The Effect of Oversight on the Quantity and Quality of Policing

Nicolás Idrobo* Dorothy Kronick† Tara Slough‡

September 25, 2024§

Abstract

Governments often impose oversight of the police. Proponents argue that oversight curbs bad behavior, while critics counter that it sparks harmful backlash. We provide evidence from the staged rollout of a new code of criminal procedure in Colombia, which introduced judicial oversight of arrests. Judicial oversight caused a 40% drop in the number of arrests, we find, and a simultaneous improvement in arrest quality. Arrests for low-level crimes like vandalism plummeted, while arrests for serious crimes (like homicide) did not decline. Colombia thus reversed the hemisphere-wide run-up in policing of minor offenses, without police backlash and likely without causing a major crime wave.

*PhD Candidate, University of Pennsylvania

†Corresponding author (kronick@berkeley.edu). Assistant Professor, Goldman School of Public Policy, University of California, Berkeley

‡Assistant Professor, NYU Department of Politics

§Special thanks to Avi Feller and Gonzalo Vazquez-Bare for extensive guidance. For comments, we thank workshop participants at AL CAPONE, the Penn Conference on Policing in Comparative Perspective, the USC Conference on New Directions in the Political Economy of Development, and the U.C. Berkeley Methods Workshop.

Governments often strengthen oversight of the police. Judicial scrutiny of stops, arrests, and use of force, for example, has trained a spotlight on police departments. Proponents of oversight argue that it works as theory would predict, prompting police agencies to replace socially inefficient activities (such as certain pretextual stops) with activities that promote public safety. Critics argue instead that oversight sparks harmful forms of police backlash and protest, including "going fetal" (refraining from all policing) or even the use of violent tactics. Empirical evidence is mixed. This debate echoes longstanding discussions of oversight and performance in policing (e.g. Wilson, 1968; Becker and Stigler, 1974), and of the ability (or inability) of low-powered incentives to improve performance in public bureaucracies more generally (e.g. McCubbins and Schwartz, 1984; McCubbins et al., 1987; Dixit, 2002).

We evaluate the effect of oversight on police behavior in Colombia. A new code of criminal procedure—rolled out in four geographic stages between 2005 and 2008—introduced judicial oversight of arrests. Under the old code, public prosecutors wrote arrest warrants and signed off on warrantless arrests; the new code transferred these powers to judges. From the perspective of the police, this shift (from prosecutors to judges) implied heightened scrutiny and additional time.

We find that officers responded to judicial oversight by making fewer arrests, and especially by eschewing arrests for minor offenses. Using difference-in-differences, we find that the arrest rate dropped more than 45% within months of the new code coming into effect. This dramatic change was driven by even larger declines in arrest rates for minor offenses like vandalism (90%) or drugs (60%). Homicide arrests, in contrast, did not decline, nor did available measures of non-arrest activity (like gun seizures). Indeed, the size of the drop in the arrest rate is negatively correlated (across crimes) with minimum sentence length: arrests declined most for crimes with the shortest sentences. Moreover, the ratio of convictions to arrests appears to have increased with the implementation of the new code,[1] suggesting that officers de-prioritized those arrests that were unlikely to lead to conviction.

Overall, then, judicial oversight of arrests in Colombia reduced arrest quantity while increasing arrest quality. Colombia, like many countries across the hemi-

[1] We say *appears to have increased* because, as we discuss below, data from the court system is of much lower quality than data from the police.

sphere, had experienced a run-up in the policing of minor offenses in the early 2000s; judicial oversight appears to have partially reversed that trend, apparently without causing harmful police backlash.

A natural question is whether the changing quantity and quality of arrests affected crime. This question is difficult to answer. For one thing, Colombia's new code of criminal procedure was a bundled treatment: judicial oversight of arrests was one component of a comprehensive change that remade the entire criminal justice system. For another, changes in crime registration preclude using police data to measure crime incidence. Using vital statistics data on homicide, we find no evidence that the new code of criminal procedure affected homicide rates—but the null is imprecisely estimated because homicide is a rare and volatile outcome. Nor do victimization data indicate an increase in crime, though the surveys are small and therefore allow for only tentative conclusions. Using data that we obtained from vehicle insurers, we do observe an uptick in vehicle theft in the months following the implementation of the new code, but that estimate, too, is imprecise. In short, we find no evidence that Colombia's new code caused a crime wave—but nor do we find strong evidence to the contrary.

One unusual and useful feature of the Colombian case is that the government introduced oversight of the police in the absence of any of the conditions that so often accompany it: a misconduct scandal, a negative shift in public sentiment, negative press coverage, or deteriorating relations between politicians and the police (Mac Donald, 2017). Instead, Colombia's new code was inspired by an international movement (Langer, 2007). In fact, the same administration that implemented the new code—that of Álvaro Uribe—also more than doubled funding for the police and hired an additional 50,000 officers, expanding the national police force by 48% (Ministerio de Defensa Nacional, 2024a,b). We are therefore able to study the effect of oversight in the absence of coincident conditions that can themselves deflate morale and affect police behavior (Ba and Rivera, 2019). In the conclusion, we reflect on how this context affected the fate of Colombia's new code of criminal procedure.

In sum, we find that judicial oversight of arrests not only reduced the number of arrests but also increased the quality of arrests, shifting the composition away from arrests for minor offenses toward arrests for more serious crimes, and away from discretionary arrests toward mandatory arrests. One general takeaway, then, is that the welfare consequences of imposing oversight depend not on the elasticity of crime with respect to policing activity overall, but on the elasticity of crime with respect to the specific activities that police endogenously curtail under the new rules.

These findings contribute to three literatures. A vast body of work considers the effect of oversight on the behavior of bureaucrats in general and on the behavior of the police in particular (e.g. Prendergast, 2001; Devi and Fryer Jr, 2020; Prendergast, 2021; Hausman and Kronick, 2023). In the paper closest to ours, Ba and Rivera (2019) study the effect of oversight on police behavior and crime in Chicago, comparing incidents in which oversight followed public outrage to incidents in which the courts imposed oversight quietly. Their findings are strikingly similar to ours: oversight alone led police to curtail unproductive activity and did not increase crime. But while their work considers oversight of misconduct specifically, we consider oversight of a staple of police activity: arrests. Empirically, our setting also allows us to exploit the staged rollout of oversight across four groups of judicial districts covering all of Colombia.

We also contribute to a large literature on the elasticity of crime rates with respect to policing (see Chalfin and McCrary, 2018), including a new crop of studies that focus on depolicing (Rushin and Edwards, 2016; Pyrooz et al., 2016; Cassell and Fowles, 2018; Rosenfeld and Wallman, 2019; MacDonald, 2019). Many scholars consider the elasticity of crime with respect to police hiring, police presence, arrest rates, and/or incarceration overall (e.g. McCrary, 2002; Di Tella and Schargrodsky, 2004; Buonanno and Raphael, 2013; Barbarino and Mastrobuoni, 2014). Other work focuses on specific types of arrests (such as misdemeanor arrests); these studies generally draw inferences from correlations or study policy interventions in a single jurisdiction (e.g. Dominguez-Rivera et al., 2019). Two exceptions are Cho et al. (2023), who use police-officer deaths (and the consequent drop in police activity) to estimate the arrest elasticity of crime, and Rivera (2024), who uses the end of a pay-per-arrest scheme in early twentieth-century Chicago to isolate the effect of arrests on crime. Our setting, in which the government introduced oversight in stages, provides additional empirical leverage.

Our work is also closely related to studies of specific police tactics, like hotspots policing (Braga et al., 2019; Collazos et al., 2021) or police militarization (Mummolo, 2018; Blair and Weintraub, 2020). But while these articles estimate the consequences of various forms of intensifying enforcement, we consider an intervention that sharply restricts a specific police activity.

Finally, the reform that we study was one of a wave of similar efforts across Latin America: a "revolution in Latin American 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 of the region's shift from inquisitorial to accusatorial criminal procedure. Ours is one of few empirical papers assessing the consequences. Magaloni and Rodriguez (2020) find that the new code of criminal procedure reduced the use of torture in Mexico, and Tiede (2012) documents a reduction in the use of pre-trial detention in Chile. Two working papers, Cattaneo et al. (2022a) and Hanson and Kronick (2024), consider Uruguay and Venezuela, respectively; Cattaneo et al. (2022a) find that the protections in the new code caused an increase in property crime, while Hanson and Kronick (2024) focus on a harmful police backlash. Additional investigation of the reform experience in other countries—such as Argentina, Peru, Bolivia, and Panama, among others—as well as harmonization of results across case studies, constitutes a promising avenue for future work.

One point of contrast between our study and most prior work on Latin America's "revolution in criminal procedure" is that we focus on the police, rather than the courts. Prior work focuses on the courts because judges and prosecutors were the primary targets of the new codes of criminal procedure; indeed, in the case that we study, the new code affected the police via empowering judges. But the courts are small relative to policing in Colombia. In 2004, for example, the year prior to the first wave of implementation of Colombia's new code of criminal procedure, Colombian police made 17,000 arrests for vandalism—but only seven people were convicted. Whole categories of low-level offenses that are judicialized in the United States—like disorderly conduct or public urination—are classified as infractions in Colombia, meaning (effectively) that they are the sole province of the police, and not subject to court proceedings. Naturally, we recognize that Colombia's new code of criminal procedure was a bundled treatment, and we discuss implications of this fact for our results. But we do view oversight and the consequent sea change in policing as the most far-reaching, visible, and immediate consequence of the new code of criminal procedure. Recognizing this fact, and considering its consequences, is one of our contributions.

1 Oversight, Depolicing, and Crime

We follow Ba and Rivera (2019) in defining an increase in oversight as any event or policy that raises the effective cost (to police) of the newly monitored activity. In the case that we study, a new code of criminal procedure raised the effective cost (to police) of making arrests; we describe the policy in more detail in the following section. Here, we outline expectations about how an increase in oversight might change the quantity and the quality of the newly monitored activity. We discuss our expectations in reference to arrests, but the logic is more general: it applies to oversight of stops or use of force, for example.

Of course, we expect police to make fewer arrests in response to new oversight. Less obvious is *which* arrests they will continue to make, and which arrests they will cut. We define arrest *quality* as the extent to which an arrest is socially productive in the sense of Owens (2019); Owens and Ba (2021): providing "more benefit to society in terms of crime reduction than they cost in terms of police legitimacy" and in terms of direct costs like police salaries.

Consider several examples. Some arrests have large deterrent and/or incapacitating effects on the most harmful crimes while simultaneously strengthening police legitimacy (to take a rare but extreme example: the arrest of a serial killer); we might think of these as the most socially productive arrests, and thus, in our terms, the highest quality. Somewhat less socially productive are arrests that have smaller deterrent and/or incapacitating effects but also do not damage a police agency's reputation; still less socially productive are arrests that do little to improve public safety but do damage police legitimacy. Still other arrests may even be criminogenic—because they incapacitate an *officer* without deterring or incapacitating would-be offenders (Rivera, 2024)—and might simultaneously incur large costs in terms of police legitimacy. We might think of this third category as the least socially productive and the lowest quality. In Owens's framework, arrests (and police activities more generally) can damage police legitimacy not only when they involve clear civil liberties violations (like arrests of non-offenders or excessive force) but also when they are merely perceived as unnecessary.

This *social productivity* characteristic of arrests is unobserved. But Owens and Ba (2021, 11) suggest that it is correlated with the severity of the offense, writing that "low-level offenses [are] the kinds of incidents for which the benefit of crime reduction is likely lower relative to the direct legitimacy cost of arrests." This assumption strikes us as especially likely to hold in the wake of a substantial increase in arrests for low-level offenses, such as those of the United States in the 1990s or Colombia just prior to the reform that we study. Just as arrests for low-level offenses are likely to be less socially productive than arrests for serious crimes, so too are discretionary arrests likely to be less socially productive than mandatory arrests. (Whether an arrest is discretionary or mandatory is related to but not entirely determined by the severity of the charge.)

We therefore consider two observable characteristics of arrests as measures of (aspects of) arrest quality. The first is the severity of the charge; we view arrests for CD piracy as less socially productive than arrests for homicide. The second is the mode of arrest: we view arrests made *in flagrancia*, i.e., when the suspect is caught in the act of committing a crime (or shortly thereafter), as likely less socially productive than arrests made with a warrant. We discuss the empirical measures of these characteristics in the following section; what is important here is to provide a sense of how we conceptualize arrest quality.

The effect of oversight on arrest quality is ambiguous ex-ante. For one thing, the *de facto* intensity of oversight itself could fall differentially on different types of arrests, even if the *de jure* policy were to apply equally to all arrests. If the cost of oversight were higher for arrests of people for minor offenses, for example—as might happen if these arrests had a weaker evidence base, and if judges sought to strengthen due process—then we might expect police to shift toward arrests for serious crime. If, on the other hand, the cost of oversight were higher for arrests of people suspected of serious crimes—as might happen if judges were to apply additional scrutiny to arrests of murder suspects, for example—then we might expect police to shift away from arrests for serious crimes. Similarly, the cost of new oversight might fall differentially on arrests with and without warrants, in either direction.

Moreover, even in the absence of differential intensity of oversight for different types of arrests, another mechanism could produce a reduction in arrest quality: protest. Critics of oversight argue that police officers' experience of the cost of oversight may differ from the actual time or disciplinary costs that reformers impose; when police officers experience oversight as an affront, they may respond by "going fetal" (e.g. Cass, 2016), i.e., halting enforcement activity almost entirely.

