

How Police and Crime Respond to Courts: Evidence from Colombia

Nicolás Idrobo* Dorothy Kronick[†] Tara Slough[‡]

August 14, 2026[¶]

Abstract

How does reform of criminal prosecution and courts affect policing and crime? We study the introduction of a new code of criminal procedure in Colombia. Though the new code primarily targeted prosecutors and judges, we find that police behavior changed in response: the arrest rate dropped 48%, and arrests for minor offenses declined even more. But neither administrative data nor survey data provide evidence of an increase in crime, and perceptions of neighborhood safety improved. We attribute this outcome to an improvement in court accuracy. Our findings underscore the value of considering strategic interaction among the police, prosecutors, and judges.

*Postdoctoral Research Associate, Department of Politics, Princeton University

[†]Associate Professor, Goldman School of Public Policy, University of California, Berkeley

[‡]Associate Professor, NYU Department of Politics

[§]Special thanks to Avi Feller and Gonzalo Vazquez-Bare for extensive guidance. For comments, we thank Guy Grossman, Roman Rivera, Carolina Torreblanca, Jessie Trudeau, Simeon Nichter, Mike Gibilisco, workshop participants at AL CAPONE, ESOC, the Penn Conference on Policing in Comparative Perspective, the USC Conference on New Directions in the Political Economy of Development, the U.C. Berkeley Methods Workshop, Tulane University, Duke University, the Center for International Security and Cooperation at Stanford, and the Universidad de Chile.

[¶]We used Codex v0.145.0 from OpenAI and Opus 5 from Anthropic to check human-written code and incorporate it into a replication package; we then re-checked all replication code ourselves. We also used Claude Sonnet 5 from Anthropic for proofreading text (which we wrote), checking the formal model (which we wrote), formatting the figures in Appendix B (which we conceived), and locating literature that we had missed (which we then read). We reviewed all Claude-suggested changes and corrections before incorporating them.

Scholars have long understood that police respond to other criminal justice institutions. [Wilson \(1968\)](#), for example, noted that police officers arrest people they expect prosecutors to indict (84) and judges to convict (137). [Chevigny \(1995\)](#), studying criminal justice in Latin America, observed that police officers used extralegal violence when they were impatient with courts (143, 181; see also [Brinks, 2007](#)); [Skolnick and Fyfe \(1993, 24\)](#), similarly, wrote that police in the United States use excessive force when they deem courts “too ponderous, too indolent, too unaware, or too constrained to deal with ‘the problem.’”

This work implies that reform targeting prosecutors and the courts could reshape police behavior, with consequences for crime. But we lack theory and evidence. Recent empirical research on policing focuses on policies internal to police forces (e.g. [Mummolo, 2018](#); [Owens et al., 2018](#); [Rivera, 2025b](#); [Dube et al., 2025](#)) or on court orders that impose direct constraints on police agencies (e.g. [Trudeau, 2022](#); [Rivera and Ba, 2023](#); [Hausman and Kronick, 2023](#)). And recent work on prosecutors and courts considers effects on crime (e.g. [Ouss and Stevenson, 2023](#); [Shem-Tov et al., 2024](#)), but seldom studies police response ([Stashko and Garro, 2021](#); [Amaral et al., 2025](#), being exceptions).

How does reform of criminal prosecution and courts affect policing and crime? We use the staged rollout of a new code of criminal procedure in Colombia to study the strategic responses of police and citizens, and the consequences for public safety. Though the new code targeted the public prosecutor’s office, the investigative police,¹ and the courts—not patrol officers—we find that patrol officers’ responses to these downstream changes reshaped law enforcement in Colombia, with unexpected effects on crime.

Colombia’s new code of criminal procedure replaced a primarily inquisitorial system, in which the eponymous inquisitor takes part both in criminal investigation and in judicial decision-making, with an accusatorial system, which separates the two: prosecution and defense investigate, judges adjudicate. The new code changed the work of public prosecutors and judges in many ways, including how they handled arrests of suspects caught in the act (*en flagrancia*). Under the old code, prosecutors could review the legality of these arrests; the new code transferred this power to judges. From the perspective of patrol officers, this change meant that each arrest took more time (because judges asked more questions) and faced more scrutiny (because judges applied stricter procedural standards). The police and prosecutors no longer “had to convince themselves, but had to convince a judge,” as one prominent criminal defense attorney put it (quoted in [Hartmann Arboleda et al., 2009, 80](#)).

We propose a theory of how such changes affect policing and crime. In our model, reform makes courts more *accurate*: more likely to convict guilty suspects and less likely to convict

¹In Colombia as in many Latin American countries, investigative police are not integrated in units alongside patrol officers; rather, they operate within dedicated units.

the innocent. All else equal, this improvement in accuracy would strengthen deterrence and therefore reduce crime. But increasing the accuracy of criminal courts also imposes costs on the police. In order to enable prosecutors and judges to make better decisions, officers must fill out additional paperwork, for example, or attend hearings. These costs constitute a ubiquitous but seldom-studied consequence of reform, and they tend to depress police effort.² That’s why the effect of reform on crime is ambiguous: accuracy strengthens people’s incentives to obey the law, but the consequent decline in police activity has a countervailing effect.

Colombia’s new code of criminal procedure sought to improve court accuracy, minimizing punishment of the innocent and facilitating conviction of the guilty. Our model implies that reform targeting court accuracy should also affect the arrest rate. Using the staged rollout of Colombia’s new code, we find that officers responded by making fewer arrests: the arrest rate dropped 48% when the new code came into effect, a fact not documented before our work.³ Because the new code primarily targeted prosecutors, investigative police, and the courts, effects on patrol officers were overlooked.

All else equal, we might expect this large drop in the arrest rate to result in higher crime rates. Indeed, studies on the police elasticity of crime generally estimate that a 1% decline in personnel causes a $\approx 0.7\%$ increase in the homicide rate and a 0.3%–1% increase in assault (Chalfin and McCrary, 2018, 179–180), meaning that it would be unsurprising if a 48% decrease in arrests raised crime by 15% or even 30%. And while the few empirical estimates of the *arrest* elasticity of crime generally suggest smaller responses (e.g. Rivera, 2024), an influential twentieth-century literature claimed that even low-level arrests could prevent serious crime. “Not every petty criminal is a serious criminal,” Kelling and Coles (1996, 146) observed, “but enough are, or have information about others who are, that contact with petty offenders alerts all criminals to the vigilance of the police.”

But we propose a mechanism through which a drop in the arrest rate might not drive an increase in crime: when it occurs as the result of an increase in court accuracy, which deters crime. While accuracy is impossible to quantify (because we do not observe true guilt or innocence), qualitative literature suggests that Colombia’s new code of criminal procedure did improve it: for example, by leveling the playing field between prosecution and defense (Hartmann Arboleda et al., 2009, 77); by facilitating punishment of suspects

²Court accuracy itself can raise police effort if officers are sufficiently public-service motivated, as we discuss below.

³For example, the thorough treatment in Hartmann Arboleda et al. (2009) does not discuss consequences for the number of arrests. The first quantitative study of Colombia’s new code of criminal procedure, Acosta Mejía et al. (2016), found large effects on crime and procedural outcomes but not on arrest. A subsequent draft, Zorro Medina et al. (2020), replicated our arrest result, citing a previous version of this article.

caught *en flagrancia* (36); by dramatically increasing resources for public defenders (34); and by shortening the time to trial (Hartmann et al. 2007, 110; Zorro Medina 2023), thereby avoiding punishment of nonoffenders in pre-trial detention. “Our starting point is that the system has obtained very positive results in a very short time,” Hartmann Arboleda et al. (2009, 38) concluded, “even in the face of huge challenges.”

Perhaps as a result, in Colombia, we find no evidence that crime rose. Using data both from vital statistics and from the police, we find no evidence that Colombia’s new code of criminal procedure affected homicide rates or assault rates. And using survey data from the Latin American Public Opinion Project, we estimate that the new code of criminal procedure likely improved perceptions of neighborhood security, and that it did not affect self-reported victimization one way or the other. In both cases, the estimates are noisy and our measures imperfect, which is why we characterize these results as *no evidence of a crime wave* rather than *evidence of no increase in crime*. But the data are good enough that, using inverted hypothesis tests, we can rule out (at 95%) homicide-rate increases greater than 4%.

A third implication of our theory is that reform targeting prosecutors and judges can affect not only the number but the composition of arrests, because the new costs of reform fall especially on unjustified or under-justified arrests, which previously escaped scrutiny. We find that arrest rates dropped more for minor offenses like vandalism or drugs than for serious crimes; homicide arrests, for example, did not decline.

Our work is closely related to recent empirical studies on Latin America’s “revolution in criminal procedure” Langer (2007). Legal scholars, along with the international financial institutions that backed these reforms in the first place, have produced scores of qualitative studies on the consequences (e.g. Alguíndigue and Pérez-Perdomo, 2008; Alguíndigue and Pérez-Perdomo, 2013; Duce and Perdomo, 2003). Among the few quantitative studies, Magaloni and Rodriguez (2020) find that Mexico’s new code of criminal procedure reduced the use of torture, Tiede (2012) documents a reduction in the use of pre-trial detention in Chile, and Cattaneo et al. (2022a) find that Uruguay’s new code caused property crime.⁴ Ours is the first of these papers to focus on police officers’ strategic response to changes in the public prosecutor’s office and the courts.

Our findings also contribute to recent literature on police behavior. Many papers consider internal policy changes and/or external mandates (Mummolo, 2018; Trudeau, 2022; Rivera and Ba, 2023; Hausman and Kronick, 2023; González, 2020; Magaloni et al., 2020). Other work considers the role of officer characteristics such as race and gender (Ba et al., 2021),

⁴One working paper, Acosta Mejía et al. (2023), also studies Colombia’s new code of criminal procedure and finds a large positive effect on crime, concluding that there is a tension “between due process protection and public safety.” We attribute the difference in our results to our correction for coincident changes in crime registration, discussed below and in Appendix C.

tendency for aggression (Rivera, 2025a), or political affiliation (Ba et al., 2023). With the exception of Stashko and Garro (2021), who study police response to prosecutor elections, none of this literature focuses on how police officers respond strategically to reform targeting other agents of the criminal justice system.

More generally, our findings relate to literature on incentives and behavior in public bureaucracies. Even though Colombia’s new code of criminal procedure had no direct effect on police wages, selection of officers, or politician oversight, we identify large changes in police behavior, underscoring the value of focusing on interaction among public-sector institutions (Grossman and Slough, 2022).

Our theory and findings have policy implications beyond Colombia. Across Latin America and the United States, various forms of depolicing arise as byproducts of efforts to minimize punishment of the innocent or to facilitate punishment of the guilty in court. Our framework clarifies that the consequences for public safety turn less on the magnitude of the decline in policing than on how that decline compares to accompanying improvements in the rate of punishing the right people. Where depolicing outpaces improvements in accuracy, we might expect a crime wave; where accuracy keeps pace, police can pull back without causing crime.

1 Theory: Courts and cops

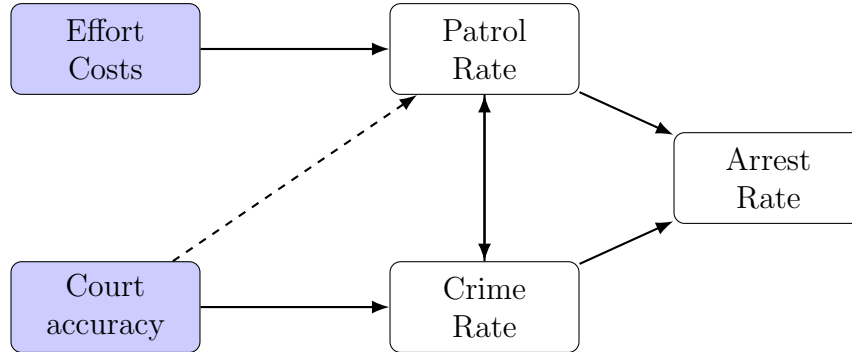
We propose a simple model for thinking about how policing and crime respond to reform of criminal prosecution and courts (formalized in Appendix A). People choose whether to work or commit crime, and the relative gains from crime vary across people. Those who commit crime incur a risk of arrest and, if arrested, a risk of conviction, in which case they pay a penalty (as in Becker, 1968). But critically, we also allow for police and courts to make mistakes: even those who do not commit crimes can be arrested and convicted.

Police choose how much to patrol. Patrol creates opportunities for making arrests but also requires costly effort. Officers pursue arrests both because of extrinsic career rewards and because they intrinsically value seeing guilty people punished. We assume that the extrinsic benefits do not depend on the true guilt or innocence of the suspect, whereas the intrinsic benefits accrue only if the suspect is actually guilty and only if the courts actually punish him; in that sense, the intrinsic benefit is related to public service motivation (e.g., Francois, 2000; Dal Bó et al., 2013; Ash and MacLeod, 2015). The model accommodates any relative weight on extrinsic vs. intrinsic benefits, and our main results hold across the full range of police motivation.

We model reform as a policy shock to two parameters of our model: the accuracy of the courts, and the cost of police effort.

Figure 1: How reform affects patrol, arrests, and crime

The shaded nodes represent the two components of reform; the white nodes are equilibrium outcomes.



We define *court accuracy* as the probability that arrested offenders are punished and arrested nonoffenders are released or acquitted.⁵ In other words, accuracy is one minus the court error rate: one minus the probability of wrongful convictions and mistaken acquittals, conditional on arrest. Reform improves court accuracy, increasing the probability of punishment for arrested offenders and lowering the probability of punishment for nonoffenders who were wrongly arrested. More accurate courts make fewer mistakes.

Improving court accuracy imposes effort costs on police. Officers must fill out additional paperwork, for example, or attend hearings. These costs should depress police effort. That accurate prosecutorial and judicial decision-making requires additional effort from police is a very general but under-recognized reality of criminal justice reform.⁶

For a given level of court accuracy and cost of police effort, officers choose how much to patrol and people choose whether to commit crime. The *patrol rate*, the *arrest rate*, and the *crime rate* emerge as equilibrium outcomes. Figure 1 summarizes how reform affects these outcomes.

Patrol rates: The reform affects the patrol rate in three ways. First, higher effort costs depress patrol effort, all else equal. Second, court accuracy directly deters crime, which weakens officers’ incentives to patrol (because patrolling yields fewer arrests).⁷ Third, accuracy increases officers’ intrinsic reward for making arrests, which can raise patrol effort if officers are sufficiently public-service motivated. Whether the third effect dominates the first two is ambiguous ex-ante.

⁵In the model, we make the simplifying assumption these probabilities are equal; in a less parsimonious model that allows them to differ, our results hold: crime and arrests respond to court accuracy in the same ways.

⁶This feature of our model echoes arguments that oversight creates paperwork burdens, depressing bureaucrats’ effort (e.g., Wang, 2022).

⁷Of course, the patrol rate and the crime rate are determined in equilibrium.

Arrest rates: Arrest rates are a function of the crime rate and the patrol rate. While the effect of the reform on arrest rates is technically ambiguous, over much of the parameter space higher court accuracy depresses the equilibrium arrest rate by raising police effort costs and directly deterring crime.⁸

Crime rates: The effect of court accuracy on crime is similarly ambiguous. Accurate prosecution and adjudication directly lower the crime rate by increasing the probability of punishment for breaking the law and decreasing the probability of wrongful punishment of law-abiding citizens. Accuracy can also lower crime indirectly if it increases the patrol rate. But accuracy also imposes new costs on police, which depress police effort and thereby can increase crime. Which of these effects dominates depends on the details. This is the central lesson of our model: reform that increases accuracy significantly while imposing minimal costs on police will lower crime; reform that increases accuracy only moderately while imposing larger costs on police will result in more crime.

2 Criminal procedure and policing in Colombia

Colombia’s new code of criminal procedure: origins and objectives. In 2004, the Colombian Congress passed a new code of criminal procedure (Ley 906). Criminal procedure defines the rules according to which criminal cases proceed through the justice system; the penal code, which is separate, defines crimes and penalties.⁹

The new code replaced what was primarily an inquisitorial system (typical of civil law) with an accusatorial system (typical of common law).¹⁰ The core distinction is that accusatorial systems separate investigation from adjudication, rather than empowering an inquisitor to do both. Under Colombia’s old code, prosecutors exercised some judicial powers (such as issuing arrest warrants and deciding on pre-trial detention) and judges could assume prosecutor powers (such as introducing evidence). The new code, in contrast, separated in-

⁸This is not a probabilistic statement: we have not made assumptions about which parameter values are more likely.

⁹Though the Ley 906 did not define or redefine sentence lengths directly (being the criminal procedure code, not the penal code), it did indirectly shorten sentence lengths by allowing for negotiated plea deals. At the same time, though, and explicitly in order to offset plea-induced sentence reductions and to incentivize plea bargaining, Colombia raised sentence lengths through a revision of the penal code (Ley 890 of 2004). Critically for our analysis, the Ley 890 did *not* redefine any of the crimes that we study (meaning, it did not change what constitutes vandalism, for example). While the Ley 890 was statutorily set to come into effect on January 1, 2005, courts later (in 2006) linked it to the rollout of the criminal procedure code ([Corte Suprema de Justicia de Colombia, Sala de Casación Penal, 2007, 2013](#)). Whether these sentence-length increases dominated the plea-induced reductions or vice-versa, or netted out, has not been studied.

¹⁰We write *primarily* inquisitorial because, in fact, Colombia’s purely inquisitorial system was replaced in 1992 with a mixed system that introduced several features of accusatorial procedure, including empowering prosecutors (rather than judges) to direct investigations. The Ley 600, passed in 2000, which we refer to as the *old code*, continued with the mixed inquisitorial/accusatorial system.

vestigation from adjudication. A new type of judge, the *juez de control de garantías*, assumed judicial powers previously exercised by prosecutors; trial judges, meanwhile, were explicitly forbidden from assuming prosecutorial powers like introducing evidence.

Inquisitorial systems are sometimes thought to be more accurate than accusatorial systems (Merryman and Pérez-Perdomo, 2018, 135). In theory, the inquisitor pursues incriminating evidence and exculpatory evidence with equal zeal, in which case the full force of state resources fall behind the search for truth, a setup possibly superior to the accusatorial tradition of pitting state prosecution against private or under-funded public defense. But in practice, throughout Latin America, inquisitors did not function as neutral investigators or arbiters; they functioned as super-powered prosecutors, leaving the accused without much of any defense at all (Langer, 2007). Paradoxically, these same systems were simultaneously seen as too slow to prosecute crime (Binder, 1993, 209, quoted in Langer, 2007, 633), in large part because cases proceeded via plodding written exchanges rather than oral hearings.

In response to these twin failures—insufficient protections for suspects alongside deficient prosecution of offenders—an international network of legal activists promoted a switch from written inquisitorial to oral accusatorial criminal procedure throughout Latin America (Langer, 2007). Colombia’s new code was a product of this process of international diffusion.

How the new code promoted accuracy. Colombia’s new code of criminal procedure was designed to improve the *accuracy* of prosecution and adjudication. By improving accuracy, we mean both reducing the rate of false convictions (Type I errors) and reducing the rate of failed prosecutions (Type II errors), conditional on arrest.

The new code sought to do both. Some of the changes primarily targeted false convictions: the introduction of public defenders for all defendants, for example, or separating investigation from adjudication, so as to level the playing field between prosecution and defense (under the old code, the prosecution “was in fact far better positioned than the accused to prevail” Hartmann Arboleda et al., 2009, 77). Other changes primarily targeted failed prosecutions: the introduction of plea bargaining, for example, or fast-track processing for suspects caught *en flagrancia*, which “contributed decisively to legitimizing the new system” Hartmann Arboleda et al. (2009, 93).

There were also components of the new code that targeted both types of errors. The replacement of written proceedings with oral trials, for example—a core component of the reform—dramatically shortened the time to case resolution (Hartmann Arboleda et al., 2009); this celerity simultaneously prevented the deterioration of evidence and witness credibility (thereby reducing failed prosecutions; indeed, celerity was one of the aspects of the reform that “found some opposition among leftists,” Hartmann Arboleda and Martínez Sánchez,

2009, 214) and also shortened de facto punishment of innocent parties in pre-trial detention (a form of Type-I error). For additional detail on how the new code changed criminal procedure in Colombia, see Appendix B.

In conceptualizing Colombia’s new code as a slate of changes aimed at improving accuracy, we are not merely simplifying for the sake of our model. Rather, we see this gloss as a useful way for researchers to understand what reformers all across Latin America sought to achieve by replacing inquisitorial with accusatorial codes of criminal procedure: to protect defendants’ rights while also fighting crime, two goals often seen as competing. Pursuing both helped generate political consensus in favor of the transition to accusatorial systems (Duce et al. 2009, 54; see also Laterzo-Tingley 2026 on the limitations of a single punitive–lenient dimension in describing criminal justice policy).

Implications for police. Colombia’s new code of criminal procedure discussed at length the roles and responsibilities of three groups of officials: judges, public prosecutors, and the judicial police, i.e. investigative police, who operate within dedicated investigative units (as is common across Latin America). Patrol officers and their supervisors, who account for the majority of all police in Colombia and indeed who are more numerous than judges, prosecutors, and judicial police combined (see Figure B.2), were almost entirely absent from the text of the new code. They are mentioned indirectly and in passing, such as when the new code specifies that a citizen stopping a crime in progress should bring the suspect “before any police authority” (Article 302), or as possible substitutes for judicial police in remote areas (Article 201). But the text of the new code did not significantly redefine the responsibilities of regular police officers.

One explanation for the near-absence of patrol police from the text of Colombia’s new code is that, relative to the inquisitorial codes replaced elsewhere in Latin America, Colombia’s “old code” was in fact recent. Whereas the new, accusatorial codes in Venezuela, Mexico, and Chile, for example, replaced inquisitorial codes that had been on the books for more than half a century (Langer, 2007), Colombia’s Ley 906 de 2004 (the new code) replaced the Ley 600 de 2000, which had already strengthened protections for suspects and defendants, and which itself built on trends that began with the 1991 constitution. For example, the “detención por sospecha” arrests (i.e. arrests “for investigative purposes”) that some accusatorial codes elsewhere in Latin America disallowed (in Venezuela, for example; Hanson and Kronick, 2026), in Colombia had been ruled unconstitutional in 1994 (Corte Constitucional de Colombia, 1994).

Despite the focus on other agents of the criminal justice system in the text of the new code, the work of patrol officers changed. The most salient change, from officers’ perspective,

was the introduction of judicial control of arrests (see Appendix B for additional detail). Under the previous code, public prosecutors (*fiscales*) could sign off on arrests in *flagrancia* (meaning, arrests without a warrant, because the suspect was caught in the act). This prosecutor power meant that arrests without warrants were typically declared legal—so much so that one former officer described the process as “legalizing” arrests, in scare quotes.

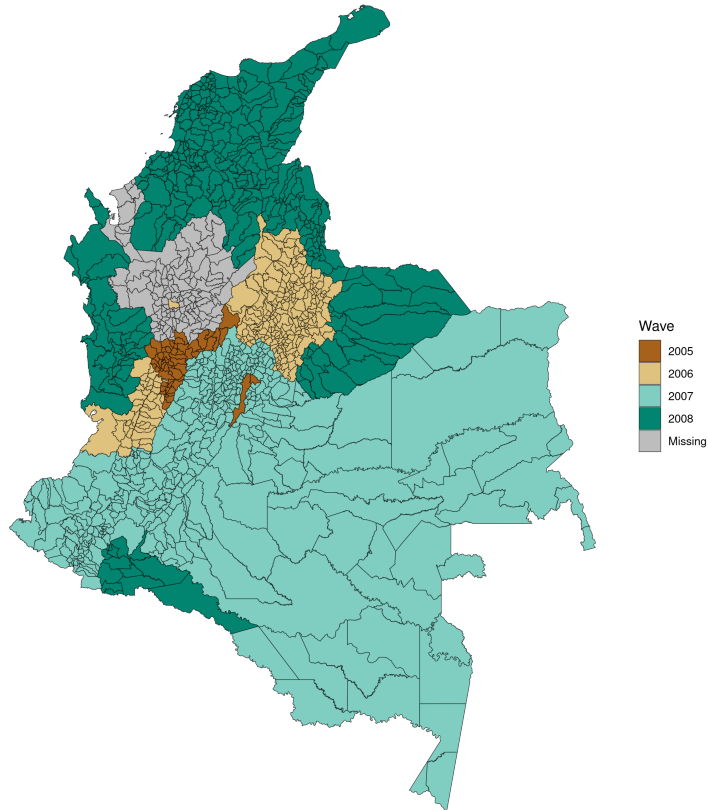
Under the new code, in contrast, *judges* had to sign off on *flagrancia* arrests. Beyond annoying the prosecutors, who “all agreed that the judges were responsible for letting subjects go” (Hartmann Arboleda et al., 2009, 85), the transfer of responsibilities from prosecutors to judges imposed substantial new time costs on the police. An officer making a *flagrancia* arrest had to bring the suspect before a prosecutor, who conducted an initial arrest-legality screening, and who then had to attend a hearing with a judge who would evaluate the legality of the arrest; though the new code did not require police officers to attend these hearings, prosecutors often called the arresting officer as a witness, so that he could describe the circumstances of the arrest. One interviewee told us that these hearings could take 45 minutes (plus travel time); a 2015 survey of 96 patrol officers found that legalizing a *flagrancia* arrest took more than five hours, on average (excluding time to travel to the prosecutor’s office; Gutiérrez and Tobón, 2019).

Moreover, many officers perceived the judges as more *garantista* than the prosecutors under the old code, meaning that the judges were more likely to declare an arrest illegal (because officers violated due process, or because available evidence did not justify the arrest). The new code may also have made prosecutors stricter, in anticipation of judicial review. Under the new code, then, *flagrancia* arrests required substantially more officer time, with a lower probability of sign-off.

Rollout. Colombia’s new code of criminal procedure was rolled out in four stages, mapped in Figure 2. The four stages divide the Colombian population in quarters, with approximately 10 million people in each wave. *De jure*, the law assigned Antioquia (the district surrounding Medellín) to the third wave (2007), but some Antioquia courts appear to have ended up *de facto* in the 2006 wave (with Medellín); for that reason, we exclude Antioquia from the primary analysis. Including it either in the 2006 or in the 2007 wave makes little difference (see Appendix Figure D.2). The judicial district of Yopal was originally assigned to the 2006 (second) wave but ended up in the 2008 wave (Piraquive Sierra, 2007).

