show an increase of 21-24 police reports per day, accounting for 23%-26% of the average annual increase. The paper also discusses which conditions are required to use the regression-discontinuity framework in applications where time periods both index the units and function as the score variable.

Keywords: crime, criminal law, criminal procedure, regression discontinuity.

[†]We thank Giorgio Chiovelli, Scott Cunningham, Jennifer Doleac, Libor Dušek, Jeff Grogger, Dorothy Kronick, Emily Owens, Zachary Peskowitz, Yotam Shem-Tov, Gonzalo Vazquez-Bare and participants at the 2021 Chicago/LSE Conference on the Economics of Crime and Justice, the 2021 Annual Meeting of the Latin American and Caribbean Economic Association, the 2021 LACEA/RIDGE Workshop on the Economics of Crime, the 2021 APSA Annual Meeting, the 2020 Annual Conference of the Latin American Society for Political Methodology, the Summer 2020 Online Seminar on the Economics of Crime, and Universidad ORT and Universidad de Montevideo Economics Department Seminar Series for thoughtful comments and suggestions.

[‡]Department of Operations Research and Financial Engineering, Princeton University.

[§]Department of Social Sciences, Catholic University of Uruguay.

 $[\]P {\rm Department}$ of Politics, Princeton University.

1 Introduction

On November 1st, 2017, a new code of penal procedure (CPP) entered into force in Uruguay. The reform implied the end of the old inquisitive and written tradition, to give rise to an accusatory, adversarial, oral and public criminal system. Under the old CPP, the inquisitorial judge led investigation and decided the appropriate punishment for a particular crime. Once the new penal procedure came into effect, judges began to serve as neutral referees to ensure the correct procedure, whereas investigation is now led by prosecutors. In other words, the investigation is exclusive responsibility of prosecutors who, representing society, must carry evidence to judges. Judges will then decide what evidence to admit into the record and what to exclude. This is the main reason why it is often argued that the new adversarial system is more fair and less prone to abuse than the old inquisitorial one.

Despite its advantages, the procedural reform could have triggered some unintended consequences in terms of public safety for Uruguay. In fact, the number of police reports substantially increased immediately after the adversarial criminal procedure code was implemented. It is worth to notice that a criminal procedure neither defines the recognized offenses nor fixes the corresponding penalties. However, Uruguay's new CPP introduced substantial changes to the adjudication process of the criminal law (e.g., plea bargain, alternatives to oral trials, exceptional use of preventive detention) that have been widely used by prosecutors and might result in lighter sentences. Moreover, prosecutors face both a totally new role under the new legal system and a significant increase in their workload. Meanwhile, the new CPP makes police officers conduct investigation under new supervision, different rules and several issues in the implementation process; all significant adjustments in the practice of policing that are expected to alter officer behavior (Mummolo, 2018; Kronick, 2019). These changes resulted in several coordination problems between prosecutors and police officers during the first months of the adversarial system, potentially affecting the probability of detection and conviction. Therefore, the sharp and sudden rise in police reports in November 2017 could have responded to changes in both the severity and the certainty of punishments (Becker, 1968).

We employ a regression-discontinuity (RD) design to identify a local causal effect in the immediate neighborhood of the date the new CPP came into effect. Our evidence suggests that a change in the adjudication process of the criminal law may have unexpectedly decrease public safety. We find evidence of a local causal effect of the implementation of Uruguay's new penal procedure on the number of crimes reported to police in Montevideo. To be more precise, RD point estimates suggest a local increase of 21 to 24 police reports per day due to the migration from an inquisitorial and written system to an accusatory, adversarial, oral and public one. These results are heterogeneous across crimes and robust to different local approximations, alternative kernel functions, data-driven bandwidth selectors, placebo cutoffs and the local randomization approach to RD. More importantly, these estimates are consistent with the view that legal codes are far from innocuous (Acemoglu and Johnson, 2005).

This paper adds directly to the strand of the economics of crime literature documenting the impact of procedural law on delinquent behavior. Most of these empirical studies have put the spotlight on the probability of detection and conviction, reporting results consistent with our main findings (Atkins and Rubin, 2003; Dalla Pellegrina, 2008; Soares and Sviatschi, 2010; Duŝek, 2015; Zorro Medina et al., 2020). For example, Atkins and Rubin (2003) estimate a statistically and economically significant increase in the number of crimes in the United States after the Supreme Court imposed the "exclusionary rule" in the 1960s. This legal rule prevents evidence obtained in violation of the Fourth Amendment from being used in a court of law, increasing the cost of police investigation. Dalla Pellegrina (2008) finds a positive effect of trial duration on crimes in Italy, supporting the hypothesis that a rise in probability of prescription increases the willingness to offend. Similarly, Soares and Sviatschi (2010) show that more efficient courts in Costa Rica are associated with lower crime rates. Duŝek (2015) exploits a criminal procedure reform in the Czech Republic and shows that simplified procedures incentivize police to reallocate enforcement efforts toward crimes that have reduced enforcement costs.

Recent studies for South America report mixed results regarding the unintended consequences on the probability of detection and conviction when an inquisitorial criminal procedure code is replaced by an accusatorial system (Langer, 2007). On the one hand, Zorro Medina et al. (2020) exploit the quasi-experimental implementation of the new criminal procedure code in Colombia and document an increase in both violent and property crimes. In spite of the fact the authors acknowledge that they are not able to identify the actual mechanism behind their estimates, they suggest that the rise in property crimes could respond to a decrease in probabilities of apprehension or detection due to the new procedural regime. As it was discussed above, these results are in line with one of the main predictions of Becker's (1968) canonical model: rational offenders respond to the probability of conviction. On the other hand, Kronick (2019) documents a large effect on arrest rates in Colombia (where the new penal procedure was rolled out in stages) and Venezuela (where the new code was enacted nationwide): 40% and 80%, respectively. Contrary to both Zorro Medina et al.'s (2020) estimates and Becker's (1968) main predictions, none of the countries experienced a jump in crime.

New penal procedures might also deter (encourage) crime by increasing (reducing) the severity of associated punishments. In particular, these reforms have usually introduced plea bargaining—among other trial-avoiding conviction mechanisms (Langer, 2021)—where defendants enter a guilty plea in exchange for potential sentence reductions or changes. Despite some exceptions (Abrams, 2011; Frazier et al., 2018), empirical evidence supports the hypothesis that sentences of defendants who plead guilty are significantly shorter than those of convicted at trial (Bushway and Redlich, 2012; Bushway et al., 2014; Ulmer and Bradley, 2006; Ulmer et al., 2010). Regarding the effects of plea discounts on criminal behavior, Zorro Medina et al.'s (2020) findings suggest that the

increase in the number of convictions under Colombia's new criminal procedure mostly respond to guilty pleas rather than to trial sentences. According to the authors, this mechanism would also explain the observed rise in crime in Colombia. As mentioned above and further discussed later in the paper, our results could also respond to the fact that criminals expect a less severe punishment due to new CPP's trial-avoiding conviction mechanisms.

Not only do new criminal procedures deter crime by affecting both the probability of detection and conviction, and the severity of punishment, but they can also increase or reduce crime control through more or less incapacitation.¹ As Zorro Medina (2020) documents, criminal procedural reforms caused a decrease in the pre-trial detention population in Latin America due to an increase in efficiency rather than higher human rights protection. For example, Zorro Medina et al. (2020) speculate that the reported increase in crime rates in Colombia could respond to stricter rules for requesting and imposing preventive detention. Far from being the exception, the percentage of pre-trial prison population in Uruguay decreased from 70% in 2015 to 20% in 2020 (Institute for Crime & Justice Policy Research, 2021).

The rest of the paper is organized as follows. Section 2 provides institutional background information on the criminal procedure in Uruguay. Section 3 describes the data on the offenses reported to police in Montevideo, whereas Section 4 discusses which conditions are required to use the RD framework in applications where time periods both index the units and function as the running variable. Section 5 presents the main results and conduct several robustness checks. Section 6 discusses the potential mechanisms linking the new CPP to an increase in the actual number of offenses.

¹Additionally, we could consider deterrence due to the swiftness of the criminal procedure: offenders discount future consequences of their behavior, including punishment (Listokin, 2007). See Zorro Medina et al. (2020) for a discussion.