Police protest is therefore another mechanism by which oversight might generate a suboptimally low arrest rate, and perhaps also lower-quality arrests.

The consequences of oversight for crime are similarly ambiguous in theory, and not only because of the ambiguity in the effect of oversight on police behavior. Consider a case in which oversight leads police to cut arrests for low-level offenses but continue to prioritize felony arrests. Some scholars would argue that this depolicing of minor offenses would lead to an increase in serious crime. "Not every petty criminal is a serious criminal," Kelling and Coles (1997) 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." In this (implicit) model, stopping and/or arresting people for (say) disorderly conduct or petty theft deters or incapacitates would-be felony offenders. Equally plausible is a model in which policing minor offenses *does not* deter or incapacity felony offenders. Similarly, some scholars argue that arrests for minor offenses reduce *fear* of crime, which in turn hampers community production of crime control (see Weisburd et al., 2015, for a summary of this view); equally plausible is the idea that these connections do not exist. Finally, if low-quality arrests entail reallocating resources *away from* the apprehension and investigation of people suspected of serious crimes, then we might expect that policing minor offenses would *raise* homicide rates—and that depolicing minor offenses could actually quell violence.

This framework clarifies that, while we expect an increase in oversight to reduce the number of arrests, the consequences for the quality of arrests—and then for crime—are ambiguous. In the remainder of the paper, we investigate these consequences empirically.

2 Police and criminal procedure in Colombia

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, brutal, militarized, under-educated and under-trained police force embroiled in an extraordinarily violent conflict 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—was "a professionalized and well-regarded police force" (Blattman et al. 2021b; see also Blattman et al. 2021a on the same point). Moncada (2009) calls this "an unexpected shift toward 'democratic policing."' Beckett and Godoy (2010), comparing Bogotá to New York City, conclude that police in Bogotá successfully reduced incivility, curbed disorder, and curtailed lethal violence without veering into Brattonesque zero tolerance.

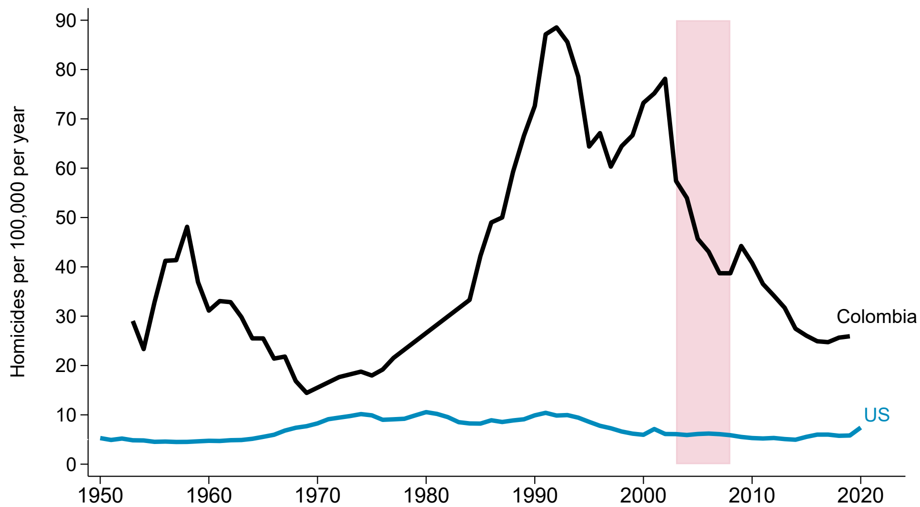
In 2002, Álvaro Uribe won Colombia's presidential election on a platform of defeating the guerilla—especially the FARC, one of whom killed Uribe's father. The flagship policy of the Uribe administration, called "Democratic Security," entailed extending police and military presence into rural areas and, in urban areas, allying with certain non-state armed groups in order to reduce violence (Acemoglu et al., 2013). Uribe's "Democratic Security" initiative dovetailed with the U.S.-funded Plan Colombia, in which Washington provided billions of dollars of aid to the Colombian military (and, by extension, the paramilitary, Dube and Naidu, 2015) to fight the guerilla and to eradicate coca cultivation (Gerez, 2023). Uribe's policies produced both a spectacular decline in lethal violence (Figure 1) and also grave human rights violations. Most notoriously, a written policy of bonuses for military officers who killed members of the guerrilla resulted in thousands of "false positives:" innocent civilians killed and dressed up as FARC (Acemoglu et al., 2020). While he was in office, Uribe's approval rating rarely dipped below 70%.

Between 2002 and 2012, the national police reoriented activities around certain principles of community policing: in particular, beat-based patrol and training in interpersonal skills, both of which had salutary effects (García et al., 2013). It is revealing that domestic criticism of this program has focused on excessive centralization (Vásquez, 2012) and ineffectiveness, rather than the persistence of widespread violence or abuse. Indeed, the laudatory tone of scholarship on the Colombian police during this period (e.g. Moncada, 2009; Beckett and Godoy, 2010; González, 2020) stands in stark contrast to work that documents Uribe's aggressive attacks on Colombian democracy (Gamboa, 2022), encouragement of grave human rights abuses by the military (Acemoglu et al., 2020), deliberate continuation of a deadly war for electoral gain (Fergusson et al., 2016), wasteful and noxious aerial fumigation of coca crops (Mejía, 2016; Mejía et al., 2017), and preferential treatment of the right-wing paramilitary (Sontag, 2016; Gerez, 2023).

At the time of the reform that we study, then, the Colombian national police was

Figure 1: Lethal violence in Colombia and the United States

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



a force of approximately 130,000 officials (309 per 100,000 people in Colombia), housed in the Ministry of Defense, with primary responsibility for patrol and arrest nationwide. Hiring, promotion, and many policy decisions were under control of the national police leadership, but city governments did share jurisdiction over the municipal divisions of the national police. Specialized units focused on coca eradication and the pursuit of drug traffickers.

A new code of criminal procedure. In 2004, the second year of Uribe's first term as president, the Colombian Congress passed legislation (Ley 906) mandating the adoption of a new code of criminal procedure. Whereas the penal code defines crimes and penalties, criminal procedure defines the rules according to which criminal cases proceed through the justice system. An international network of legal activists, with support from international financial institutions, promoted this change throughout Latin America (Langer, 2007); Colombia's new code of criminal procedure was a product of this process of international diffusion, not a response to a domestic scandal about police misconduct or court malfunction.

From the perspective of the police, the most salient change was the introduction of judicial oversight of arrests. Under the previous code, public prosecutors (*fiscales*) could write arrest warrants and sign off on warrantless arrests. This prosecutor power meant that police officers could obtain warrants quickly and easily, and that warrantless arrests were almost always 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* wrote arrest warrants, and *judges* had to sign off on warrantless arrests. Beyond annoying the prosecutors, this major transfer of responsibilities (from prosecutors to judges) imposed substantial new time costs on the police. Under the old code, a police officer making a warrantless arrest had to bring the suspect to a prosecutor for a cursory sign-off process that could conclude in fifteen minutes. Under the new code, in contrast, an officer making a warrantless arrest had to bring the suspect first to a prosecutor and then to a hearing before a judge. These hearings often lasted 45 minutes or more. Moreover, many officers perceived the judges as more *garantista* than the prosecutors, 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).[2] Under the new code, then, warrantless arrests required substantially more officer time, with a lower probability of sign-off.

The introduction of judicial oversight of arrests was far from the only change in Colombia's new code of criminal procedure. Indeed, the new code replaced what was primarily an inquisitorial system (typical of civil law) with an accusatorial system (typical of common law). Under the inquisitorial system, proceedings were written rather than oral, and judges—the eponymous inquisitors—participated in the process of investigation and also rendered decisions. The accusatorial system replaced written with oral proceedings, removed the judge from the process of gathering evidence, and introduced juries. Scholars have studied the effects of the new code on the functioning of the courts. Hartmann Arboleda (2016), for example, documents a decline in the use of pre-trial detention; Mejía (2016) show that the new code significantly reduced the time to trial and increased the use of plea bargaining.

In that sense, Colombia's new code of criminal procedure was a bundled treatment: it changed the rules governing indictment, pre-trial detention, investigation, plea bargaining, trials, the process of sentencing (though not sentences themselves), and punishment. Our focus on judicial oversight of arrests is motivated by sub-

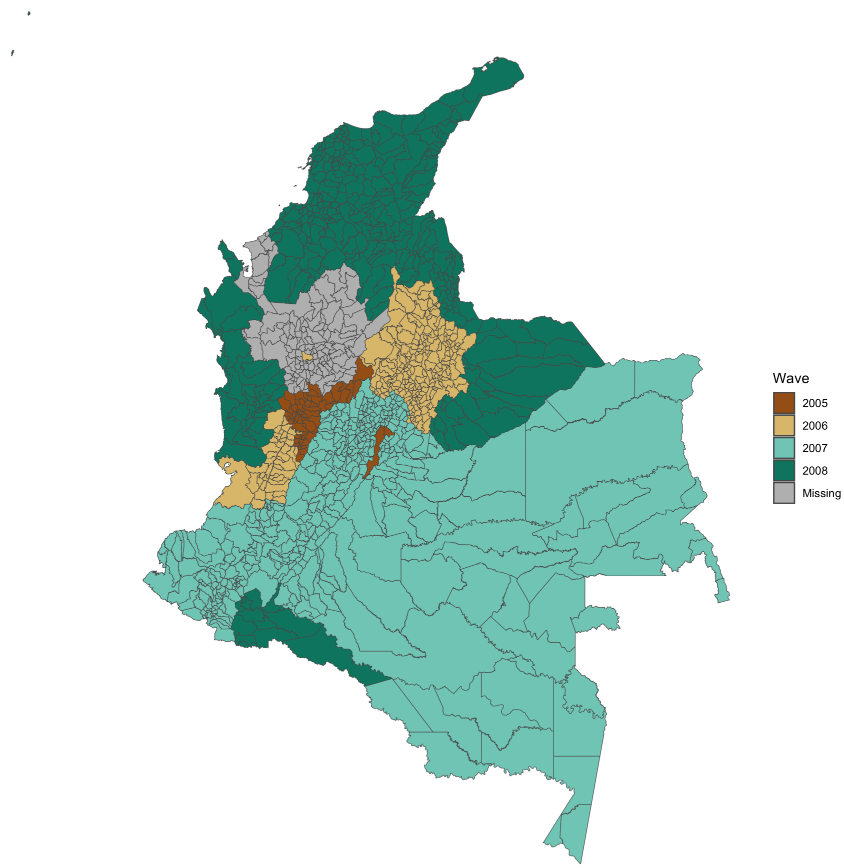
[2]In response to requests for data on the number and outcome of warrantless-arrest proceedings, both the public prosecutor's office and the judiciary provided numbers that are so incomplete as to be uninformative.

stantive and practical considerations. Substantively, contact with the criminal justice system begins and ends with arrests for the majority of suspects. In 2004, as noted above, there were only seven convictions for vandalism—even though the police arrested more than 17,000 people for this crime. Many offenses that are classified as misdemeanors in the United States are mere infractions in Colombia, meaning that they are not judicialized. Practically, the information systems collating data in the public prosecutor's office (Fiscalía) and the court system—SIAN, SIJUF, and SIERJU—are far more problematic and less complete than their counterpart in the police (SIEDCO, which we discuss below). For example, SIERJU (the court-system information system) did not begin recording outcomes of cases under the new code of criminal procedure until one year after it first came into effect. We discuss below the partial data that do exist from these systems, along with implications for our interpretation of our primary findings.

Colombia's new code of criminal procedure was rolled out in four stages. The first stage, shown in Figure 2 in brown, included Bogotá and the coffee belt; these were the judicial districts deemed most prepared for implementing the new code. In these places, the new code came into effect on January 1, 2005, just four months after the bill passed the legislature. One year later, on January 1, 2006, the new code came into effect in the city of Medellín and in the Andean judicial districts adjacent to the coffee belt, including the city of Cali; like the judicial districts of the first wave, these districts had relatively high implementation capacity. On January 1, 2007, the new code reached the more rural judicial districts of the eastern plains and the Amazon. *De jure*, the law assigned Antioquia (the district surrounding Medellín, north of the coffee belt) 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). Another difference between the text of the code and implementation in practice occurred in the judicial district of Yopal, which was originally assigned to the 2006 (second) wave but ended up in the 2008 wave (Piraquive Sierra, 2007). Finally, on January 1, 2008, the new code reached the Caribbean lowlands—where the biggest city is Barranquilla, and which includes tourist regions like Cartagena—part of the Pacific coast, and much of the region that borders Venezuela. The four stages divide the Colombian population in quarters, with approximately 10 million people in each wave.

Figure 2: Rollout Map

The new code of criminal procedure came into effect on January 1, 2005 in the judicial districts of the first wave, and on January 1 of 2006, 2007, and 2008, respectively, in the second, third, and fourth waves.