Policing in Colombia. In the mid-1990s, ten years prior to the reform that we study, far-reaching legislation transformed the Colombian national police. What had been a corrupt, militarized, under-educated and under-trained police force embroiled in a violent conflict

Figure 2: Rollout Map



with drug cartels became a less corrupt, less violent, partially demilitarized, better-educated and better-trained agency newly focused on citizen security ([González, 2019](#)). It remained deeply flawed ([Llorente, 2005](#)). But, by the 2000s, the Colombian national police—the only major police force in the country—had experienced “an unexpected shift toward democratic policing” ([Moncada, 2009](#)). Even as President Álvaro Uribe sent the Colombian military (and specific specialized police units) out to fight the guerrilla, producing both a spectacular decline in lethal violence (Figure F.1) and also grave human rights violations (e.g. [Acemoglu et al., 2020](#)), the national police reoriented activities around certain principles of community policing ([García et al., 2013](#)). This is all to say that, by the time of the reform that we study, the Colombian national police was “a professionalized and well-regarded police force” ([Blattman et al., 2021](#)).

3 Data

To analyze the effects of the new code of criminal procedure, we use data from the police, from vital statistics, and from surveys.

Police data. The Colombian national police began collecting systematic data on operations and on reported crimes in 1958, via paper forms that regional offices filled out and remitted to Bogotá for tabulation, and later via Excel. Only in 2003 did the police develop and introduce dedicated software for counting crimes and documenting police activity: the System for Statistical Information on Crimes, Infractions, and Operations, or SIEDCO.

Via a freedom-of-information request, we obtained SIEDCO’s count of the number of arrests by suspected crime (such as theft or assault), by type of arrest (with a warrant or in *flagrancia*), in each municipality ($N = 1,102$) in each month between 2003 and 2011. This count includes all arrests, regardless of whether the arrest was subsequently declared legal or illegal, and regardless of whether the suspect was ultimately indicted.

We also obtained SIEDCO’s count of crimes in each municipality in each month, via freedom-of-information request. But these data suffer from three types of changes in crime registration. First, in some cities, police introduced new sources of crime counts into SIEDCO, creating the illusion of a huge increase in crime (see Appendix C; [Policía Nacional de Colombia 2005](#), 36). Second, in Bogotá specifically, police had previously counted thefts in SIEDCO only when the value of the stolen goods exceeded ten times the monthly minimum wage; coincident with the implementation of the new code, police began to record thefts below that threshold ([Cámara de Comercio de Bogotá, 2006](#)). And third, the incorporation of a new modes of crime reporting likely increased the number of citizen reports included in SIEDCO crime counts ([Rodríguez-Ortega et al., 2018](#); [Rodríguez Cuadrado, 2008](#), 116).

For two major crimes—assault (*lesiones personales*) and mugging (*hurto a personas*)—we are able to partially address the first problem (the introduction of new sources into SIEDCO) by excluding these sources from our count of crimes. The resulting measure is likely an undercount, but does not suffer from the problem of a dramatic change in crime registration (see Appendix C for details). For other crimes, such as car theft, the partial correction illustrated in Appendix C is not possible because some cities already incorporated data from external sources beginning in 2003.¹¹

These changes in registration likely did not affect SIEDCO’s count of homicides. We therefore use SIEDCO’s count of homicides in each municipality in each month as one of two measures of intentional lethal violence. [Reed and Ball \(2016\)](#) document certain forms of under-registration in the police count of homicides; for that reason, we also use counts from vital statistics.

¹¹We also obtained data on car theft from the association of vehicle insurers. But because only one-fifth of vehicles are insured, these data are extremely sparse; three-quarters of all reported thefts occur in just three municipalities (Bogotá, Cali, and Medellín). We discuss these data in Appendix H.

Vital statistics. As a second source of data on homicides, we use the mortality microdata published by Colombia’s national statistics institute. Death certificate coverage improved quickly in Colombia in the period prior to the reform that we study, climbing from an estimated 80% in the early 1990s to 99% by 2009 ([World Bank, 2026](#)). Like the police count of homicides, the vital statistics count is incomplete; future work could supplement our analysis by conducting case studies that use press reports and NGO counts to construct more complete homicide series (as in [Guberek et al., 2010](#), for Casanare).

Survey data. We also use data from the Latin American Public Opinion Project (LAPOP) to measure self-reported crime victimization and perceptions of neighborhood security. The LAPOP data for Colombia identify each respondent’s municipality, allowing us to locate respondents within judicial districts and thereby within the correct survey wave. It is not a panel; rather, the LAPOP data provide a repeated cross-section. We aggregate these data to the level of municipality-year or, for some specifications, judicial district-year. Each municipality-year observation consists of an average of the responses of six to 231 respondents. Our repeated cross sections consist of 49 municipalities and 26 judicial districts, for a total of 297 municipality-years or 156 judicial district-years. LAPOP is representative of the adult population, not of all municipalities; as with our analysis of arrest rates, we report both unweighted and population-weighted estimates.

4 The effect of the new code on arrest rates

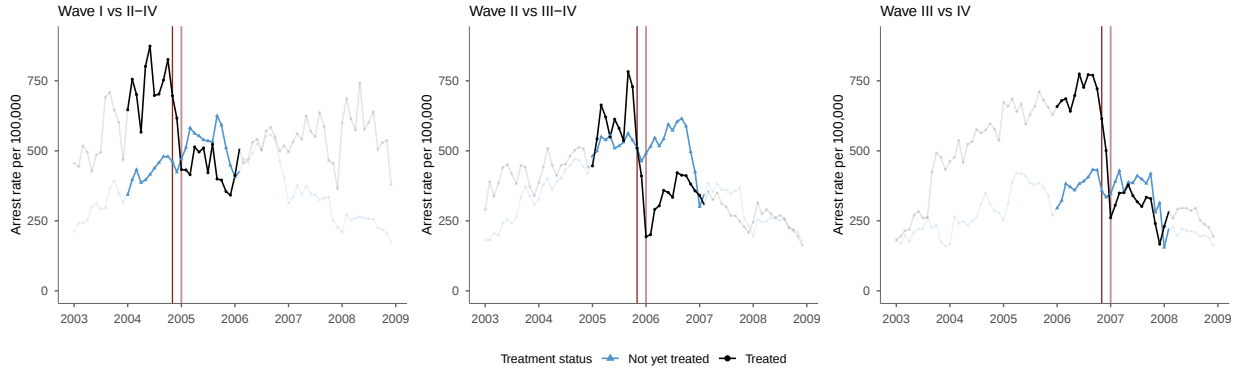
Though Colombia’s new code of criminal procedure primarily targeted the functioning of prosecutors, courts, and judicial police, we find that patrol officers responded to new scrutiny from judges by sharply reducing the number of arrests.

In 2003–2004, the two years before the new code of criminal procedure came into effect in the first judicial districts (see [Figure 2](#) above), Colombia’s national arrest rate was approximately 750 per 100,000 per year—about one-third of current arrest rates in the United States, and similar to arrest rates in many countries in Latin America and Europe. When the new code came into effect, the arrest rate plummeted. In Bogotá, for example, the arrest rate declined by 50% within months, from approximately 1,300 per 100,000 per year in 2003–2004 to 600 per 100,000 per year in 2005 (see [Appendix Figure D.1](#)). In Cali, the third-largest city in Colombia, part of the second wave; Neiva, a mid-size city in the third wave; and Valledupar, in the fourth wave; arrest rates similarly collapsed.

These examples are not atypical. Indeed, arrest rates fell in 78% of Colombia’s 1,101 municipalities in the twelve months following the implementation of the new code, compared

Figure 3: The Effect of the New Code on Arrest Rates

The black lines plot average municipal arrest rates in Waves I, II, and III, respectively. The blue lines plot average municipal arrest rates in not-yet-treated waves.



to the twelve months prior (Appendix Figure D.3).

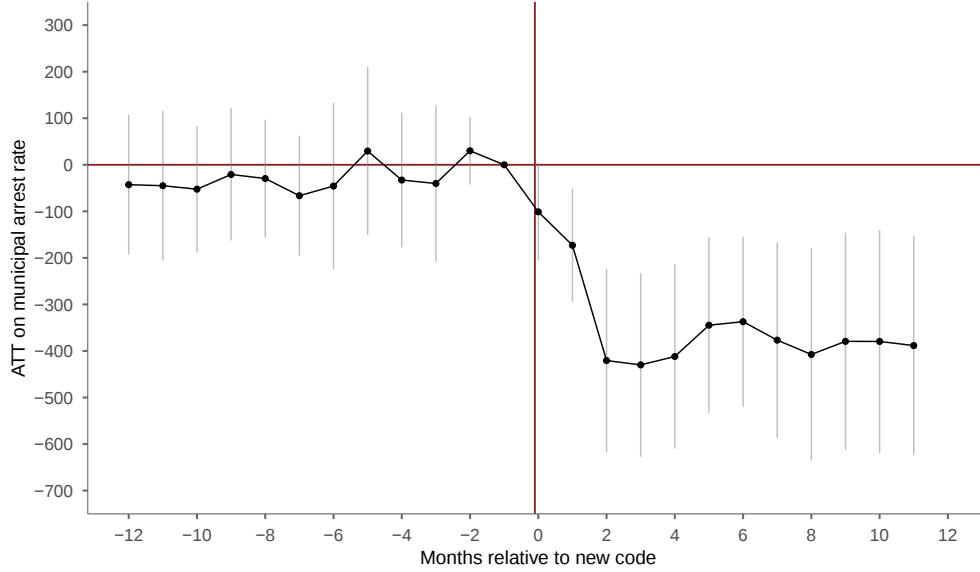
To estimate the effect of the new code on arrest rates, we use difference-in-differences. Figure 3 plots the raw data by wave. Though the data are monthly, we annualize the rates ($\times 12$) so that the scale is more familiar. The left panel compares the mean arrest rate (across municipalities) in the first wave—Bogotá and the coffee belt, where the new code came into effect in January, 2005—against the mean arrest rate in Waves II, III, and IV. In 2004, arrest-rate trends were similar in these two groups (Wave I vs. Waves II–IV); then, when the new code came into effect in the judicial districts of Wave I, arrest rates diverged: they fell sharply in Wave-I municipalities while continuing to climb elsewhere. A similar pattern emerges in Waves II and III.

These figures also indicate that arrest rates began to fall one or two months prior to the implementation date of the new code. The second, lighter red vertical lines in each figure mark January of 2005, 2006, 2007, and 2008 (when the new code came into effect in each wave); the first, darker red vertical lines mark the previous November. This anticipation likely stems from two forces. First, there is seasonality in police activity in general: Figure 3 reveals a November–December drop in arrests even in untreated units. But unlike a *purely* seasonal phenomenon, arrests in treated places don’t bounce all the way back up after the holidays. Second, Jairo Sánchez Otalora, special delegate of the Public Prosecutor’s Office to the implementation committee for the new criminal procedure code, explained to us that:

Everything was done at the last minute (*sobre el tiempo*). Since the priority was training prosecutors and judges, training for police was lower priority and didn’t start until the second half of the year. From that point on, a great deal of expectation and fear was generated among police officers regarding how to carry

Figure 4: Arrest Rates Drop As the New Code Comes into Effect

Estimates of Equation 1. Confidence intervals use 1,000 bootstrap replications at the judicial-district level.



out lawful arrests and regarding possible sanctions if they were not done properly.

That is why arrests dropped in the months leading up to implementation.

The November–December pre-trend would be concerning if it might signal endogeneity rather than anticipation: if, for example, judicial districts where arrests were falling for other reasons decided to adopt the new code because of those same underlying policy trends, our estimates would be invalid. But the rollout was fixed in statute rather than determined by local decision-making. We therefore accommodate this anticipation by coding November (rather than January) as the treatment date.

To estimate the average treatment effect on the treated (ATT) at each month *relative* to when the new code came into effect, we use the estimator proposed by [De Chaisemartin and d’Haultfoeuille \(2020\)](#) (other estimators produce similar results; see Appendix D). Let i index municipalities, G index reform waves, F_G denote the month in which the new code comes into effect in wave G (shifted to accommodate anticipation, such that $F_1 = \text{November, 2004}$, for example), S_G contain the municipalities in wave G , and $Y_{i,t}$ denote outcome Y in municipality i at time t . For event month $h \in \{0, \dots, 11\}$, let C_G contain the municipalities in waves

Table 1: The Effect of the New Code on Arrest Rates

	Municipal level			Judicial district level					
	Arrest rate			Homicide		Drug		Vandalism	
	Rate	Rate	Log	Rate	Log	Rate	Log	Rate	Log
<i>Panel A. Unweighted</i>									
Effect of new code	-345.81 (84.80)	-378.98 (90.28)	-0.47 (0.09)	-1.28 (2.68)	-0.02 (0.17)	-72.79 (42.82)	-0.62 (0.18)	-64.35 (24.46)	-1.03 (0.27)
Observations	31,056	972	972	972	972	972	972	972	972
Change (%): log-transformed			-37.6		-1.7		-46.1		-64.3
Change (%): avg. unit	-30.3	-36.2		4.5		-36.2		-67.8	
Pre-period mean	571.29	744.22	6.48	18.06	2.76	159.22	4.55	54.64	3.21
<i>Panel B. Population-weighted</i>									
Effect of new code	-431.28 (86.37)	-430.42 (83.42)	-0.52 (0.09)	-0.77 (1.77)	-0.02 (0.10)	-91.34 (53.70)	-0.71 (0.21)	-92.77 (34.64)	-1.27 (0.33)
Observations	31,056	972	972	972	972	972	972	972	972
Change (%): log-transformed			-40.3		-2.3		-50.9		-71.8
Change (%): aggregate rate	-47.8	-47.8		-4.1		-43.6		-113.7	
Change (%): avg. unit	-35.5	-36.8		-0.1		-37.1		-66.2	
Pre-period mean	901.58	900.53	6.68	18.69	2.87	209.52	4.96	81.57	3.76

Estimates of Equation 1, aggregated over twelve months. Standard errors (in parentheses) come from a district-level pairs bootstrap with 1,000 replications. Antioquia is excluded. Rates are annualized arrests per 100,000 residents; log outcomes are $\log(\text{rate} + 1)$. Percentage changes use the 12 months before treatment as the baseline. For log outcomes, they're $100 \times (e^b - 1)$. For rate outcomes, the avg. unit row reports the average proportional change after dividing each unit's drop by its own pre-treatment mean. In Panel B, the aggregate-rate row divides the population-weighted ATT by the population-weighted pre-treatment mean, which is the change in the aggregate per-capita arrest rate.

$G + 1$ through 4, which have not yet adopted the new code. We calculate:

$$\begin{aligned}
\widehat{\text{DID}}_{G,h} &= \frac{1}{|S_G|} \sum_{i \in S_G} (Y_{i,F_G+h} - Y_{i,F_G-1}) - \frac{1}{|C_G|} \sum_{j \in C_G} (Y_{j,F_G+h} - Y_{j,F_G-1}), \\
\widehat{\text{ATT}}_h &= \sum_{G=1}^3 w_G \widehat{\text{DID}}_{G,h}, \quad w_G = \frac{|S_G|}{\sum_{R=1}^3 |S_R|},
\end{aligned} \tag{1}$$

Equation 1 shows the unweighted estimator. Each effect compares the cumulative change since the last pre-treatment month in newly treated municipalities with the change over the same months in not-yet-treated municipalities. As with the descriptive graphs, we annualize the monthly rates ($\times 12$) for interpretability. .

Figure 4 plots the estimates. In the months prior to the introduction of the new code, the point estimates are substantively close to and statistically indistinguishable from zero (the p -value for the joint test that all pre-treatment ATTs are zero is 0.37), indicating that arrest rates followed similar trends in treated and control municipalities (consistent with Figure 3). Those trends diverged sharply as the new code came into effect: Figure 4 reveals a large and negative ATT on the arrest rate in the months around the introduction of the new code.

Figure 5: Discretionary Arrests Decline, Homicide Arrests Do Not

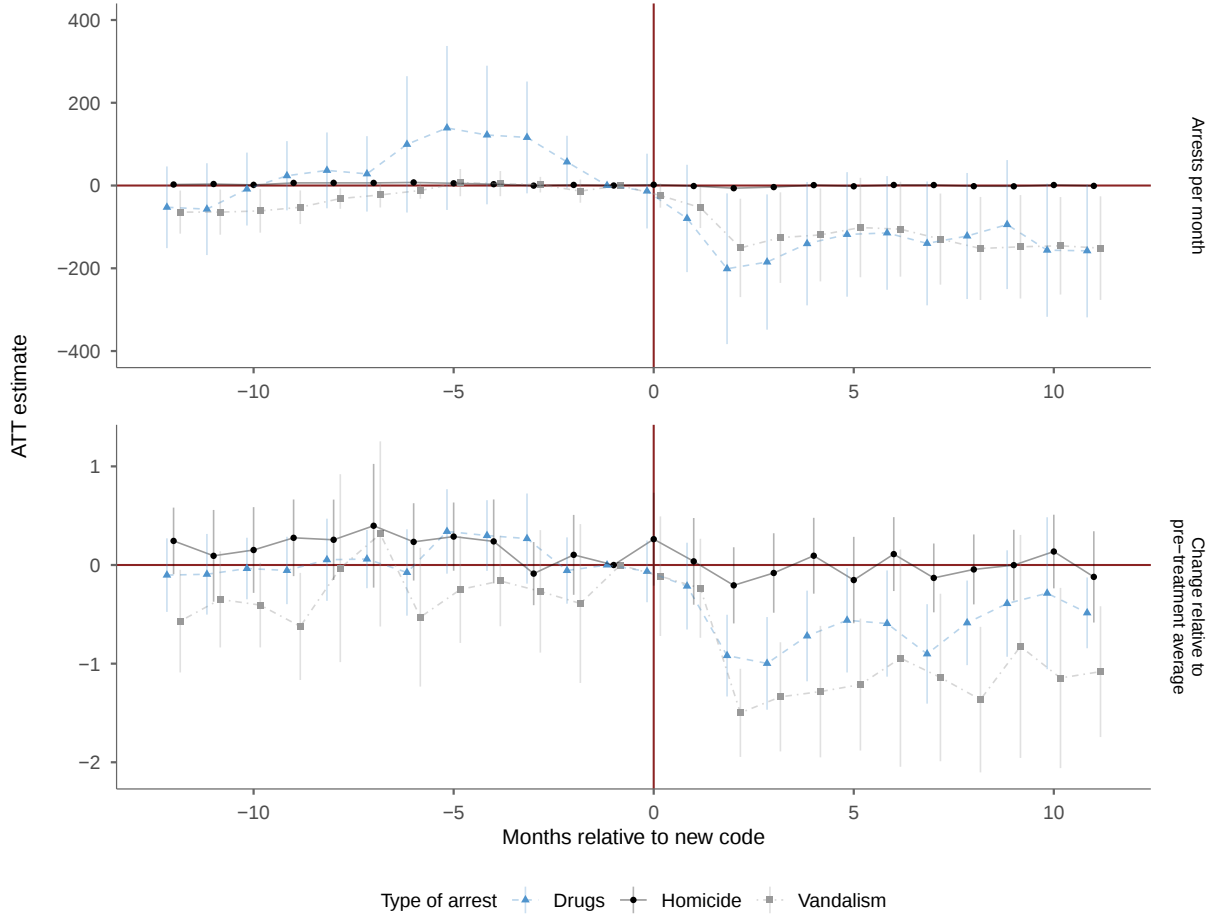


Table 1 reports the overall ATT. The average municipal arrest rate dropped by 346 arrests per 100,000 per year when the new code came into effect (Column 1, Panel A), a 30% decline in the average district relative to the twelve months before treatment.¹² Weighting the observations by population (Column 1, Panel B) produces an estimate of -431 arrests per 100,000 per year, a 48% decline relative to a twelve-month pre-treatment average of 901. When aggregating at the judicial-district level rather than the municipal level, we see effects of similar magnitude (Column 2). Another way to see this is to use the *log* of the arrest rate as the dependent variable (Column 3); because there are many zeros in the municipality-month panel, we aggregate to the judicial-district-month level in order to study log arrest rates (Chen and Roth, 2023). The relative effect sizes using the log-transformed outcome are similar.

¹²Since larger cities have higher arrest rates in the pre-treatment period and exhibit larger drops in arrests, simply dividing by the pre-period mean (571) overstates the magnitude of the treatment effect in the average municipality.

The overall drop in arrest rates was driven by even larger declines in arrest rates for minor offenses.¹³ Though all major arrest categories in our data (except homicide) include a mix of minor offenses and serious crimes—theft, for example, the largest arrest category, includes everything from very minor shoplifting (“ham was stolen, candy was stolen” Hartmann Arboleda et al., 2009, 64) to large-scale burglary, meaning that we cannot classify *theft* either as a minor offense or as a serious crime—two of the top five largest categories pre-reform are dominated by minor offenses: vandalism and drugs. For that reason, we focus on vandalism and drug arrests in the main text (see Appendix D for additional discussion).

Figure 5 compares changes in arrest rates for three crime types. This figure suggests that any declines in homicide arrest rates are substantively small and statistically indistinguishable from zero; drug and vandalism arrest rates, in contrast, fell precipitously (see also Table 1). This pattern—larger declines in arrests for minor offenses than in homicide arrests—generalizes beyond the specific categories of drug and vandalism arrests, as we discuss in Appendix D. Another empirical result echoes it: arrests with warrants declined less than arrests made *in flagrancia* (Appendix E). Because minor offenses are more likely to produce *flagrancia* arrests than arrests with warrants, we interpret this result as additional evidence that the composition of arrests shifted toward serious crimes.

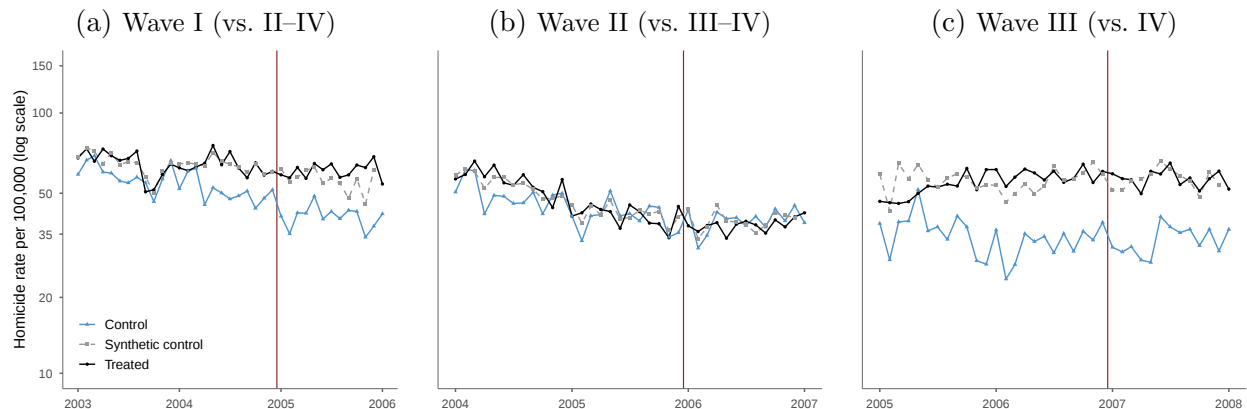
These differential changes in arrest rates did not occur because the new code imposed crime-specific restrictions on arrests. Rather, they emerged endogenously as the police re-optimized in response to downstream changes. Homicide arrests required significant effort and faced some scrutiny even under the old code; whereas for drug-possession arrests, vandalism arrests, and other arrests *in flagrancia* or for minor crimes, the new regime implied dramatic changes in officer time and effort—and, as a result, dramatic changes in arrest quantity.

5 The effect of the new code on crime

Colombia’s new code of criminal procedure affected crime rates through multiple channels. We group these mechanisms into two categories: (1) those that improved the accuracy of criminal prosecution and adjudication, raising the probability of punishment for guilty arrestees and lowering it for innocent ones; and (2) those that reduced the probability of arrest. All else equal, the increase in accuracy would likely reduce crime, while the decline in arrests would typically increase it. We find that the new code had no detectable effect on

¹³The Colombian penal code does not divide crimes into categories like *felony* and *misdemeanor*; instead, there are *crimes* (listed in the penal code) and *infractions* (or *contravenciones*), which are violations of the Police Code and do not lead to arrest. Infractions include many activities that would be classified as misdemeanors in the United States, such as public urination or disorderly conduct.

Figure 6: The Effect of the New Code on Homicide Rates



homicide, assault, mugging, or self-reported crime victimization.

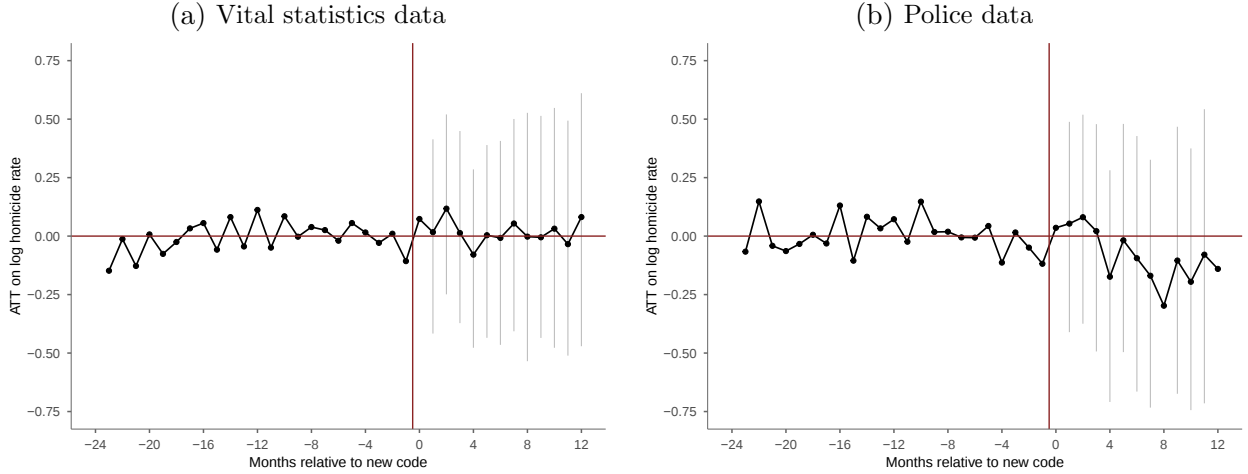
No detectable effect on homicide rates. Estimating the effect of the new code on homicide rates is more difficult than estimating the effect of the new code on arrest rates. For one thing, we do not observe any obvious patterns in raw homicide-rate trends in any specific cities. Whereas the raw arrest-rate trends reveal dramatic drops in many cities at the moment when the new code comes into effect (see Appendix Figure D.1 for examples), raw homicide-rate trends do not reveal any obvious changes (Appendix Figure F.2). For another, the national homicide rate was plummeting throughout the period of the rollout of the new code (Figure F.1); it declined at different rates in different places, making it difficult to account for secular trends and/or adjust for divergent pre-trends. Moreover, relative to arrest, homicide is a rare outcome, meaning that the municipality-month panel is sparse. Because there are many zeros in the municipality-month panel, we do not analyze the natural log of the outcome (Chen and Roth, 2023), but the distribution of municipal homicide rates themselves is of course highly skewed. For that reason, we consider two panels: a municipality-month panel in which the outcome is the (highly skewed) homicide rate, and a judicial-district-month panel, in which the outcome is either the homicide rate itself or the natural log of the homicide rate.¹⁴ Finally, as we discuss below, lethal violence in Colombia takes many forms, only some of which might plausibly respond to the new code.