2 Institutional Background

Uruguay has been one of the last countries in Latin America to reform its code of penal procedure. A wave of reforms started in the early 1990s and included most of the countries in the region with the exception of Brazil and Cuba, that still preserve their old systems (Fandiño and González Postigo, 2020). As Langer (2007) argues, this generalized transition towards an accusatorial or adversarial system represents the deepest transformation that Latin American criminal procedures have undergone in the last two centuries.

Latin America's penal process used to be mostly governed by inquisitorial and written systems that were originally adopted in either the 19th century or the beginning of the 20th century. Langer (2007) notices that these systems generally shared two key characteristics that were also present in the Uruguayan case. First, criminal procedures were divided into two written phases: the pretrial investigation phase and the verdict and sentencing phase. Given its written nature, the chief support of this process was a dossier compiled by the judge and the police, including the evidence that would be evaluated during the verdict phase. Second, the pretrial investigation was led by the judge. As Langer (2007) explains, not only did the judge perform an investigatory and prosecutorial role, but also an adjudicatory one. In few words, in the verdict phase a judge evaluated the evidence collected by the same judge during the pretrial phase. Furthermore, investigation was kept in secret from the defendant and pretrial detention was the norm rather than the exception. The increasing recognition of human rights (1970s) and the transition to democracy (1980s and 1990s) fueled the perception among local actors that minimum due process standards were fare from met in Latin America (Maier and Struensee, 2000; Langer, 2007).

With the advent of the accusatory, adversarial, oral and public system in Uruguay and the rest of Latin America, regional criminal process has been resting on three well defined stages: the formalization of the investigation, a preliminary hearing, and a trial phase (Fandiño and González Postigo, 2020). During the first procedural stage, the General Prosecution Office has to decide to undertake criminal prosecution against the defendant or not (i.e., the judicialization of the case). If the investigation was formalized, the second procedural stage consists of a preliminary hearing aimed to discuss the outcomes of the investigation conducted by the police and led by a prosecutor. Finally, there is a trial phase where an oral and public discussion takes places for those cases that could not be solved through an alternative outcome. As Fandiño and González Postigo (2020) notice, Uruguay has a classic adversarial procedure where defense and prosecution arguments are presented at a hearing.²

The implementation of Uruguay's adversarial CPP was nowhere near immediate or simple. In spite of the fact the new law of criminal procedure was passed on December 2014, it was not until November 2017 that the adversarial code came into effect.³ The relevance and magnitude of such a reform made local authorities to take almost three years to bring the new code into force. The adversarial system represented major changes for several of the agents involved in the adjudication process of criminal law, in particular prosecutors and police. As it was explained above, prosecutors play a total new role under the adversarial system, relative to what they used to do under the old inquisitorial CPP. In fact, they are now in charge of the criminal investigation, representing both a change in their tasks and a significant increase in their workload.⁴ Regarding the police, the adversarial systems make officers to conduct criminal investigation under both new rules

²This adversarial proceeding is used in several countries in Latin America, including Argentina, Chile, Ecuador, Mexico and Uruguay. However, there are other two models used in Latin America: the bureaucratic adversarial process (e.g., Costa Rica and Guatemala) and the written adversarial process (e.g., Brazil and Cuba). See Fandiño and González Postigo (2020) for a discussion.

³Law No. 19293, passed on December 19th, 2014.

⁴Since late 2017, the prosecutors' union has expressed opposition to how the new CPP impacted on their workload. For example, some prosecutors have reported to deal with several hundreds cases simultaneously. As a result, not a few prosecutors have taken mental health days due to work-related stress during the first year of the implementation of the new CPP. However, the highest point of tension was reached on May 10th, 2019 when a prosecutor suffered heart attack and died at her office. For a detailed coverage of all these events, please see Solomita (2019).

and new supervision (i.e., a prosecutor instead of the judge assigned to the case).⁵

3 Data

We collect data on the offenses reported to police in Montevideo from the Ministry of Interior of Uruguay.⁶ We first focus on the immediate years before and after the new criminal procedure entered into force. Since the new CPP came into effect on November 1st, 2017, we start by building a two-year symmetric window from November 1st, 2016 trough October 31st, 2018. Table 1 reports the average number of crimes reported daily to police during the defined timeline, as well as the corresponding figures for the years before and after the switch to the new CPP. As it can be noticed, an average of 330 crimes were reported to police every day and just six crime categories account for almost 9 of every 10 reports processed in Montevideo: theft ($\approx 46\%$ of all offenses reported to police), robbery ($\approx 16\%$), domestic violence ($\approx 10\%$), threat ($\approx 7\%$), damage ($\approx 6\%$) and personal injury ($\approx 3\%$).

[INSERT TABLE 1 HERE]

Table 1 illustrates a significant increase in the number of police reports in Montevideo during the first year of the new CPP. When we compare the last year of the old inquisitorial system to the first year of the new adversarial system, the total number offenses reported daily to police increases in 93 incidents (from 284 to 377 per day). Not only is this difference in the number of police reports statistically significant at the 0.01 level, but it is also relevant in magnitude: it represents a 33% increase. This upward trend is also reflected in the six most frequent crimes reported in Montevideo. Regarding property crimes, the number of thefts, robberies and damages reported to police every day increased

⁵The new CPP represented a major change in police culture for several officers ("Policía deberá ingeniársela para adaptarse al nuevo CPP", 2017).

⁶Capital and largest city of Uruguay: 40% of Uruguay's population live in Montevideo.

by 33%, 49% and 43%, respectively. In spite of the fact the reports on the most frequent crimes against the person (i.e., domestic violence, threat and personal injury) exhibited more modest increases (i.e., 5%-15%), the rise is also statistically significant at the 0.01 level. However, these preliminary results are suggestive but not conclusive evidence that Montevideo's crime wave responds to the modifications suffered by Uruguay's criminal process.

When we visually inspect the data, it could be argued that the increase in police reports was strongly associated to the implementation of a new penal procedure. Figure 1 plots the number of times per day these six incidents were reported to police in Montevideo for our complete data set (i.e., January 2014 trough June 2019), highlighting the date the new CPP entered into force. According to the graph, a sudden increase in the number of reports takes place immediately after the new penal procedure was implemented.

[INSERT FIGURE 1 HERE]

When we look at specific types of crime (Figure 2, Plots (a) to (f)), the pattern displayed in Figure 1 seems to be mainly driven by property crimes: theft, robbery and damage. The pattern is not that clear for crimes against the person. However, we will need to be more careful when collecting visual evidence on the discontinuities generated by the implementation of the adversarial system.

[INSERT FIGURE 2 HERE]

4 Research Design

Our goal is to evaluate the effect of Uruguay's new CPP on the number of crimes reported to the police in Montevideo. In our setup, multiple cross-sectional units are observed in multiple time periods, a treatment is introduced at the same time for all units, and inferences are based on a before-and-after comparison. This setup shares the basic features of event study (ES) designs in finance (MacKinlay, 1997) and interrupted time-series (ITS) designs in program evaluation (St. Clair et al., 2016). Common analyses of this kind of design are based on linear panel data models where the outcome for unit i in period t is regressed on a unit's fixed effect, a time fixed effect, time trends, unit-level covariates, and time-indexed treatment indicators (see, for example, Duggan et al., 2016; St. Clair et al., 2014). These estimation strategies are valid under appropriate parametric assumptions.

4.1 A Regression Discontinuity Framework for ES/ITS Designs

Instead of using regression-based models, we apply recent developments in the methodological literature of regression discontinuity (RD) designs to ES/ITS designs. These tools allow us to define a local causal parameter, avoid parametric assumptions, and specify the required assumptions for identification in terms of potential outcomes rather than estimating equations.

The standard RD design is a cross-sectional setting where all units receive a score (also known as the *running variable*), and a treatment is assigned based on whether this score exceeds a known, fixed cutoff: all units whose score is above the cutoff are assigned to the treatment condition, and all units whose score is below it are assigned to the control condition. In the so-called *sharp* RD design, compliance with the treatment assignment is perfect, so that all units above the cutoff receive the treatment and none of the units below the cutoff receive it. In the so-called *fuzzy* RD design, compliance with treatment assignment is imperfect, and there may be treated and untreated units both above and below the cutoff. For a practical introduction to standard RD designs, see Cattaneo et al. (2020).