3 Data

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

Police data. The Colombian national police began collecting systematic data on operations and on reported crimes in 1958, when a team of criminologists and sociologists developed a system of paper forms that regional offices filled out and remitted to Bogotá for tabulation. This system eventually migrated to Excel, where it remained until 2003. As late as 2002, then, regional offices would email

Excel sheets to the (then newly established) Unit of Criminal Analysis, where they were eventually consolidated into national tallies. One officer who worked in the unit at that time described the process as slow and error-ridden. These problems inspired analysts within the police to request financing and support from the Inter-American Development Bank for the development of dedicated software for counting crimes and documenting police activity. The result was the System for Statistical Information on Crimes, Infractions, and Operations, or SIEDCO. By the end of 2003, SIEDCO was collecting information from all regional police offices in the country.

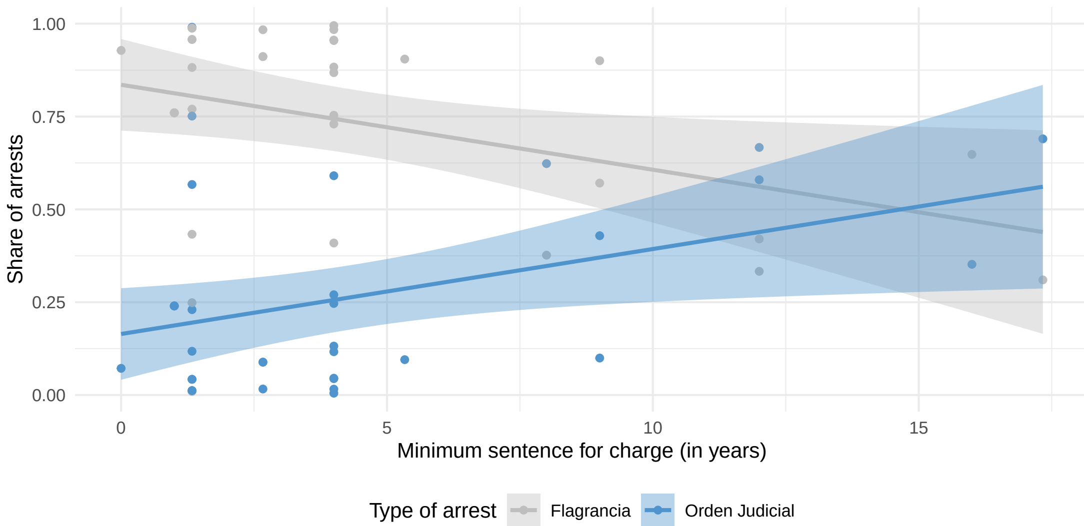
Via a freedom-of-information request, we obtained SIEDCO's count of the number of arrests by suspected crime (such as theft or assault), by type of arrest (with a warrant or in *flagrancia*), in each municipality ($N = 1,102$) in each month between 2003 and 2011. This count includes all arrests, regardless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

As noted above, we are interested in arrest *quality*, which we define as the extent to which an arrest is socially productive in the sense of Owens and Ba (2021): the benefits in terms of public safety exceed costs, including in terms of police legitimacy. Arrest quality is unobserved, but we assume (again for reasons stated above) that it is correlated both with the severity of the charge and with the mode of arrest: we assume that arrests for more serious crimes (like homicide) are typically higher quality than arrests for minor crimes (like CD piracy); similarly, we assume that arrests made with a warrant are typically higher quality than arrests made *in flagrancia*. To measure crime severity, we use the minimum sentence in Colombia's penal code.

Our two measures of quality—crime severity (as proxied by minimum sentence length) and mode of arrest—are correlated but not perfectly so. Using data from 2003 and 2004, the two years prior to the first wave of the rollout of Colombia's new code of criminal procedure, we estimate the proportion of arrests made with warrants (*orden judicial*) and without warrants (*in flagrancia*) for each crime for which there were at least 1,000 arrests in at least one of these two years. Figure 3 shows that arrests for crimes with shorter penalties were typically made without warrants (*in flagrancia*), while arrests for crimes with longer penalties were more likely to have warrants. Figure B.17 plots the mix of modes for each

Figure 3: Two Measures of Arrest Quality

This figure plots the relationship between mode of arrest (with or without a warrant) and crime severity in the *pre* period (specifically, in the two years prior to the introduction of the new code of criminal procedure).



crime separately. These figures reveal that, while mode-of-arrest is correlated with crime severity, these two measures also contain separate information.

We also obtained SIEDCO's count of crimes in each municipality in each month (also via freedom-of-information request). But changes in the process of crime registration make it difficult to use these data to study trends in property crime or in violent crime (other than homicide).

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 death certificates (i.e., vital statistics).

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 (Cendales and Pardo, 2018). Following González Mejías and Kronick (2023), we use the mortality microdata to construct two measures of homicides: (1) the number of deaths actually classified as homicides by the preparers of death certificates (i.e. those with ICD-10 codes X85–Y09) and (2) the deaths classified as homicides *plus* gun and sharp-object deaths "of unknown intent," i.e., gun and sharp-object deaths that are not known to be suicides or accidents. Like the police count of homicides, the vital statistics count is incomplete; future work could supplement our analysis by conducting case studies that use multiple sources—not only police and vital statistics data but also press reports and NGO counts—to construct more complete homicide series (as in Guberek et al., 2010, for Casanare).

4 The effect of oversight on arrest rates and arrest quality

In this section we establish that the new code of criminal procedure caused a large and immediate drop in arrest rates, and that arrest *quality* simultaneously improved. Officers made fewer discretionary arrests without reducing the number of arrests for serious crimes like homicide.

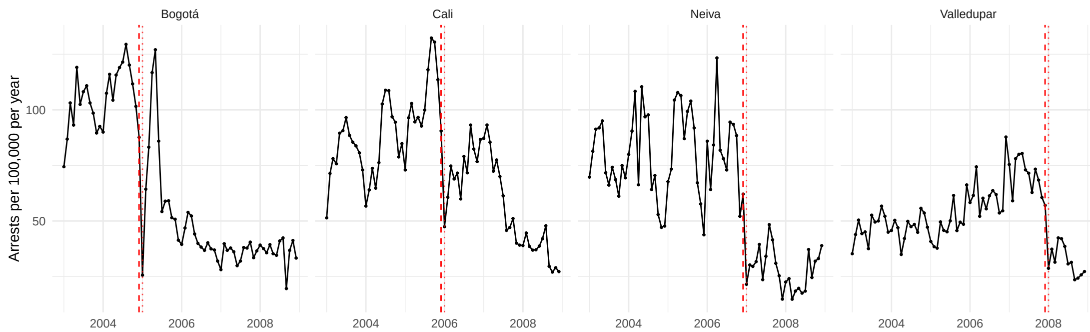
Arrest quantity. That the new code caused a decline in arrest rates is not entirely surprising, but nor is it widely known. It is unsurprising because, as we explain above, the new code made arrests more costly for the police: judges, rather than prosecutors, issued arrest warrants and signed off on warrantless arrests. But despite the straightforward nature of the connection between the new code and a drop in arrest rates, we are the first to document it. Previous literature on Colombia's new code of criminal procedure focuses on how it affected pre-trial detention (Hartmann Arboleda, 2016) and the duration and outcome of court cases (Mejía et al., 2016).[3] The new code of criminal procedure is not seen or studied as a police reform. But it did reshape policing.

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 half 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

[3]A subsequent draft of Mejía et al. (2016), Zorro Medina et al. (2020), replicates our arrest result, citing a previous version of this working paper.

Figure 4: 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).



months, from approximately 1,300 per 100,000 per year in 2003–2004 to 600 per 100,000 per year in 2005 (see Figure 4). In Cali, part of the second wave; Neiva, a mid-size city (population 400,000) in the third wave; and Valledupar, a city of half a million people in the fourth wave; arrest rates similarly collapsed.

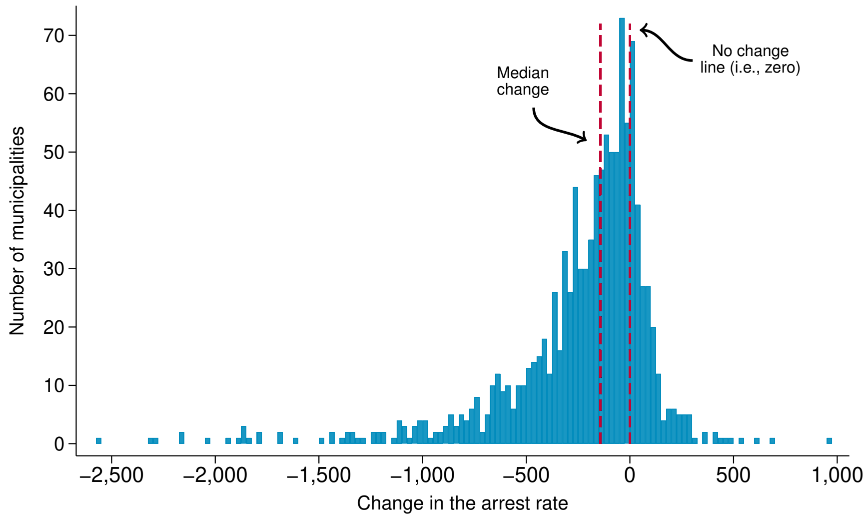
These examples are not atypical. Indeed, arrests rates fell in 78% of Colombia's 1,101 municipalities in the twelve months following the implementation of the new code, compared to the twelve months prior (Figure 5). The municipalities in which the arrest rate did *not* decline generally started out with much lower and more variable arrest rates, meaning that the pre–post estimates are noisier.

The examples in Figure 4 suggest 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 makes sense given that the implementation committee held trainings, including mock hearings, for police and prosecutors in the months leading up to the rollout (Contraloría Generalde la Nación, 2010). We accommodate this anticipation by coding November (rather than January) as the treatment date for our analysis of arrest rates.

To estimate the effect of the new code on arrest rates, we use a difference-in-differences approach. The left panel of Figure 6 plots 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,

Figure 5: Arrest Rates Decline in 78% of Municipalities

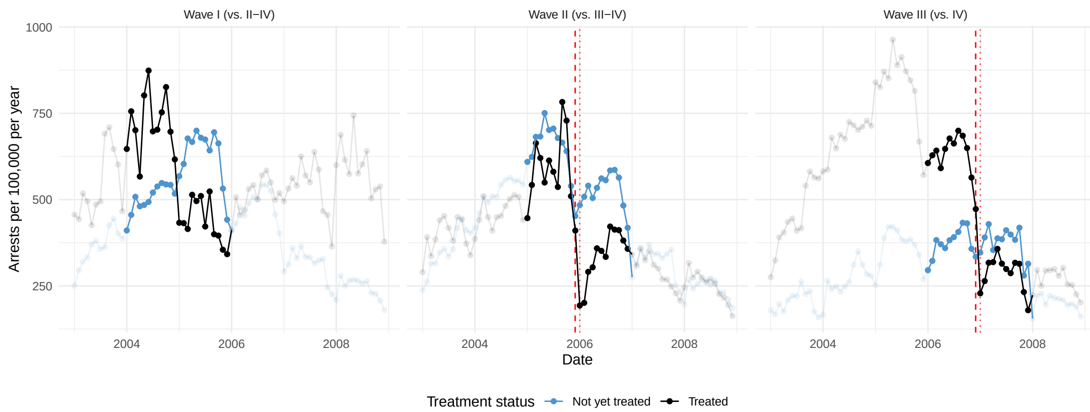
This figure plots the distribution of pre–post changes in the arrest rate across Colombian municipalities, comparing the twelve months before the reform comes into effect to the twelve months after.



and IV, where the new code had not yet come into effect. 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 in the country. This difference-in-differences suggests that the new code caused a drop in arrest rates in Bogotá and the coffee belt. A similar pattern emerges in Waves II and III. In 2005, arrest-rate trends in the Andean municipalities that constitute Wave II were parallel to arrest-rate trends in the more outlying areas of Waves III and IV; then, in when the new code came into effect in the Wave-II municipalities, arrest rates fell sharply, while they held steady in the places where the old code was still in effect (Waves III and IV; see the second panel of Figure 6). In 2006, similarly, arrest-rate trends in the municipalities of Wave III moved together with arrest-rate trends in the municipalities of Wave IV; as the new code came into effect in Wave III, arrest rates there fell more than 50% in a span of months, while arrest rates in the as-yet-untreated Wave IV remained unchanged (Figure 6, third panel). When the new code came into effect in Wave IV, at the beginning of 2008, arrest rates dropped sharply there, too (though there is no valid control group for Wave IV because the rest of the country had already implemented the new code).

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

The black lines plot arrest rates in the judicial districts of Waves I, II, and III, respectively. The blue lines plot arrests rates in not-yet-treated judicial districts.