The black line in Figure 6a marks the average (log) homicide rate across judicial districts in the first wave of the rollout of the new code (Bogotá and the coffee belt); the blue line

¹⁴At the judicial-district-month level, fewer than half of one percent of the observations are zeros, meaning that the estimates are not sensitive to whether we add 1 or 0.5 or 1.5 or another small constant to the outcome before taking the natural log.

Figure 7: No Evidence that the New Code Affected Homicide Rates

This figure plots synthetic-control estimates of the horizon-specific ATT on the judicial-district-level log homicide rate ([Cattaneo et al., 2022b](#)).



in that same figure marks the average (log) homicide rate in the judicial districts in Waves II–IV. They are not parallel: in the twelve months prior to the first-wave implementation of the code (i.e. the twelve months of 2004), homicide rates declined less in Bogotá and the coffee belt than they did everywhere else. A similar pattern emerges in 2005 and 2006: the gap between homicide rates in about-to-be-treated and later-treated judicial districts changes over the course of each year.

We address these not-parallel pre-trends in two ways. First, we estimate a difference-in-differences (again using the [De Chaisemartin and d’Haultfoeuille \(2020\)](#) estimator described above) that includes unit-specific linear trends; the residual trends are closer to parallel. Second, we use the methods proposed by [Cattaneo et al. \(2022b\)](#) to create synthetic control groups whose pre-trends are much closer to those of the treated units.

Table 2 reports the difference-in-differences (estimates of Equation 1). The unweighted estimates (Panel A) reveal no evidence of an increase in homicide rates when the new code came into effect. In the municipality-month panel, the difference-in-differences point estimate is negative and statistically indistinguishable from zero; using the judicial-district-month panel, the point estimates are positive but very imprecisely estimated. Using inverted hypothesis tests, we also calculate the smallest positive effect that we can rule out with 95% confidence; the magnitude varies across specifications (at the judicial-district level, it’s 4% using the vital statistics data and 7% using the police data), but it is in all cases tiny relative to the overall decline in the homicide rate during this period (Appendix Figure F.1).

Table 2: The Effect of the New Code on Homicide Rates

	Municipal level		Judicial-district level			
	Police Data	Vital Statistics	Police Data		Vital Statistics	
	Rate	Rate	Log	Log	Log	Log
Panel A: Unweighted						
Effect of new code	-13.78 (9.47)	-4.93 (7.14)	0.11 (0.09)	0.08 (0.11)	0.02 (0.07)	0.01 (0.10)
Observations	31,056	31,668	972	972	972	972
Effect (% change)	-25.2	-10.0	11.6	8.3	2.0	1.0
Pre-period mean	54.75	49.31	3.83	3.83	3.78	3.78
Unit-specific trends				✓		✓
Panel B: Weighted by population						
Effect of new code	-2.10 (3.27)	0.33 (2.59)	-0.02 (0.07)	-0.11 (0.09)	0.01 (0.06)	-0.05 (0.09)
Observations	31,056	31,668	972	972	972	972
Effect (% change)	-3.8	0.7	-2.0	-10.4	1.0	-4.9
Pre-period mean	55.62	49.74	3.86	3.86	3.76	3.76
Unit-specific trends				✓		✓

Estimates of Equation 1. Standard errors use 1,000 bootstrap replications at the judicial-district level. Percentage changes are calculated by dividing the estimate by the pre-period mean in the rate columns, or $100 \times (e^{\hat{\beta}} - 1)$ in the log columns.

The population-weighted estimates in Panel B are also of substantive interest. If the new code had caused an increase in homicide rates in Bogotá and one or two other large judicial districts but not elsewhere, we would want to know—but this effect might not show up in the unweighted estimates in Panel A. Using the municipality–month panel with homicide rates as the dependent variable (Columns 1–2), the estimates are substantively very close to zero. Using the judicial-district–month panel, the point estimates are very close to zero in the absence of unit-specific trends, negative with unit-specific trends, and in all cases imprecisely estimated (Columns 3–6). Taken together, these results provide no evidence that the new code adversely affected homicide rates.

The gray lines in Figures 6a–6c plot the synthetic controls for the judicial districts in each wave of the rollout, and Figure 7 reports the aggregate synthetic control ATT estimates. Neither suggests any evidence of an increase in the homicide rate.

In some parts of Colombia during this period, the dynamics of violence so tied to war that

we might not *expect* the new code to affect the homicide rate. In Saravena, for example, “the Colombian Sarajevo” (León, 2005), conflict between the ELN and the FARC—and between these groups and the Colombian military—produced more than 100 violent deaths per 100,000 per year. It is hard to imagine how the *de jure* introduction of a new code of criminal procedure would affect lethal violence in places like Saravena. In cities, in contrast, much of the lethal violence of the mid-2000s stemmed from community violence and conflict among local gangs, which could plausibly respond to the new code through the channels outlined above. One might then wonder whether the apparent null effect of the new code on homicide rates is driven by the inclusion of war zones, in which we would not expect the new code to affect violence. We consider this possibility in three ways: by excluding municipalities like Saravena, where the war accounted for a significant fraction of all violent deaths, which does not much affect the results (Appendix Figure F.3); by plotting the homicide rate in Colombia’s 18 largest cities, which reveals no evidence of an apparent effect of the new code except perhaps in Ibagué and Montería, cases that we believe merit further investigation (Appendix Figure F.4); and by considering synthetic control difference-in-differences for each judicial district separately (Appendix Figure F.5). We do not observe any judicial district in which the introduction of the new code appears to have affected the homicide rate.

These findings provide no evidence that the new code increased the Colombian homicide rate overall or in cities (where we would be more likely to observe a nonzero effect).

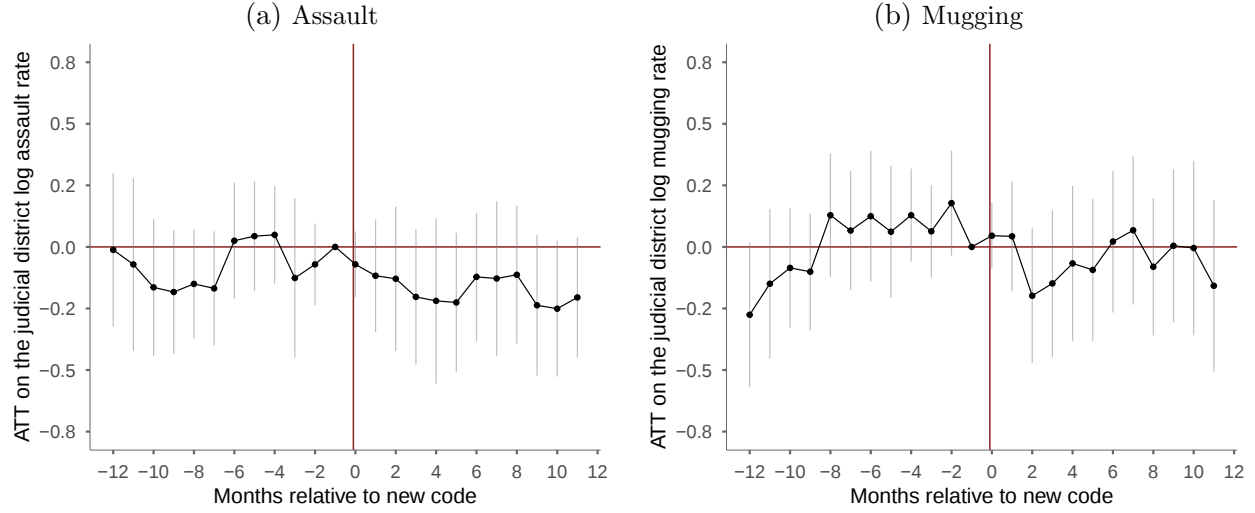
No detectable effect on assault or mugging. We find no evidence that assault or mugging increased as a result of Colombia’s new code of criminal procedure. Figure 8 plots the horizon-specific ATT of the new code on log assault (left panel) and mugging (right panel) rates, after removing new reporting sources that created artificial jumps in these counts (see Section 3 and Appendix C). Neither set of estimates reveals an increase following the implementation of the new code. Because the municipal-level panel is very sparse, i.e., has many zeros, we aggregate to the judicial-district level in order to take the log of these crime rates.

These findings contrast with those of Acosta Mejía et al. (2023), who find large and precisely estimated effects of the new code on these two crimes. We attribute the difference to our use of data that excludes new sources of crime reporting, which created the appearance of sharp jumps in crime in months that, in some places, coincided with the rollout of the new code (see Appendix C for details).

No detectable effect on self-reported crime victimization; improvement in perceptions of security. We use data from the Latin American Public Opinion Project

Figure 8: No detectable effect of the new code on assault or mugging

These figures plot the horizon-specific ATT of the new code on (a) log assault (*lesiones*) rates and (b) log mugging rates (*hurto personas*) at the judicial-district level (De Chaisemartin and d'Haultfoeuille, 2020), after excluding external sources (Appendix C). Confidence intervals use 1,000 bootstrap replications at the judicial-district level.



(LAPOP) to study the effect of the new code on self-reported crime victimization and on perceptions of neighborhood security. Unlike the data on arrests and crime analyzed above, the LAPOP data are annual, meaning that our ability to study pre-trends or the evolution of treatment effects is more limited; moreover, sample sizes are small in some municipalities. But surveys have two advantages. First, they allow us to study trends in victimization from crimes *not* reported to the police; second, they allow us to study victimization overall, rather than only for specific crimes.

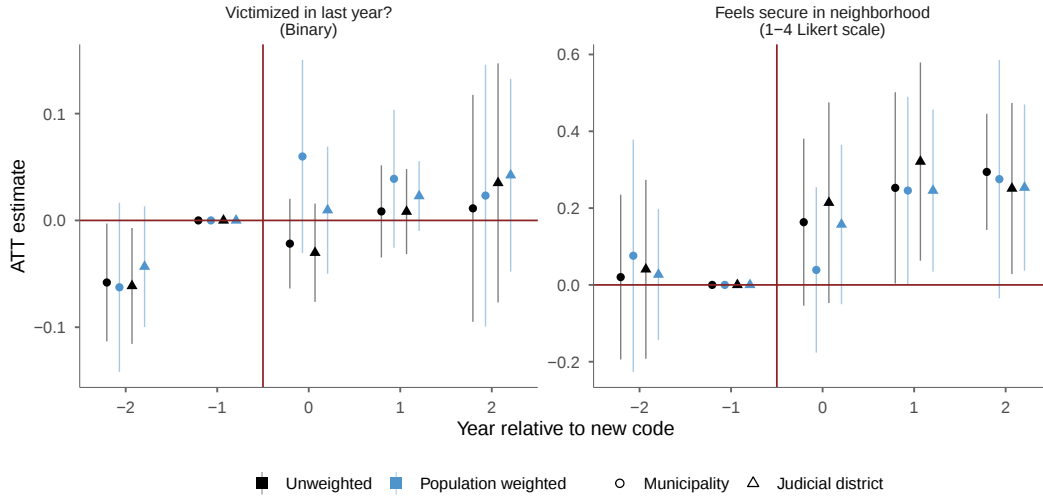
If Colombia's new code of criminal procedure had caused a major crime wave, we would expect it to show up in self-reported crime victimization. Instead, we see null effects. The estimates in the left panel of Figure 9 do not suggest that crime increased because of the new code; rather, the post-treatment estimates look like a continuation of pre-treatment trends. When we aggregate these effects in Table 3, the *sign* of the effect changes across estimators (De Chaisemartin and d'Haultfoeuille 2020 vs. Callaway and Sant'Anna 2021) and across weighting choices, and none of the estimates is statistically distinguishable from zero.¹⁵

These survey data also suggest an *improvement* in perceptions of neighborhood security

¹⁵The substantively large (though statistically insignificant) positive point estimate in Column (1), Panel B is largely driven by a single municipality, Bogotá, where self-reported victimization jumped from 24% in 2004 to 34% in 2005, before returning to 24% in 2006. If we exclude only Bogotá, that estimate falls by roughly 70% (from 0.045 to 0.013).

Figure 9: No detectable effect on self-reported victimization; improvement in feelings of security

Horizon-specific ATTs using (De Chaisemartin and d’Haultfoeuille, 2020). Confidence intervals use 1,000 bootstrap replications at the judicial-district level.



(right panel, Figure 9). Here, the pre-trends in perceptions are closer to parallel, and the overall effect is more precisely estimated (Table 3). Across specifications, the improvement ranges from 0.16 to 0.25 points on a 1–4 scale, or at least 40% of the pre-treatment standard deviation. This finding is hard to reconcile with claims that Colombia’s new code of criminal procedure sparked a crime wave: it would be odd if perceptions of neighborhood security improved even as security itself were deteriorating.

Summary. Overall, available data on homicide, assault, mugging, and self-reported crime victimization reveal no evidence of an increase in crime following the implementation of Colombia’s new code of criminal procedure. Given the limitations of the crime data (discussed in detail above), we interpret these results with caution, writing *no evidence of an increase in crime* rather than *evidence of no increase in crime*; future work could pursue additional sources of crime data, such as shrinkage rates from large grocery-store chains. Yet the crime data that we do have are consistent with the possibility that any criminogenic consequences of the decline in arrests were counterbalanced by the crime-reducing effects of an increase in prosecutorial and judicial accuracy.

Table 3: Effect of the new code on self-reported victimization and feelings of security

	De Chaisemartin and d’Haultfoeuille (2020)				Callaway and Sant’Anna (2021)			
	Victimized		Neighborhood Security		Victimized		Neighborhood Security	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Unweighted								
Post-reform	-0.008 (0.019)	-0.009 (0.023)	0.208 (0.096)	0.253 (0.110)	-0.027 (0.020)	-0.034 (0.024)	0.168 (0.106)	0.188 (0.133)
Effect / pre-reform mean (%)	-5.5	-6.3			-18.5	-23.9		
Effect / pre-reform SD	-0.07	-0.11	0.54	0.89	-0.24	-0.40	0.43	0.66
Panel B: Weighted by population								
Post-reform	0.045 (0.035)	0.019 (0.023)	0.163 (0.099)	0.203 (0.087)	0.026 (0.032)	-0.002 (0.023)	0.156 (0.082)	0.192 (0.093)
Effect / pre-reform mean (%)	21.7	11.7			12.6	-1.2		
Effect / pre-reform SD	0.46	0.24	0.62	0.72	0.27	-0.02	0.60	0.68
Cross-sectional unit	Mun.	JD	Mun.	JD	Mun.	JD	Mun.	JD
N Observations	297	156	297	156	294	156	294	156
N Cross-sectional units	51	26	51	26	49	26	49	26
Outcome scale	[0, 1]	[0, 1]	[1, 4]	[1, 4]	[0, 1]	[0, 1]	[1, 4]	[1, 4]
Pre-reform mean (2004), unweighted	0.146	0.142	2.797	2.787	0.146	0.142	2.797	2.787
Pre-reform mean (2004), population weighted	0.207	0.163	2.541	2.700	0.207	0.163	2.541	2.700
Pre-reform std. dev. (2004), unweighted	0.112	0.083	0.387	0.284	0.112	0.083	0.387	0.284
Pre-reform std. dev. (2004), population weighted	0.097	0.081	0.262	0.284	0.097	0.081	0.262	0.284

Difference-in-differences estimates of the effect of the new code on self-reported crime victimization and perceptions of neighborhood security (De Chaisemartin and d’Haultfoeuille, 2020), averaging the first three post-treatment horizons. Columns 5–8 use the method proposed by Callaway and Sant’Anna (2021). Standard errors use 1,000 bootstrap replications clustered by judicial district. Percentage changes are shown only for victimization because the neighborhood-security scale has no meaningful zero. All outcomes are estimated in levels. Mun. and JD denote municipality and judicial district.

6 Discussion

Most directly, our results illuminate the results of an unusual policy experiment in Colombia. Between 2005 and 2008, a new code of criminal procedure took effect sequentially in four geographic blocs. The new code sought to improve the accuracy of prosecutorial and judicial decisionmaking, and in doing so required more effort from police. As a result, we find, the police sharply curtailed the number of arrests, and in particular the number of arrests for low-level offenses like vandalism and drug possession. We find no evidence of a consequent increase in crime. Court accuracy may have compensated for the drop in the arrest rate.

While our empirical analysis focuses on Colombia, our theory structures thinking about other cases. It suggests that we should expect to see similar results where prosecutorial and judicial accuracy improved sufficiently relative to the reduction in police activity, while we might expect crime to rise where police activity plummeted alongside meager improvements in accuracy. Chile provides an example of the former, Venezuela and Mexico of the latter.

Chile rolled out a new code of criminal procedure between 2000 and 2005, replacing an inquisitorial system with an accusatorial system; as in Colombia, the new code bolstered court accuracy, seeking to strengthen suspects’ and defendants’ rights while also facilitating

conviction of offenders (Bellalta, 2007, 29). Studies of the Chilean experience reach relatively rosy conclusions. Tiede (2012) finds that Chile’s new code reduced pre-trial detention while improving citizens’ perceptions of security. Earlier studies, analyzing administrative data, found that the new code appeared to cause a sharp drop in arrests and a sharp increase in crime reports (Azócar and Undurraga, 2005). But scholars attribute the apparent increase in crime reports (and part of the drop in arrests) to changes in registration and perhaps in reporting behavior,¹⁶ concluding that “crime did not increase” (Bellalta, 2007, 81). At the same time, one report on the functioning of the courts found that Chile’s new code shortened the time to case resolution and raised the ratio of convictions to arrests (Marangunic and Foglesong, 2004, 10).

These accounts paint a picture strikingly similar to our view of the Colombian case, in which a new code of criminal procedure produced fewer arrests and no change in self-reported victimization, while at the same time courts became faster and perhaps better-placed to convict guilty parties.¹⁷ We therefore tentatively view Chile as a second case in which court accuracy improved sufficiently to compensate for a reduction in policing, protecting rights and public safety at the same time.

In Mexico and Venezuela, in contrast, available data suggest that new codes of criminal procedure caused even larger drops in arrests and that homicide rates did increase (Amuedo-Dorantes and Ibarra-Caton, 2025 and Hanson and Kronick, 2026, respectively)¹⁸—perhaps because the accuracy of criminal prosecution and adjudication did not improve in the ways that reformers had hoped. In Mexico, in fact, Magaloni and Salmón (2026, 11) find that police and prosecutors began fabricating evidence in response to the new code. In Venezuela, similarly, the new code was implemented with such anemic fiscal and political commitment that legal scholar Rogelio Pérez-Perdomo complained of “magical legalism,” the fantasy that *de jure* changes alone could remake an entire criminal justice system (Pérez-Perdomo, 2011). Venezuela’s new code failed to relax a disastrous constraint on *de facto* accuracy: Venezuela

¹⁶Bellalta (2007, 68) claim that, in some cases of arrests in *flagrancia*, police began recording the incident as a crime report rather than an arrest, “because it would give prosecutors more time to bring the suspect to trial.” This reclassification would create *both* the false appearance of an increase in crime *and* the false appearance of a decline in arrests. (We confirmed with interviewees that this type of reclassification did not occur in Colombia.) Azócar and Undurraga (2005, 22) also provide qualitative evidence of other mechanisms driving the drop in arrests, such as police anticipation of *garantismo* on the part of judges.

¹⁷While available data do not allow us to follow cases through the Colombian criminal justice system as researchers did in Chile, we note that the convicted prison population was higher in 2008 than in 2004 (Figure D.6), despite sharp drops in arrests.

¹⁸Amuedo-Dorantes and Ibarra-Caton (2025) claim a 57% decline in arrests in Mexico and Hanson and Kronick (2026) an 80% decline in Venezuela, though the former is based on data that likely capture only a subset of arrests and the latter is based on a single time series. That said, the direction of the arrest-rate claim in Amuedo-Dorantes and Ibarra-Caton (2025), on Mexico, is consistent with qualitative evidence in Magaloni and Salmón (2026, 10).

had very few prosecutors per capita, a bottleneck that undermined criminal procedure as a whole. This aspect of the Venezuelan case echoes the Mexican experience of “chronic underinvestment in investigative capacity” (Magaloni and Salmón, 2026, 3). Chile, in contrast, built the public prosecutor’s office from scratch alongside the implementation of the new code; Colombia more than doubled the number of investigative police (Contraloría General de la República, 2010, 23).

To put the point another way, we conceive of criminal procedure reform as a vector with two components: an increase in court accuracy, and consequent effort costs for police. Reformers in all four cases sought to improve accuracy; in all four cases, the pursuit of accuracy created new demands on police officer time. What differed, in our reading, was the relative magnitude of these changes, or the direction of the reform vector, which determines the consequences for crime (Appendix Figure A.1). In Mexico and Venezuela, accuracy improved little relative to the new costs borne by police; in Colombia and Chile, reform pointed in a more productive direction.

Though the United States never implemented wholesale criminal-procedure reform on the scale of the Latin American codes that we study, changes to U.S. criminal procedure in the 1960s were so sweeping that they too earned the moniker *revolution* (Pye, 1968). Some of these changes concerned police directly, such as the introduction of Miranda rights or the decision to allow stop-and-frisk; others primarily concerned the courts but had upstream consequences for police, such as barring illegally obtained evidence from trials or allowing judges to authorize wiretapping. As in Latin America, critics claimed that new costs for police depressed police activity and thereby caused crime (e.g. Atkins and Rubin, 2003, 157). Our theory provides a way to think about these changes. If the new rules prevented many false convictions and facilitated many convictions of true offenders, remaking criminal procedure in the United States could have reduced crime even as police pulled back from certain activities. If instead these changes merely “handcuffed the cops” (Cassell and Fowles, 1998) without reducing errors in court, we might expect crime to rise.

Our findings speak to the value of studying interactions among different institutions of criminal justice, and indeed among bureaucracies more generally. Despite the clear implications of Colombia’s new code of criminal procedure for patrol officers, previous work did not focus on the implications for officer behavior. Scholars studying other criminal justice policy changes that do not target the police, such as the elimination of cash bail, should also consider police response; similarly, prosecutors respond to new constraints on judge, and public bureaucracies respond to reform of the other bureaucracies with which they interact. We view these interactions as a valuable topic for future work.

References

- Daron Acemoglu, Leopoldo Fergusson, James Robinson, Dario Romero, and Juan F Vargas. The perils of high-powered incentives: evidence from colombia’s false positives. *American Economic Journal: Economic Policy*, 12(3):1–43, 2020.
- Camilo Acosta Mejía, Daniel Mejía Londoño, and Angela Zorro Medina. Certainty vs. severity revisited: Evidence for colombia. *Working Paper*, 2016.
- Camilo Acosta Mejía, Daniel Mejia Londoño, and Angela Zorro Medina. On the tension between due process protection and public safety: The case of an extensive procedural reform in colombia. *Documento CEDE*, (32), 2023.
- Carmen Alguíndigue and Rogelio Pérez-Perdomo. The Inquisitor Strikes Back: Obstacles to the Reform of Criminal Procedure. *Southwestern Journal of Law and Trade in the Americas*, 2008. URL <http://bit.ly/2bk4GQS>.
- Carmen Alguíndigue and Rogelio Pérez-Perdomo. Revolución y proceso penal en Venezuela: 1999-2012. *Anales de la Universidad Metropolitana*, 13(2):119–144, 2013. URL <http://dialnet.unirioja.es/servlet/articulo?codigo=4709912&info=resumen&idioma=ENG>.
- Francesca A Amaral, Aurélie Ouss, and Dalila I Ozier. Prosecutor-driven reform and racial disparities. *Criminology & Public Policy*, 2025.
- Catalina Amuedo-Dorantes and Marilyn Ibarra-Caton. How mexican judicial reforms may have fueled crime: Arrest trends and trust erosion. *Criminology & Public Policy*, 2025.
- Elliott Ash and W. Bentley MacLeod. Intrinsic motivation in public service: Theory and evidence from state supreme courts. *The Journal of Law and Economics*, 58(4):863–913, 2015.
- Raymond A. Atkins and Paul H. Rubin. Effects of criminal procedure on crime rates: Mapping out the consequences of the exclusionary rule. *The Journal of Law and Economics*, 46(1):157–179, 2003.
- María José Azócar and Beltrán Undurraga. Reforma procesal penal y seguridad ciudadana: Análisis de cuatro indicadores. Documentos de Trabajo ICSO 4, Instituto de Investigación en Ciencias Sociales (ICSO), Universidad Diego Portales, Santiago, 2005.

- Bocar Ba, Haosen Ge, Jacob Kaplan, Dean Knox, Mayya Komisarchik, Gregory Lanzalotto, Rei Mariman, Jonathan Mummolo, Roman Rivera, and Michelle Torres. Political diversity in us police agencies. *American Journal of Political Science*, 2023.
- Bocar A Ba, Dean Knox, Jonathan Mummolo, and Roman Rivera. The role of officer race and gender in police-civilian interactions in chicago. *Science*, 371(6530):696–702, 2021.
- Gary S Becker. Crime and punishment: An economic approach. *Journal of political economy*, 76(2):169–217, 1968.
- Alejandra Mohor Bellalta. *El nuevo procedimiento penal en Chile: una aproximación empírica desde la seguridad ciudadana*. RIL editores, 2007.
- Alberto M. Binder. Perspectivas de la reforma procesal penal en américa latina. In *Justicia penal y Estado de derecho*, page 201. Ad-Hoc, Buenos Aires, 1993.
- Christopher Blattman, Donald P Green, Daniel Ortega, and Santiago Tobón. Place-based interventions at scale: The direct and spillover effects of policing and city services on crime. *Journal of the European Economic Association*, 19(4):2022–2051, 2021.
- Kirill Borusyak, Xavier Jaravel, and Jann Spiess. Revisiting event-study designs: Robust and efficient estimation. *The Review of Economic Studies*, 2024.
- D.M. Brinks. *The Judicial Response to Police Killings in Latin America: Inequality and the Rule of Law*. Cambridge University Press, 2007. ISBN 9781139466509. URL <https://books.google.com/books?id=d0zdpw-hZZUC>.
- Brantly Callaway and Pedro H.C. Sant’Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230, 2021.
- Cámara de Comercio de Bogotá. Balance del año 2005. Observatorio de Seguridad de Bogotá, 2006.
- Paul G. Cassell and Richard Fowles. Handcuffing the cops? a thirty-year perspective on miranda’s harmful effects on law enforcement. *Stanford Law Review*, 50(4):1055–1145, 1998.
- Matias D Cattaneo, Carlos Diaz, and Rocio Titunik. Breaking the code: Can a new penal procedure affect public safety? Working Paper, 2022a.