4.2 A Review of Standard Cross-Sectional RD Setup

We start by briefly reviewing the standard RD setup, which we then modify to accommodate the ITS/ES design. In the standard cross-sectional RD design, there are n units, indexed by i = 1, ..., n, and every unit receives a score S_i . The treatment indicator is denoted by D_i , and is assigned based on the score S_i according to the sharp RD rule

$$D_i = \mathbb{1} (S_i \ge c) = \begin{cases} 0 & \text{if } S_i < c \\ 1 & \text{if } S_i \ge c, \end{cases}$$

where c is the cutoff, and $\mathbb{1}(\cdot)$ is the indicator function.

Each unit *i* is assumed to have two potential outcomes, $Y_i(1)$ and $Y_i(0)$, that capture the two possible outcomes under the treated and untreated conditions, respectively.

The random variable that captures the individual "treatment effect" for unit i is defined as the difference between both potential outcomes,

$$\tau_i = Y_i(1) - Y_i(0).$$

Since all units with scores above the cutoff are treated and all units with scores below the cutoff are untreated, the observed outcome is given by

$$Y_i = Y_i(1) \cdot D_i + Y_i(0) \cdot (1 - D_i) = \begin{cases} Y_i(0) & \text{if } D_i = 0\\ Y_i(1) & \text{if } D_i = 1. \end{cases}$$

The observed data is $\{Y_i, D_i, S_i\}_{i=1}^n$, which we assume to be an iid random sample from a larger population. Only one of the potential outcomes is observed for each unit, leading to the *the fundamental problem of causal inference* (Holland, 1986). The classic RD design addresses this fundamental problem by comparing treated units that are "slightly

above" the cutoff to control units that are "slightly below". The intuition is that, under appropriate assumptions, a relatively small window around the threshold will contain units that, on average, will be similar in terms of observed and unobserved characteristics, except for the treatment condition. By this rationale, the units slightly above the cutoff would have the same outcome as the units slightly below, if they were under the control condition.

4.3 Local Randomization vs. Continuity-Based Approaches

We review two different approaches to the analysis and interpretation of RD designs in the cross-sectional setting: the *continuity-based approach*, first formalized by Hahn et al. (2001) and discussed in Imbens and Lemieux (2008) and more recently in Cattaneo et al. (2020), and the *local randomization approach*, first formalized by Cattaneo et al. (2015) in a Fisherian randomization framework and discussed more recently by Cattaneo et al. (2021). These two frameworks are discussed in the review by Cattaneo et al. (2017).

The continuity-based approach seeks to approximate unknown regression functions based on smoothness conditions. Define the conditional expectation functions

$$\mu_1(s) = \mathbb{E}\left[Y_i(1)|S_i = s\right]$$
$$\mu_0(s) = \mathbb{E}\left[Y_i(0)|S_i = s\right],$$

and the observed regression function

$$\mu(s) = \mathbb{E}\left[Y_i | S_i = s\right] = \begin{cases} \mu_1(s) & \text{if } s \ge c \\ \mu_0(s) & \text{if } s < c. \end{cases}$$

The conditional average treatment effect given ${\cal S}_i$ is

$$\tau(s) = \mathbb{E}\left[Y_i(1) - Y_i(0)|S_i = s\right] = \mu_1(s) - \mu_0(s).$$

The continuity-based framework defines its target parameter as the average treatment effect at the cutoff, $\tau(c)$, and relies on the assumption that $\mu_1(s)$ and $\mu_0(s)$ are continuous on the score at the time cutoff, s = c.

Assumption 1: Continuity. The functions $\mu_1(s)$ and $\mu_0(s)$ are continuous on the score s at the cutoff c: $\lim_{s\to c} \mu_i(s) = \mu_i(c)$ for i = 0, 1.

Under this assumption, $\tau(c)$ is identifiable (Hahn et al., 2001) by

$$\tau(c) = \lim_{s \downarrow c} \mu(s) - \lim_{s \uparrow c} \mu(s).$$

An alternative framework is provided by the *local randomization approach*. Instead of relying on the continuity of $\mu_0(s)$ and $\mu_1(s)$, this approach assumes that the treatment is randomly assigned in a small neighborhood of the threshold. The parameter of interest in this framework is different from $\tau(c)$, because interest lies not on the average treatment effect at the cutoff point, but rather on the average treatment effect on a small interval or window around the cutoff. Let W_c be a window around c, defined by $W_c = [c - h, c + h]$ for h > 0. The local randomization parameter of interest is

$$\tau(W_c) = \mathbb{E}\left[Y_i(1) - Y_i(0) | S_i \in W_c\right]$$

Cattaneo et al. (2015) identify two conditions that must hold in a neighborhood of the cutoff for this local randomization interpretation: first, the distribution of the score cannot depend on the potential outcomes; second, the score can affect the potential outcomes only via the treatment, but not directly. We state these conditions below.

Assumption 2: Local Randomization. There exists a neighborhood $W_c = [c - h, c + h]$ such that for all i with $S_i \in W_c$: (a) $\mathbb{E}[Y_i(1)|S_i = s, S_i \in W_c] = \mathbb{E}[Y_i(1)|S_i \in W_c]$, and $\mathbb{E}[Y_i(0)|S_i = s, S_i \in W_c] = \mathbb{E}[Y_i(0)|S_i \in W_c]$. (b) $Y_i(1, s) = Y_i(1)$ and $Y_i(0, s) = Y_i(0)$ for all s.

The first condition formalizes the idea that the score can be considered as good as randomly assigned near the cutoff, requiring the that conditional expectation of the potential outcomes given that the score is in the window W_c is not a function of the particular value that the score takes inside this window. The second assumption is an exclusion restriction requiring the potential outcomes be affected only by placement above or below the cutoff, but not by the score itself. Note that this exclusion restriction was explicitly adopted in the notation we used above for the continuity-based framework. However, because the continuity-based parameter of interest is the average treatment effect at a point, this condition is not truly required or binding, because even if the score affects the potential outcomes directly, the parameter of interest fixes the score at the single value c in both regression functions.

Under the local randomization assumption, we have

$$\tau(W_c) = \mathbb{E}\left[Y_i | S_i \ge c, S_i \in W_c\right] - \mathbb{E}\left[Y_i | S_i < c, S_i \in W_c\right]$$

and we can estimate the effect by averaging the observed outcomes in a window around the cutoff c.

4.4 RD Designs with Time as Running Variable

The ES/ITS setup introduces a time dimension. The same cross-sectional units are observed in multiple time periods indexed by t, with $t = 1, 2, ..., \underline{t}$, and the treatment goes into effect at a precise and known moment in time. For example, in our application, the new CPP begins precisely at 12:00 AM on November 1st, 2017. An important feature of this assignment is that the treatment is "triggered" at the same specific moment for all units, t_0 , so that all periods before t_0 are untreated and all periods after t_0 (including it) are treated.

We modify our notation above to accommodate the time dimension. The potential outcomes under treatment and control for period t are, respectively, $Y_i(1,t)$ and $Y_i(0,t)$. The moment in time when treatment was into effect is denoted by t_0 . The treatment indicator is now denoted by $D_i(t)$, defined by

$$D_i(t) = \begin{cases} 0 & \text{if } t < t_0 \\ 1 & \text{if } t \ge t_0. \end{cases}$$

We can therefore define the score, denoted by $S_i(t)$, as

$$S_i(t) = t - t_0,$$

which normalizes the cutoff to zero and leads to the RD treatment assignment rule $D_i(t) = 1$ ($S_i(t) \ge 0$). Since the treatment is assigned at the same time for all units, $D_i(t)$ and $S_i(t)$ are constant for all *i* for a given *t*. We can therefore drop the *i* subscript and write D(t) and S(t). Moreover, because all units are observed in every period, the time variable is discrete and has "mass points"—for every time period $t, t = 1, \ldots$, there are *n* cross-sectional observations that share the same value of the score S(t).

We define the parameter of interest as

$$\tau = \mathbb{E}[Y_i(1,t) - Y_i(0,t) \mid S(t) = 0] = \mathbb{E}[Y_i(1,t) - Y_i(0,t) \mid t = t_0]$$

In order to be able to identify τ , we need the average untreated potential outcome the

day the treatment becomes effective to be the same as the average untreated potential outcome the day before.