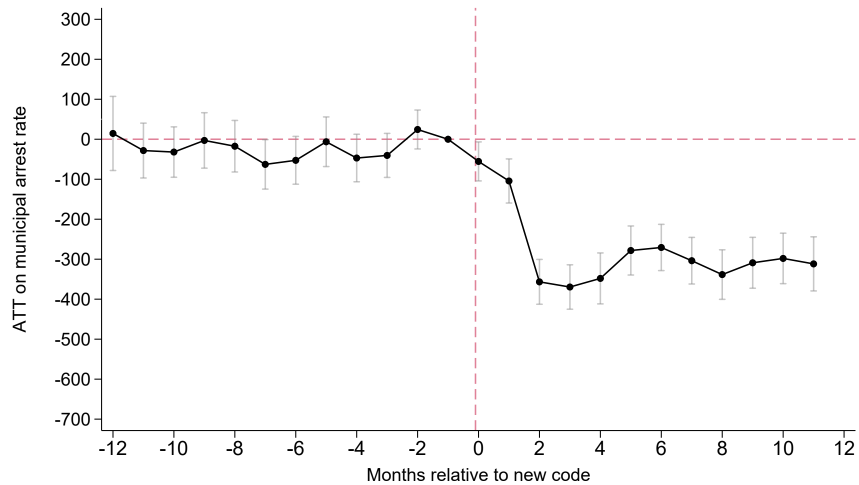


To estimate the average treatment effect on the treated (ATT) at each horizon (meaning each time period *relative* to the month in which the new code came into effect), we use the estimator proposed by De Chaisemartin and d'Haultfoeuille (2020). Letting i index judicial districts, G index the wave of the reform ($G \in \{1, 2, 3, 4\}$), F_G denote the month in which the new code comes into effect in wave G (such that F_1 = January, 2005, for example), S_G denote the set of districts that belong to wave G, $Y_{i,t}$ denote outcome Y in district i at time t, and $h \in \{-12, -11 \ldots, 12\}$ index months relative to the implementation date of the reform, we calculate:

$$\widehat{\text{ATT}}_h = \sum_{G=1}^{3} w_G \left(\sum_{i \in S_G} \frac{Y_{i,F_G+h} - Y_{i,F_G-1}}{|S_G|} - \sum_{\substack{i \in S_R \\ R>G}} \frac{Y_{i,F_G+h} - Y_{i,F_G-1}}{|S_R|} \right) \tag{1}$$

where $w_G = |S_G| / \sum_g |S_g|$ are the weights given to each wave, based on the number of treated units relative to the total. The first term in the difference captures how the outcome (arrest rates) changes between the month before the new code comes into effect $(F_G - 1)$ and h months before or after the new code comes into effect, *in newly treated judicial districts*; the second term compares how that same outcome changes over the same time period in not-yet-treated districts. This approach compares changes in arrest rates in treated units to changes in arrest rates in not-yet-treated units, avoiding the "forbidden comparison" of newly

Figure 7: The Effect of the New Code on Arrest Rates: Event Study



Using the estimator proposed by De Chaisemartin and d'Haultfoeuille (2020).

treated vs. earlier-treated units (for reviews, see Steigerwald et al., 2021; Roth et al., 2023; de Chaisemartin and D'Haultfoeuille, 2023). The estimators proposed by Borusyak et al. (2024) and Callaway and Sant'Anna (2021) provide very similar results and are reported in Appendix A.

Figure 7 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. These estimates indicate that, consistent with Figure 6, arrest rates followed similar trends in two groups of municipalities: those in their final year before the arrival of the new code, and those at least two years out from the arrival of the new code. That the point estimates are slightly negative—rather than exactly zero or zero on average—reflects the fact that the pre-trends are not *exactly* parallel (as is also evident from close inspection of Figure 6). Regardless, the striking finding in Figure 6 is the large and negative ATT in the months following the introduction of the new code.

As an estimate of the overall (not horizon-specific) ATT, De Chaisemartin and d'Haultfoeuille (2020) propose using the change in the first post-treatment period, i.e., setting $h = 1$ in Equation 1 above. Given that the dynamic treatment effects appear stable in Figure 7, this strikes us as a reasonable approach. Again, other approaches (Borusyak et al., 2024; Callaway and Sant'Anna, 2021; Sun and Abraham, 2021) produce similar estimates.

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

This table reports estimates of Equation 1, with $h = 1$.

	Municipal level		Judicial district level							
	Arrest rate				Homicide		Drug		Vandalism	
	Rate	Rate	Rate	Log	Rate	Log	Rate	Log	Rate	Log
Effect of new code	-345.11	-430.89	-378.98	-0.47	-1.28	-0.02	-72.79	-0.62	-64.35	-1.03
	(26.20)	(42.22)	(91.89)	(0.09)	(2.49)	(0.16)	(39.94)	(0.18)	(25.69)	(0.27)
Observations	31,092	31,092	972	972	972	972	972	972	972	972
Pre-period mean	570.78	901.25	744.22	6.48	18.06	2.76	159.22	4.55	54.64	3.21
Population weights		✓								

Standard errors clustered at the municipal level for columns 1 and 2, and at the judiciary district for the rest.

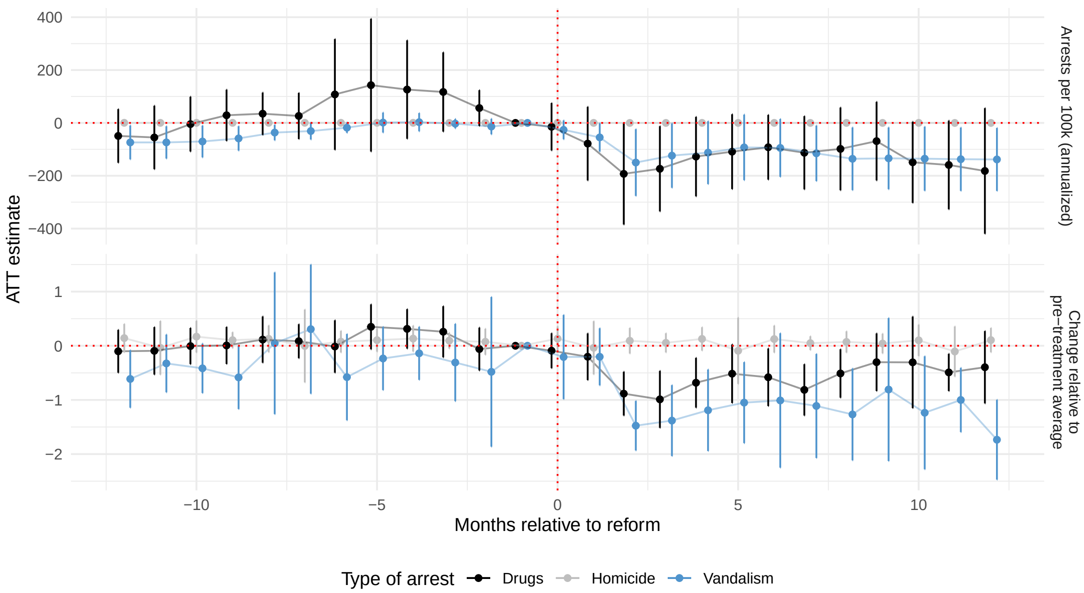
Table 1 reports the results. The average municipal arrest rate dropped by 345 arrests per 100,000 per year when the new code came into effect (Column 1); relative to the average municipal arrest rate of 570 in the *pre* period (specifically, the twelve months prior to the arrival of the new code), this estimate represents a 60% decline. Weighting the observations by population (Column 2) produces an estimate of -431 arrests per 100,000 per year, relative to a base of 901, or a 47% decline (this is the figure that we highlight in the abstract and introduction). Another way to see this is to use the *log* of the arrest rate as the dependent variable (Column 4); 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).

Arrest quality. The overall drop in arrest rates was driven by even larger declines in arrest rates for minor offenses like vandalism and drugs.[4] The blue line in the lower panel of Figure 8 plots the mean vandalism arrest rate in each month *as a fraction of* the mean vandalism arrest rate in the twelve months prior to the introduction of the new code, revealing that the mean vandalism arrest rate drops by nearly 90%. (We normalize relative to the pre-period rate in order to illustrate comparisons across crimes with very different initial arrest rates). Drug arrests (black line), too, decline more than 50% when the new code

[4]The Colombian penal code does not divide crimes into categories like *felony* and *misdemeanor*. Instead, the police distinguish between *crimes*, i.e., those listed in the penal code, and *infractions* (or *contravenciones*, which are violations of the Police Code. Infractions include many activities that would be classified as misdemeanors in the United States, such as public urination or disorderly conduct. Infractions in and of themselves typically do not lead to arrests. By "minor offenses," then, we mean those crimes for which conviction would entail noncarceral punishment or short sentences.

Figure 8: Discretionary Arrests Decline, Homicide Arrests Do Not

Each line marks the crime-specific arrest rate in each month before and after the implementation of the new code, scaled relative to the mean arrest rate (for that category) in the twelve months prior to implementation.

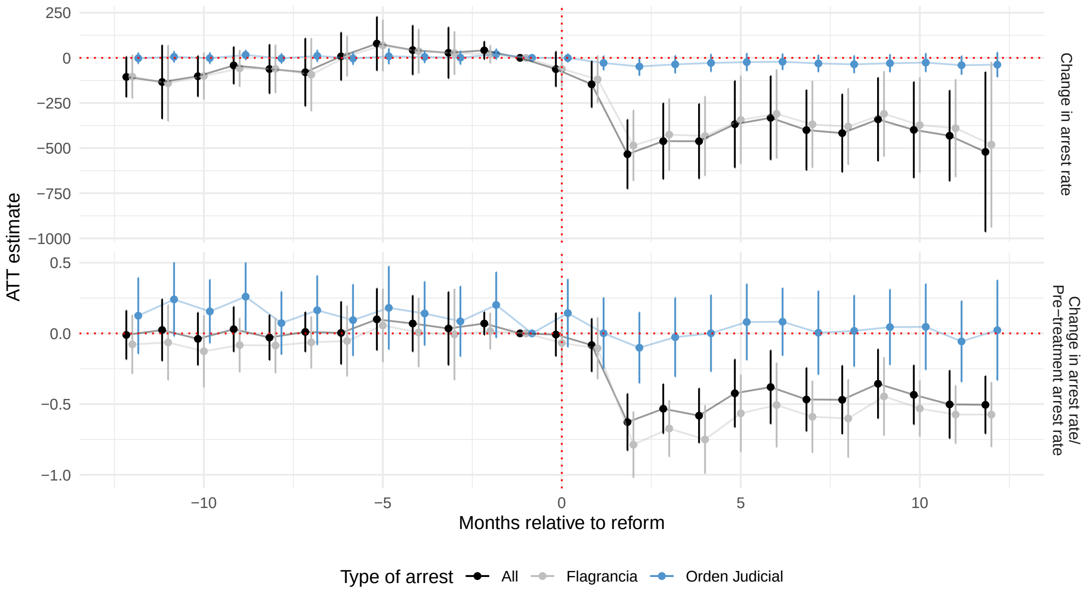


comes into effect. Homicide arrests, in contrast, do not drop as a result of the reform. Columns (6)–(10) of Table 1 confirm that similar patterns emerge using the difference-in-differences estimator described above. The declines in homicide arrest rates are substantively small and statistically indistinguishable from zero. Drug arrest rates, in contrast, fell by approximately half, and vandalism arrests collapsed almost entirely.

This pattern—larger declines in arrests for minor offenses than in homicide arrests—extends beyond the specific categories of drug and vandalism arrests. Figure 10 plots the size of the drop in the arrest rate against the minimum sentence length in Colombia's penal code (one measure of crime severity). For most crimes, there are too few arrests to estimate a treatment effect; for the nineteen charges with sufficient numbers of arrests, we observe that the arrest rate generally declines *less* for more serious crimes. Similarly, Figure B.16a plots estimates of the ATT on the homicide *share* of all arrests, in each month relative to the introduction of the new code. (To avoid zeros in the denominator, we again aggregate to the judicial-district–month level.) The mean (across judicial districts) homicide share of all arrests doubled, from approximately 1% in the pre-period to 2% after the new

Figure 9: Arrests Without Warrants Decline More than Arrests With Warrants

Each line marks the crime-specific or mode-specific arrest rate in each month before and after the implementation of the new code, scaled relative to the mean arrest rate (for that category) in the twelve months prior to implementation.

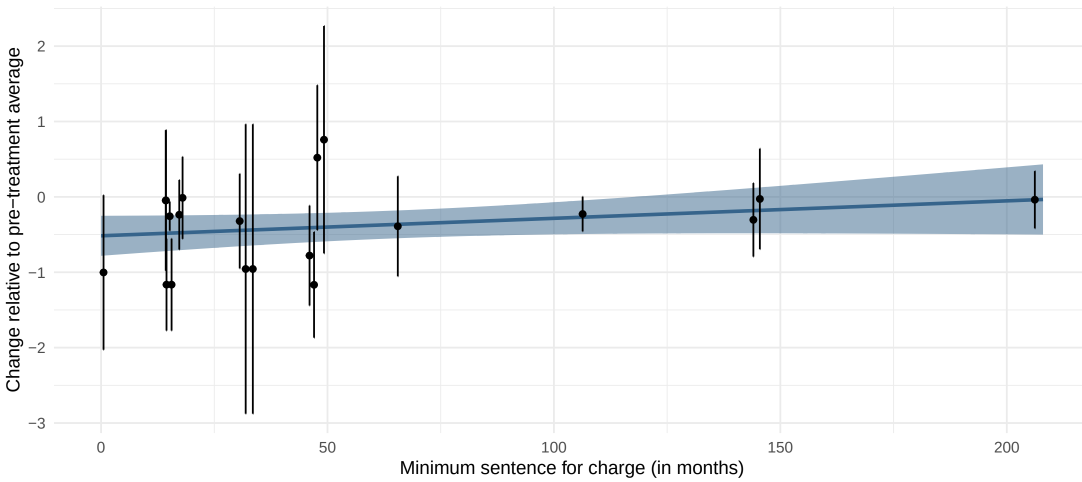


code came into effect. This result is especially striking given that the homicide rate itself was falling.