- Matias D Cattaneo, Yingjie Feng, Filippo Palomba, and Rocio Titiunik. Uncertainty quantification in synthetic controls with staggered treatment adoption. *arXiv preprint arXiv:2210.05026*, 2022b.
- Aaron Chalfin and Justin McCrary. Are u.s. cities underpoliced? theory and evidence. *The Review of Economics and Statistics*, 100(1):167–186, 2018.
- Jiafeng Chen and Jonathan Roth. Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics*, page qjad054, 2023.
- Paul Chevigny. *The Edge of the Knife: Police Violence in the Americas*. The New Press of New York, 1995.
- Congreso de Colombia. Ley 600 de 2000 (código de procedimiento penal), 2000.
- Congreso de Colombia. Ley 906 de 2004 (código de procedimiento penal), 2004.
- Contraloría General de la República. Evaluación sobre la implementación del sistema penal oral acusatorio en Colombia. Technical report, Contraloría General de la República, Dirección de Estudios Sectoriales, Dirección de Justicia y Seguridad, Bogotá, 2010.
- Corte Constitucional de Colombia. Sentencia c-024 de 1994, 1994. M.P. Alejandro Martínez Caballero.
- Corte Constitucional de Colombia. Sentencia c-591 de 2005, 2005. M.P. Clara Inés Vargas Hernández.
- Corte Suprema de Justicia de Colombia, Sala de Casación Penal. Sentencia de casación, radicación 26065, 2007. M.P. Álvaro Orlando Pérez Pinzón.
- Corte Suprema de Justicia de Colombia, Sala de Casación Penal. Sentencia de casación, radicación 33254, 2013. M.P. José Leonidas Bustos Martínez.
- Ernesto Dal Bó, Federico Finan, and Martín A. Rossi. Strengthening state capabilities: The role of financial incentives in the call to public service. *The Quarterly Journal of Economics*, 128(3):1169–1218, 2013.
- Clément De Chaisemartin and Xavier d’Haultfoeuille. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996, 2020.
- Departamento Nacional de Planeación. Cifras de justicia: jurisdicción penal, 1996–2007. Technical report, Departamento Nacional de Planeación, Dirección de Justicia y Seguridad, Bogotá, 2009.

- Oeindrila Dube, Sandy Jo MacArthur, and Anuj K Shah. A cognitive view of policing. *The Quarterly Journal of Economics*, 140(1):745–791, 2025.
- M Duce and R P Perdomo. Citizen Security and Reform of the Criminal Justice System in Latin America. In Hugo Fruhling, Joseph Tulchin, and Heather Golding, editors, *Crime and Violence in Latin America: Citizen Security, Democracy, and the State*. Woodrow Wilson Center Press, 2003. URL http://books.google.com/books?hl=en&lr=&id=H7Ge7kk0C1kC&oi=fnd&pg=PA69&dq=rogelio+perez+perdomo+codigo+penal&ots=6DkyqSSxQT&sig=gf_fdhBjj281fGunX132GZ0t6TU.
- Mauricio Duce, Claudio Fuentes, and Cristián Riego. La reforma procesal penal en américa latina y su impacto en el uso de la prisión preventiva. In *Prisión preventiva y reforma procesal penal en América Latina: Evaluación y perspectivas*. Centro de Estudios de Justicia de las Américas (CEJA), Santiago, 2009.
- Patrick Francois. ‘public service motivation’ as an argument for government provision. *Journal of Public Economics*, 78(3):275–299, 2000.
- Juan García, Daniel Mejía, and Daniel Ortega. Police reform, training and crime: experimental evidence from colombia’s plan cuadrantes. *Documento CEDE*, (2013-04), 2013.
- Yanilda González. The social origins of institutional weakness and change: Preferences, power, and police reform in latin america. *World Politics*, 71(1):44–87, 2019.
- Yanilda María González. *Authoritarian police in democracy: Contested security in Latin America*. Cambridge University Press, 2020.
- Guy Grossman and Tara Slough. Government responsiveness in developing countries. *Annual Review of Political Science*, 25:131–153, 2022.
- Tamy Guberek, Daniel Guzmán, Megan Price, Kristian Lum, and Patrick Ball. To count the uncounted: An estimation of lethal violence in casanare. *A Report by the Benetech Human Rights Program*, 2010.
- Isabel Gutiérrez and Santiago Tobón. El gasto fiscal de la guerra contra los portadores de drogas ilícitas: una aproximación para colombia. *Lecturas de Economía*, (91):79–107, 2019.
- Rebecca Hanson and Dorothy Kronick. Official vigilantism. *World Development*, 198:106978, 2026. doi: 10.1016/j.worlddev.2025.106978.

- Mildred Hartmann, Carolina Villadiego, and Cristián Riego. La reforma procesal penal en Colombia. In Cristián Riego R., editor, *Reformas Procesales Penales en América Latina: Resultados del Proyecto de Seguimiento, IV Etapa*, pages 101–143. Centro de Estudios de Justicia de las Américas (CEJA), Santiago, Chile, 2007.
- Mildred Hartmann Arboleda and Wilson Alejandro Martínez Sánchez. La detención preventiva y la reforma procesal penal en colombia. In *Prisión preventiva y reforma procesal penal en América Latina: Evaluación y perspectivas*. Centro de Estudios de Justicia de las Américas (CEJA), Santiago, 2009.
- Mildred Hartmann Arboleda, Carlos Andrés Gómez, and Camilo Alberto Ortiz. *Estudio empírico del funcionamiento del sistema acusatorio*. Editorial Tadeo Lozano, 2009.
- David Hausman and Dorothy Kronick. The illusory end of stop and frisk in chicago? *Science Advances*, 9(39):eadh3017, 2023.
- George L Kelling and Catherine M Coles. *Fixing broken windows: Restoring order and reducing crime in our communities*. The Free Press, 1996.
- Máximo Langer. Revolution in latin american criminal procedure: Diffusion of legal ideas from the periphery. *The American Journal of Comparative Law*, 55(4):617–676, 2007.
- Isabel Laterzo-Tingley. The public security issue area. *Political Research Quarterly*, 2026. doi: 10.1177/10659129261463236.
- Juanita León. *País de plomo: Crónicas de guerra*. Aguilar, 2005.
- María Victoria Llorente. ¿desmilitarización en tiempos de guerra? la reforma policial en colombia. *Seguridad y reforma policial en las Américas. Experiencias y desafíos. México: Siglo XXI*, pages 192–216, 2005.
- Beatriz Magaloni and Luis Rodriguez. Institutionalized police brutality: Torture, the militarization of security, and the reform of inquisitorial criminal justice in mexico. *American Political Science Review*, 114(4):1013–1034, 2020.
- Beatriz Magaloni and Esteban Salmón. Fabricated justice: How due process reform enables evidence manipulation. *World Development*, 199:107222, 2026.
- Beatriz Magaloni, Edgar Franco-Vivanco, and Vanessa Melo. Killing in the slums: Social order, criminal governance, and police violence in rio de janeiro. *American Political Science Review*, 114(2):552–572, 2020.

- Antonio Marangunic and Todd Foglesong. Charting justice reform in chile: A comparison of the old and new systems of criminal procedure. Technical report, Vera Institute of Justice, New York, December 2004.
- John Henry Merryman and Rogelio Pérez-Perdomo. Criminal procedure. In *The Civil Law Tradition: An Introduction to the Legal Systems of Europe and Latin America*, chapter 17, pages 127–135. Stanford University Press, Stanford, CA, 4 edition, 2018. doi: 10.1515/9781503607552-019.
- Eduardo Moncada. Toward democratic policing in colombia? institutional accountability through lateral reform. *Comparative Politics*, 41(4):431–449, 2009.
- Jonathan Mummolo. Modern police tactics, police-citizen interactions, and the prospects for reform. *The Journal of Politics*, 80(1):1–15, 2018.
- Aurelie Ouss and Megan Stevenson. Does cash bail deter misconduct? *American Economic Journal: Applied Economics*, 15(3):150–182, 2023.
- Emily Owens, David Weisburd, Karen L Amendola, and Geoffrey P Alpert. Can you build a better cop? experimental evidence on supervision, training, and policing in the community. *Criminology & Public Policy*, 17(1):41–87, 2018.
- Rogelio Pérez-Perdomo. *Justicia e Injusticias en Venezuela: Estudios de Historia Social del Derecho*. Quid Pro Books, 2011.
- F. O. Piraquive Sierra. Distrito judicial de yopal. *Derecho y Realidad*, 5:87–114, 2007.
- Policía Nacional de Colombia. Análisis: regiones de policía. *Revista Criminalidad*, 48:36, 2005.
- A. Kenneth Pye. The warren court and criminal procedure. *Michigan Law Review*, 67(2): 249–268, 1968.
- Michael Reed and Patrick Ball. El registro y la medición de la criminalidad. el problema de los datos faltantes y el uso de la ciencia para producir estimaciones en relación con el homicidio en colombia, demostrado a partir de un ejemplo: el departamento de antioquia (2003-2011). *Revista Criminalidad*, 58(1):9–23, 2016.
- Roman Rivera. Performance pay and multitasking police, 2024.
- Roman Rivera. Are “bad” cops better police? the trade-off between officer aggression and public safety. Working Paper, 2025a.

- Roman Rivera. Do peers matter in the police academy? *American Economic Journal: Applied Economics*, 17(2):127–164, 2025b.
- Roman G Rivera and Bocar A Ba. The effect of police oversight on crime and misconduct allegations: Evidence from chicago. *Review of Economics and Statistics*, pages 1–45, 2023.
- Sneider Rivera et al. La impunidad en el sistema penal colombiano: informe de consultoría. Technical report, Proyecto Fortalecimiento del Sector Justicia para la Reducción de la Impunidad en Colombia (Comisión Europea), 2010.
- José Ferney Rodríguez Cuadrado. La historia y el presente de las cifras delictivas y contravencionales en Colombia: Un nuevo conocimiento. *Revista Criminalidad*, 50(1):109–118, 2008.
- Jair David Rodríguez-Ortega, Daniel Mejía-Londoño, Lorena del Pilar Caro-Zambrano, Mauricio Romero-Hernández, and Franney Campos-Méndez. Implicaciones del proceso de integración de los registros administrativos de criminalidad entre el spoa de la fiscalía general y el siedco de la policía nacional de colombia, y la puesta en marcha del aplicativo “¡ adenunciar!” sobre las cifras de criminalidad. *Revista Criminalidad*, 60(3):9–27, 2018.
- Yotam Shem-Tov, Steven Raphael, and Alissa Skog. Can restorative justice conferencing reduce recidivism? evidence from the make-it-right program. *Econometrica*, 92(1):61–78, 2024.
- Jerome H. Skolnick and James Fyfe. *Above the Law: Police and the Excessive Use of Force*. The Free Press, 1993.
- Allison Stashko and Haritz Garro. Prosecutor elections and police accountability. *Working Paper*, 2021.
- Lydia Brashear Tiede. Chile’s criminal law reform: Enhancing defendants’ rights and citizen security. *Latin American Politics and Society*, 54(3):65–93, 2012.
- Jessie Trudeau. Limiting aggressive policing can reduce police and civilian violence. *World development*, 160:105961, 2022.
- Rodrigo Uprimny Yepes, Sergio Chaparro Hernández, and Luis Felipe Cruz Olivera. *Delitos de drogas y sobredosis carcelaria en Colombia*. Documentos Dejusticia 37. Dejusticia (Centro de Estudios de Derecho, Justicia y Sociedad), Bogotá, 2017.
- Erik H. Wang. Frightened mandarins: The adverse effects of fighting corruption on local bureaucracy. *Comparative Political Studies*, 55(11):1807–1843, 2022.

J.Q. Wilson. *Varieties of Police Behavior: The Management of Law and Order in Eight Communities*. Harvard Paperback. Harvard University Press, 1968. ISBN 9780674045200. URL <https://books.google.com/books?id=yzIXXFDkotgC>.

World Bank. Completeness of death registration with cause-of-death information (%), colombia. World Development Indicators, series SP.REG.DTHS.ZS (source: World Health Organization), 2026. URL <https://data.worldbank.org/indicator/SP.REG.DTHS.ZS?locations=C0>. Accessed 2026.

Angela Zorro Medina. The failed war on pre-trial detention: Evidence from a quasi-experimental reform. *Available at SSRN 3686855*, 2020.

Angela Zorro Medina. Neoliberalism vs. bureaucracy: Evidence from mass-incarceration in colombia. *Available at SSRN 4494325*, 2023.

Angela Zorro Medina, Camilo Acosta Mejía, and Daniel Mejía Londoño. The unintended consequences of the us adversarial model in latin american crime. *Available at SSRN 3686828*, 2020.

Appendix

A	Theory	2
A.1	Benchmarks: No patrolling or surveillance state	4
A.2	Equilibrium characterization	4
A.3	Equilibrium Quantities of Interest	6
A.4	Mapping to intervention and empirical implications	6
B	Comparing criminal procedure under the old and new codes	12
C	Changes in crime registration	18
D	Additional results on the decline in arrests	23
E	Arrests in <i>flagrancia</i> vs. arrests with warrants	31
F	Additional evidence of the effect of the new code on homicide rates	33
G	LAPOP sample	39
H	Vehicle theft	40

A Theory

Consider a unit mass of individuals, indexed by i . Citizens choose whether to work in the labor market or commit a crime, denoted by $d_i \in \{W, C\}$. By forgoing a crime, a citizen takes their market income, which we normalize to 0. By committing a crime without being caught, a citizen extracts a benefit $\delta_i \sim U[a, b]$. Given our normalization of income, we interpret δ_i as the benefit of crime relative to working in the market for individual i . We denote the pdf and cdf of δ as f_δ and F_δ throughout.

Police choose a patrol rate $\lambda \in [0, 1]$. Patrolling requires costly effort. For convenience, set the cost of patrolling to be $\frac{c\lambda^2}{2}$.¹⁹ Police value patrolling either because it: (1) allows them to demonstrate some output—arrests—for promotion or (2) because they value taking the guilty off the street. The latter could be conceptualized as a taste for order or some type of public service motivation. We weight these two benefits by $\mu \in [0, 1]$ (arrests) and $1 - \mu$ (public service motivation), respectively.

Neither the police nor the courts are perfect in determining whether an individual committed a crime, however. We assume that, conditional on reaching the court, court actors (prosecutors, judges etc.) accurately punish a guilty suspect with probability $\rho_c \in (\frac{1}{2}, 1)$ and inaccurately punish an innocent suspect with probability $1 - \rho_c$.²⁰ Similarly, assume that when patrolling, police correctly arrest a guilty citizen with probability $\rho_p \in (\frac{1}{2}, \frac{1}{2-\rho_c}]$ and incorrectly arrest an innocent citizen with probability $1 - \rho_p$. Note that this upper bound on ρ_p ensures that the police are not substantially more accurate than the courts. One can view ρ_p and ρ_c as measures of the quality of police and the court, respectively. If punished, the citizen pays a penalty $P > 0$.

Each citizen’s utility is therefore given by:

$$U_i(d_i) = \begin{cases} 0 & \text{if } d_i = W \text{ and not punished} \\ \delta_i & \text{if } d_i = C \text{ and not punished} \\ -P & \text{if subject is punished} \end{cases}$$

¹⁹More generally, one could set some cost function $\kappa(\lambda)$, where $\kappa'(\lambda) > 0$ and $\kappa''(\lambda) > 0$.

²⁰We use the word “punish” as a catchall that could include lengthy pre-trial detention or conviction.

Each citizen's expected utility, is therefore given by:

$$E[U_i(d_i)] = \begin{cases} - \overbrace{P\lambda(1-\rho_p)(1-\rho_c)}^{\text{Patrolled, wrongly arrested, and wrongly punished}} & \text{if } d_i = W \\ \delta_i \left(\underbrace{1-\lambda}_{\text{Not patrolled}} + \underbrace{\lambda}_{\text{Patrolled}} \left[\underbrace{(1-\rho_p)}_{\text{Not arrested}} + \underbrace{\rho_p(1-\rho_c)}_{\text{Arrested, but not punished}} \right] - \underbrace{P\lambda\rho_p\rho_c}_{\text{Patrolled, arrested, and punished}} \right) & \text{if } d_i = C \end{cases}$$

Note that in equilibrium, some proportion of the population will commit crime, i.e. $\Pr(d_i = C)$. We can therefore write the police officer's objective as:

$$\max_{\lambda \in [0,1]} \lambda \left(\underbrace{\mu [\Pr(d_i = C)\rho_p + (1 - \Pr(d_i = C))(1 - \rho_p)]}_{\text{Arrests}} + (1 - \mu) \underbrace{[\Pr(d_i = C)\rho_p\rho_c]}_{\text{Guilty punished}} \right) - \frac{c\lambda^2}{2}$$

To avoid corner solutions, we impose two parametric assumptions on the support of δ_i that are consistent with empirical observation: $a < 0$ and $b \geq \frac{2P}{1-\rho_c\rho_p}$. These assumptions imply that the market wage is greater than the benefit from committing crime for some individual ($a < 0$), and that perfect deterrence of crime is not possible ($b \geq \frac{2P}{1-\rho_c\rho_p}$).

The sequence is as follows:

1. All individuals decide whether to commit or forego crime.
2. The police determine their patrol rate and patrol.
3. Nature determines whether a patrolled subject is arrested, and, conditional on arrest, whether they are punished. Utilities are realized.

We characterize the subgame perfect Nash equilibrium in pure strategies. Individuals' decisions are a choice $d_i : \{W, C\}^i$. The police's patrol rate is given by $\lambda : \{W, C\}^i \rightarrow [0, 1]$.

Comments on the model:

We make two comments on the set-up of the model. First, the accuracy parameters, ρ_c and ρ_p do not vary on the basis of whether the individual commits a crime or not. This ensures that the Type-I error rate and Type-II error rates are equivalent and that changes in court accuracy (ρ_c) affect both types of errors (equally). This is a simplifying assumption that is consistent with our substantive discussion of the reform as enhancing both dimensions of accuracy. A less parsimonious model that incorporates different Type-I and Type-II error

rates produces qualitatively similar predictions and does not provide additional empirical leverage because we do not have a good way to calibrate differences these error rates (because the true crime rate is not observable). For this reason we opt for the (slightly) more restrictive but parsimonious parameterization.

Second, we allow police to value both making arrests (weighted by μ) and punishing the guilty (weighted by $1 - \mu$) to reflect the empirical reality that officers' motives vary (Dal Bó et al., 2013; Francois, 2000). By allowing officers (or forces) to weigh both motives, we show that our results are not confined to a given assumption about officer motivation.

A.1 Benchmarks: No patrolling or surveillance state

First consider crime rates under the two possible corner solutions to the police's patrol rate decision: no patrolling ($\lambda = 0$) or a surveillance state in which everyone is patrolled ($\lambda = 1$). When $\lambda = 0$, it is trivial to see that citizens commit crime if their benefit from crime exceeds the market wage, e.g., when $\delta_i > 0$. The crime rate is therefore given by $1 - F_\delta(0)$.

In a surveillance state ($\lambda = 1$), a citizen commits a crime if:

$$\begin{aligned} \delta_i ((1 - \rho_p) + \rho_p(1 - \rho_c)) - P\rho_p\rho_c &> -P(1 - \rho_p)(1 - \rho_c) \\ \delta_i &> \frac{P(\rho_p + \rho_c - 1)}{1 - \rho_p\rho_c} \end{aligned}$$

The crime rate is $1 - F_\delta(\frac{P(\rho_p + \rho_c - 1)}{1 - \rho_p\rho_c})$. It is clear from inspection that $\frac{P(\rho_p + \rho_c - 1)}{1 - \rho_p\rho_c} > 0$. Therefore, the crime rate is lower under the surveillance state than under no patrolling.

A.2 Equilibrium characterization

The above benchmarks help us think about the officer's patrolling decision. Specifically, the share of the population that commits crime can be expressed as a function of the patrol rate $Pr(d = C) = 1 - F_\delta(g(\lambda))$. Substitution allows us to rewrite the officer's objective as:

$$\max_{\lambda \in [0,1]} \lambda [F(g(\lambda))[\mu(1 - 2\rho_p + \rho_p\rho_c) - \rho_c\rho_p] + \rho_p(\mu + \rho_c - \mu\rho_c)] - \frac{c\lambda^2}{2} \quad (2)$$

To characterize $g(\lambda)$, consider the population's decisions of whether to commit crime. A citizen commits crime if:

$$E[u_i(c)] - E[u_i(w)] > 0,$$

which we can write in terms of the threshold $\delta_i^*(\lambda)$:

$$\delta^*(\lambda) \equiv \frac{P\lambda(\rho_p + \rho_c - 1)}{1 - \lambda\rho_p\rho_c}. \quad (3)$$

The citizen commits crime if $\delta_i > \delta^*(\lambda)$.

Obviously, this expression nests the two corner cases derived above. Note that $\frac{\partial \delta^*}{\partial \lambda} = \frac{P(\rho_p + \rho_c - 1)}{(\lambda\rho_c\rho_p - 1)^2} > 0$, under the above parametric assumptions. This means that as the rate of patrols increases, the marginal citizen that is indifferent to crime/foregoing crime must have a higher net benefit from committing crime (δ_i). It is straightforward to see that $\delta^* = g(\lambda)$.

Substituting (3) into (2) and maximizing yields:

$$\begin{aligned} \frac{\mu(1 - 2\rho_p + \rho_p\rho_c) - \rho_c\rho_p}{b - a} \left(\frac{P\lambda^*(\rho_p + \rho_c - 1)}{1 - \lambda^*\rho_p\rho_c} - a + \frac{P\lambda^*(\rho_p + \rho_c - 1)}{(\lambda^*\rho_c\rho_p - 1)^2} \right) + \\ \rho_p(\mu + \rho_c - \mu\rho_c) - c\lambda^* = 0 \end{aligned} \quad (4)$$

To show that this equilibrium is unique, note the left hand side of this expression of this expression is strictly decreasing in λ^* . The following expression gives the second order condition with respect to λ^* :

$$\underbrace{2P}_{>0} \underbrace{\frac{\mu(1 - 2\rho_p + \rho_p\rho_c) - \rho_c\rho_p}{b - a}}_{<0} \underbrace{\frac{\rho_p + \rho_c - 1}{(1 - \lambda^*\rho_c\rho_p)^3}}_{>0} - c < 0$$

Proposition 1. *In the unique subgame perfect Nash Equilibrium, police patrol at rate:*

$$\lambda = \begin{cases} 0 & \text{if } \lambda^* < 0 \\ \lambda^* & \text{if } \lambda^* \in [0, 1] \\ 1 & \text{if } \lambda^* > 1, \end{cases}$$

where λ^* solves (4). Individuals commit crime if $\delta_i > \frac{P\lambda^*(\rho_p + \rho_c - 1)}{1 - \lambda^*\rho_p\rho_c}$ and forego crime otherwise.

A.3 Equilibrium Quantities of Interest

This equilibrium allows us to characterize two equilibrium outcomes that we observe empirically. First, the crime rate is given by:

$$\mathcal{C} = 1 - F_\delta \left(\frac{P\lambda^*(\rho_p + \rho_c - 1)}{1 - \lambda^*\rho_p\rho_c} \right). \quad (5)$$

Note that the equilibrium crime rate is *decreasing* in the patrol rate λ^* :

$$\frac{\partial \mathcal{C}}{\partial \lambda^*} = \frac{P(\rho_c + \rho_p - 1)}{(a - b)(\lambda^*\rho_c\rho_p - 1)^2} < 0,$$

since $b > a$, which demonstrates the deterrent effect of patrolling on crime.

Second, given the crime rate, the arrest rate is:

$$\mathcal{A} = \lambda^* \left[\rho_p \left(1 - F_\delta \left(\frac{P\lambda^*(\rho_p + \rho_c - 1)}{1 - \lambda^*\rho_p\rho_c} \right) \right) + (1 - \rho_p) F_\delta \left(\frac{P\lambda^*(\rho_p + \rho_c - 1)}{1 - \lambda^*\rho_p\rho_c} \right) \right] \quad (6)$$

$$= \lambda^* \left[\rho_p + (1 - 2\rho_p) F_\delta \left(\frac{P\lambda^*(\rho_p + \rho_c - 1)}{1 - \lambda^*\rho_p\rho_c} \right) \right] \quad (7)$$

How does the arrest rate vary in the patrol rate? Let $x = \frac{P\lambda^*(\rho_p + \rho_c - 1)}{1 - \lambda^*\rho_p\rho_c}$.

$$\begin{aligned} \frac{\partial \mathcal{A}}{\partial \lambda^*} &= (\rho_p + (1 - 2\rho_p)F_\delta(x)) + \lambda(1 - 2\rho_p)F'(x)\frac{\partial x}{\partial \lambda^*} \\ &= \rho_p + (1 - 2\rho_p) \left(\frac{x - a}{b - a} + \frac{\lambda^*}{b - a} \frac{\partial x}{\partial \lambda^*} \right) \\ &= \rho_p + \frac{1 - 2\rho_p}{b - a} \left(\frac{P\lambda^*(\rho_p + \rho_c - 1)}{1 - \lambda^*\rho_p\rho_c} - a + \frac{\lambda^*P(\rho_p + \rho_c - 1)}{(1 - \lambda^*\rho_p\rho_c)^2} \right) \\ &= \underbrace{\frac{\rho_p b - a(1 - \rho_p)}{b - a}}_{>0} + \underbrace{\frac{(1 - 2\rho_p)P\lambda^*(\rho_p + \rho_c - 1)(2 - \lambda^*\rho_p\rho_c)}{(b - a)(1 - \lambda^*\rho_p\rho_c)^2}}_{<0} \end{aligned}$$

This expression can be positive or negative, which reflects the idea that arrest would mechanically increase in police effort (if the crime rate were fixed). However, police patrolling effort deters crime (as demonstrated above). These effects are countervailing.

A.4 Mapping to intervention and empirical implications

Our intervention manipulates two attributes of the above model. First, it (weakly) increases the cost for patrolling, c . As effort becomes more costly, police patrol less. This reduces deterrence of crime, (weakly) increasing the crime rate.

Second, it induces greater court/prosecutor scrutiny over arrests. One way to interpret this is an increase in ρ_c —the courts become more accurate in assessing the true status of the defendant.