Assumption 3: ES/ITS Local Randomization. $\mathbb{E}[Y_i(0,t)|S(t) = -1] = \mathbb{E}[Y_i(0,t)|S(t) = 0].$

Although Assumption 3 is strong, it is weaker than the standard local randomization condition in Assumption 2. When time is a discrete variable, a local randomization approach does not require to focus on a neighborhood or window around the cutoff; instead, we can focus on the discontinuity point $t = t_0$, just like in a continuity-based analysis. When we define the parameter of interest as τ , we do not need to invoke the exclusion restriction in Assumption 2(b) because both expectations are conditional on $t = t_0$. Assumption 3 alone is sufficient to achieve identification of our parameter of interest, τ :

$$\mathbb{E}[Y_i(t) \mid S(t) = 0] - \mathbb{E}[Y_i(t) \mid S(t) = -1] = \mathbb{E}[Y_i(1,t) \mid S(t) = 0] - \mathbb{E}[Y_i(0,t) \mid S(t) = -1]$$
$$= \mathbb{E}[Y_i(1,t) \mid S(t) = 0] - \mathbb{E}[Y_i(0,t) \mid S(t) = 0]$$
$$= \tau$$

In our application, t_0 is November 1st, 2017 and time periods are measured in days. This means that $\mathbb{E}[Y_i(0,t)|S(t) = -1]$ is the expected untreated potential outcome on October 31st, 2017, one day before the treatment is introduced, and $\mathbb{E}[Y_i(0,t)|S(t) = 0]$ is the average untreated potential outcome the day after, November 1st, 2017.

The local randomization framework interprets the choice of the particular date when the treatment is introduced as random. In other words, $D_i(T)$ is the treatment indicator for unit *i*, and this depends on the random variable *T* that denotes the date when the treatment becomes effective. In our local randomization interpretation, the cutoff t_0 is the realized value of the random variable *T*. And in general, we imagine that the event T = t is random. But this is simple a heuristic device, as Assumption 3 does not restrict the assignment mechanism but rather imposes the comparability condition that is needed.

Note that given Assumption 3, our definition of our treatment effect of interest τ at t_0 , and the assumption that we have n units for every time period, estimation can proceed using only observations for two periods: t_0 and $t_0 - 1$. Often, however, researchers have access to many prior periods before the cutoff. We now discuss how this information can be used to both asses the plausibility of Assumption 3 and relax it.

4.5 Time Confounders

If another variable changes discontinuously between $t_0 - 1$ and t_0 , Assumption 3 will not hold. In our crime application this can happen, for example, if $t_0 - 1$ falls on a Sunday and t_0 on a Monday, and the average number of crimes on Sunday might be higher than on Monday due to higher crowds during the weekend.⁷ In particular, this day-of-the-week effect is very common in event studies (e.g., C-section births versus vaginal births).

Although a direct test of this assumption is not feasible, its plausibility can be investigated using prior periods. Let f(t) be a function that for every date t returns the day of the week associated with t, and F a random variable that indicates which day of the week occurs. To test whether there is a potential bias due to time confounders, we estimate

$$\mathbb{E}[Y_i(0,t)|F = f(t_0)] - \mathbb{E}[Y_i(0,t)|F = f(t_0-1)]$$

using data from the multiple prior periods available in the dataset prior to $t_0 - 1$, that is, $\{Y_i(t)\}_{t \leq t_0-2}$. We consider two different strategies to evaluate the plausibility of Assumption 3. In Section 5.5, we conduct robustness exercises by employing alternative cutoffs under the local-randomization approach for the following calendar effects: (i) on Novem-

⁷Crimes generally form daily and weekly patterns. For empirical evidence and a discussion regarding the temporal concentration of crime, see Prieto Curiel (2021).

ber 1st but for adjacent years (i.e., potential turn-of-month effect), and on the same day of the week but for adjacent months (i.e., potential day-of-the-week effect).

5 Estimation Results

Treatment is given by the implementation of the new CPP and begins when the score, $S(t) = t - t_0$, reaches zero, that is, when t is equal to the cutoff $t_0 = \{12\text{AM}, 11/01/2017\}$. The cornerstone of this quantitative method is the ability to compare dates that are close enough to both sides of that temporal threshold. In our case, the key identification assumption would imply that offenses, Y_t , dated November 1st, 2017 and October 31st, 2017 were reported under identical conditions except for the treatment status: only the former of the two will be prosecuted following the rules of the new adversarial system.⁸

We start by defining the continuity-based RD treatment effect at the temporal threshold is

$$\tau(0) = \mathbb{E}\left[Y_i(\text{new CPP}, t) - Y_i(\text{old CPP}, t) | S(t) = 0\right].$$

We estimate this parameter with low-degree local polynomials, as suggested by the literature (Cattaneo et al., 2020; Gelman and Imbens, 2019).

5.1 Visual Evidence

Exploratory data analysis is key to provide a transparent validation of our RD design. Following Calonico et al. (2015), Figure 3 presents RD plots that approximate the underlying variance of the data based on an evenly-spaced binning method. This data-driven selector avoids subjective and less reliable benchmarks for graphical analysis. Plots (a) to (d) illustrate different RD plots when the population conditional expectation functions

⁸Since we are using the number of police reports as the best available proxy for the actual number of committed crimes, we also need to assume that the propensity to report crimes did no change due to the new crime-related policy. This is an usual identification assumption researchers make when working with this type of data.

are approximated by global polynomials of order zero to order three, respectively.

[INSERT FIGURE 3 HERE]

The different approximations suggest the presence of a discontinuity at the threshold. In other words, even when we increase the order of the global polynomial used to approximate the conditional expectation functions, we observe a jump when the running variable takes the cutoff value. This visual evidence indicates a jump (i.e., an increase in the number of police reports for the six most frequent crimes in Montevideo) due to treatment (i.e., a switch from an inquisitorial penal procedure to an adversarial one), rather than a nonlinearity in the counterfactual conditional mean function.

5.2 Continuity-Based Approach: *Estimates*

Table 2 exhibits our main RD estimation results using the continuity-based framework. We compute the local treatment effect, $\hat{\tau}$, by employing a local-linear approximation of the conditional expectation functions and a triangular kernel. For inference, we use robust bias-corrected confidence intervals (Calonico et al., 2014, 2017).

[INSERT TABLE 2 HERE]

We start by reporting global linear estimates; Columns (i) and (ii) present the results for the largest available symmetric window around the cutoff and a two-year bandwidth also centered at the cutoff, respectively. According to these estimations, there is an average treatment effect of 50-61 crimes per day due to the implementation of the migration to an adversarial penal system. As discussed above, however, these global estimates do not enjoy desirable properties when the effect of interest is local.

To implement local polynomial methods, we select the bandwidth, \hat{h} , following recommendations by Cattaneo et al. (2020), who suggest the use of two data-driven procedures:

the MSE-optimal and the CER-optimal bandwidths. The mean squared error (MSE) criterion is the most widely used procedure and its goal is to balance the tradeoff between bias and variance of RD point estimators, $\hat{\tau}$. These bandwidths are optimal for point estimation. Meanwhile, CER-optimal bandwidths minimize the coverage error probability of confidence intervals. To be more precise, given an α -level confidence interval for a parameter associated with $\hat{\tau}$, the CER-optimal procedure makes the coverage probability as close as $1 - \alpha$ as possible. These bandwidths are optimal for inference. Panel A shows the results when we use a local-linear method and triangular kernel, for MSE and CERoptimal bandwidths. Columns (1) and (2) present the results for a symmetric bandwidth (i.e., the same number of observations before and after the cutoff), whereas columns (3) and (4) report results when the data-driven procedures employ different selectors at both sides of the threshold (i.e., different number of observations before and after the cutoff). Results show an increase of 21-24 police reports per day at the moment the new CPP entered into force, accounting for 23%-26% of the average annual increase of 93 police reports per day (Table 1). All the results are statistically significant at the 0.05 level.

5.3 Continuity-Based Approach: Robustness

This section of the paper is devoted to robustness analysis. We will present results comparable to those in Panel A of Table 2 but modifying some of the key ingredients of our RD design: local approximation methods, kernel functions, control variables and placebo cutoffs. We also report the RD estimation results but using the above discussed local randomization framework. As it will be noticed, the conclusions of the paper survive the different estimation results presented in the current section.