Another way to study the changing quality of arrests is to consider the share of arrests made with a warrant (as opposed to in *flagrancia*). 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 B.17). Indeed, the warrant share of arrests increased with minimum sentence length in the period prior to the new code (Figure 10). Overall, in the pre-period, 75% of arrests were made *en flagrancia* and 25% were made with warrants.

When the new code came into effect, arrests made with warrants declined by approximately 20% (Figure 9), 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 (across judicial districts) warrant *share* of arrests increased by 10 percentage points, from approximately 25 to 35 (Figure B.16b). This pat-

Figure 10: Arrest Rate Declines Less for More Serious Crimes

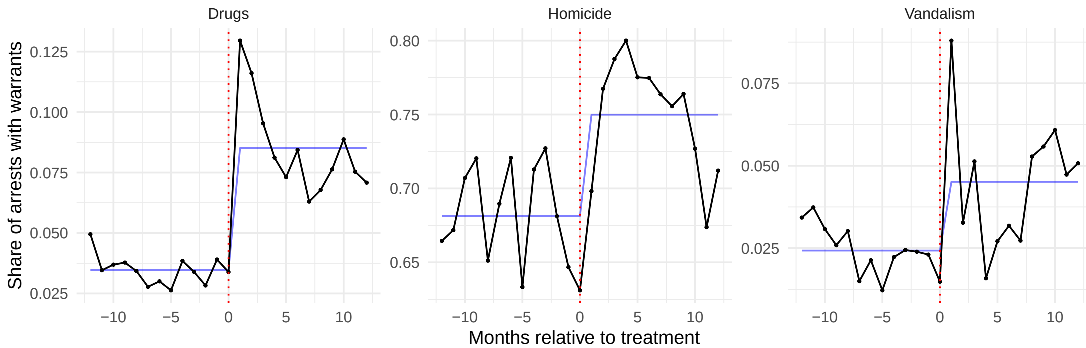


Using the estimator proposed by De Chaisemartin and d'Haultfoeuille (2020).

tern is consistent with the idea that police reallocated arrest effort away from discretionary arrests toward arrests of people suspected of serious crimes.

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. Among drug arrests, for example, the warrant share of arrests more than doubles, from 3.5% to 8.5% (Figure 11). As drug arrests fell overall, then, police prioritized drug arrests serious enough to elicit warrants. Among homicide arrests, too, the warrant share rises by seven percentage points (from 68% to 75%)—meaning that the quality of homicide arrests may have improved even as the quantity of homicide arrests held steady. Among vandalism arrests, too, the warrant share nearly doubled (from 2.4% to 4.5%). This shift toward warrant arrests is general. 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, pre-treatment). The mean increase was 6.7 percentage points over a baseline of 26%; the maximum increase was 27.5 percentage points over a baseline of 44% (for sexual acts with a minor); and the maximum decrease was 3.7 percentage points on a baseline of 28.6% (for assault). (At the level of municipality-month or even judicial-district–month, there are too many zeros in the crime-specific arrest rates to estimate ATTs on the warrant share of arrests; for that reason, we focus on the descriptive analysis in Figure 11.)

Figure 11: Warrant Share of Arrests Increases *Within* Crimes



If these shifts away from arrests for low-level offenses and away from arrests *in flagrancia* (without warrants) indeed implied prioritizing more-justified arrests, we might expect that new-regime arrests would be more likely to produce convictions. Our ability to evaluate the effect of the new code on conviction rates is limited because of problems with data from the courts: Colombia's *Consejo Superior de la Judicatura* informed us that they have no records from the year 2006; that, for the year 2005, they only have records corresponding to cases considered under the old code; that they could provide only annual (not monthly) data; and, moreover, that they could not provide microdata that would allow us to follow individual cases through the system. These limitations imply (among other issues) that we do not observe immediate pre–post values for any of the four waves of the rollout and that we cannot evaluate whether any individual arrest ends up producing a conviction. Still, we are able to calculate the ratio of convictions to arrests in each year, in each rollout wave, for each crime. Table 2 reports the results, shading treated values in blue. These values suggest that drug and homicide arrests were somewhat *more* likely to produce convictions under the new code than under the old code.

These differential changes in arrest rates did not occur because the new code imposed crime-specific or mode-specific restrictions on arrests. Rather, they emerged endogenously as the police re-optimized in response to the new rules. The advent of SIEDCO in 2003, the police in-house software for counting crimes and police activity (described above), created a Compstat-esque environment in which station commanders were evaluated in part on the basis of crime rates—with special emphasis on homicide rates. We might then expect that, facing higher costs for all arrests, commanders would direct officers to prioritize those arrests that were

Table 2: Convictions:Arrests Rise Under the New Code

This table reports the ratio of convictions to arrests in each stage of the rollout of the new code, in each available year. Blue marks the treated years.

	Drug Possession and Sale				Homicide			
Year	Wave I	Wave II	Wave III	Wave IV	Wave I	Wave II	Wave III	Wave IV
2003	0.47	0.29	0.25	0.22	0.83	0.48	0.68	0.55
2004	0.25	0.25	0.19	0.16	0.86	0.46	0.55	0.50
2005	.	0.11	0.14	0.09	.	0.43	0.49	0.57
2006	.	.	.	.	.	.	.	.
2007	0.55	0.26	0.48	0.08	1.35	0.59	0.83	0.63
2008	0.52	0.23	0.41	0.29	1.22	0.54	0.73	0.69
2009	0.29	0.26	0.37	0.20	1.01	0.52	0.73	0.53

We obtained these data via freedom-of-information request. The Consejo Superior de la Judicatura informed us that the data for 2006 are not available; for 2005, they only have data on cases considered under the old code.

most likely to reduce crime in general and homicide in particular. These details are specific to the Colombian case, but the point is general: oversight changes not only the quantity but also the quality of police activity. Understanding the shift in composition is essential to evaluating the welfare consequences.

5 The effect of the new code on crime

Colombia's new code of criminal procedure changed the criminal justice system in many ways. In studying the effect of the new code on police officers' choices of arrests, we could be confident that one specific change was most salient: the introduction of judicial oversight of arrests. But in turning to the effect of the new code on crime rates, we can no longer maintain a focus on judicial oversight alone; the new code affected crime rates not only through oversight and the consequent change in arrest rates but also through new rules for pre-trial detention, plea bargains, confessions and criminal investigations, and many other parts of the process—all of which could plausibly affect people's decisions over whether to commit crimes. For that reason, we interpret all estimates in this section as the effect of the new code overall, a bundled treatment, rather than as the effect of judicial oversight of arrests.

We first explain why changes in the method of crime registration in Colombia preclude using police data to measure property crime or violent crime other than homicide. We then introduce measures of homicide, finding no evidence that the new code caused an increase in homicide rates—but also that we lack the statistical power to estimate a precise null. Finally, we investigate data on vehicle theft from Colombia's association of insurers, finding suggestive evidence that the new code *did* cause a short-lived increase in vehicle theft. If so, we might interpret this result as evidence that the new code did weaken deterrence and/or incapacitation of people who would steal cars and motorcycles.

A change in crime registration. Ideally, we would use police counts of each crime as a measure of (reported) crime incidence. In practice, though, changes in crime registration preclude using the police data for this purpose. SIEDCO—the police agency's information system, which we introduced above—incorporated new sources of information in ways that create the appearance of huge jumps in crime incidence.[5] Consider, for example, the number of assaults in the city of Buga (Figure 12). SIEDCO's count of assaults makes it look as if assaults quintupled in January of 2006, just as the new code of criminal procedure came into effect. But data that we obtained from the police during fieldwork reveal that, in fact, this jump actually derives from the introduction of data from the public prosecutor's office (*fiscalía*).

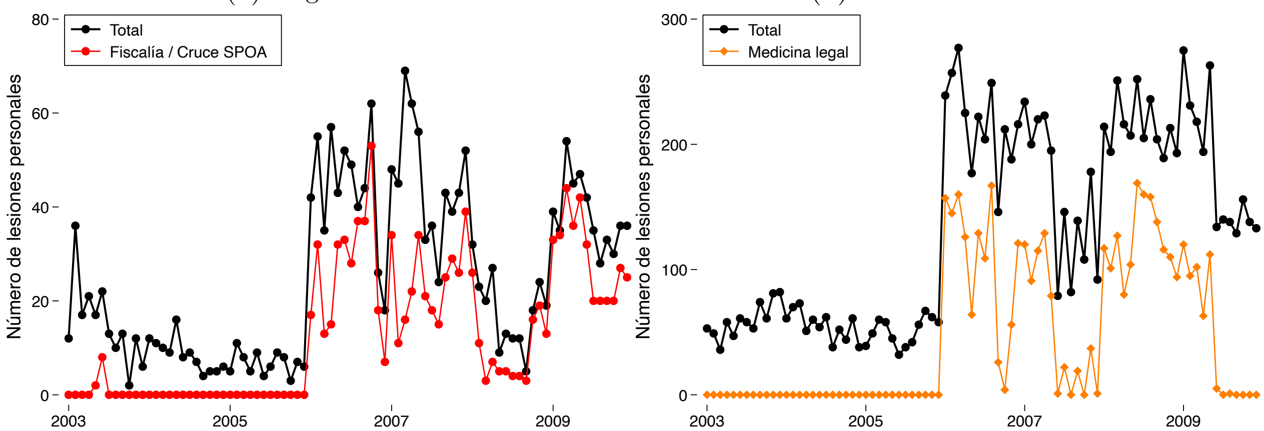
These changes are not perfectly correlated with the rollout of the new code, but nor are they uncorrelated (they do not occur all at one point in time, for example). In Villavicencio, where the new code came into effect in 2007, what looks like a quadrupling of the assault rate is actually the result of incorporating data from another source: forensic medicine (Figure 12). In other cities, other changes in crime registration similarly create a false impression of dramatic changes in crime incidence. In Bogotá, for example, in January of 2005—coincident with the arrival of the new code—SIEDCO lowered the dollar-value threshold for recording theft, creating the false appearance of a large jump in theft rates (Cámara de Comercio de Bogotá, 2006).

For these reasons, we do not use police data to measure crime incidence. Acosta et al. (2023), "On the tension between due process protection and public safety," use these data to claim that Colombia's new code of criminal procedure caused a

[5] We document these changes in in-progress work with additional coauthors (Idrobo, Kronick, Norza, and Zorro-Medina, n.d.).

Figure 12: Changes in Crime Registration Preclude Using Police Crime Data

(a) Buga (b) Villavicencio



large increase in crime.[6] Like us, they find a null effect on homicide.[7] But they estimate a 25% increase in assault. These apparent changes may stem at least in part from changes in crime registration.

The effect of the new code on homicide rates. Estimating the effect of the new code on homicide rates is much 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 nearly every city at the moment when the new code comes into effect (see Figure 13 for examples), raw arrest-rate trends do not reveal any obvious changes. In Bogotá, Medellín, and Valledupar, for example, homicide rates appeared to continue along the path set under the old code (Figure 13). For another, the national homicide rate was plummeting throughout the period of the rollout of the new code (Figure 1); unsurprisingly, it declined at very different rates (and sometimes nonlinearly) 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

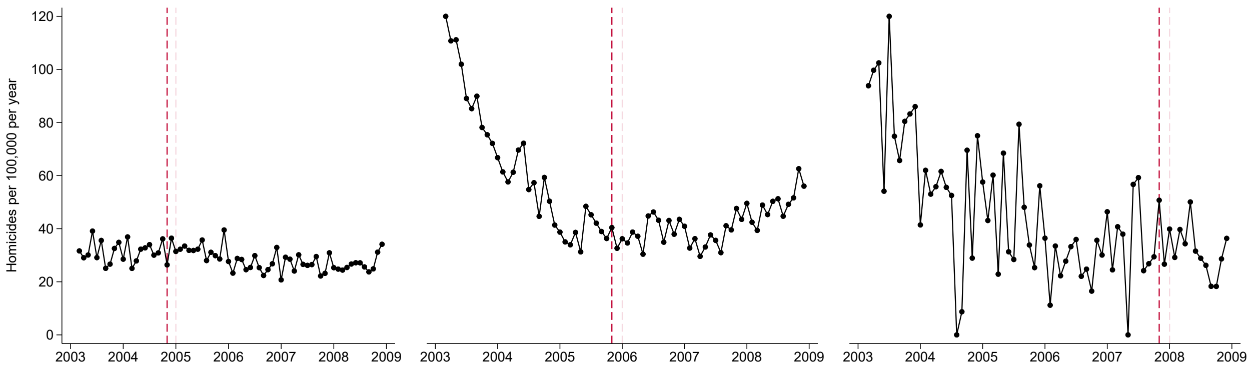
[6] Mejía et al. (2016) was the first quantitative study of Colombia's new code of criminal procedure. In that working paper, the authors documented that the new code reduced the time to trial but also reduced the probability of conviction, and they claimed that these changes caused a large increase in crime. They did not focus on arrests. A subsequent draft, Zorro Medina et al. (2020), replicated our finding on arrests, as does the most recent draft (Acosta et al., 2023).

[7] They highlight a small positive point estimate in the main text using TWFE, but the results in their appendix, using new estimators, are null.

Figure 13: 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.