We begin by characterizing comparative statics with respect to police effort and the two equilibrium outcomes (arrest rate and crime rate) for each parameter in isolation. We then characterize comparative statics with respect to a simultaneous shock to both c and ρ_c , which we view as consistent with our intervention.

Proposition 2. *An increase in c , the cost of patrolling:*

1. *Weakly decreases the rate of patrol.*
2. *Weakly increases the crime rate.*
3. *Can increase or decrease the arrest rate.*

Proof: Inspection of (4) shows that the LHS expression is decreasing in c . For the equality to hold, it must therefore be the case that the patrol rate λ^* weakly decreases as c increases. This inequality is weak due to the possibility of corner patrol rates. With respect to the crime rate, c enters (5) implicitly through λ^* and $\frac{\partial \lambda^*}{\partial c} \leq 0$. Therefore:

$$\frac{\partial \mathcal{C}}{\partial c} = -\frac{\frac{\partial \lambda^*}{\partial c}}{b-a} \cdot \frac{P(\rho_p + \rho_c - 1)}{(1 - \lambda^* \rho_p \rho_c)^2} \geq 0.$$

With respect to the arrest rate, note that:

$$\frac{\partial \mathcal{A}}{\partial c} = \underbrace{\frac{\frac{\partial \lambda^*}{\partial c}}{b-a}}_{<0} \left[\underbrace{(b-a)\rho_p}_{>0} + \underbrace{(1-2\rho_p) \left(\frac{P\lambda^*(\rho_p + \rho_c - 1)(2 - \lambda^* \rho_p \rho_c)}{(1 - \lambda^* \rho_p \rho_c)^2} - a \right)}_{>0} \right],$$

which is negative if $(b-a)\rho_p > 1 - 2\rho_p) \left(\frac{P\lambda^*(\rho_p + \rho_c - 1)(2 - \lambda^* \rho_p \rho_c)}{(1 - \lambda^* \rho_p \rho_c)^2} - a \right)$, positive if $(b-a)\rho_p < 1 - 2\rho_p) \left(\frac{P\lambda^*(\rho_p + \rho_c - 1)(2 - \lambda^* \rho_p \rho_c)}{(1 - \lambda^* \rho_p \rho_c)^2} - a \right)$, and zero otherwise. ■

Proposition 3. *An increase in ρ_c , the accuracy of the court:*

1. *Can increase or decrease the rate of patrol.*
2. *Can increase or decrease the arrest rate*
3. *Can increase or decrease the crime rate.*

Proof: Let $G(\lambda^*, \rho_c)$ denote left-hand side of the first-order condition in (4). Since (4) defines λ^* implicitly, the implicit function theorem gives:

$$\frac{\partial \lambda^*}{\partial \rho_c} = - \frac{\partial G / \partial \rho_c}{\partial G / \partial \lambda^*}.$$

The denominator is the second-order condition,

$$\frac{\partial G}{\partial \lambda^*} = \frac{2P(\mu(1 - 2\rho_p + \rho_p\rho_c) - \rho_c\rho_p)(\rho_p + \rho_c - 1)}{(b - a)(1 - \lambda^*\rho_p\rho_c)^3} - c < 0,$$

which is negative, as established above. The numerator is:

$$\begin{aligned} \frac{\partial G}{\partial \rho_c} = & \frac{\rho_p(\mu - 1)}{b - a} \left(\frac{P\lambda^*(\rho_p + \rho_c - 1)(2 - \lambda^*\rho_p\rho_c)}{(1 - \lambda^*\rho_p\rho_c)^2} - a \right) + \\ & \frac{\mu(1 - 2\rho_p + \rho_p\rho_c) - \rho_c\rho_p}{b - a} \cdot \frac{P\lambda^*[(2 - \lambda^*\rho_p\rho_c)(1 - \lambda^*\rho_p\rho_c) + \lambda^*\rho_p(\rho_p + \rho_c - 1)(3 - \lambda^*\rho_p\rho_c)]}{(1 - \lambda^*\rho_p\rho_c)^3} + \\ & \rho_p(1 - \mu), \end{aligned}$$

which is ambiguous in sign. Because $\partial G / \partial \lambda^* < 0$, the sign of $\partial \lambda^* / \partial \rho_c$ is equivalent to the sign of $\partial G / \partial \rho_c$, which is ambiguous. For example, suppose that $a = -1$, $b = 3$, $\rho_c = \frac{17}{32}$, $\rho_p = \frac{273}{544}$, $P = 1$, and $c = 1$. When $\mu = \frac{1}{2}$, $\lambda^* \approx 0.350$ and $\partial \lambda^* / \partial \rho_c \approx 0.159 > 0$; in contrast, when $\mu = 1$, $\lambda^* \approx 0.501$ and $\partial \lambda^* / \partial \rho_c \approx -0.001 < 0$. Consider now the effect of ρ_c on arrest rates:

$$\frac{\partial \mathcal{A}}{\partial \rho_c} = \frac{\partial \lambda^*}{\partial \rho_c} \left[\rho_p + \underbrace{\frac{(1 - 2\rho_p)}{b - a}}_{<0} \left(\underbrace{\frac{P\lambda^*(\rho_p + \rho_c - 1)(2 - \lambda^*\rho_p\rho_c)}{(1 - \lambda^*\rho_p\rho_c)^2} - a}_{>0} \right) \right] + \underbrace{\frac{\lambda^{*2}(1 - 2\rho_p)P[1 + \lambda^*\rho_p(\rho_p - 1)]}{(b - a)(1 - \lambda^*\rho_p\rho_c)^2}}_{<0}.$$

The sign of this first-order condition depends on the sign of $\frac{\partial \lambda^*}{\partial \rho_c}$ which is ambiguous, as is the sign of the term in hard brackets. Finally, consider the effect of ρ_c on crime rates.

$$\frac{\partial \mathcal{C}}{\partial \rho_c} = \frac{P \left[\frac{\partial \lambda^*}{\partial \rho_c}(\rho_p + \rho_c - 1) + \lambda^*(1 + \lambda^*(\rho_p^2 - \rho_p)) \right]}{(a - b)(\lambda^*\rho_c\rho_p - 1)^2}.$$

When $\frac{\partial \lambda^*}{\partial \rho_c} > \frac{-\lambda^*(1 + \lambda^*(\rho_p^2 - \rho_p))}{\rho_p + \rho_c - 1}$, the crime rate is decreasing in ρ_c . However, this cannot be ensured without additional parametric restrictions. ■

Proposition 4. *A simultaneous increase in c , the cost of patrolling, and ρ_c , the accuracy of the court:*

1. *Can increase or decrease the patrol rate.*
2. *Can increase or decrease the rate of arrest.*
3. *Can increase or decrease the crime rate.*

Proof: The contour plots in Figures A.1 and A.2 establish that a simultaneous increase in c and ρ_c can increase or decrease the patrol rate, the crime rate, and the arrest rate. By showing contour plots for two varying sets of parameters, we demonstrate that this ambiguity is not non-generic.²¹ ■

²¹It is straightforward, if cumbersome, to demonstrate this ambiguity algebraically.

Figure A.1: Contour plots to show how patrol rates, crime rates, and arrest rates change in the cost of effort (c) and the accuracy of the court (ρ_c). For all plots, $a = -25$, $b = 25$, $\rho_p = 0.7$, $P = 1$, and $\mu = 0.5$.

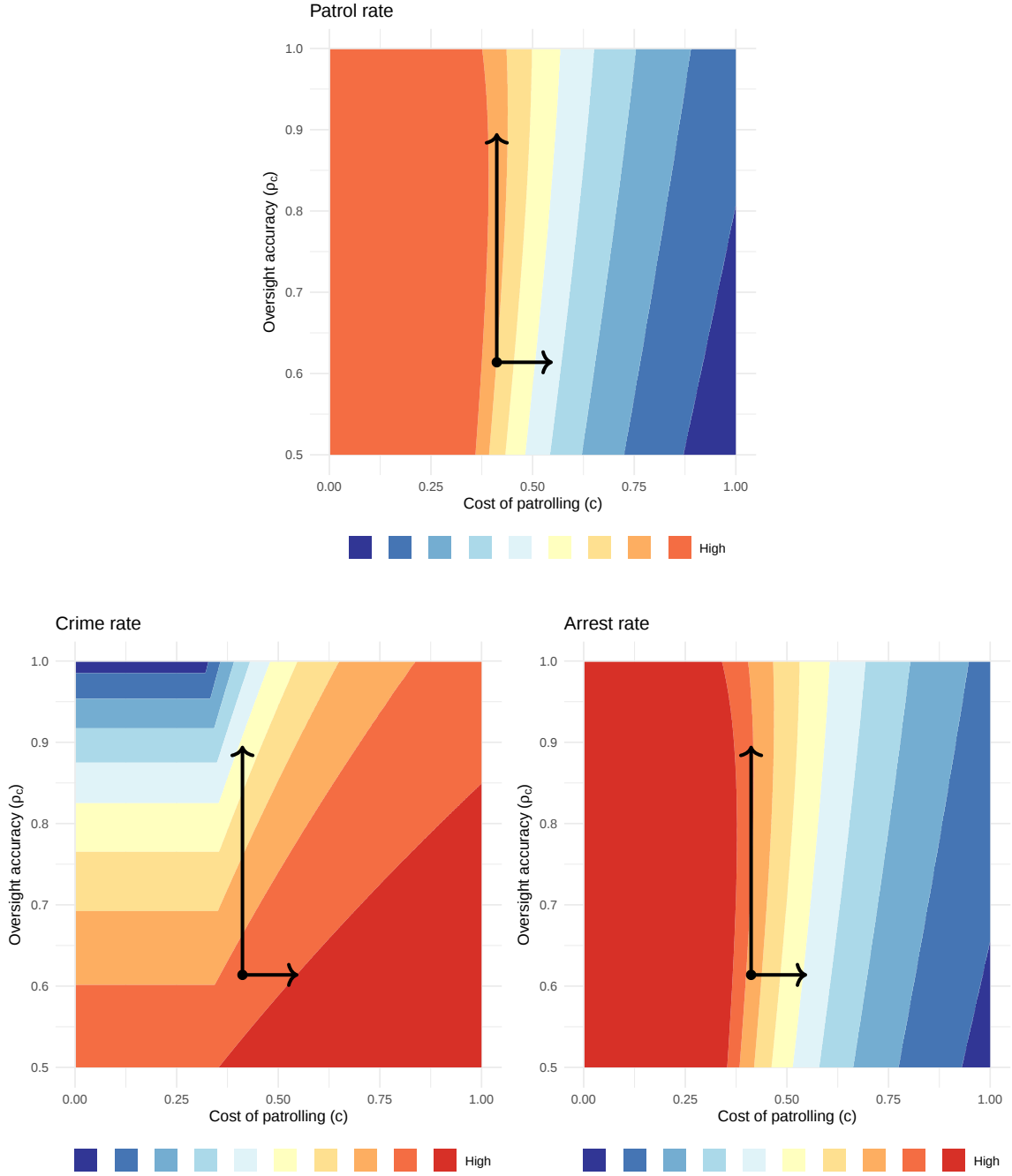
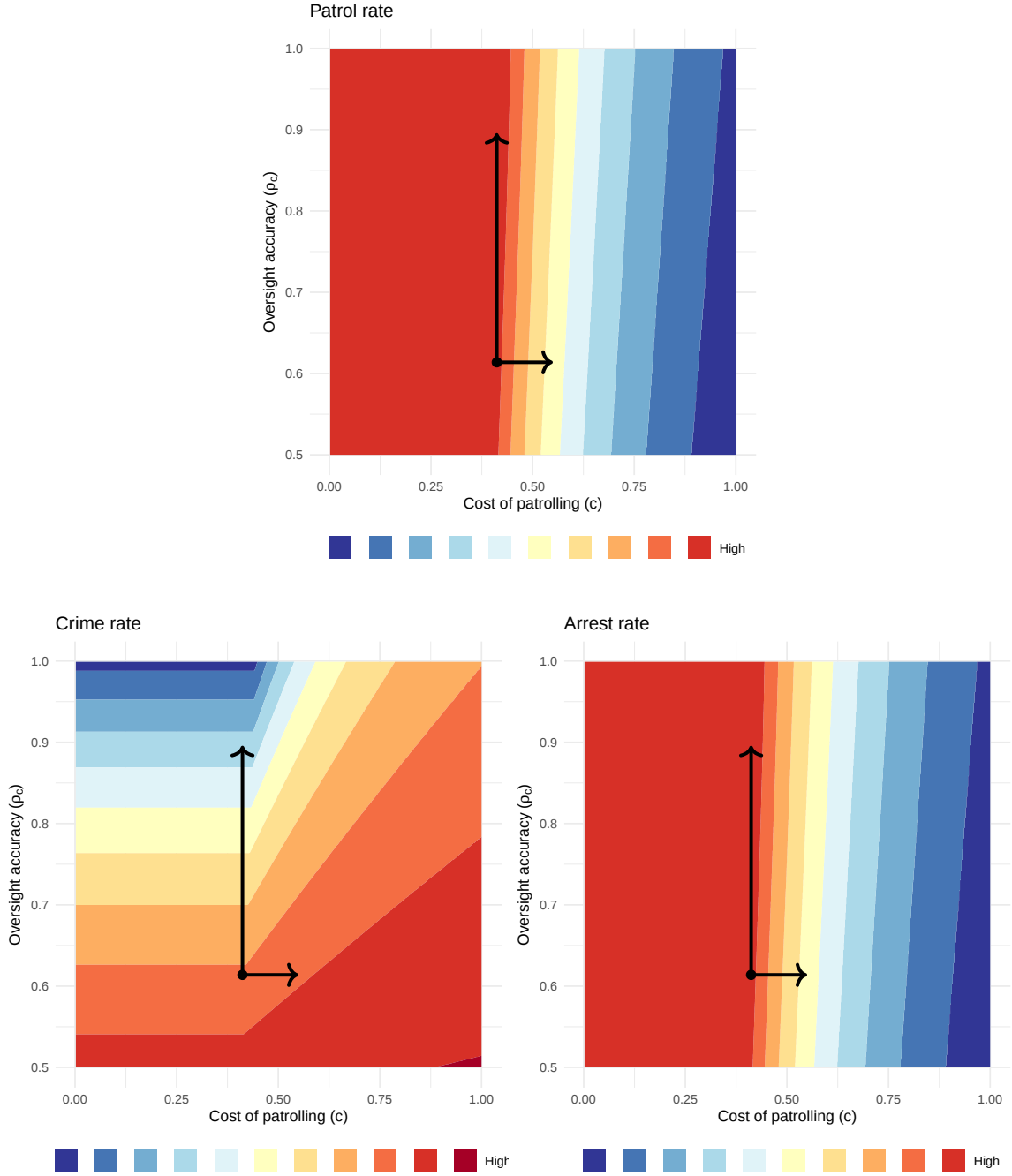


Figure A.2: Contour plots to show how patrol rates, crime rates, and arrest rates change in the cost of effort (c) and the accuracy of the court (ρ_c). For all plots, $a = -25$, $b = 25$, $\rho_p = 0.6$, $P = 0.25$, and $\mu = 0.25$.



B Comparing criminal procedure under the old and new codes

Most cases that go anywhere in Colombia begin with a suspect arrested *en flagrancia*. Under both the old code and the new code, *flagrancia* accounted for the vast majority of arrests (around 75%).²² The remaining one-quarter of arrests began either with citizen crime reports or with government-initiated investigations that ended up producing warrants. As in most criminal justice systems, the vast majority of such reports and investigations do not produce arrests: in 2004, there were more than 800,000 crime reports (excluding *flagrancia* incidents) and fewer than 60,000 arrests with warrants. Given our focus in the paper on the behavior of patrol officers, who are responsible for nearly all *flagrancia* arrests but few warrant arrests (which are carried out primarily by Judicial Police), and given the weight of *flagrancia* in the system overall, we focus on *flagrancia* cases in comparing criminal procedure under Ley 600 (the old code) to criminal procedure under Ley 906 (the new code).

Consider in particular cases that began with a suspect arrested *en flagrancia* for alleged drug possession, vandalism, or theft. Together, these three categories accounted for two-thirds of *flagrancia* arrests in 2004, the year prior to the first wave of rollout of the new code.

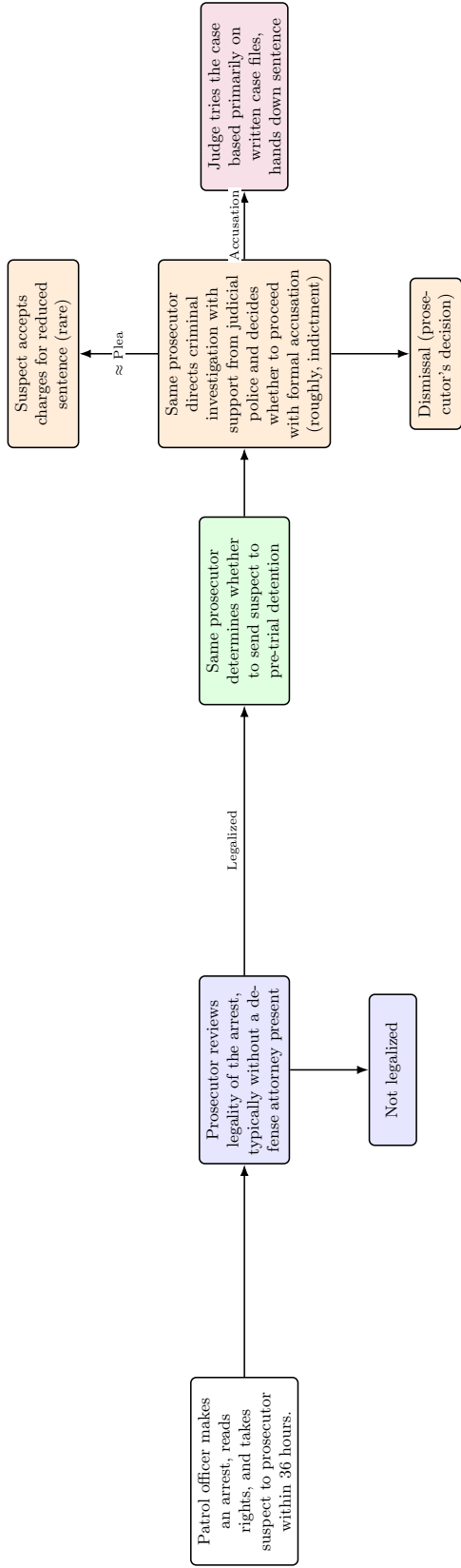
De jure, the patrol officers who made these arrests were bound by strict procedural guarantees even under Colombia’s previous code of criminal procedure (Ley 600 of 2000). The code itself required officers to read the suspect a list of rights and to bring him before a prosecutor within 36 hours, for example, and Colombian law (outside the code) prohibited planting evidence or using unnecessary force. *De jure*, the prosecutor responsible for legalizing the arrest was meant to evaluate whether the arresting officer had followed these rules. But *de facto*, prosecutors often conducted only a cursory review; one officer described this to us as “legalizing” the arrest, in scare quotes. Typically, no defense attorney accompanied the suspect to this initial meeting with a prosecutor.

That same prosecutor would then decide whether the suspect would be held in pre-trial detention; for a long list of crimes, pre-trial detention was the norm rather than the exception, so much so that it was said that *un auto de detención no se le niega a nadie* (Hartmann Arboleda and Martínez Sánchez, 2009, 218).

The prosecutor then directed the criminal investigation—often while the suspect was in pre-trial detention—with support from the judicial police, assembling evidence in a written case file (the *expediente*). Because criminal procedure was written, not oral, the case file

²²We don’t have data that allows us to calculate the fraction of *convictions* that began with *flagrancia* arrests under the old code, but Rivera et al. (2010, 144) find that it was 78% under the new code.

Flagrancia cases under *Ley 600* (the “old code”)



Flagrancia cases under *Ley 906* (the “new code”)

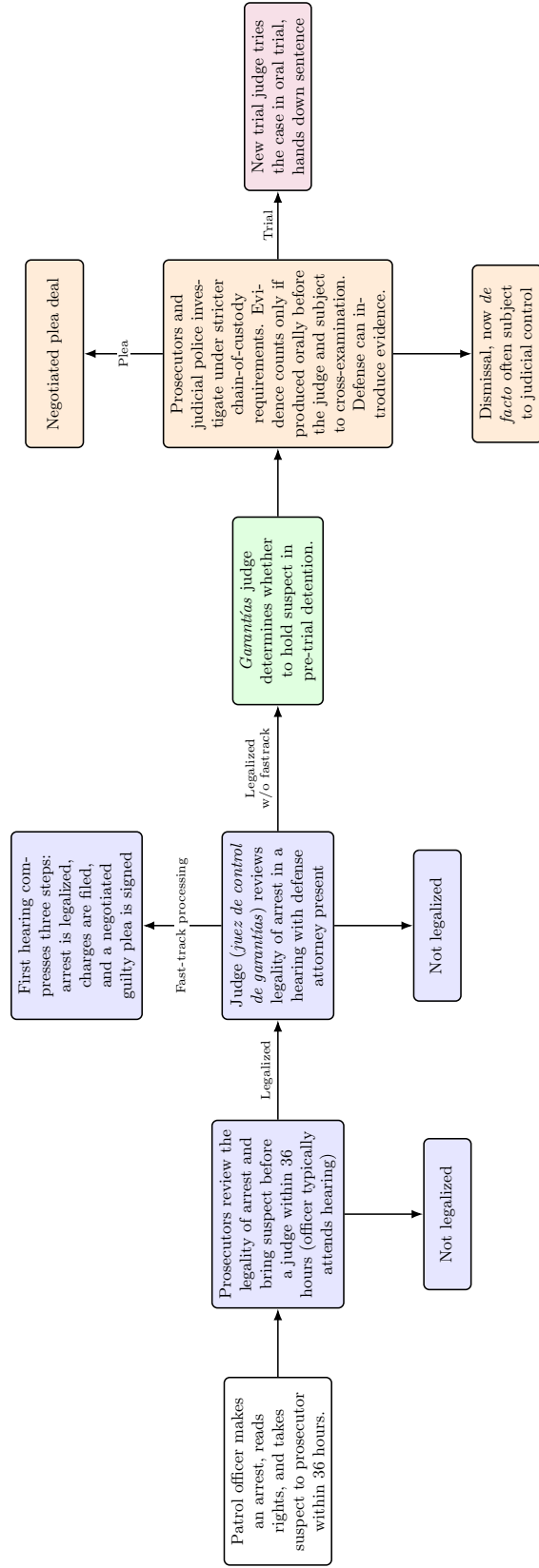


Figure B.1: Criminal procedure for *flagrancia* cases under *Ley 600* (the old code) and *Ley 906* (the new code). Boxes are shaded by stage: legalization (blue), pre-trial detention (green), investigation and case disposition (orange), and trial (purple).

was not merely a compilation of notes, reminders, and evidence: it was a complete, polished narrative, and putting it together was “engorroso” (cumbersome [Hartmann Arboleda et al., 2009](#), 51). That is part of why cases dragged on for so long: 567 days for theft, for example ([Hartmann et al., 2007](#), Table 5, p. 110).

During the investigation, the suspect’s defense attorney had the *de jure* right to consult the case file, respond, and ask that the prosecutor incorporate additional evidence. But in practice, all defense attorneys worked pro-bono (“sin remuneración alguna, incluso subsidiando los gastos con fondos propios”), and “it was not clear whether they had incentives to do a good job” ([Hartmann Arboleda and Martínez Sánchez, 2009](#), 34). (Technically, during the investigation, the suspect could preemptively accept all potential charges for a reduced sentence; in practice, this rarely happened.) And in any case, the prosecutor could often deny the defense’s request to include a given piece of evidence or witness testimony. The prosecutor owned the case file.

At the conclusion of the investigation, the prosecutor would decide whether to proceed with a formal accusation (roughly but not quite *indictment*) or to dismiss the case. The prosecutor did not need judicial approval for dismissal.

If the prosecutor decided to issue a formal accusation, the case would proceed to trial (except in rare cases when suspects chose to accept all charges in exchange for a slightly reduced sentence). Though trials under the previous code did include some oral elements—the judge questioned the suspect, for example, and could introduce new evidence, and the prosecutor and defense attorney could make statements—ultimately the judge’s decision was driven primarily by the huge written case file that was submitted during the formal accusation (indictment).

Under the new code, a suspect caught *en flagrancia* enjoyed nearly identical *de jure* procedural protections during the arrest: the police were to read him a list of rights (the same four as under the previous code, plus one), bring him before a prosecutor within 36 hours, and refrain from planting evidence, for example. But *de facto*, the rules changed. Prosecutors could no longer legalize arrests on their own; rather, prosecutors reviewed the legality of the arrest and then brought the suspect (also within 36 hours) before a new type of judge, the *juez de control de garantías*, who conducted her own review of the legality of the arrest. Though the text of the code did not require the arresting officer to attend these hearings, prosecutors often brought them along to answer questions ([Gutiérrez and Tobón, 2019](#)).

Critically, the new code also required suspects to have a defense attorney present at the arrest-legalization hearing ([Congreso de Colombia, 2004](#), art. 8; recall, under the old code, prosecutors reviewed the legality of the arrest without a judge and without a defense

attorney). Unlike the unpaid defense attorneys who operated under the old system, the new code was implemented together with significant new resources for the public defender’s office; [Hartmann Arboleda et al. \(2009\)](#) concluded that these new attorneys made a big difference: “the value-add that the accusatorial system has had for defense attorneys is that, according to our observation of hearings, the suspect participates much more than in the [previous system], in the sense that attorneys are concerned with ensuring that he understands what’s happening and that his decisions be well-informed and independently taken” (34).

The transfer of arrest-legalization responsibility from prosecutors to judges had two major consequences. First, judges were thought to be more *garantista* than prosecutors under the old code, meaning, more likely to declare an arrest illegal because of procedural violations (perhaps in part because of the presence a defense attorney); prosecutors likely also became more *garantista* in anticipation of the judicial screen. Though prosecutors and the press sometimes criticized *jueces de control de garantías* for being too strict, others interpreted this same behavior as evidence that they were doing their jobs. As one legal scholar and lawyer told [Hartmann Arboleda et al. \(2009, 81\)](#), “My impression of the *jueces de garantía* has been not good but excellent. Of all the actors in the new system, they are the ones who have best understood their role.”

The second consequence was that determining the legality of an arrest took much longer under the new code than under the old code. A survey of 96 patrol officers conducted in 2015 found that this part of the process took more than five hours, on average (excluding time to travel to the prosecutor’s office; [Gutiérrez and Tobón, 2019](#)). This finding echoes qualitative observations by [Hartmann Arboleda et al. \(2009, 38\)](#): “On many occasions we saw that judges were inclined to take the time necessary in hearings to ensure that the defendant understood proceedings being brought against him.”