(a) Polynomial Order

Table 2 presented RD point estimates that result from fitting a local-linear polynomial. In order to show robustness to the polynomial order, we will also show results using local constant and local quadratic polynomials.⁹

[INSERT TABLES 3 AND 4 HERE]

According to Tables 3 and 4, the results remain qualitatively unchanged when we change the method we use to approximate the conditional expectation functions within the selected bandwidths. Except for the case of the local-quadratic method under a CER-optimal bandwidth with two selectors, all the RD point estimates are positive and statistically significant for estimators with robust bias-corrected confidence intervals. As expected, the lower the degree of the local polynomial, the larger the magnitude of the point estimate. However, the first two columns of Panel B in Tables 3 and 4 exhibit coefficients that are both statistically significant and comparable in magnitude to those in the respective columns of Panel A.

(b) Kernel Function

The kernel function assigns non-negative weights to each observation t based on the distance between its score, S_t , and the cutoff, c. In spite of the fact that the triangular kernel exhibits desirable asymptotic optimality properties, there are other alternatives to be considered. Tables 5 and 6 replicate the estimations for the same local-linear method as in Table 2, but employing different kernel functions: Epanechnikov and uniform. Whereas the triangular and Epanechnikov kernels give linear and quadratic decaying weight to observations within interval, using a uniform kernel is similar to estimating a simple

⁹Third and higher-order polynomials in RD analysis suffer from well known problems, see Gelman and Imbens (2019).): (i) noisy estimates (i.e., weights with unattractive properties), (ii) results sensitive to the degree of the polynomial (without a method for choosing the order in a way desirable for the causal inference), and (iii) the coverage of the confidence intervals is usually poor (i.e., fail to include zero with high probability).

linear regression without weights for the observations within the bandwidth of size h. As expected, the results are not very sensitive to the selection of the kernel function.

[INSERT TABLES 5 AND 6 HERE]

(c) Placebo Cutoffs

A very standard and useful falsification test consists in analyzing treatment effects at artificial cutoff values. Recall that the RD design relies in the assumption of lack of abrupt changes in the regression functions at the cutoff in the absence of treatment. Since it is impossible to test that key assumption, we could alternatively test if the estimable regression functions are continuous away from the cutoff. Finding that these functions are not continuous at points other than the cutoff would cast doubts on our RD design. We set placebo cutoffs 30, 45, 60 and 75 days before and after the actual threshold.¹⁰ Tables 7 and 8 present the results for different artificial cutoffs when we use a linear method, a triangular kernel and a MSE-optimal bandwidth, and a linear method, a triangular kernel and a CER-optimal bandwidth, respectively. As expected, none of the RD point estimates are statistically significant.

[INSERT TABLES 7 AND 8 HERE]

5.4 Local-Randomization Approach: *Estimates*

As explained above, we can also adopt the local-randomization approach to RD when the time invariance assumption holds. A crucial step is always to define the window W_c where the local randomization framework will be invoked. In order to chose the size of W_c in a data-driven way, we could take advantage of the information provided by relevant covariates and conduct a balance test comparable to those implemented in randomized

¹⁰Artificial cutoffs are set at least one month away from the actual cutoff so we have a minimum of thirty observations on each side of the placebo cutoff. These falsification tests include observations from only one side of the real cutoff.

control trials. However, this approach requires the econometrician to know at least one of those covariates, something that is not always feasible. Alternatively, we can exploit the temporal dimension of our RD design and use the most frequent incidents reported to police in Montevideo (i.e., theft, domestic violence and robbery; almost 75% of the total) exactly a year before the new CPP entered into force (i.e., around November 1st, 2016). We set a standard significance level of $\alpha^* = .15$ and start with a symmetric window that has at least 10 observations on either side of the 2016 cutoff. According to plot (a) in Figure 4, the minimum p-value is above 0.15 for all windows smaller than [-22, 22], rapidly decreasing to zero for windows of 50 units or more. As expected, plot (b) shows that we never fail to reject the null hypothesis that both sides of the different windows are balanced when we conduct the same tests but for the relevant cutoff.

[INSERT FIGURE 4 HERE]

Table 9 reports the estimation results for both the Fisherian (finite sample) and the Neymann (large sample) inference approaches. In the first column we have the differencein-means for the data-driven window, $W_c^* = [-22, 22]$. The local treatment effect is 26.364, considerably similar to the continuity-based local linear point estimates that were reported in Table 2 (from 21.5 to 24.0 police reports). We are able to reject the null hypothesis that the new CPP has of no effect on the number of police reports at 0.01 level, using both the Fisherian and the Neyman inference approaches. In columns (3) to (4) we repeat the estimation but for narrower windows and the obtained results are both comparable in magnitude and statistically significant.

[INSERT TABLE 9 HERE]

5.5 Local-Randomization Approach: *Robustness*

As it was discussed in Section 4, the local-randomization framework assumes the temporal threshold when the treatment is introduced, t_0 , to be random. Assumption 3 does not

restrict the assignment mechanism but rather imposes the comparability condition that is needed. In other words, if something different from the treatment changes discontinuously between $t_0 - 1$ and t_0 , Assumption 3 will not hold. We test for two different calendar effects in order to evaluate the plausibility of this assumption.

We start by considering the 62 neighborhoods of Montevideo as our units of analysis. So far we have considered the city as a whole, hence the one-to-one relationship between the sample size and the width of the temporal window, W_c , around the cutoff. This new approach allows us to get arbitrary close to the threshold without dramatically decreasing the size of our sample.

Figure 5 reports the main estimation results. We start with a five-day bandwidth on either side of the cutoff and then we increase it by five days until we have a sixty-day bandwidth on either side. Reported differences in means range from 0.20 to 0.65 incidents a day. Since this the local treatment effect in police reports for the average neighborhood, these results are equivalent to 12-40 reports per day for the entire city of Montevideo. As expected, these results are in line with those reported in Table 9. For instance, when we set a twenty-day bandwidth on either side of W_c we get an estimate of 0.4 for the average neighborhood and 25 for the city, comparable to the 26.4 reported in column (2) of Table 9.

[INSERT FIGURE 5 HERE]

(a) Turn-of-Month Effect

We repeat the exercise reported in Figure 5 but for the first day of November in 2016, 2018 and the 2016-2018 period. Results are reported in Figure 6 and in line with Assumption 3. It is worth noticing that these estimates are well below those reported in Figure 5 and not statistically different from zero in most of the cases.

[INSERT FIGURE 6 HERE]

(b) Day-of-the-Week Effect

The day on November 1st, 2017, was a Wednesday. In order to rule out a potential bias due to weekly patterns in crime, we repeat the analysis for the first Wednesdays in October and December during the five years prior to the implementation of the new CPP (i.e., 2012-2016). Results are reported in Figure 7. In spite of the fact that there seems to be a calendar effect associated with the first Wednesday in October, results are well below those reported for November 1st, 2017 (see Figure 5). On the other hand, difference in means are negative for the first Wednesday in December. According to these results, Assumption 3 also seems plausible.

[INSERT FIGURE 7 HERE]

6 Policy Discussion

Continuity-based RD point estimates suggest a local increase of 21 to 24 police reports per day due to the migration from an inquisitorial and written system to an accusatory, adversarial, oral and public one (see Figure 8). In spite of the fact that a criminal procedure will neither define the recognized offenses nor fix the corresponding penalties, our evidence suggests that a change in the adjudication process of the criminal law could unexpectedly decrease public safety. In fact, this increase in the number of police reports was not anticipated as a potential consequence of the implementation of the new CPP in Uruguay. Our findings are consistent with those of Zorro Medina et al. (2020), who document that the procedural reform in Colombia caused an increase in overall crime of 22%-34%.¹¹ In contrast to their study, our design allows us to identify local effects in narrower periods of time (i.e., a few weeks).

¹¹As it was discussed before, the previous empirical evidence also finds that these reforms can be associated with lower crime rates due to a reduction in the procedural length. For examples, see Dalla Pellegrina (2008) for Italy, Soares and Sviatschi (2010) for Costa Rica, and Duŝek (2015) for Czech Republic.

[INSERT FIGURE 8 HERE]

These results could respond to both more lenient punishments and lower probability of being caught and punished, see Becker (1968). First, the new adversarial system seems to produce less severe penalties than the old inquisitorial one. In this case, the new CPP introduces procedural alternatives to oral trials that may result in lighter sentences and have been widely used by prosecutors. In fact, they were employed in 90% of solved cases.¹² Procedural alternatives help lawyer and defendant to resolve their case faster and face more lenient sentence, while public prosecutors avoid lengthy and demanding criminal trials. Moreover, preventive detention (i.e., while the process lasts until there is a sentence) ceased to be the norm, as was the case under the inquisitorial system (we will return to this point). Therefore, some criminals would expect a less severe punishment if they were caught. Second, the new legal system could represent a lower probability of conviction. Prosecutors face (i) a totally new role under the new legal system, and (ii) a significant increase in their workload. Recall that under the inquisitorial system, the judge leads the investigation and decides the appropriate punishment, whereas judges serve as neutral referees under the adversarial penal procedure. Consequently, the investigation is exclusive responsibility of prosecutors who must carry evidence to judges. Moreover, the police conducts investigation under new supervision (prosecutors instead of judges), different rules and several issues in the implementation process (at least during the first months of the new CPP).