(a) Bogotá (b) Medellín (c) Valledupar



is sparse. Because there are many zeros in the municipality–month panel, we are hesitant to 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.[8] Finally, as we discuss below, violence in Colombia takes many forms—only some of which might plausibly respond to a change in arrest rates.

In the Data section above, we discussed two sources of data on homicide rates: police and vital statistics. We focus on the vital statistics data here in the main text, repeating all of the analysis with the police data in the Appendix. The results are similar.

The black line in Figure 14a 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 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),

[8]At 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.

homicide rates declined less in Bogotá and the coffee belt than they did everywhere else. A similar pattern emerges in 2005 (the year prior to the second wave of the rollout) and 2006 (the year prior to the third wave): 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 Ben-Michael et al. (2022) and Cattaneo et al. (2022b) to create synthetic control groups whose pre-trends are much closer to those of the treated units.

Table 3 reports the difference-in-differences (estimates of Equation 1, for $h = 1$). The unweighted estimates (i.e., those that weight each municipality or judicial district equally), which we report in 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.

The population-weighted estimates in Panel B are also of substantive interest. Different judicial districts have very different sizes; the population of the judicial district of Bogotá, for example, was more than 6.5 million in 2003, while the population of the smallest judicial district, the Archipielago de San Andrés y Providencia, was just over 69,000. 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. For that reason, we estimate population-weighted versions of the difference-in-differences. Panel B reports the results. 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 14a–14c plot the synthetic controls for the judicial dis-

Table 3: The Effect of the New Code on Homicide Rates

This table reports estimates of Equation 1, with $h = 1$.

	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	-13.67	-4.87	0.11	0.08	0.02	0.01
	(9.52)	(5.36)	(0.09)	(0.10)	(0.07)	(0.09)
Observations	31,092	31,704	972	972	972	972
Pre-period mean	54.70	49.26	3.71	3.71	3.67	3.67
Unit-specific trends				✓		✓
Panel B: Weighted by population						
Effect of new code	-2.06	0.36	-0.02	-0.11	0.01	-0.05
	(2.83)	(1.89)	(0.06)	(0.07)	(0.04)	(0.06)
Observations	31,092	31,704	972	972	972	972
Pre-period mean	55.60	49.72	3.86	3.86	3.76	3.76
Unit-specific trends				✓		✓

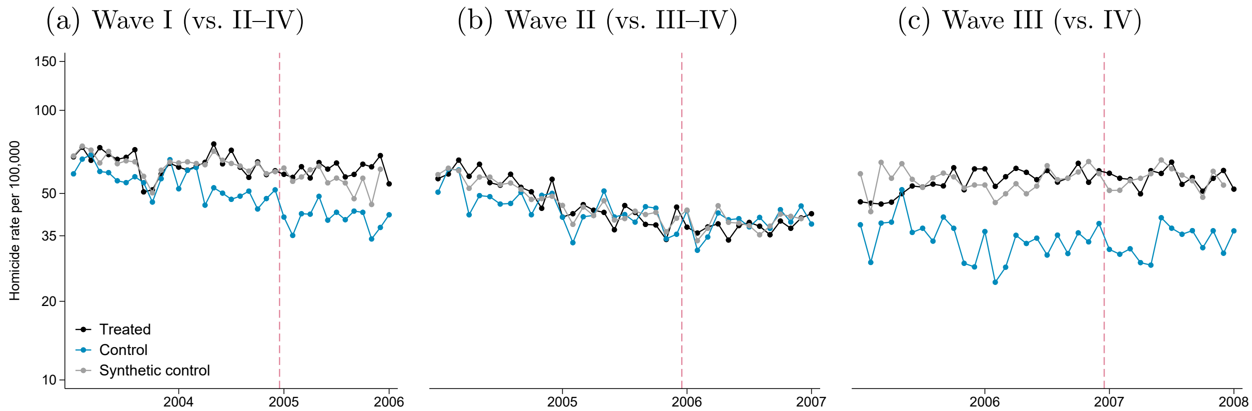
Standard errors clustered at the municipal level for Columns 1 and 2 and at the judicial-district level for Columns 3–6.

tricts in each wave of the rollout. In Waves II and III, homicide rates in the newly treated districts follow the same trend as homicide rates in the synthetic control districts. In Wave I, homicide rates in the newly treated districts do appear to increase slightly more than homicide rates in the synthetic control districts, but we do not observe any such separation when either (a) using the other measure of homicide rates (the police count) or (b) studying homicide rates in levels (rather than logs). Moreover, when we aggregate the results across all waves, the overall synthetic-control estimate of the ATT at each time horizon is substantively close to zero, using both the vital statistics and police counts of homicides (Figure 15).

In some parts of Colombia during this period, police presence was so tenuous and 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, a city of approximately 80,000 people near the border with Venezuela, conflict between the ELN and the FARC—and between these groups and the Colombian military—produced the five or ten deaths every month (a rate of more than 100 per 100,000 per

Figure 14: The Effect of the New Code on Homicide Rates

The black lines plot arrest rates in the judicial districts of Waves I, II, and III, respectively. The blue lines plot arrests rates in not-yet-treated districts. The gray lines plot synthetic controls.

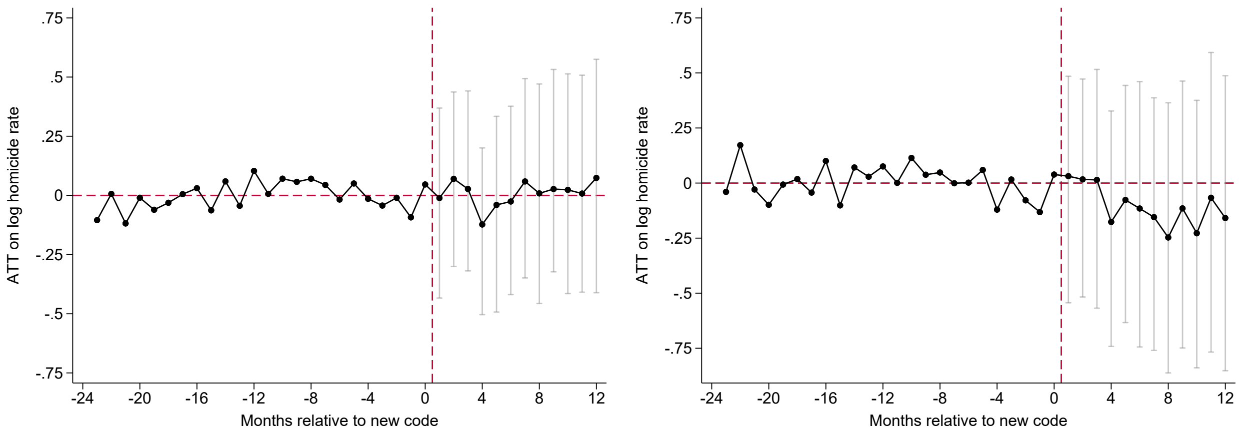


year) that led journalists call Saravena "the Colombian Sarajevo" (León, 2005); Uribe's Democratic Security initiative did establish police presence in Saravena, but a *rehabilitation zone* designation exempted their behavior from normal due process requirements (such as arrest warrants). It is hard to imagine how the *de jure* introduction of judicial oversight of arrests and/or the consequent change in arrest rates would affect lethal violence in places like Saravena. In the large cities of Bogotá or Medellín, in contrast, much of the lethal violence of the mid-2000s stemmed from conflict among small and consolidating local gangs as well as from community violence (such as bar fights and brawls). It is this violence could plausibly respond to arrest rates through the channels outlined above: deterrence, incapacitation, and/or the reallocation of police time across activities.

One might then wonder whether the apparent null effect of the new code on homicide rates (presented above) 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 four ways. First, we use the carefully constructed CERAC data set on deaths in the Colombian civil war to identify municipalities like Saravena, where the guerrilla, paramilitaries, and the armed forces were responsible for a significant fraction of all lethal violence (in Saravena in 2003, for example, war casualties amounted to 77% of all violent deaths). One limitation of this data set is that the data are annual (as Bazzi et al., 2022, note); for that reason, we use it merely to restrict the sample to municipalities with no or few war casualties, which does not much affect the results (Section D). Second, in Appendix C, we simply plot

Figure 15: The Effect of the New Code on Homicide Rates: Event Study

(a) Vital statistics data (b) Police data



Using the estimator proposed by Cattaneo et al. (2022b).

the homicide rate in Colombia's 18 largest cities; these raw time trends reveal no evidence of an obvious effect of the new code except perhaps in Ibagué and Montería, cases that we believe merit further investigation. Third, in Appendix E, we plot synthetic control difference-in-differences for each judicial district separately; we do not observe a larger effect in urban than in rural districts, nor indeed do we observe any judicial district in which the introduction of the new code appears to have affected the homicide rate. Third, we estimate the overall difference-in-differences with synthetic control, restricting the sample to municipalities with population greater than 100,000. The result remains null. We interpret these findings as evidence that the new code not only did not affect the Colombian homicide rate overall, but also that it did not affect the homicide rate in cities (where, in our view, we would be more likely to observe a nonzero effect).

The effect of the new code on vehicle theft. Finally, we consider property crime. As explained above, changes in crime registration preclude using police data to estimate the effect of the new code on crimes other than homicide. As a partial work-around, we obtained data from Fasecolda, Colombia's national association of insurance companies. These data have the same panel structure as the police data: they record the number of insured vehicles and the number of stolen vehicles in each municipality in each month between 2003 and 2011.[9]

[9] The data actually record the number of insured cars and motorcycles separately; separating the two counts does not change our main takeaway. For parsimony, we combine the two counts

But because only 20% of vehicles are insured, and because insured vehicles are so highly geographically concentrated, even the judicial-district–month panel (to say nothing of the municipality–month panel) is extremely sparse: nearly 50% of observations have zero vehicle thefts. Time trends even in the judicial-district–month panel are extremely noisy. In fact, 46% of all insured vehicles are located in Bogotá alone, and 78% are located in Bogotá, Cali, and Medellín. The time trends from these three cities suggest that the new code may have caused an uptick in vehicle theft, as we show in Appendix F. In future work, we plan to investigate the effect of the new code on shoplifting using data from retailers.

6 Discussion

We study an unusual policy experiment in Colombia. Between 2005 and 2008, a new code of criminal procedure took effect in four successive geographic blocs. The new code significantly raised the cost (to police) of making arrests; as a result, the police sharply curtailed the number of arrests, and in particular the number of arrests for low-level offenses like vandalism and drug possession. The quantity of arrests declined; the quality of arrests improved. But unlike other cases of dramatic declines in specific police activities, the drop in arrests in Colombia occurred alongside a general run-up in police funding, personnel, street presence, police stations, and extension into every municipality in the country, under an avowedly pro-police president. There were no loud calls for defunding the police.

Ex ante, one might think that such an environment would be inhospitable for what was in many ways a progressive criminal justice reform. One might think that it would be precisely a setting in which police would feel empowered to resist any constraints on their power or discretion, a setting in which we would be most likely to observe the pattern that Brinks (2007) documented in his classic study of police violence in Latin America: that "when formal rights do indeed run in favor of the hitherto powerless, we should expect more resistance from the otherwise powerful." Instead, the Colombian national police appear to have reallocated effort away from arrests for minor offenses toward other activities—with the result that homicide rates continued to plummet. This outcome is perhaps especially remarkable given that, at the time of the reform, the Colombian police were just one decade removed from the most brutal and brutalizing period of the conflict

in the main text.

with drug traffickers.

The contrast with neighboring Venezuela suggests possible explanations. In 1999, the Venezuelan government implemented a criminal procedure reform similar to the one that we study: a switch from inquisitorial to accusatorial criminal procedure, and one that imposed new constraints on police powers of arrest. Ex ante, one might have expected that Venezuela in 1999 would implement reform better than Colombia in 2005–2008: Hugo Chávez, then the new president, had railed against human rights abuses and initially embraced the protections instituted by the new code of criminal procedure; moreover, Jorge Rosell, a Venezuelan jurist, was one of the key figures in promotion of criminal procedure reform across all of Latin America. But instead, the police rebelled against the new code, both in public statements and by using violence to punish the suspected criminals whom they could no longer arrest or detain (Alguíndigue and Pérez-Perdomo, 2008; Antillano, 2010; Hanson and Kronick, 2024). Rosell later wrote that the entire reform experience in Venezuela went so awry as to leave him heartbroken (Rosell Senhenn, 2014).

We cannot draw conclusions from just two cases, but we do note that the comparison echoes Hanson (2017), who argues that when police feel abandoned or attacked by the state they are *more* likely to abuse their power. Perhaps high-level support for the Colombian police paradoxically enabled the reform to sidestep the perverse consequences observed in other incidents of oversight and depolicing. It is a possibility that reformers and activists in the United States might consider.

References

Daron Acemoglu, James A Robinson, and Rafael J Santos. The monopoly of violence: Evidence from colombia. *Journal of the European Economic Association*, 11(suppl_1):5–44, 2013.

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, Daniel Mejia, 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.

Andrés Antillano. ¿Qué conocemos de la violencia policial en venezuela? Las investigaciones e hipótesis sobre el uso de la fuerza física por la policía. *Espacio Abierto*, 2010.