The changes described thus far might all seem to be on the side of preventing one type of error (punishment of the innocent). The next major shift in the process pushed in the other direction, seeking to avoid failed prosecution of guilty parties: fast-track processing for *flagrancia* cases. Many *flagrancia* arrest-legalization hearings ended up also including an indictment *and* a plea bargain, effectively ending the case almost as soon as it had begun ([Hartmann Arboleda et al., 2009, 70](#)). The rise of plea bargaining was among the most profound shifts from the old system to the new ([Zorro Medina, 2023](#)).

For all other cases (i.e., those not resolved on the fast track), the same judge who ruled on the legality of the arrest would also decide whether the suspect would be held in pre-trial detention ([Hartmann Arboleda and Martínez Sánchez, 2009](#)). Qualitative accounts suggest that the use of pre-trial detention dropped as a result (as one attorney put it, “Tanto la Policía como los fiscales perdieron el poder de estar privando de la libertad a las personas porque sí,

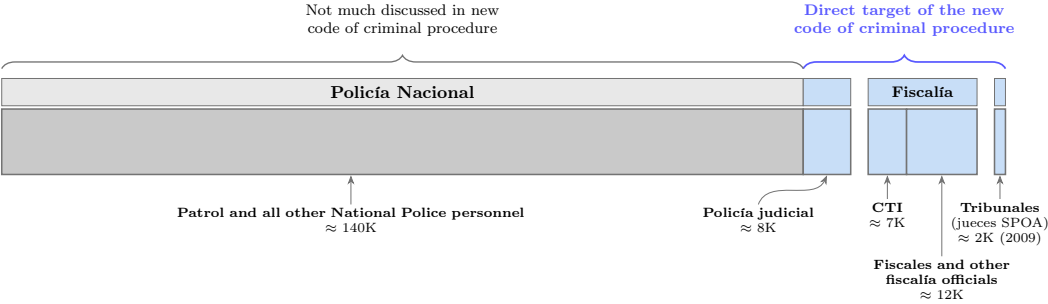
como venían acostumbrados” [Hartmann Arboleda et al. 2009](#), 81), though the quantitative evidence is less clear ([Hartmann Arboleda and Martínez Sánchez, 2009](#); [Zorro Medina, 2020](#)).

The new code, together with a Constitutional Court decision that interpreted it ([Corte Constitucional de Colombia, 2005](#)), also ended up restricting prosecutors’ power to dismiss cases without judicial approval. Under the old code, prosecutors decided whether to dismiss cases or proceed to indictment ([Congreso de Colombia, 2000](#), arts. 327 and 39; the old code enumerated five grounds for prosecutor-driven dismissal, including *ausencia de responsabilidad*); the new code appeared to restrict the valid grounds for prosecutor-driven dismissal ([Congreso de Colombia, 2004](#), art. 79), and the court decision left many prosecutors with the impression that every dismissal required a hearing before a judge (to protect the victim). [Hartmann Arboleda et al. \(2009, 41–44\)](#) concludes that, *de jure*, judicial approval was only required if the victim contested the prosecutor’s decision, as some prosecutors understood—but widespread confusion about this point led to underuse of dismissal.

The adversarial, oral trials held under the new system bore little resemblance to their old-system counterparts. Under the new code, a dedicated trial judge heard oral arguments from both the prosecution and the defense in open court. Rather than place a written, prosecutor-owned case file at the center of judicial decision-making, the new trials were meant to establish *igualdad de armas*, putting both sides on equal footing before the judge. In observing many such trials, Hartmann found that orality prevailed not just *de jure* but *de facto* ([Hartmann Arboleda et al., 2009](#), 74), and that time-to-sentence fell dramatically as a result (see also [Departamento Nacional de Planeación, 2009](#), 27). Of course, the long written-file tradition did not vanish overnight. Some litigants still read from prepared scripts rather than argue orally ([Hartmann Arboleda et al., 2009](#), 75); worse, some trial judges asked to see the prosecutor’s evidence before the hearing, ([Hartmann Arboleda et al., 2009](#), 58). But overall, “in practice the system has functioned orally from the beginning, meaning that the official record is an audio recording, and decisions are made in the context of a public, oral hearing” (p. 74). [Hartmann Arboleda et al. \(2009, 34\)](#) concludes that orality was “the most positive feature of the functioning of the new system.”

Figure [B.1](#) summarizes these changes. The following figure visualizes the relative sizes of the judicial police, prosecutors, and judges, relative to the rest of the police:

Figure B.2: Snapshot of relative sizes of judicial police, prosecutors, judges, and all other police



Totales: Policía Nacional 145 871 · Fiscalía 18 708 · Tribunales 1 881. **Fuentes:** Policía Nacional, Dir. Talento Humano/OBTAH (total 2008, cruzado con MinDefensa); presentación DIJIN “Policía Judicial y SPA” (8 155 = DIJIN 1 265 + SIJIN 5 572 + DIASE/DIRAN/etc.); CEJ, *Balance 5 años* (policía judicial 8 116 en 2007; CTI). Fiscalía: FGN *Informe de Gestión 2008–2009* (vía CEJ); “fiscales y personal” = planta no-CTI (de los cuales ~3 995 fiscales). Tribunales: CSJ (planta SPOA 1 881 en 2009).

C Changes in crime registration

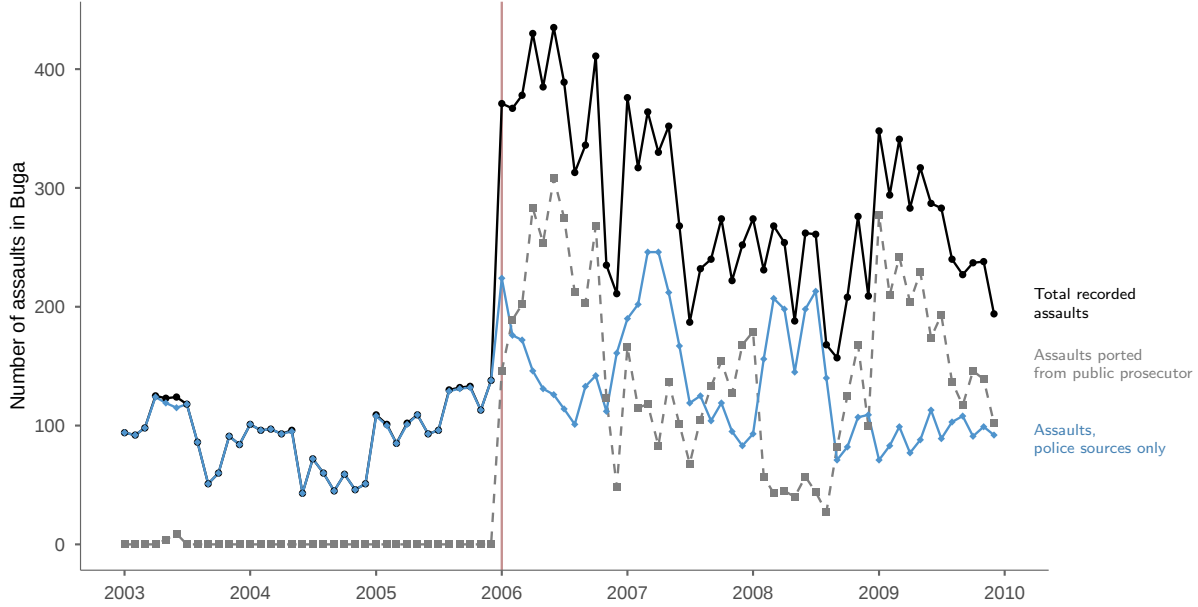
In Colombia, the National Police is not the only institution to which citizens report crimes. Many crimes are reported instead to the Public Prosecutor’s office (*Fiscalía*), and still others are reported to security secretariats within city governments. Prior to the rollout of the new code, the police database SIEDCO generally included only crimes reported to the police; over the course of the rollout, SIEDCO began incorporating crimes reported to other institutions, such as the Public Prosecutor. We write *generally* because these changes were far from uniform: in some places and for some crimes, officers incorporated other institutions’ crime reports into SIEDCO by hand even before the new code; similarly, there were places in which SIEDCO began ingesting other institutions’ crime reports *en masse* at moments that did not correspond to the reform rollout. But the coincidence of reform rollout and the incorporation of other institutions’ crime reports occurred in enough places that using total SIEDCO crime counts to estimate the effect of the new code on crime produces misleading results.

We are not the first to notice this pattern or its consequences for considering the effects of the new code. Analysis in the *Revista Criminalidad* ([Policía Nacional de Colombia, 2005, 36](#)), a magazine produced by the police and primarily for the police, concluded that:

It wasn’t a coincidence that there was an increase in [mugging in three departments where the new code had just come into effect]. We expected that, beginning with the implementation of the [new code], statistics from various state agencies would be consolidated into a unified system . . . so that the sum total of crime reports received by each institution, would be consolidated into total statistics for the region as a whole. **In short, our read of the data is to conclude that there was no increase in crime, but rather bigger and better crime registration, which produced the increases reported above.** (Emphasis added).

The data that we obtained from the police, which include *medio de conocimiento*, i.e. source of crime report, allow us to document this phenomenon quantitatively. In the mid-size city of Buga, for example, SIEDCO previously recorded only those assaults reported to the police; coincident with the new code, SIEDCO also incorporated assaults reported to the public prosecutor’s office. Ignoring this change in SIEDCO crime registration makes it look as if the number of assaults in Buga tripled within one month under the new code (see Figure C.1). The blue line in this figure plots our attempt to correct for this change by using only crime counts from internal police sources.

Figure C.1: Partial correction for a change in crime registration

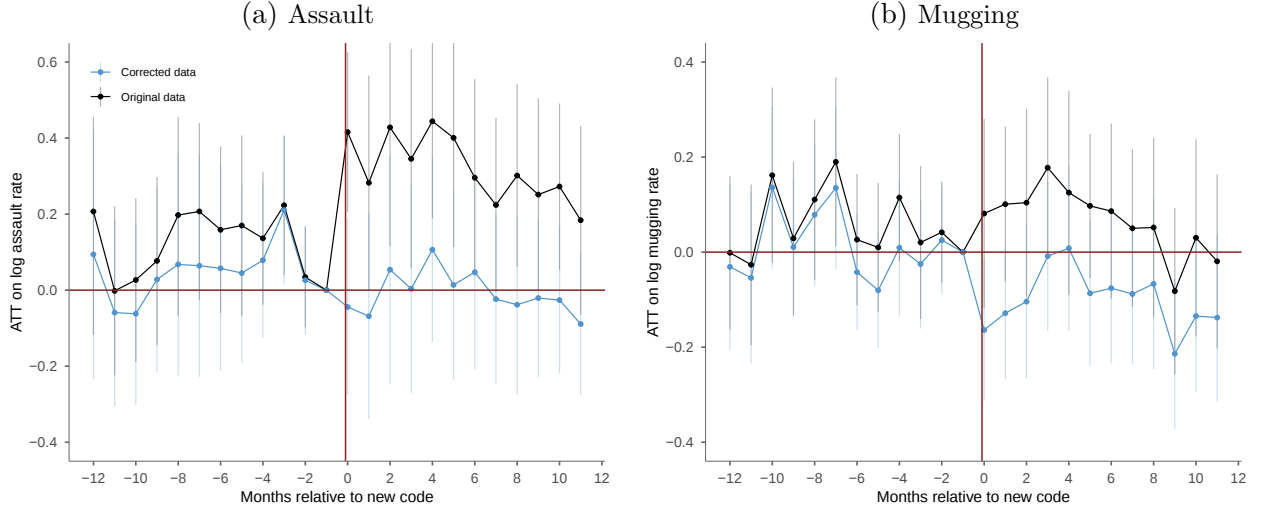


The black line plots the total number of assaults recorded in the police database SIEDCO. The gray line plots the number of assaults recorded in SIEDCO that were originally reported to the public prosecutor; in 2006, in Buga, the police began incorporating the public prosecutor’s data into SIEDCO, creating the illusion of a sharp increase in assault. The blue line plots a count of assaults using internal police sources only.

In fact, we consider two different definitions of “police sources only.” The first includes only sources that appear to be internal to police in all judicial districts, of which the biggest categories are *por información policiva*, *vigilancia* and *inspección policía*. This is the measure that we use in Figure C.2, which shows that the original result is an artifact of using the original SIEDCO counts without accounting for the incorporation of other institutions’ crime reports.

The other measure adds two additional sources: *querella* and *por denuncia*. It appears that these sources are in some places used to indicate crime reports received directly by the police, and in other places are used to indicate crime reports imported from Fiscalía. In Antioquia, for example, the *denuncia/querella* trend looks like the dotted gray line in Figure C.1: it’s zero under the old code, then jumps up right in the month when the new code comes into effect. In Antioquia, unlike in Buga, there’s no source called *fiscalía*, and one officer who worked in the data unit told us that the labels *denuncia/querella* could be used to indicate crime reports coming from *fiscalía*. That’s why we do not include *denuncia/querella* in our principal measure in Figure C.2: it should create the artificial appearance of a jump in assaults. The problem with excluding *denuncia/querella* is that, in four judicial districts (Bogotá, Cúcuta, Sincelejo, and Villavicencio), it does *not* enter at the same time as the

Figure C.2: No detectable effect of the new code on assault or mugging



The black lines use the original crime counts analyzed by [Acosta Mejía et al. \(2023\)](#); the blue lines use counts restricted to internal police sources, correcting for the change in crime registration described in this appendix. Both series report municipality-level horizon-specific ATTs estimated using [De Chaisemartin and d’Haultfoeuille \(2020\)](#), the official implementation dates, and a sample that excludes Antioquia. Confidence intervals use 1,000 bootstrap replications clustered by judicial district.

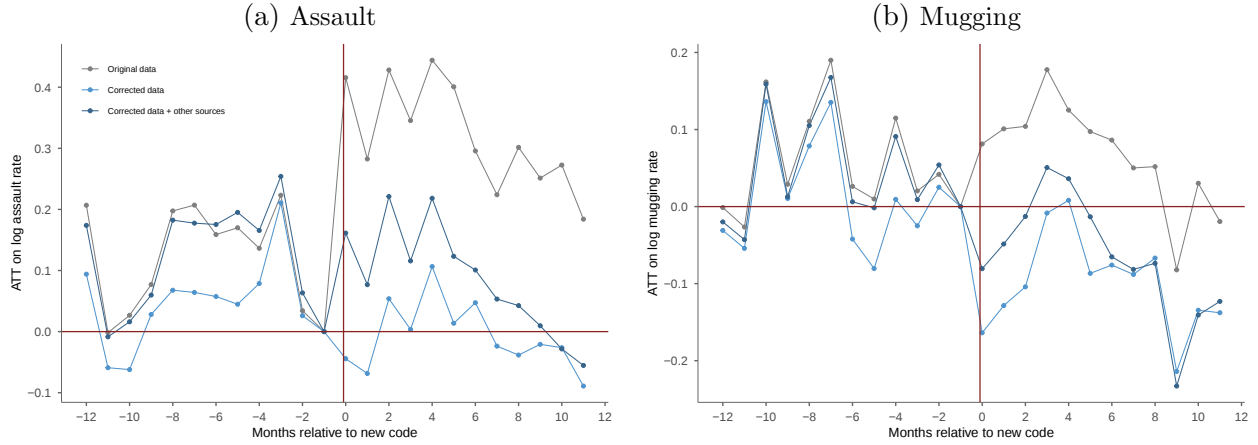
new code but rather exists throughout the 2003–2009 period. Excluding *denuncia/querella* in these places therefore strikes us as problematic.

For that reason, we repeat the analysis with this second measure. Figure C.3 compares the original counts, our principal police-only measure, and the expanded measure that also includes *denuncia/querella*. All three series use the same [De Chaisemartin and d’Haultfoeuille \(2020\)](#) estimator, so differences across the lines arise from the crime-count definitions rather than from different event-study constructions. The expanded measure is best interpreted as a sensitivity check because, in some places, *denuncia/querella* appears to include reports imported from the public prosecutor’s office.

Table C.1 reports the corresponding pooled estimates using the method proposed by [De Chaisemartin and d’Haultfoeuille \(2020\)](#). The estimates based on the corrected crime counts are substantially smaller than those based on the original counts.

As mentioned in the main text, the incorporation of other institutions’ *denuncias* is not the only change in crime registration that coincided with the new code. The other two were: (1) lowering the monetary threshold for *theft* in Bogotá specifically and (2) making it easier for citizens to report crimes. Another article in *Revista Criminalidad* highlighted the consequences of the latter for interpreting the effect of the new code on crime rates ([Rodríguez Cuadrado, 2008](#)):

Figure C.3: Effect of the new code on crime: Alternate measures



The gray lines use the original crime counts, the lighter blue lines use our principal police-only measure, and the darker blue lines add *denuncia/querrela* to the police-source definition. All series report municipality-level horizon-specific ATTs estimated using [De Chaisemartin and d'Haultfoeuille \(2020\)](#), the official implementation dates, and a sample that excludes Antioquia.

The increase in crimes in 2006 is explained by, among other causes, the entry into force of the [new code], which fortified the crime-reporting stations (*centros de atención al ciudadano para recepción de denuncias*), making them more visible and accessible to the community. Moreover, the government established more Mobile Immediate Attention Centers (CAIs); with these measures, the community was more motivated to report crimes of which they were victims, or of which they had knowledge.

Our disaggregation of SIEDCO crime reports (into preexisting and new sources) does not address either of these problems, meaning that, if anything, our proposed measure of assault and/or mugging could still create the artificial appearance of a rise in crime coincident with the introduction of the new code. In other words, the theft threshold in Bogotá and/or spread of crime-reporting stations should tend to

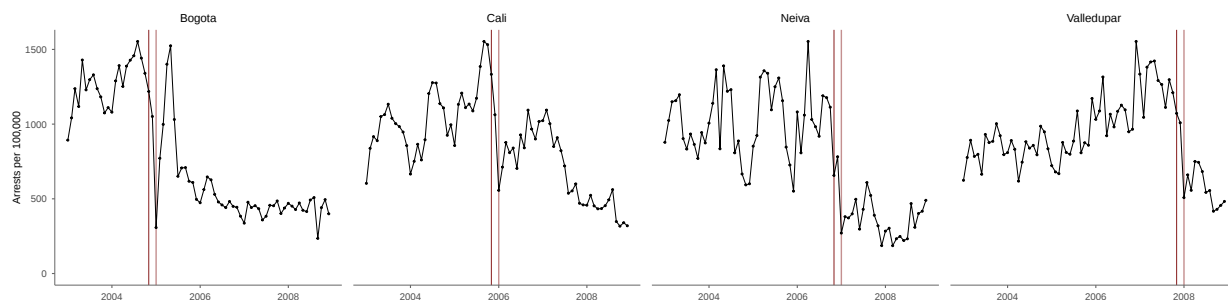
Table C.1: Estimates of the effect of the new code on assault and mugging rates

	Assault rate		Mugging rate	
	Original data	Corrected	Original data	Corrected
Panel A: Principal measure of “police sources only”				
Effect of new code	0.321 (0.112)	-0.007 (0.103)	0.067 (0.075)	-0.100 (0.062)
Effect (% change)	37.9	-0.7	6.9	-9.5
Pre-period mean	0.816	0.706	0.497	0.404
Observations	82,656	82,656	79,044	79,044
Panel B: Expanded measure of “police sources only”				
Effect of new code	0.321 (0.112)	0.087 (0.132)	0.067 (0.075)	-0.066 (0.066)
Effect (% change)	37.9	9.1	6.9	-6.4
Pre-period mean	0.816	0.771	0.497	0.477
Observations	82,656	82,656	79,044	79,044

Estimates use [De Chaisemartin and d’Haultfoeuille \(2020\)](#) at the municipality level, the official implementation dates, and the first 12 post-treatment months; Antioquia is excluded. Standard errors use 1,000 bootstrap replications clustered by judicial district. The “Effect (% change)” row expresses each estimate as a percentage change in the crime rate; because all columns are estimated in logs, it is $100 \times (e^{\hat{\beta}} - 1)$.

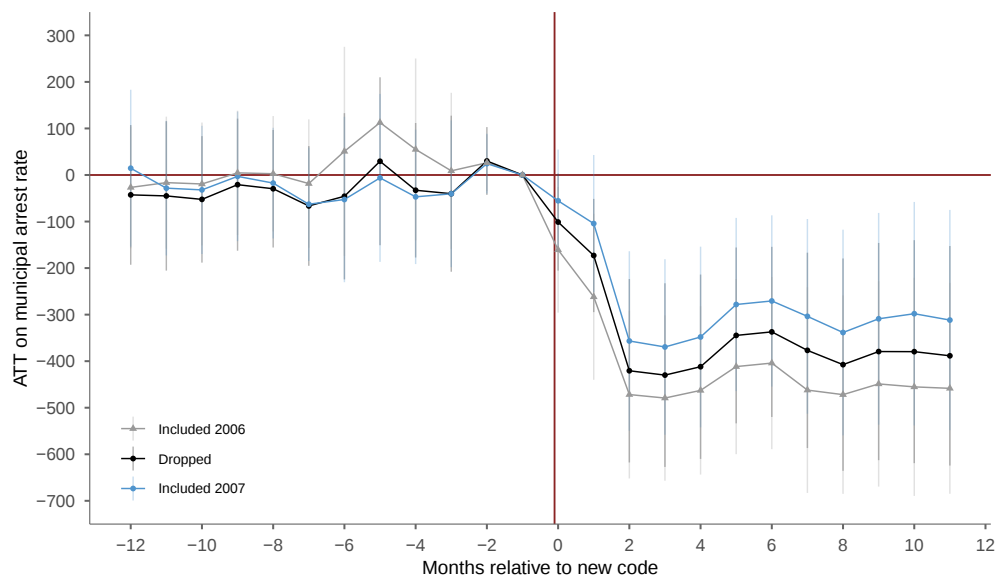
D Additional results on the decline in arrests

Figure D.1: The New Code and Arrest Rates in Four Cities (Examples)



This figure plots raw arrests per 100,000 people per year in four cities (one in each wave). The pale red lines mark the Januaries when the new code came into effect in each place; the darker red lines mark the preceding Novembers (when we observe anticipation effects).

Figure D.2: Including or dropping Antioquia does not substantively change the results.



Our main results in Figure 4 drop the Antioquia judicial district given uncertainty about the timing of reform implementation. Its inclusion does not substantively change our findings with respect to the decline in arrests. Using the estimator proposed by [De Chaisemartin and d'Haultfoeuille \(2020\)](#). Confidence intervals use 1,000 bootstrap replications at the judicial-district level.

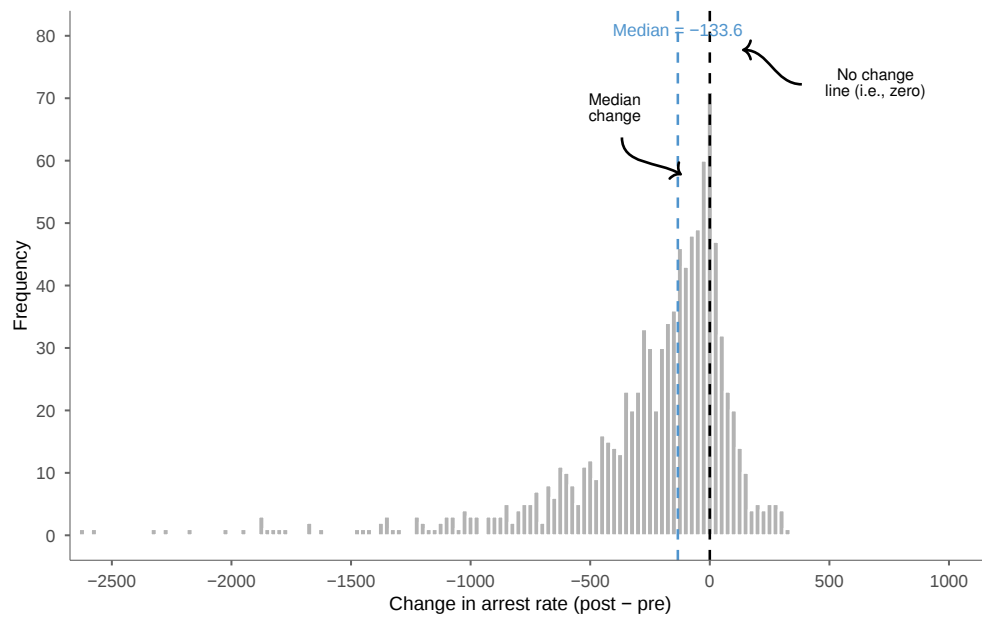
Table D.1: Estimates of the decline in arrests, using different estimators

	Municipal level				Judicial district level					
	Arrest rate				Homicide		Drug		Vandalism	
	Rate	Rate	Rate	Log	Rate	Log	Rate	Log	Rate	Log
Estimates from De Chaisemartin and d’Haultfoeulle (2020) (reported in main text)										
Effect of new code	-345.81 (89.30)	-431.28 (85.95)	-378.98 (91.66)	-0.47 (0.09)	-1.28 (2.75)	-0.02 (0.17)	-72.79 (44.25)	-0.62 (0.18)	-64.35 (23.73)	-1.03 (0.27)
Observations	31,056	31,056	972	972	972	972	972	972	972	972
Change (%): log-transformed				-37.6		-1.7		-46.1		-64.3
Change (%): aggregate rate		-47.8								
Change (%): avg. unit	-30.3	-35.5	-36.2		4.5		-36.2		-67.8	
Pre-period mean	571.29	901.58	744.22	6.48	18.06	2.76	159.22	4.55	54.64	3.21
Estimates from Borusyak et al. (2024)										
Effect of new code	-295.03 (24.80)	-414.90 (104.49)	-473.11 (133.49)	-0.48 (0.10)	-3.78 (2.22)	-0.13 (0.10)	-143.32 (62.52)	-0.58 (0.17)	-79.57 (30.45)	-1.19 (0.31)
Observations	41,314	41,314	1,303	1,303	1,303	1,303	1,303	1,303	1,303	1,303
Change (%): log-transformed				-38.1		-12.6		-44.0		-69.7
Change (%): aggregate rate		-46.3								
Change (%): avg. unit	-14.1	-30.7	-37.2		-9.4		-43.4		-32.9	
Pre-period mean	571.14	896.35	738.69	6.47	17.92	2.76	158.58	4.54	54.80	3.20
Estimates from Callaway and Sant’Anna (2021)										
Effect of new code	-340.33 (26.56)	-400.87 (44.78)	-369.81 (103.92)	-0.45 (0.09)	-0.86 (3.03)	0.03 (0.20)	-70.53 (43.85)	-0.60 (0.19)	-60.93 (24.41)	-1.03 (0.28)
Observations	53,556	53,556	1,694	1,694	1,694	1,694	1,694	1,694	1,694	1,694
Change (%): log-transformed				-36.2		3.3		-45.2		-64.2
Change (%): aggregate rate		-44.4								
Change (%): avg. unit	-30.8	-31.5	-34.5		10.4		-33.0		-73.6	
Pre-period mean	570.78	903.15	744.22	6.48	18.06	2.76	159.22	4.55	54.64	3.21
Population weights	✓									

The point estimates from De Chaisemartin and d’Haultfoeulle (2020) reproduce those in Table 1; standard errors use 1,000 bootstrap replications at the judicial-district level. For Callaway and Sant’Anna (2021), columns 1–2 report municipal-level standard errors and columns 3–10 use 1,000 bootstrap replications at the judicial-district level. For Borusyak et al. (2024), columns 1–2 report municipal-level standard errors and columns 3–10 use 1,000 bootstrap replications at the judicial-district level. The alternative-estimator samples end before 2008, omit post-treatment observations from Wave IV, and retain at most the first 12 treated months; the Borusyak et al. (2024) specifications are further restricted to 2004–2007. Percentage changes are calculated as in Table 1. For log outcomes, they equal $100 \times (e^{\hat{\beta}} - 1)$. For rate outcomes, the avg. unit row reports the estimate after dividing each unit’s rate by its own pre-treatment mean. The aggregate-rate row divides the population-weighted municipal estimate in Column 2 by its population-weighted pre-treatment mean.