Figure 9 provides some visual evidence in line with the hypothesis that Uruguay's switch from an inquisitorial to and adversarial system has been far from innocuous in terms of crime incentives. For the first time in more than a decade, the average number of people in prisons decreased in 2018, from 11,005 to 10,179 inmates (i.e., a 7.5%). According to

¹²Most of these solved cases are the result of the use of the abbreviated process (i.e., a type of plea bargaining introduced by the new penal procedure) that implies an agreement between defendant and prosecutor whereby the former pleads guilty to a particular charge in return for a more lenient sentence from the prosecutor.

Chart (a), this sharp reduction in prison population starts with the implementation of the new penal procedure and contrasts with the already documented sudden rise in the number of crimes reported to police. Meanwhile, the number of criminal imputations (i.e., the number of formal accusations made by public prosecutors) also experienced a strong month-to-month decrease in November 2017, from 1,001 to 584 cases (i.e., a 42% reduction). Chart (b) illustrates the evolution of imputations relative to police reports. As it can be noticed, the average ratio for the first two months of the new CPP is 3.2% (i.e., November and December 2017), well below the corresponding figure for the rest of 2017 and the last months of the old CPP: 5.9%.

[INSERT FIGURE 9 HERE]

The change in the criminal procedure could also impose heterogeneous effects across different types of offenses. To be more precise, less severe crimes could have increased due to a change in the crime incentives (i.e., less deterrence), but also responding to an increase in the number of active criminals (i.e., less incapacitation). The reduction in prison population could be particularly explained by the new constraints imposed on the use of preventive detention by prosecutors.¹³ As mentioned above, preventive prison was used extensively under the inquisitorial system, in particular for cases of recidivism (55% of the cases). This was basically the norm for both misdemeanors (e.g., thefts) and felonies (e.g., robberies). Under the adversarial system, preventive detention is applicable only in cases in which there is enough evidence suggesting the defendant could intend to escape, obstruct the investigation or pose a risk to society (i.e., severe and usually violent crimes). Therefore, a convicted offender who re-offends by committing a lesser crime could be immediately released under the adversarial system, whereas she would probably

¹³There are other potential explanations that are also consistent with Figure 9. For instance, there is evidence that police officers may deviate their tactics towards more inappropriate ones when norms that restrict their actions are imposed, resulting in reductions of imprisonment rates (Hausman and Kronick, 2021). As it was previously discussed, the new CPP introduced substantial changes in the work of police officers, including new tasks, different supervision and a substantial increase in the workload.

serve preventive detention under the old CPP. To test this hypothesis, we can empirically evaluate the local impact of the new CPP on the number of reported thefts and compare it to how the new penal procedure affected robberies (i.e., violent thefts) and domestic violence (i.e., the most frequent crime against the person).

Figure 10 presents plots for global polynomials of order three that approximate the underlying variance of the data based on an evenly-spaced binning method. Plots (a) to (d) illustrate different RD plots for total crime (i.e., six most frequent crimes), thefts, robberies and domestic violence, respectively. These three individual categories account for more than 7 of every 10 crimes reported in Montevideo (see Table 1). As expected, the evolution in theft reports strongly resembles that of the total number of police reports. In spite of the fact that the visual inspection of these RD plots is not enough to reject an actual "jump" at the threshold, we could have just a non-linearity at the cutoff for both robberies and domestic violence. In fact, when we compare our main results reported in Table 2 to the corresponding RD estimates for these three individual offenses we obtain heterogeneous effects, as expected.

[INSERT FIGURE 10 HERE]

According to Table 10, the local treatment effect of the new CPP on the number of thefts reported to police in Montevideo (Panel B) is consistent to our main RD estimation results (Panel A) that were originally reported in Table 2. However, Tables 11 and 12 indicate no local effect of the new CPP on robberies and domestic violence, respectively. This evidence is consistent with the hypothesis of an immediate increase in the number of active criminals (i.e., thieves) due to the constrains imposed on the use of preventive detention by the new CPP.

[INSERT TABLES 10, 11 AND 12 HERE]

References

- Abrams, D. S. (2011), "Is Pleading Really a Bargain?" Journal of Empirical Legal Studies, 8, 200–221.
- Acemoglu, D., and Johnson, S. (2005), "Unbundling Institutions," Journal of Political Economy, 113, 949–995.
- Atkins, R. A., and Rubin, P. H. (2003), "Effects of Criminal Procedure on Crime Rates: Mapping Out the Consequences of the Exclusionary Rule," *The Journal of Law and Economics*, 46, 157–179.
- Becker, G. S. (1968), "Crime and Punishment: An Economic Approach," Journal of Political Economy, 76, 169–217.
- Bushway, S. D., and Redlich, A. D. (2012), "Is Plea Bargaining in the "Shadow of the Trial" a Mirage?" Journal of Quantitative Criminology, 28, 437–454.
- Bushway, S. D., Redlich, A. D., and Norris, R. J. (2014), "An Explicit Test of Plea Bargaining in the "Shadow of the Trial"," *Criminology*, 52, 723–754.
- Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2017), "rdrobust: Software for Regression-Discontinuity Designs," *The Stata Journal*, 17, 372–404.
- Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014), "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 82, 2295–2326.
- (2015), "Optimal Data-Driven Regression-Discontinuity Plots," Journal of the American Statistical Association, 110, 1753–1769.
- Cattaneo, M. D., Frandsen, B., and Titiunik, R. (2015), "Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate," *Journal of Causal Inference*, 3, 1–24.
- Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2020), A Practical Introduction to Regression Discontinuity Designs: Foundations, Cambridge Elements: Quantitative and Computational Methods for Social Science, Cambridge University Press.
- (2021), A Practical Introduction to Regression Discontinuity Designs: Extensions, Cambridge Elements: Quantitative and Computational Methods for Social Science, Cambridge University Press, to appear.

- Cattaneo, M. D., Titiunik, R., and Vazquez-Bare, G. (2017), "Comparing Inference Approaches for RD Designs: A Reexamination of the Effect of Head Start on Child Mortality," *Journal of Policy Analysis and Management*, 36, 643–681.
- Dalla Pellegrina, L. (2008), "Court Delays and Crime Deterrence," European Journal of Law and Economics, 26, 267–290.
- Duggan, M., Garthwaite, C., and Goyal, A. (2016), "The Market Impacts of Pharmacentrical Product Patents in Developing Countries: Evidence from India," *American Economic Review*, 106, 99–135.
- Duŝek, L. (2015), "Time to Punishment: The Effects of a Shorter Criminal Procedure on Crime Rates," International Review of Law and Economics, 43, 134–147.
- Fandiño, M., and González Postigo, L. (2020), "Adversarial Criminal Justice in Latin America: Comparative Analysis and Proposals," Justice Studies Center of the Americas (JSCA).
- Frazier, A., Shockley, K., Keenan, J. M., Wilford, M. M., and Gonzales, J. E. (2018), "When a Plea is no Bargain at All: Comparing Sentencing Outcomes for Massachusetts Defendants in Non-Sexual and Sexual Crimes," *Alb. L. Rev.*, 82, 775.
- Gelman, A., and Imbens, G. W. (2019), "Why High-Order Polynomials Should Not be Used in Regression Discontinuity Designs," *Journal of Business & Economic Statistics*, 37, 447–456.
- Hahn, J., Todd, P., and van der Klaauw, W. (2001), "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69, 201–209.
- Hausman, D., and Kronick, D. (2021), "When Police Sabotage Reform by Switching Tactics," *SSRN*.
- Holland, P. W. (1986), "Statistics and Causal Inference," Journal of the American statistical Association, 81, 945–960.
- Imbens, G., and Lemieux, T. (2008), "Regression Discontinuity Designs: A Guide to Practice," Journal of Econometrics, 142, 615–635.
- Institute for Crime & Justice Policy Research (2021), *The World Prison Brief*, https://www.prisonstudies.org/.