Bocar A Ba and Roman Rivera. The effect of police oversight on crime and allegations of misconduct: Evidence from chicago. *U of Penn, Inst for Law & Econ Research Paper*, (19-42), 2019.

Alessandro Barbarino and Giovanni Mastrobuoni. The incapacitation effect of incarceration: Evidence from several italian collective pardons. *American Economic Journal: Economic Policy*, 6(1):1–37, 2014.

Samuel Bazzi, Robert A Blair, Christopher Blattman, Oeindrila Dube, Matthew Gudgeon, and Richard Peck. The promise and pitfalls of conflict prediction: evidence from colombia and indonesia. *Review of Economics and Statistics*, 104 (4):764–779, 2022.

Gary S Becker and George J Stigler. Law enforcement, malfeasance, and compensation of enforcers. *The Journal of Legal Studies*, 3(1):1–18, 1974.

Katherine Beckett and Angelina Godoy. A tale of two cities: a comparative analysis of quality of life initiatives in new york and bogotá. *Urban Studies*, 47 (2):277–301, 2010.

Eli Ben-Michael, Avi Feller, and Jesse Rothstein. Synthetic controls with staggered adoption. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(2):351–381, 2022.

Robert A Blair and Michael Weintraub. Mano dura: An experimental evaluation of military policing in cali, colombia. *Unpublished working paper*, 2020.

Christopher Blattman, Gustavo Duncan, Benjamin Lessing, and Santiago Tobón. Gang rule: Understanding and countering criminal governance. Technical report, National Bureau of Economic Research, 2021a.

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, 2021b.

Kirill Borusyak, Xavier Jaravel, and Jann Spiess. Revisiting event-study designs: Robust and efficient estimation. *The Review of Economic Studies*, 2024.

Anthony A Braga, Brandon S Turchan, Andrew V Papachristos, and David M Hureau. Hot spots policing and crime reduction: An update of an ongoing systematic review and meta-analysis. *Journal of experimental criminology*, 15: 289–311, 2019.

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.

Paolo Buonanno and Steven Raphael. Incarceration and incapacitation: Evidence from the 2006 italian collective pardon. *American Economic Review*, 103(6): 2437–2465, 2013.

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.

John Cass. Chicago cop is the face of the ferguson effect, 2016.

Paul G Cassell and Richard Fowles. What caused the 2016 chicago homicide spike: An empirical examination of the aclu effect and the role of stop and frisks in preventing gun violence. *U. Ill. L. Rev.*, page 1581, 2018.

Matias D Cattaneo, Carlos Dıaz, and Rocıo Titiunik. 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.

Ricardo Cendales and Constanza Pardo. Quality of death certification in colombia. *Colombia Médica*, 49(1):121–127, 2018.

Aaron Chalfin and Justin McCrary. Are us cities underpoliced? theory and evidence. *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.

Sungwoo Cho, Felipe Gonçalves, and Emily Weisburst. The impact of fear on police behavior and public safety. Technical report, National Bureau of Economic Research, 2023.

Daniela Collazos, Eduardo García, Daniel Mejía, Daniel Ortega, and Santiago Tobón. Hot spots policing in a high-crime environment: An experimental evaluation in medellin. *Journal of Experimental Criminology*, 17:473–506, 2021.

Contraloría Generalde la Nación. Evaluación sobre la implementación del sistema penal oral acusatorio en colombia. *Contraloría Delegada, Sector Defensa, Justicia y Seguridad*, 2010.

Clément De Chaisemartin and Xavier d'Haultfoeuille. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110 (9):2964–2996, 2020.

Clément de Chaisemartin and Xavier D'Haultfoeuille. Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: a survey. *The Econometrics Journal*, 26(3):C1–C30, 2023.

Tanaya Devi and Roland G Fryer Jr. Policing the police: The impact of" pattern-or-practice" investigations on crime. Technical report, National Bureau of Economic Research, 2020.

Rafael Di Tella and Ernesto Schargrodsky. Do police reduce crime? estimates using the allocation of police forces after a terrorist attack. *American Economic Review*, 94(1):115–133, 2004.

Avinash Dixit. Incentives and organizations in the public sector: An interpretative review. *Journal of human resources*, pages 696–727, 2002.

Patricio Dominguez-Rivera, Magnus Lofstrom, and Steven P Raphael. The effect of sentencing reform on crime rates: Evidence from california's proposition 47. IZA Discussion Paper, 2019.

Oeindrila Dube and Suresh Naidu. Bases, bullets, and ballots: The effect of us military aid on political conflict in colombia. *The Journal of Politics*, 77(1): 249–267, 2015.

Leopoldo Fergusson, James A Robinson, Ragnar Torvik, and Juan F Vargas. The need for enemies. *The Economic Journal*, 126(593):1018–1054, 2016.

Laura Gamboa. *Opposition at the Margins*. Cambridge University Press, 2022.

Juan García, Daniel Mejia, and Daniel Ortega. Police reform, training and crime: experimental evidence from colombia's plan cuadrantes. *Documento CEDE*, (2013-04), 2013.

Julian E Gerez. Political forbearance and intensification of counternarcotics enforcement. 2023.

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.

Josbelk González Mejías and Dorothy Kronick. The problem with venezuelan homicide data, and a solution. In D. Smilde, V. Zubillaga, and R. Hanson, editors, *The Paradox of Violence in Venezuela: Revolution, Crime, and Policing During Chavismo*, Pitt Latin American Series. University of Pittsburgh Press, 2023. ISBN 9780822947127. URL https://books.google.com/books?id=4ryZzgEACAAJ.

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.

Rebecca Hanson. Unruling the law: Democratic policing, socialist revolution, and violent pluralism in Venezuela. *Doctoral Dissertation*, 2017.

Rebecca Hanson and Dorothy Kronick. Official vigilantism. Working Paper, 2024.

MIldred Hartmann Arboleda. La detención preventiva y la reforma procesal penal en colombia. 2016.

David Hausman and Dorothy Kronick. The illusory end of stop and frisk in chicago? *Science Advances*, 9(39):eadh3017, 2023.

Nicolás Idrobo, Dorothy Kronick, Ervyn Norza, and Ángela Zorro-Medina. El efecto de los cambios en registro sobre el número de delitos registrados en siedco en el período 2003–2009, n.d.

George L Kelling and Catherine M Coles. *Fixing broken windows: Restoring order and reducing crime in our communities*. Simon and Schuster, 1997.

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.

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.

Heather Mac Donald. *The war on cops: How the new attack on law and order makes everyone less safe*. Encounter Books, 2017.

John M MacDonald. De-policing as a consequence of the so-called ferguson effect, 2019.

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.

Justin McCrary. Using electoral cycles in police hiring to estimate the effect of police on crime: Comment. *American Economic Review*, 92(4):1236–1243, 2002.

Mathew D McCubbins and Thomas Schwartz. Congressional oversight overlooked: Police patrols versus fire alarms. *American journal of political science*, pages 165–179, 1984.

Mathew D McCubbins, Roger G Noll, and Barry R Weingast. Administrative procedures as instruments of political control. *The Journal of Law, Economics, and Organization*, 3(2):243–277, 1987.

Camilo Acosta Mejía, Daniel Mejía Londoño, and Angela Zorro Medina. Certainty vs. severity revisited: Evidence for colombia. *Working Paper*, 2016.

Daniel Mejía. Plan colombia: an analysis of effectiveness and costs. *Foreign Policy at Brookings*, 17, 2016.

Daniel Mejía, Pascual Restrepo, and Sandra V Rozo. On the effects of enforcement on illegal markets: evidence from a quasi-experiment in colombia. *The World Bank Economic Review*, 31(2):570–594, 2017.

Ministerio de Defensa Nacional. Respuesta a derecho de petición 497659-20240408, 2024a.

Ministerio de Defensa Nacional. Respuesta a derecho de petición 497659-20240408, 2024b.

Eduardo Moncada. Toward democratic policing in colombia? institutional accountability through lateral reform. *Comparative Politics*, 41(4):431–449, 2009.

Jonathan Mummolo. Militarization fails to enhance police safety or reduce crime but may harm police reputation. *Proceedings of the national academy of sciences*, 115(37):9181–9186, 2018.

Emily Owens. Economic approach to de-policing. *Criminology & Pub. Pol'y*, 18: 77, 2019.

Emily Owens and Bocar Ba. The economics of policing and public safety. *Journal of Economic Perspectives*, 35(4):3–28, 2021.

F. O. Piraquive Sierra. Distrito judicial de yopal. *Derecho y Realidad*, 5:87–114, 2007.

Canice Prendergast. Selection and oversight in the public sector, with the los angeles police department as an example, 2001.

Canice Prendergast. 'drive and wave': The response to lapd police reforms after rampart. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2021-25), 2021.

David C Pyrooz, Scott H Decker, Scott E Wolfe, and John A Shjarback. Was there a ferguson effect on crime rates in large us cities? *Journal of criminal justice*, 46:1–8, 2016.

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.

Jorge L. Rosell Senhenn. La reforma procesal penal en venezuela ¿en qué fallamos? *Ciencias Penales desde el Sur: Segundo Congreso Latinoamericano de Derecho Penal y Criminología*, 2014.

Richard Rosenfeld and Joel Wallman. Did de-policing cause the increase in homicide rates? *Criminology & Public Policy*, 18(1):51–75, 2019.

Jonathan Roth, Pedro H.C. Sant'Anna, Alyssa Bilinski, and John Poe. What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*, 2023.

Stephen Rushin and Griffin Edwards. De-policing. *Cornell L. Rev.*, 102:721, 2016.

Deborah Sontag. The secret history of colombia's paramilitaries and the u.s. war on drugs. The New York Times Magazine, 2016.

Douglas G. Steigerwald, Gonzalo Vazquez-Bare, and Jason Maier. Measuring heterogeneous effects of environmental policies using panel data. *Journal of the Association of Environmental and Resource Economists*, 8(2):277–313, 2021.

Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2): 175–199, 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.

JC Ruiz Vásquez. Community police in colombia: an idle process. *Policing and society*, 22(1):43–56, 2012.

David Weisburd, Joshua C Hinkle, Anthony A Braga, and Alese Wooditch. Understanding the mechanisms underlying broken windows policing: The need for evaluation evidence. *Journal of research in crime and delinquency*, 52(4): 589–608, 2015.

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.

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

Appendix

A Robustness of the Arrest Estimates

Table A.1

	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'Haultfoeuille (2020) (reported in main text)										
Effect of new code	-345.11	-430.89	-378.98	-0.47	-1.28	-0.02	-72.79	-0.62	-64.35	-1.03
	(26.20)	(42.22)	(91.89)	(0.09)	(2.49)	(0.16)	(39.94)	(0.18)	(25.69)	(0.27)
Observations	31,092	31,092	972	972	972	972	972	972	972	972
Pre-period mean	570.78	901.25	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	-414.90	-473.11	-0.48	$-3.78^{\S}$	$-0.13^{\S}$	-143.32	-0.58	-79.57	-1.19
	(24.80)	(36.64)	(120.89)	(0.09)	(2.27)	(0.10)	(58.22)	(0.16)	(29.95)	(0.28)
Observations	41,314	41,314	1,303	1,303	1,303	1,303	1,303	1,303	1,303	1,303
Estimates from Callaway and Sant'Anna (2021)										
Effect of new code	-340.33	-337.90	-369.81	-0.45	-0.86	0.03	-70.53	-0.60	-60.93	-1.03
	(26.56)	(.)	(94.28)	(0.09)	(2.89)	(0.18)	(41.01)	(0.19)	(21.70)	(0.27)
Observations	53,556	53,556	1,694	1,694	1,694	1,694	1,694	1,694	1,694	1,694
Population weights		✓								

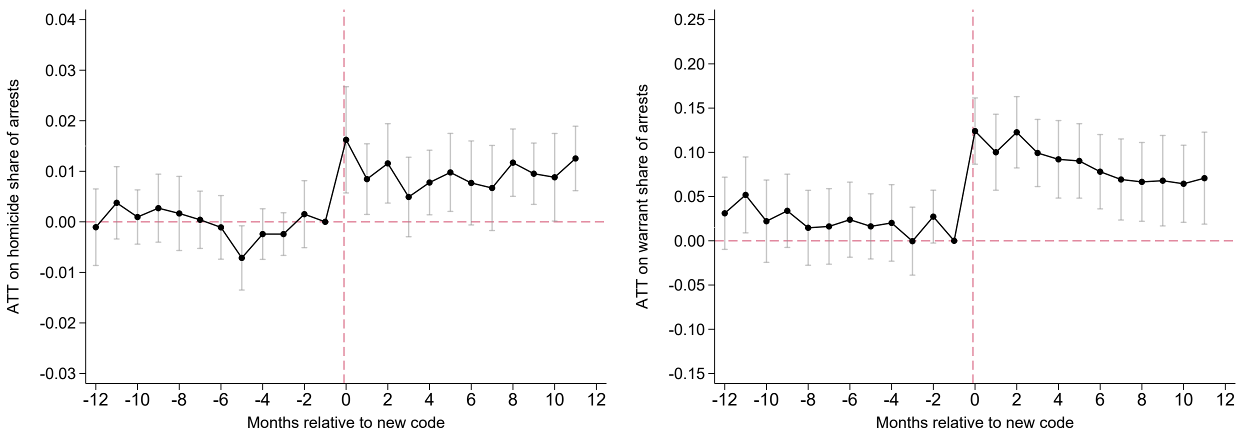
Standard errors clustered at the municipal level for columns 1 and 2, and at the judiciary district for the rest.