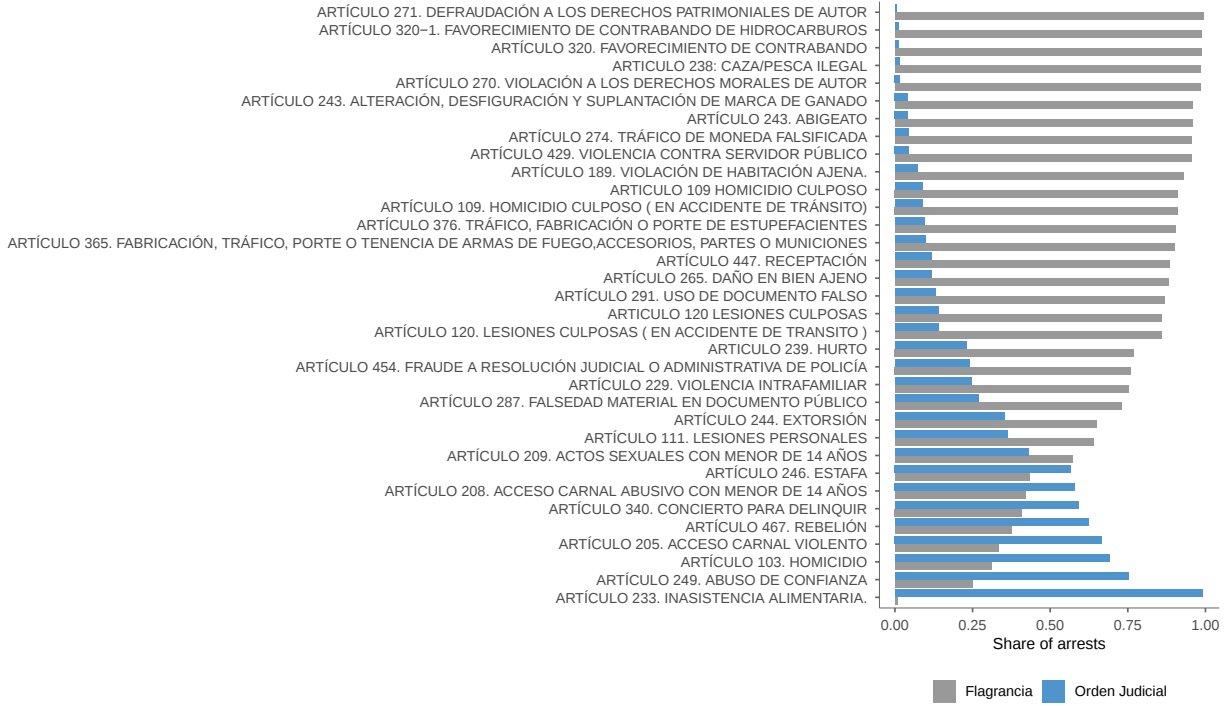
All judicial-district specifications using Borusyak et al. (2024) (Columns 3–10) include unit-specific linear trends because the pre-treatment estimates rely on the earliest periods as a reference group. Excluding the trends yields a somewhat larger negative point estimate of -0.25 in Column 6.

Figure D.3: Arrest Rates Decline in 78% of Municipalities



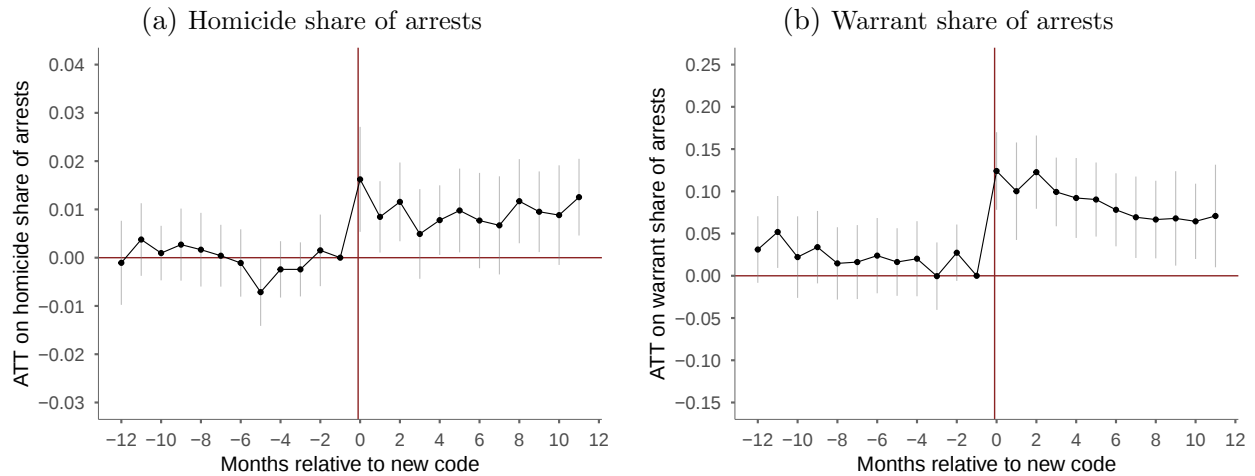
This figure plots the middle 98% of the distribution of pre-post changes in the arrest rate across Colombian municipalities, comparing the twelve months before the reform comes into effect to the twelve months after. The 78% reported in the title is calculated from the full sample.

Figure D.4: Warrants Seldom Used for Minor Offenses.



This figure plots the share of arrests for common charges with and without a warrant nation-wide in 2003 and 2004 (during the pre-treatment period). A charge is common if there were more than 1,000 arrests for that charge nationwide in any year between 2003 and 2008.

Figure D.5: The Effect of the New Code on the Composition of Arrests



Reform increased the share of arrests for homicide (by reducing arrests for more minor offenses) and increased the share of arrests made with warrants (by reducing the share of warrantless arrests). Estimates use the method proposed by [De Chaisemartin and d'Haultfoeuille \(2020\)](#) and the official implementation dates. The homicide-share panel is weighted by population; the warrant-share panel is unweighted. Confidence intervals use 1,000 bootstrap replications at the judicial-district level.

Figure D.6: Prison Population During the Rollout of the New Code

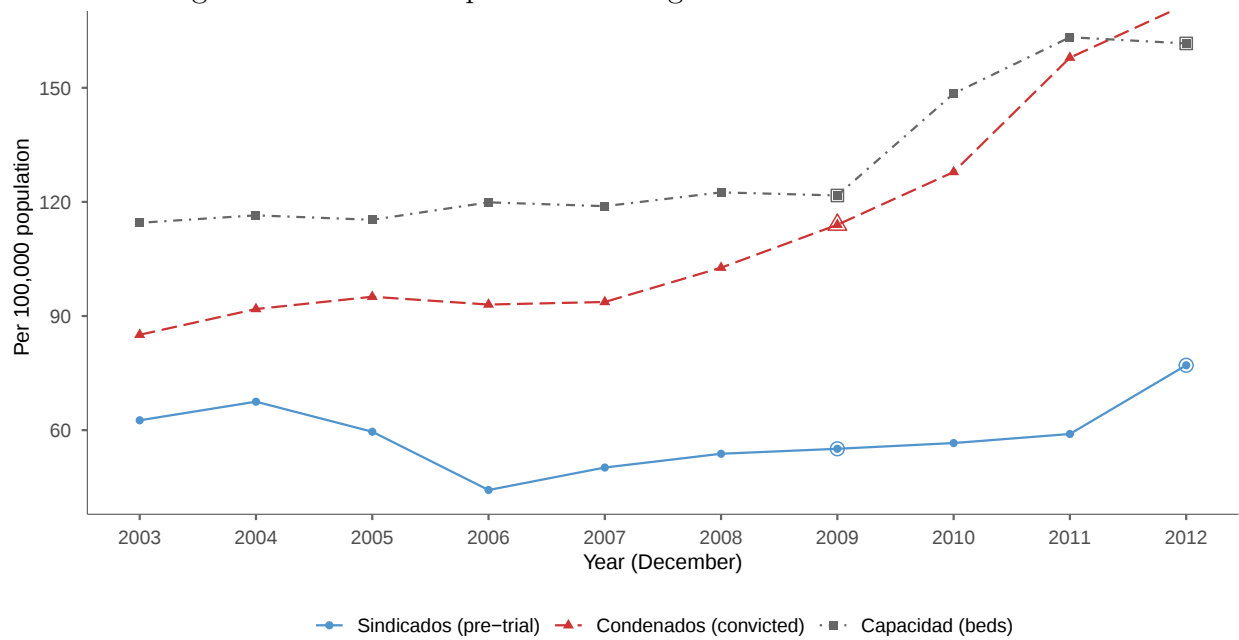


Figure D.7 plots the relationship between minimum-sentence length (an imperfect measure of crime severity) and the normalized decline in the arrest rate, for the eighteen arrest categories with at least 1,000 arrests in at least one year between 2003 and 2008. The negative correlation suggests that the arrest rate declined more for less-serious than for more-serious offenses.

Table D.2 reports the offense-by-offense estimates that underlie Figure D.7. The negative relationship between the size of the arrest-rate decline and offense severity is driven by offenses at both extremes. At the low-severity end, arrests for unlawful entry into a dwelling (breaking and entering) and for vandalism fell by well over half of their pre-reform baseline rates. At the high-severity end, arrests for homicide and for sexual abuse of a minor under fourteen were essentially unchanged (normalized ATTs near zero).

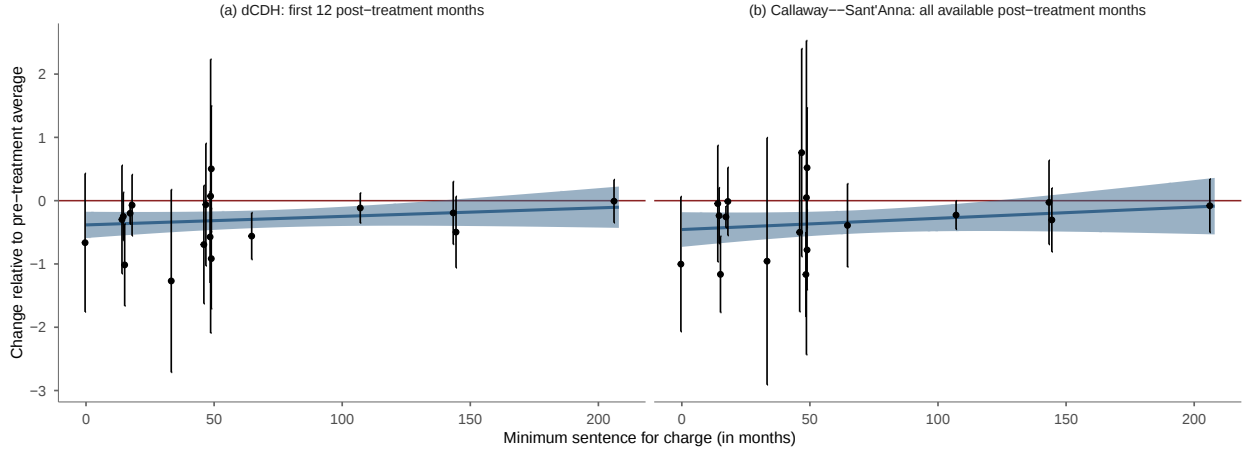
Unlike our main results, these crime-specific estimates are more sensitive to choice of estimator, which is why we report results from both De Chaisemartin and d’Haultfoeuille (2020) and Callaway and Sant’Anna (2021). Here, we normalize the outcome first (meaning, divide by the pre-treatment mean in each unit), and *then* calculate the ATT, as in the bottom panel of Figure 5; the percent changes differ slightly from those of Table 1, where we calculate the ATT in levels and then normalize ex-post, but the order is the same: vandalism arrests decline more than drug arrests which decline more than homicide arrests. In other words, the results are robust to different ways of calculating relative change.

Minimum sentence is at best noisy signal of crime severity. Drug arrests (Article 376) enter Figure D.7 at a statutory minimum sentence of sixty-four months, which considerably overstates the severity of most cases: as we note in the main text, arrests under this article are overwhelmingly for simple possession rather than trafficking or production. One study of drug prosecution in Colombia concluded that

according to a sample of drug-conviction case files from 2011 to 2014, 30.71% of *convictions* [emphasis added] were for the offense “carrying on one’s person” (*llevar consigo*) exclusively—that is, for simple drug possession, without any proof that the substances were being sold, transported, stored, or carried for purposes other than personal use. In some of these cases, people were even convicted for possessing quantities below or slightly above the personal dose (Uprimny Yepes et al., 2017, 67).

In other words, the *de facto* minimum sentence for drug arrests is far less than 64 months. Measured by *de facto* rather than *de jure* severity, the drug point would move to the left in Figure D.7, making the slope steeper. The same is true of vandalism arrests. The statutory minimum is 16 months, but in practice, almost no one serves time for vandalism: in 2004, there were 17,000 vandalism arrests but just 7,000 vandalism convictions. For these reasons,

Figure D.7: Arrest Rate Declines Less for More Serious Crimes



Each point is the normalized ATT for a common charge (as defined in Figure D.4) or homicide, plotted against the offense's minimum statutory prison sentence. Panel (a) uses the method proposed by [De Chaisemartin and d'Haultfoeuille \(2020\)](#), not-yet-treated districts, and the first 12 post-treatment months. Panel (b) uses the method proposed by [Callaway and Sant'Anna \(2021\)](#), Wave IV as the never-treated comparison group, and all available post-treatment months: 36, 24, and 12 months for Waves I–III, respectively. Outcomes in both panels are divided by each district's pre-reform mean for the offense. Homicide combines penal-code Articles 103–106 and 135, as in Table D.2. The line and shaded 95% confidence interval come from a random-effects meta-regression of all 18 offense-specific ATTs, including homicide, on minimum sentence. Differences across panels therefore reflect both the estimator and the associated comparison-group and post-treatment aggregation choices.

we believe that Figure D.7 understates the relationship between severity and decline in arrests, but we use the statutory minimum sentence for transparency and consistency.

Table D.2: The Effect of the New Code on Arrest Rates, by Offense

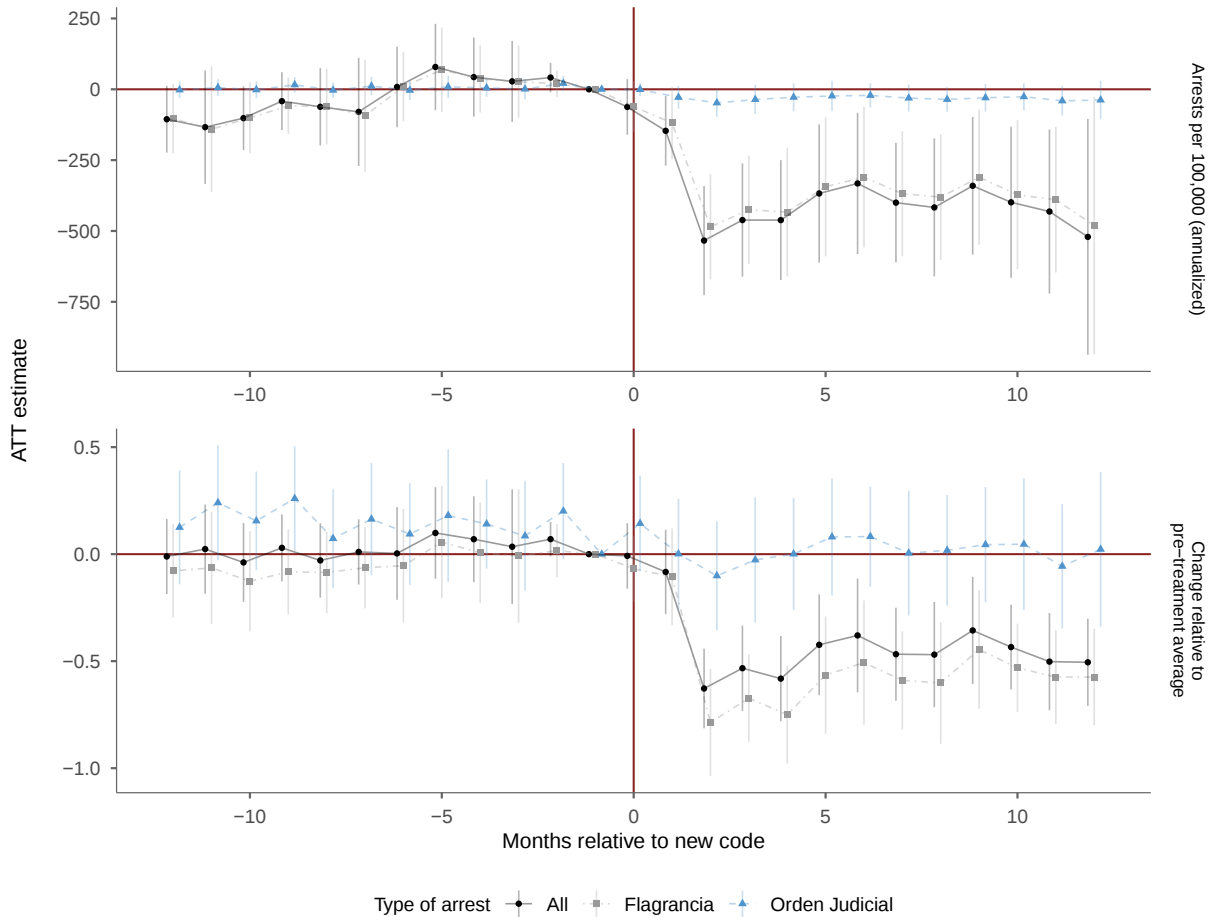
Offense	Min. sentence	De Chaisemartin and d'Haultfoeuille (2020)		Callaway and Sant'Anna (2021)		Pre-period mean
		ATT	Normalized ATT	ATT	Normalized ATT	
Unlawful entry into a dwelling	0	−3.9 (1.5)	−0.66 (0.56)	−3.3 (1.5)	−1.00 (0.54)	2.4
Failure to provide family support	16	−8.6 (4.8)	−0.07 (0.25)	−5.9 (5.8)	−0.01 (0.27)	19.9
Theft	16	−82.5 (37.2)	−0.20 (0.09)	−139.3 (85.2)	−0.26 (0.09)	145.7
Fraud	16	−2.0 (1.2)	−0.25 (0.19)	−2.6 (1.0)	−0.24 (0.23)	6.6
Breach of trust	16	−4.2 (3.3)	−0.30 (0.44)	−5.0 (4.5)	−0.05 (0.47)	3.3
Vandalism	16	−64.3 (25.1)	−1.02 (0.33)	−63.2 (19.3)	−1.16 (0.31)	33.6
Infringement of authors' moral rights	32	−9.0 (5.3)	−1.27 (0.74)	−11.1 (6.5)	−0.96 (1.00)	7.2
Domestic violence	48	−37.9 (17.1)	−0.57 (0.37)	−59.2 (32.4)	−1.17 (0.34)	32.5
Trafficking in counterfeit currency	48	−4.9 (2.0)	−0.92 (0.40)	−4.1 (1.5)	−0.78 (0.32)	3.7
Forgery of a public document	48	−0.8 (0.7)	0.50 (0.51)	−1.3 (0.8)	0.52 (0.48)	2.8
Criminal conspiracy	48	−0.9 (2.5)	−0.06 (0.49)	3.8 (4.1)	0.76 (0.84)	6.5
Violence against a public servant	48	−3.1 (1.2)	−0.69 (0.48)	−3.3 (1.4)	−0.50 (0.64)	3.1
Receiving stolen property	48	−8.0 (3.3)	0.07 (1.10)	−7.5 (3.3)	0.05 (1.27)	7.3
Drug [†]	64	−73.0 (42.3)	−0.56 (0.19)	−92.7 (53.8)	−0.39 (0.34)	110.1
Illegal firearms possession or trafficking	108	−7.9 (4.3)	−0.12 (0.12)	−14.3 (5.6)	−0.23 (0.11)	38.1
Rape	144	−3.0 (1.4)	−0.50 (0.29)	−2.7 (1.1)	−0.30 (0.26)	5.0
Sexual abuse of a minor under 14	144	−1.7 (1.0)	−0.19 (0.25)	−0.9 (1.1)	−0.03 (0.34)	4.6
Homicide	208	−1.3 (2.7)	−0.01 (0.17)	−3.1 (3.5)	−0.08 (0.22)	15.9

Each row reports judicial-district-month estimates for one offense. The specification based on [De Chaisemartin and d'Haultfoeuille \(2020\)](#) compares treated districts with not-yet-treated districts and averages the first 12 post-treatment months. The specification based on [Callaway and Sant'Anna \(2021\)](#) uses Wave IV as the never-treated comparison group and averages all available post-treatment months: 36, 24, and 12 months for Waves I–III, respectively. The *ATT* columns use annualized arrests per 100,000 residents. The *Normalized ATT* columns divide that rate by each district's pre-reform mean for the offense, so estimates are proportional changes from the baseline. The final column reports the corresponding average pre-reform arrest rate. The minimum sentence is the lowest statutory prison sentence, in months, for the penal-code article. Standard errors are in parentheses. For [De Chaisemartin and d'Haultfoeuille \(2020\)](#), they use a district bootstrap; for [Callaway and Sant'Anna \(2021\)](#), they use a district-clustered multiplier bootstrap, each with 1,000 replications. Offenses are ordered by minimum sentence. [†] Although drug offenses (Art. 376) carry a high statutory minimum sentence, arrests under this article are overwhelmingly for simple possession or small-quantity retail by low-level actors rather than large-scale trafficking or production.

E Arrests in *flagrancia* vs. arrests with warrants

Prior to the new code, arrests for more serious crimes were much more likely to be made with a warrant than arrests for low-level crimes: whereas just 25% of homicide arrests were made *en flagrancia* (no warrant), more than 98% of arrests for selling pirated CDs were made *en flagrancia* (Appendix Figure D.4). Overall, in the pre-period, 75% of arrests were made *en flagrancia* and 25% were made with warrants.

Figure E.1: Arrests Without Warrants Decline More than Arrests With Warrants



The upper panel plots horizon-specific ATTs for annualized rates of all arrests, arrests in *flagrancia*, and arrests with warrants. The lower panel divides each district's mode-specific arrest rate by its pre-treatment mean. Estimates use the method proposed by Callaway and Sant'Anna (2021). Confidence intervals use 1,000 bootstrap replications clustered by judicial district.

When the new code came into effect, arrests made with warrants declined by approximately 20% (Figure E.1), while arrests in *flagrancia* (i.e., in which a suspect is caught committing a crime) declined much more, by approximately 50%. As a result, the mean

Figure E.2: Warrant Share of Arrests Increases *Within* Crimes



Figure E.1 above shows that arrests with warrants declined much less than arrests without warrants. Here, we show that this pattern is driven not only by the changing composition of arrests across crimes, but also by changes *within* each crime category.

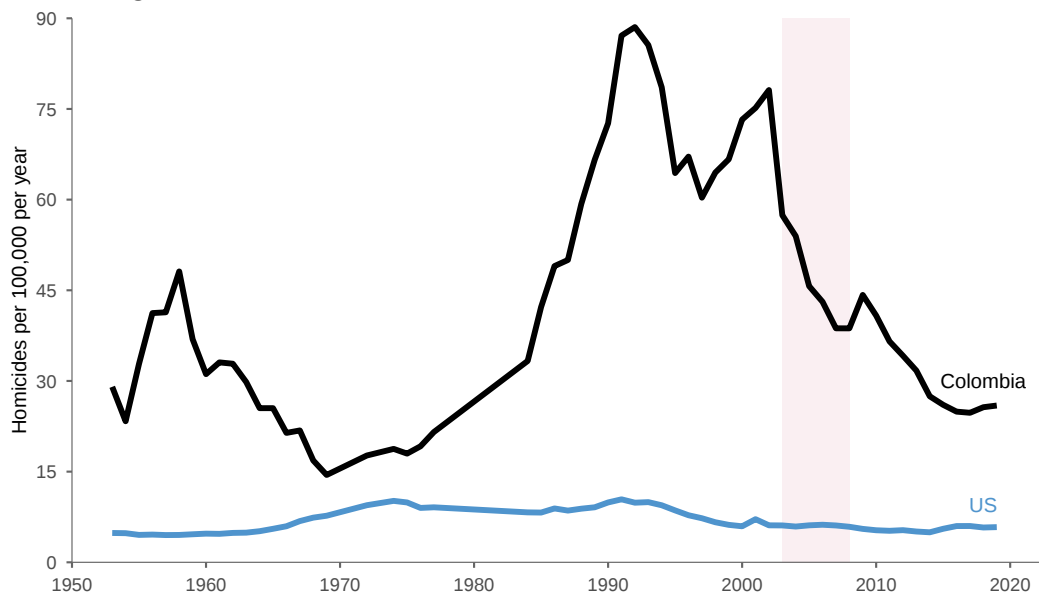
(across judicial districts) warrant *share* of arrests increased by 10 percentage points, from approximately 25 to 35 (Figure D.5b).