Kronick, D. (2019), "The Legal Origins of State Violence," Unpublished Manuscript.

- Langer, M. (2007), "Revolution in Latin American Criminal Procedure: Diffusion of Legal Ideas from the Periphery," *The American Journal of Comparative Law*, 55, 617–676.
- Langer, M. (2021), "Plea Bargaining, Conviction without Trial, and the Global Administratization of Criminal Convictions," Annual Review of Criminology, 4, 377–411.
- Listokin, Y. (2007), "Crime and (with a Lag) Punishment: The Implications of Discounting for Equitable Sentencing," Am. Crim. L. Rev., 44, 115.
- MacKinlay, A. C. (1997), "Event Studies in Economics and Finance," Journal of Economic Literature, 35, 13–39.
- Maier, J., and Struensee, E. (2000), "Introducción: Las Reformas Procesales Penales en América Latina," J. Maier, K. Ambos, & J. Woischnik (Edits.), Las Reformas Procesales Penales en América Latina, 17–34.
- Mummolo, J. (2018), "Modern Police Tactics, Police-Citizen Interactions and the Prospects for Reform," *The Journal of Politics*, 80, 1–15.
- "Policía deberá ingeniársela para adaptarse al nuevo CPP" (2017), *Diario El Observador*, November, 22th.
- Prieto Curiel, R. (2021), "Weekly Crime Concentration," Journal of Quantitative Criminology, 1–28.
- Soares, Y., and Sviatschi, M. M. (2010), "Does Court Efficiency Have a Deterrent Effect on Crime? Evidence for Costa Rica," *Unpublished Manuscript*.
- Solomita, M. (2019), "Fiscales Agobiados: Turnos Interminables, Licencias por Estrés y Jubilaciones Anticipadas," *Diario El País*, Suplemento Qué Pasa, May 18th.
- St. Clair, T., Cook, T. D., and Hallberg, K. (2014), "Examining the Internal Validity and Statistical Precision of the Comparative Interrupted Time-Series Design by Comparison with a Randomized Experiment," *American Journal of Evaluation*, 35, 311–327.
- St. Clair, T., Hallberg, K., and Cook, T. D. (2016), "The Validity and Precision of the Comparative Interrupted Time-Series Design: Three Within-Study Comparisons," *Journal of Educational and Behavioral Statistics*, 41, 269–299.

- Ulmer, J. T., and Bradley, M. S. (2006), "Variation in Trial Penalties among Serious Violent Offenses," *Criminology*, 44, 631–670.
- Ulmer, J. T., Eisenstein, J., and Johnson, B. D. (2010), "Trial Penalties in Federal Sentencing: Extra-Guidelines Factors and District Variation," *Justice Quarterly*, 27, 560– 592.
- Zorro Medina, A. (2020), "The Failed War on Pre-Trial Detention: Evidence from a Quasi-Experimental Reform," SSRN.
- Zorro Medina, A., Acosta, C., and Mejia, D. (2020), "The Unintended Consequences of the U.S. Adversarial Model in Latin American Crime," *SSRN*.

Tables and Figures

	Two-Year Window	Old CPP^a	New CPP^b	Δ
	11/01/10 to 10/31/18	(a)	(0)	(b)-(a)
The ft	152.30	130.38	174.22	43.84^{***}
	46%	(0.944)	(1.221)	(1.120)
Robbery	51.85	41.66	62.05	20.39***
	16%	(0.425)	(0.615)	(0.748)
Domestic Violence	33.59	32.79	34.39	1.61^{***}
	10%	(0.379)	(0.380)	(0.537)
Threat	22.36	21.21	23.52	2.32***
	7%	(0.319)	(0.343)	(0.468)
Damage	21.14	17.39	24.89	7.50***
	6%	(0.249)	(0.279)	(0.374)
Personal Injury	10.18	9.49	10.87	1.38^{***}
	3%	(0.182)	(0.184)	(0.259)
Total Police Reports	330.34	283.92	376.76	92.84***
Ĩ	100%	(1.529)	(1.995)	(2.513)
Days	730	365	365	

Table 1: Crime in Montevideo, 2016-2018 (offenses reported to police; average number per day)

Percentage of the total number of reports in *italic*; standard errors in parentheses. *** Difference between (b) and (a) is significant at the 0.01 level.

Source: Ministry of Interior of Uruguay.

 $^a {\rm Last}$ year of the inquisitorial system: November 1st, 2016 to October 31st, 2017. $^b {\rm First}$ year of the adversarial system: November 1st, 2017 to October 31st, 2018.



Figure 1: Crime in Montevideo, 2014-2019



Figure 2: Crime in Montevideo, 2014-2019 (by type of incident)



Figure 3: RD Plots (top-six police reports in Montevideo; sample average within bin)

				PANI	EL A	
	(i)	(ii)	(1)	(2)	(3)	(4)
Treatment Effect, $\hat{\tau}$ Robust Bias-Correction	60.629*** (0.000)	50.455*** (0.000)	21.528** (0.032)	20.636** (0.045)	24.037*** (0.003)	23.131*** (0.007)
Bandwidth Selection, \hat{h}						
Data-Driven Procedure	No	No	Yes	Yes	Yes	Yes
Mean Squared Error (MSE)-Optimal	-	-	Yes	No	Yes	No
Coverage Error Rate (CER)-Optimal	-	-	No	Yes	No	Yes
Number of Selectors	-	-	One	One	Two	Two
Days Before Cutoff	607	365	106	71	178	121
Days After Cutoff	607	365	106	71	132	90
Number of $Observations$	1.214	730	212	142	310	211

Table 2: RD Point Estimates, $\hat{\tau}$ (top-six reports; linear method, triangular kernel)

Probability value (p-value) in parentheses;

*, ** and *** indicate significance at the 0.1, 0.05 and 0.01 levels, respectively.

Reports include top-six crime types (\approx 90% of total) for Montevideo.

	PAN	EL A: Local	-Linear Met	hod	PANE	L B: Local-	Constant Me	thod
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Local Treatment Effect, $\hat{\tau}$	21.528**	20.636**	24.037^{***}	23.131***	22.992**	22.992**	38.682^{***}	38.682***
Robust Bias-Correction	(0.032)	(0.045)	(0.003)	(0.007)	(0.032)	(0.032)	(0.000)	(0.000)
Bandwidth Selection, \hat{h}								
Mean Squared Error (MSE)-Optimal	Yes	No	Yes	No	Yes	No	Yes	No
Coverage Error Rate (CER)-Optimal	No	Yes	No	Yes	No	Yes	No	Yes
Number of Selectors	One	One	Two	Two	One	One	Two	Two
Days <u>Before</u> Cutoff	106	71	178	121	28	28	103	103
Days <u>After</u> Cutoff	106	71	132	90	28	28	38	38
Number of Observations	212	142	310	211	56	56	141	141

Table 3: RD Point Estimates, $\hat{\tau}$ (top-six reports; triangular kernel)

Probability value (p-value) in parentheses;

*, ** and *** indicate significance at the 0.1, 0.05 and 0.01 levels, respectively.

Reports include top-six crime types (\approx 90% of total) for Montevideo.

PANEL A: Local-Linear Method PANEL B: Local-Quadratic Method (2)(3)(4)(1)(2)(3)(4)(1)Local Treatment Effect, $\hat{\tau}$ 21.528** 20.636** 24.037*** 23.131*** 17.880^{*} 18.340^{*} 20.038** 15.994 Robust Bias-Correction (0.032)(0.045)(0.003)(0.007)(0.052)(0.097)(0.038)(0.157)Bandwidth Selection, \hat{h} Mean Squared Error (MSE)-Optimal Yes No Yes NoYes No Yes No Coverage Error Rate (CER)-Optimal No No No Yes No Yes Yes Yes $Number \ of \ Selectors$ One One TwoTwoOne One TwoTwoDays Before Cutoff 10671178121161104249161Days After Cutoff 106711329016110414795Number of $\overline{Observations}$ 212142310 211322 208396 256

Table 4: RD Point Estimates, $\hat{\tau}$ (top-six reports; triangular kernel)

Probability value (p-value) in parentheses;

*, ** and *** indicate significance at the 0.1, 0.05 and 0.01 levels, respectively.

Reports include top-six crime types (\approx 90% of total) for Montevideo.

	PA	NEL A: Tri	angular Kerr	iel	PANEL B: Epanechnikov Kernel			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Local Treatment Effect, $\hat{\tau}$	21.528**	20.636**	24.037***	23.131***	23.044**	21.193**	29.311***	27.633***
Robust Bias-Correction	(0.032)	(0.045)	(0.003)	(0.007)	(0.021)	(0.041)	(0.000)	(0.001)
Bandwidth Selection, \hat{h}								
Mean Squared Error (MSE)-Optimal	Yes	No	Yes	No	Yes	No	Yes	No
Coverage Error Rate (CER)-Optimal	No	Yes	No	Yes	No	Yes	No	Yes
Number of Selectors	One	One	Two	Two	One	One	Two	Two
Days Before Cutoff	106	71	178	121	97	66	205	140
Days After Cutoff	106	71	132	90	97	66	118	80
Number of Observations	212	142	310	211	194	132	323	220

Table 5: RD Point Estimates, $\hat{\tau}$ (top-six reports; linear method)

Probability value (p-value) in parentheses;

 $^{\ast},$ ** and *** indicate significance at the 0.1, 0.05 and 0.01 levels, respectively.

Reports include top-six crime types (\approx 90% of total) for Montevideo.

Table 6: RD Point Estimates, $\hat{\tau}$ (top-six reports; linear method)

	PA	NEL A: Tri	angular Kerr	iel	PANEL B: Uniform Kernel			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Local Treatment Effect, $\hat{\tau}$	21.528**	20.636**	24.037***	23.131***	28.060***	14.531	32.652***	34.227***
Robust Bias-Correction	(0.032)	(0.045)	(0.003)	(0.007)	(0.006)	(0.176)	(0.000)	(0.000)
Bandwidth Selection, \hat{h}								
Mean Squared Error (MSE)-Optimal	Yes	No	Yes	No	Yes	No	Yes	No
Coverage Error Rate (CER)-Optimal	No	Yes	No	Yes	No	Yes	No	Yes
Number of Selectors	One	One	Two	Two	One	One	Two	Two
Days Before Cutoff	106	71	178	121	78	53	202	143
Days After Cutoff	106	71	132	90	78	53	91	81
Number of $\overline{Observations}$	212	142	310	211	156	106	293	224

Probability value (p-value) in parentheses;

*, ** and *** indicate significance at the 0.1, 0.05 and 0.01 levels, respectively.

Reports include top-six crime types (\approx 90% of total) for Montevideo.

	PANE	L A: Days I	Before Actua	l Cutoff	PANI	PANEL B: Days After Actual Cutoff				
	-75	-60	-45	-30	30	45	60	75		
Local Treatment Effect, $\hat{\tau}$ Robust Bias-Correction	13.206 (0.119)	-10.173 (0.391)	-6.487 (0.127)	11.575 (0.361)	7.946 (0.250)	-22.811 (0.708)	1.579 (0.988)	5.321 (0.842)		
Days Before Cutoff	124	86	96	116	29	44	40	35		
Days After Cutoff	75	60	45	30	42	52	41	36		
Number of Observations	216	141	146	153	71	96	80	70		

Table 7: Placebo Cutoffs I (linear method; triangular kernel; MSE-optimal bandwidth)

Probability value (p-value) in parentheses; *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 levels, respectively.

Table 8: Placebo Cutoffs II (linear method; triangular kernel; CER-optimal bandwidth)

	PANE	L A: Days <i>E</i>	Before Actua	l Cutoff	PAN	NEL B: Days .	After Actual	Cutoff
	-75	-60	-45	-30	30	45	60	75
Local Treatment Effect, $\hat{\tau}$ Robust Bias-Correction	12.467 (0.134)	-16.001 (0.271)	-7.116 (0.181)	12.315 (0.286)	14.209 (0.252)	-10.382 (0.983)	-7.254 (0.670)	-8.740 (0.544)
Days <u>Before</u> Cutoff	86	60	67	81	30	37	29	35
Days After Cutoff	75	60	45	30	30	37	29	35
Number of $Observations$	161	120	112	111	60	74	58	70

Probability value (p-value) in parentheses; *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 levels, respectively.

Figure 4: Window Size Selection (top-three police reports in Montevideo)



Table 9: RD Point Estimates, $\hat{\tau}$ (top-six reports; local randomization approach)

	(1)	(2)	(3)	(4)
Local Treatment Effect, $\hat{\tau}$	26.364***	26.400***	23.056**	18.500^{*}
Fisherian Approach	(0.000)	(0.004)	(0.024)	(0.054)
Neymann Approach	(0.001)	(0.002)	(0.013)	(0.049)
Days Before Cutoff	22	20	18	16
Days <u>After</u> Cutoff	22	20	18	16

Probability value (p-value) in parentheses;

*, ** and *** indicate significance at the 0.1, 0.05 and 0.01 levels, respectively. Reports include top-six crime types ($\approx 90\%$ of total) for Montevideo.





Figure 6: Randomization Inference (Turn-of-Month Effect) (neighborhood level; uniform kernel; 1,000 reps; 95% CI)









Figure 9: Police Reports, Prison Population and Criminal Imputations





(b) Criminal Imputations



Figure 10: RD Plots by Offense Type (sample average within bin; polynomial of degree three)

		PANEL A:	Total Crime			PANEL	B: Theft	
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Local Treatment Effect, $\hat{\tau}$	21.528** (0.032)	20.636** (0.045)	24.037*** (0.003)	23.131*** (0.007)	10.999* (0.061)	13.428^{**} (0.025)	17.821*** (0.001)	18.039*** (0.002)
Bandwidth Selection, \hat{h}								
Mean Squared Error (MSE)-Optimal	Yes	No	Yes	No	Yes	No	Yes	No
Coverage Error Rate (CER)-Optimal	No	Yes	No	Yes	No	Yes	No	Yes
Number of Selectors	One	One	Two	Two	One	One	Two	Two
Days Before Cutoff	106	71	178	121	151	103	204	139
Days <u>After</u> Cutoff	106	71	132	90	151	103	101	69
Number of Observations	212	142	310	211	301	206	305	208

Table 10: RD Point Estimates, $\hat{\tau}$, by Offense Type (local-linear method; triangular kernel)

Robust probability values in parentheses; *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 levels, respectively.

Table 11: RD Point Estimates, $\hat{\tau}$, by Offense	Type
(local-linear method; triangular kernel)	

]	PANEL A: 2	Total Crime		PANEL B: Robbery				
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	
Local Treatment Effect, $\hat{\tau}$	21.528** (0.032)	20.636** (0.045)	24.037*** (0.003)	23.131*** (0.007)	0.204 (0.824)	0.636 (0.951)	-0.034 (0.594)	0.269 (0.897)	
Bandwidth Selection, \hat{h}									
Mean Squared Error (MSE)-Optimal	Yes	No	Yes	No	Yes	No	Yes	No	
Coverage Error Rate (CER)-Optimal	No	Yes	No	Yes	No	Yes	No	Yes	
Number of Selectors	One	One	Two	Two	One	One	Two	Two	
Days Before Cutoff	106	71	178	121	98	67	156	106	
Days After Cutoff	106	71	132	90	98	67	95	65	
Number of $Observations$	212	142	310	211	196	134	251	171	

Robust probability values in parentheses; *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 levels, respectively.

		PANEL A:	Total Crime		PANEL B: Domestic Violence				
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	
Local Treatment Effect, $\hat{\tau}$	21.528** (0.032)	20.636** (0.045)	24.037*** (0.003)	23.131*** (0.007)	1.092 (0.906)	-1.017 (0.553)	1.050 (0.986)	-0.398 (0.689)	
Bandwidth Selection, \hat{h}									
Mean Squared Error (MSE)-Optimal	Yes	No	Yes	No	Yes	No	Yes	No	
Coverage Error Rate (CER)-Optimal	No	Yes	No	Yes	No	Yes	No	Yes	
Number of Selectors	One	One	Two	Two	One	One	Two	Two	
Days <u>Before</u> Cutoff	106	71	178	121	85	58	188	129	
Days After Cutoff	106	71	132	90	85	58	75	51	
Number of Observations	212	142	310	211	170	116	263	180	

Table 12: RD Point Estimates, $\hat{\tau}$, by Offense Type (local-linear method; triangular kernel)

Robust probability values in parentheses; *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 levels, respectively.