What's more, the shift toward warrant arrests is not driven entirely by the shift toward arrests for more serious offenses. The warrant share of arrests also increases sharply *within* crime (Figure E.2). The warrant share of arrests increased for 25 of the 30 crimes with the highest arrest rates prior to the new code (i.e., the most common charges).

Unlike flagrancia arrests, which are typically executed by patrol police, arrests with warrants are typically carried out by judicial police; judicial-police work changed directly as a result of the new code, not merely in response to downstream changes. For that reason, we can't interpret the decline in arrests with warrants as the consequence of additional effort and a stricter judicial screen alone; arrests with warrants could have declined because, for example, judicial police had to spend more time on other activities. But still, we cautiously interpret this result as a consequence of the fact that the reform had more bite for *flagrancia* arrests than for arrest with warrants: the change in police effort and downstream scrutiny was larger for the former than for the latter.

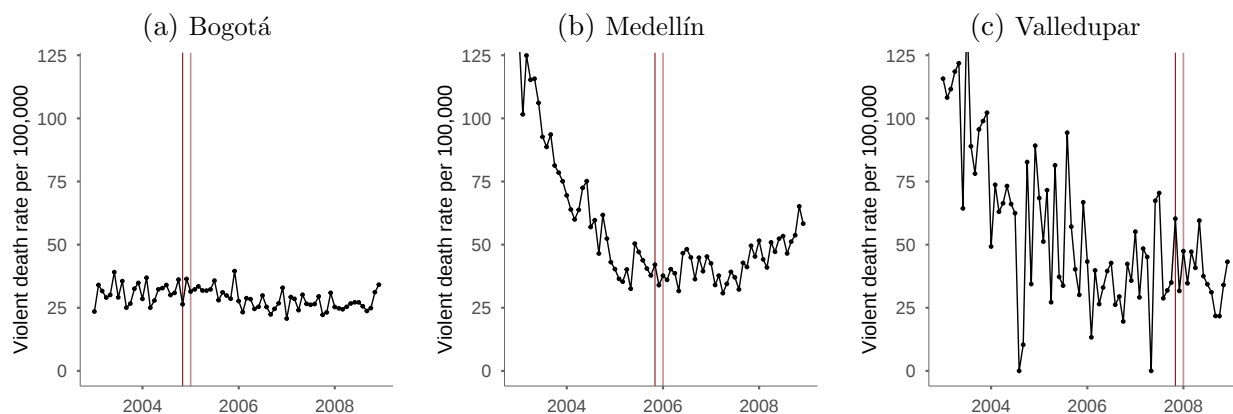
F Additional evidence of the effect of the new code on homicide rates

Figure F.1: Lethal violence in Colombia and the United States



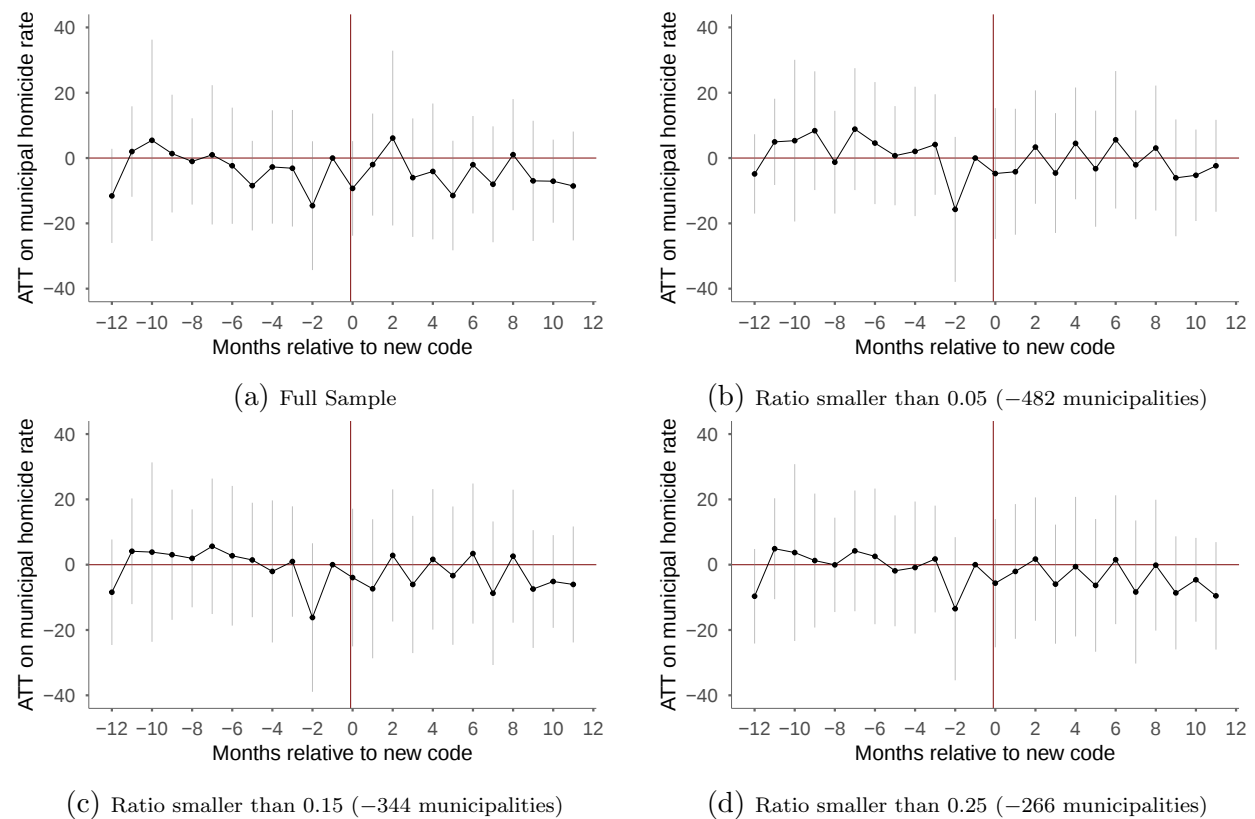
The red shaded region marks 2005–2008, the rollout period of Colombia's new code of criminal procedure.

Figure F.2: The New Code and Homicide Rates in Three Cities



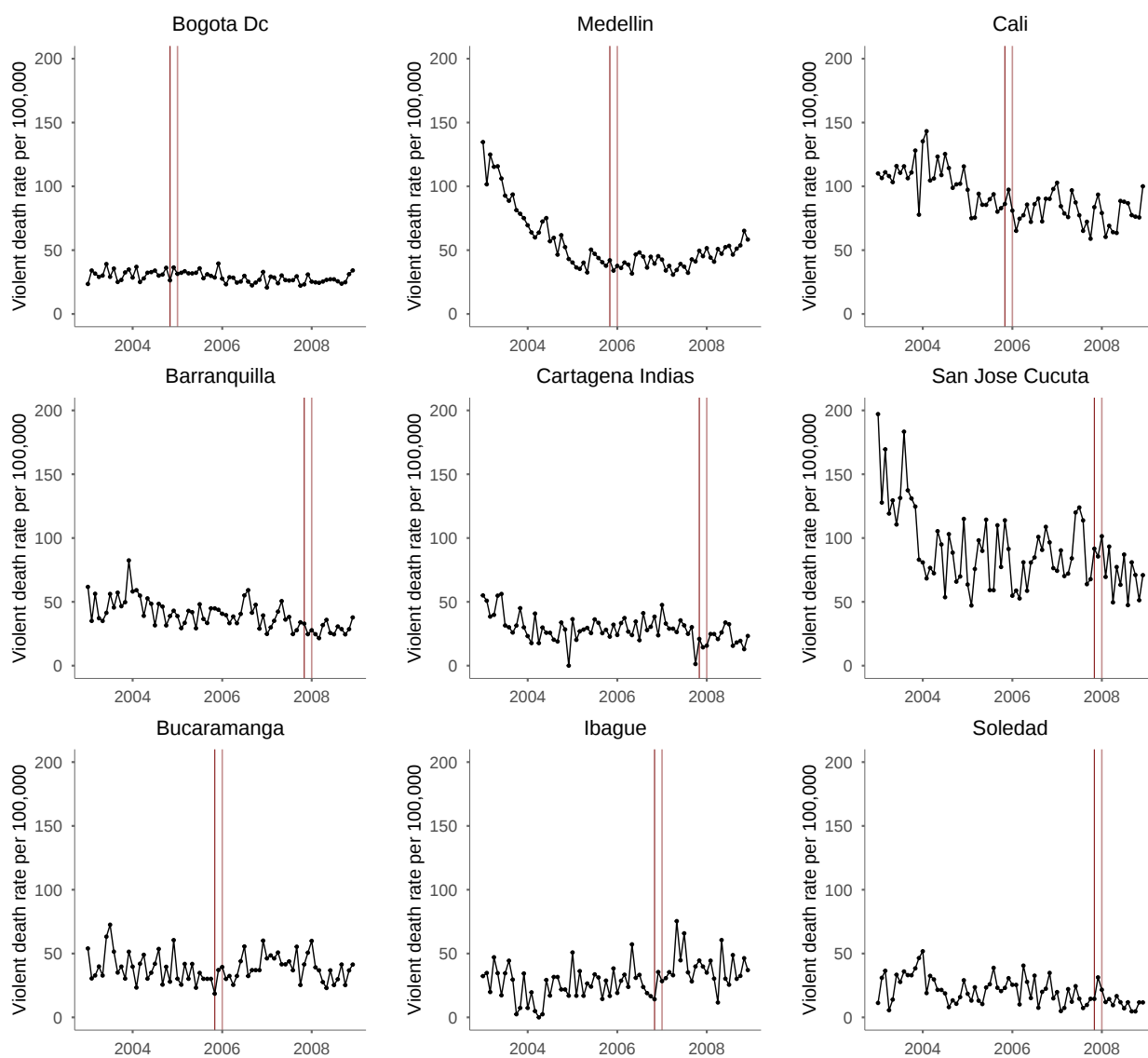
These figures plot the homicide rate in three cities. The pale red lines mark the Januaries when the new code came into effect in each city; the darker red lines mark the previous Novembers, in light of the fact that arrest rates began falling in anticipation.

Figure F.3: Effect of the new code on homicide rates, excluding municipalities with significant war violence



Each panel plots horizon-specific ATTs on municipal homicide rates using the method proposed by [De Chaisemartin and d'Haultfoeuille \(2020\)](#); Antioquia is excluded. Confidence intervals use 1,000 bootstrap replications clustered by judicial district. Panels (b)–(d) progressively exclude municipalities where war deaths account for more than the stated share of violent deaths.

Figure F.4: Homicide Rates in Colombia's Largest Cities



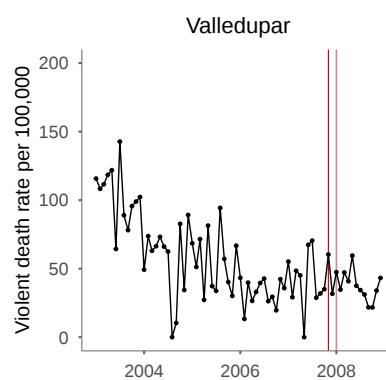
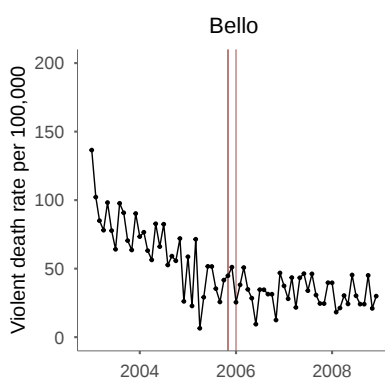
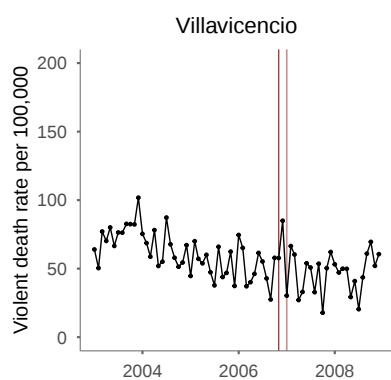
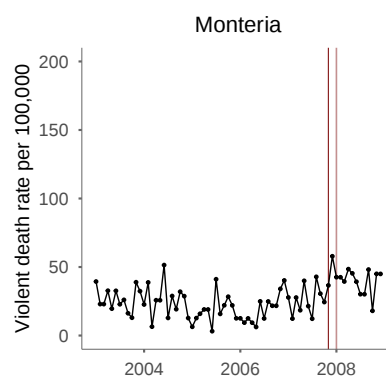
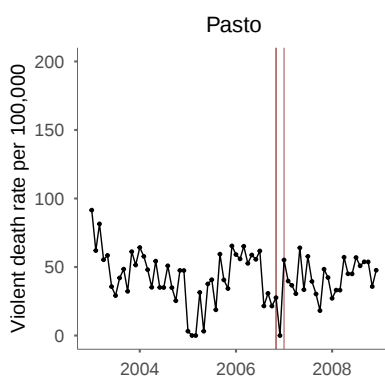
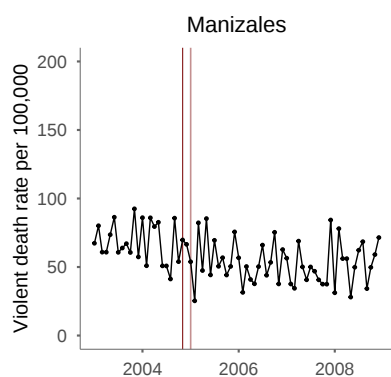
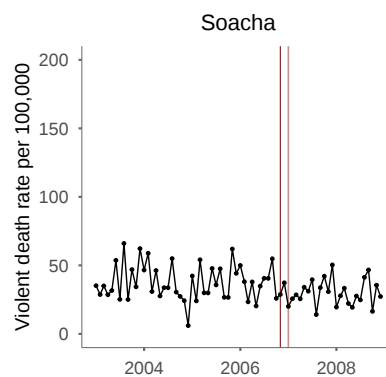
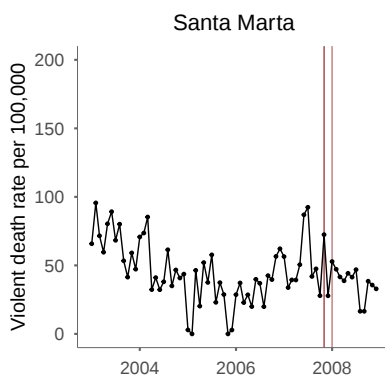
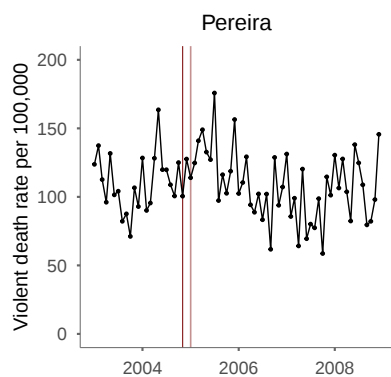
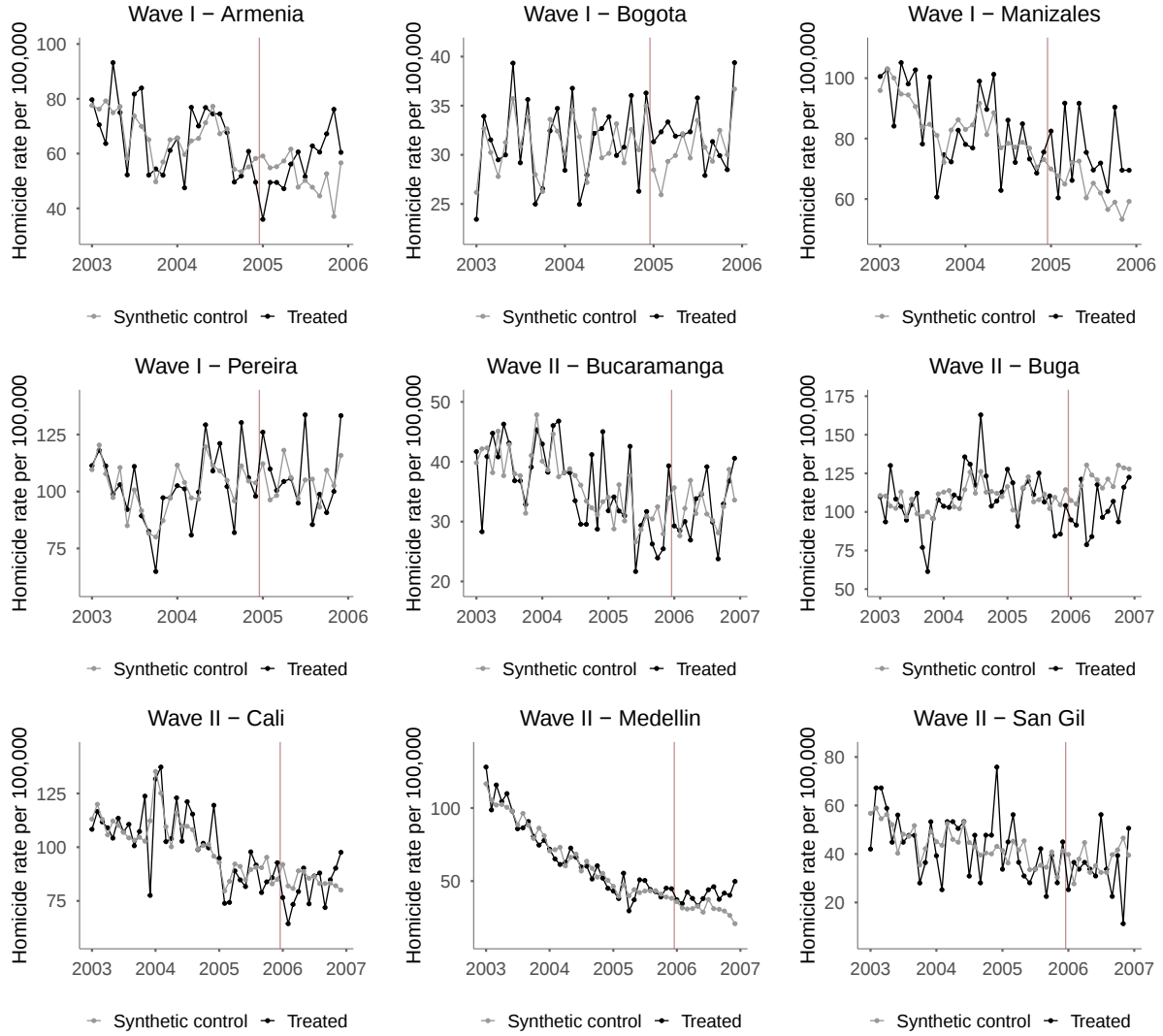
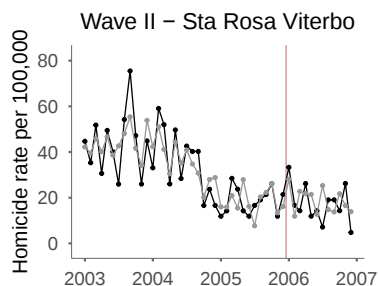
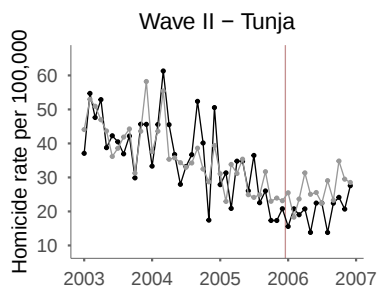


Figure F.5: Synthetic Control by Judicial District

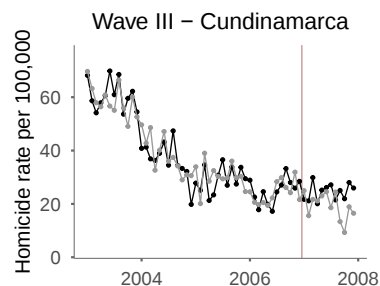




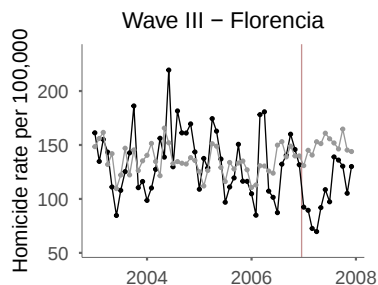
— Synthetic control — Treated



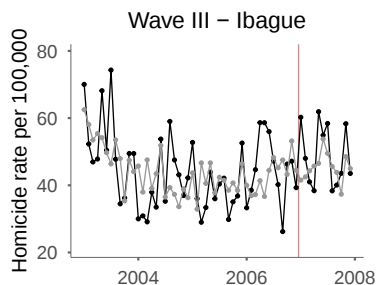
— Synthetic control — Treated



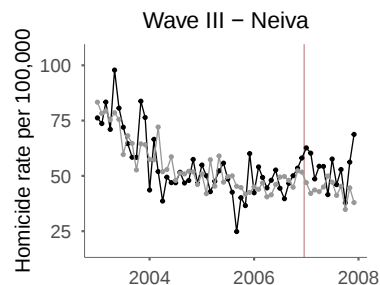
— Synthetic control — Treated



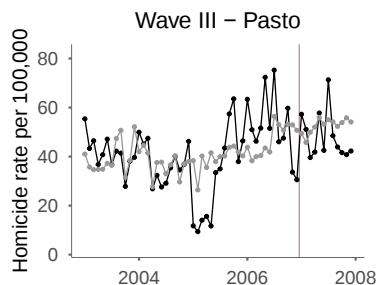
— Synthetic control — Treated



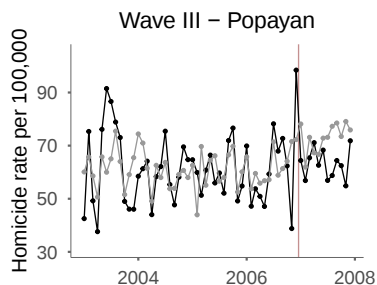
— Synthetic control — Treated



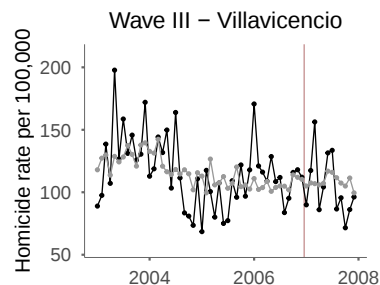
— Synthetic control — Treated



— Synthetic control — Treated



— Synthetic control — Treated



— Synthetic control — Treated

G LAPOP sample

The Latin American Public Opinion Project (LAPOP) survey data are representative of the *adult* population at the national level, rather than a representative sample of all municipalities. By 2005, Colombia was highly urban, with nearly 60% of the population living in the 5% of municipalities over 100,000 population. Thus, to achieve a representative survey of the population, the LAPOP team stratified on (1) region and (2) population in selecting municipalities (above and below a population threshold of 300,000), as we document in Table G.1, which covers our multi-year sample. By stratifying on region, LAPOP achieves substantial geographic coverage. However, to represent the urban population, LAPOP oversamples large municipalities. These municipalities are richer, more educated, and generally have higher state capacity than the many small municipalities in Colombia. Since we collapse data to the municipality-survey wave or judicial-district survey wave, we note that the sample sizes range from 6 to 231 respondents per municipality per wave.²³

Table G.1: LAPOP inclusion of municipalities by region and size. The regional designations are those used by LAPOP.

Strata	Category	2005 Census		LAPOP (pooled 2004-2009)	
		# Munis.	Pop. share	Respondent share	Pr(Muni. in LAPOP)
Region	Ant. Terr. Nacionales	117	0.05	0.07	0.04
	Atlántica	194	0.21	0.20	0.07
	Bogotá	1	0.16	0.16	1.00
	Central	262	0.24	0.25	0.06
	Oriental	366	0.16	0.14	0.02
	Pacífica	177	0.17	0.18	0.05
Population	≤10k	443	0.06	0.09	0.01
	10-100k	618	0.36	0.37	0.04
	100k-1m	52	0.29	0.25	0.40
	≥1m	4	0.29	0.29	1.00

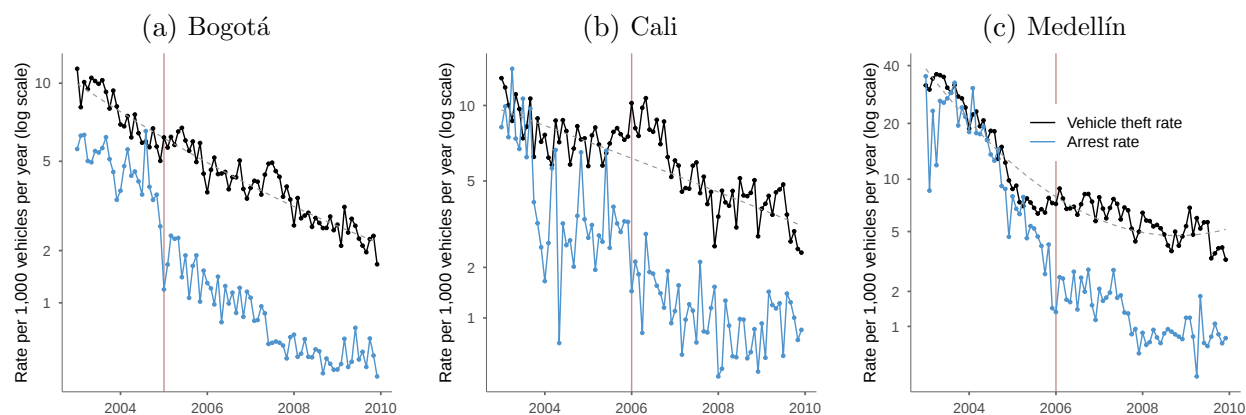
To provide a connection between our estimates with administrative data and those with LAPOP data (with a non-random sample of municipalities), we add a population-weighted version of the LAPOP analyses in Table 3 and Figure 9. These estimates correspond to our population-weighted administrative crime analyses in Table 1 and 2.

²³In the 2009 wave, only six respondents were surveyed in one municipality, but this is the only municipality-year observation with fewer than 12 respondents.

H Vehicle theft

The data on vehicle theft (from insurers) is so sparse that just three cities contain 78% of insured vehicles: Bogotá, Cali, and Medellín. The vehicle theft rates in those three cities (vehicle thefts per 1,000 insured vehicles per year), which we plot in Figure H.1, leave open the question of whether the new code caused an increase in vehicle theft. The panel is so sparse that difference-in-differences estimates are highly unstable across reasonable specifications.

Figure H.1: The Effect of the New Code on Vehicle Theft



The black lines mark the rate of vehicle thefts per 1,000 insured vehicles per year, according to data that we obtained from the Colombian association of insurers (Fasecolda). Gray dashed lines mark polynomial fits to the theft-rate trend in all years except the first year of the new code. Blue lines mark the rate of vehicle-theft arrests.