§ When studying the homicide arrest rate, estimates from Borusyak et al. (2024)'s estimator do not appear to follow parallel pre-trends; moveover, because the pre-trend estimates rely on the earliest time periods as a reference group, we include unit-specific linear trends in Panel B, Columns 5–6. Excluding them yields a somwhat larger negative point estimate of -0.25 (for Column 6).

B Arrest Quality: Additional Results

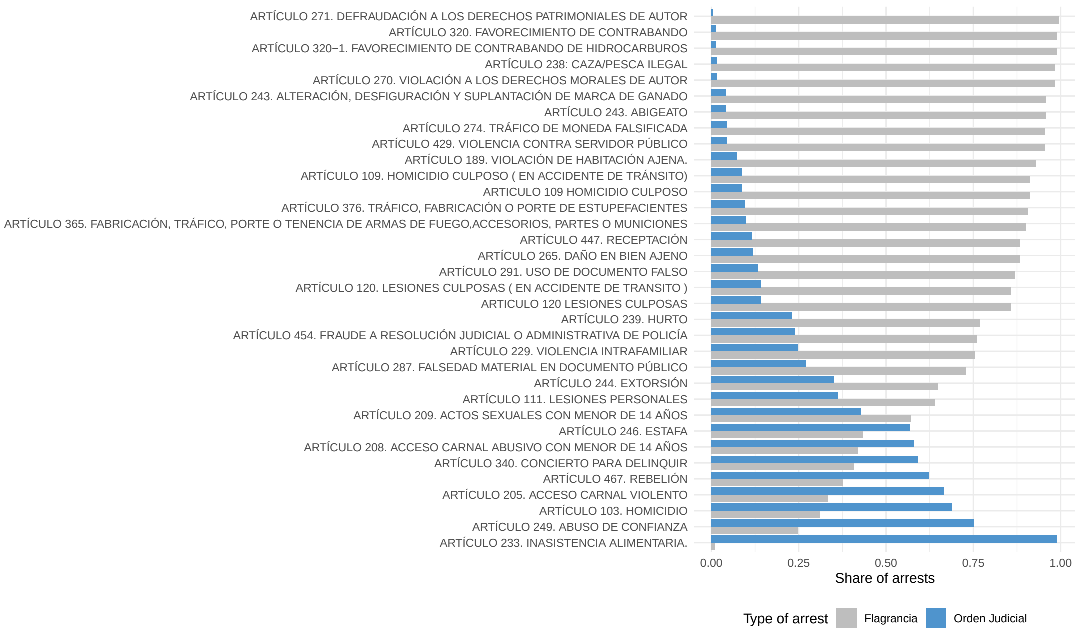
Figure B.16: The Effect of the New Code on the Composition of Arrests

(a) Homicide share of arrests (b) Warrant share of arrests

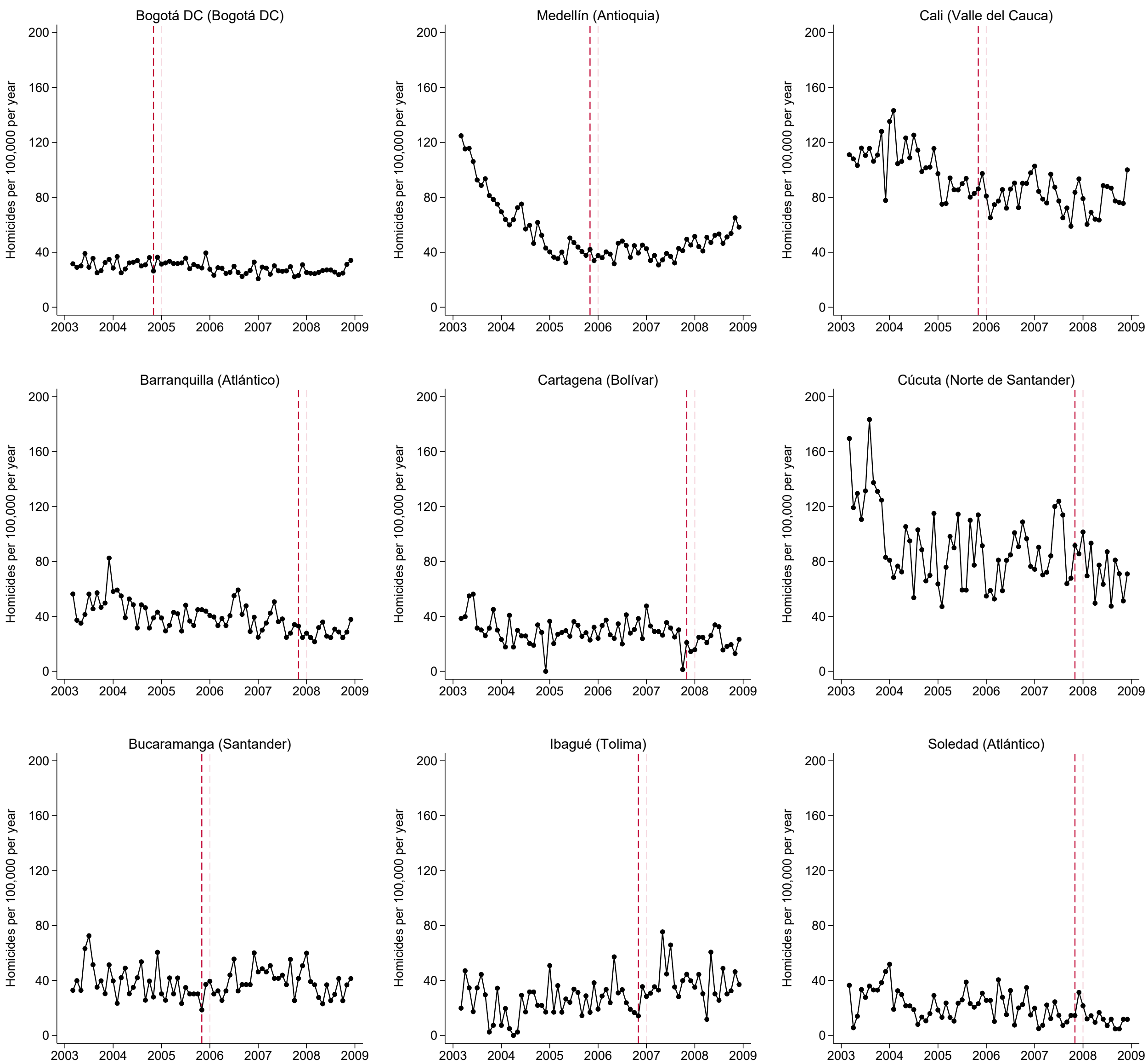


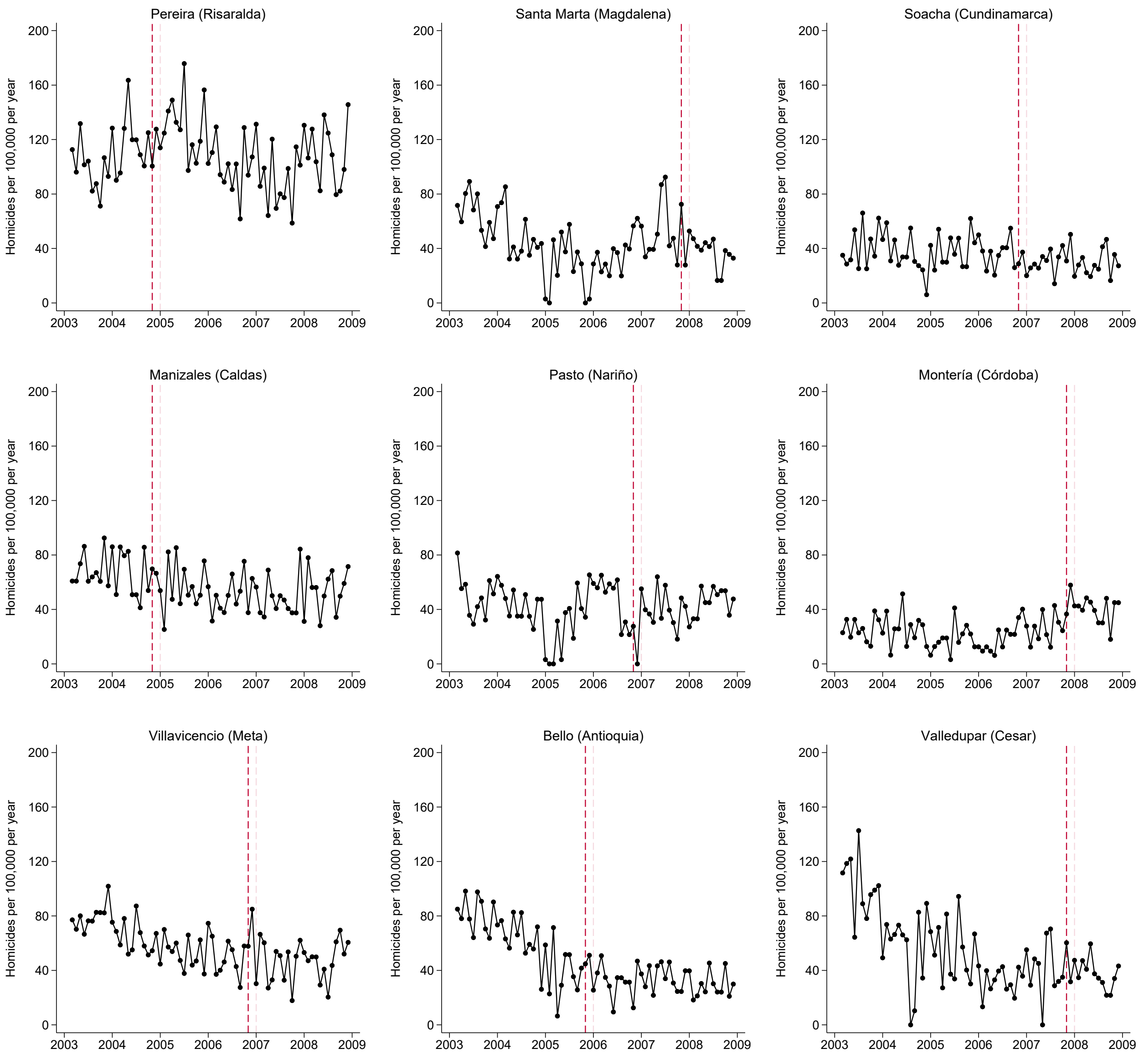
Using the estimator proposed by De Chaisemartin and d'Haultfoeuille (2020).

Figure B.17: Warrants Seldom Used for Minor Offenses

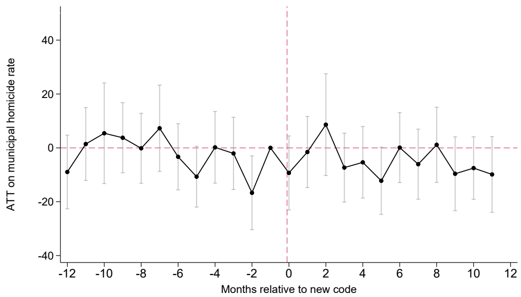


C Homicide Rate in the Largest Cities

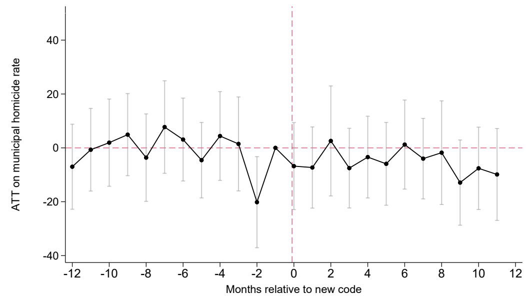



Pereira (Risaralda)
Santa Marta (Magdalena)
Soacha (Cundinamarca)
Manizales (Caldas)
Pasto (Nariño)
Montería (Córdoba)
Villavicencio (Meta)
Bello (Antioquia)
Valledupar (Cesar)
Homicides per 100,000 per year
0
40
80
120
160
200
2003
2004
2005
2006
2007
2008
2009

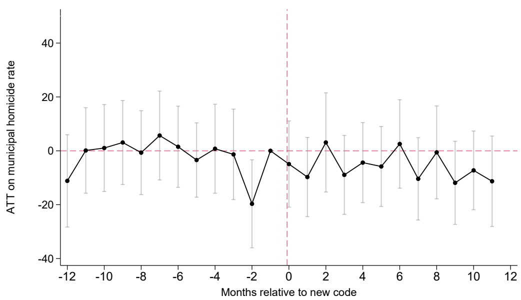
D Homicide Estimates Excluding War Deaths



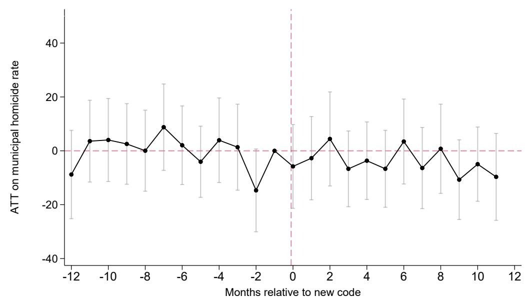
(a) Full Sample



(b) Ratio smaller than 0.05 (−482 municipalities)

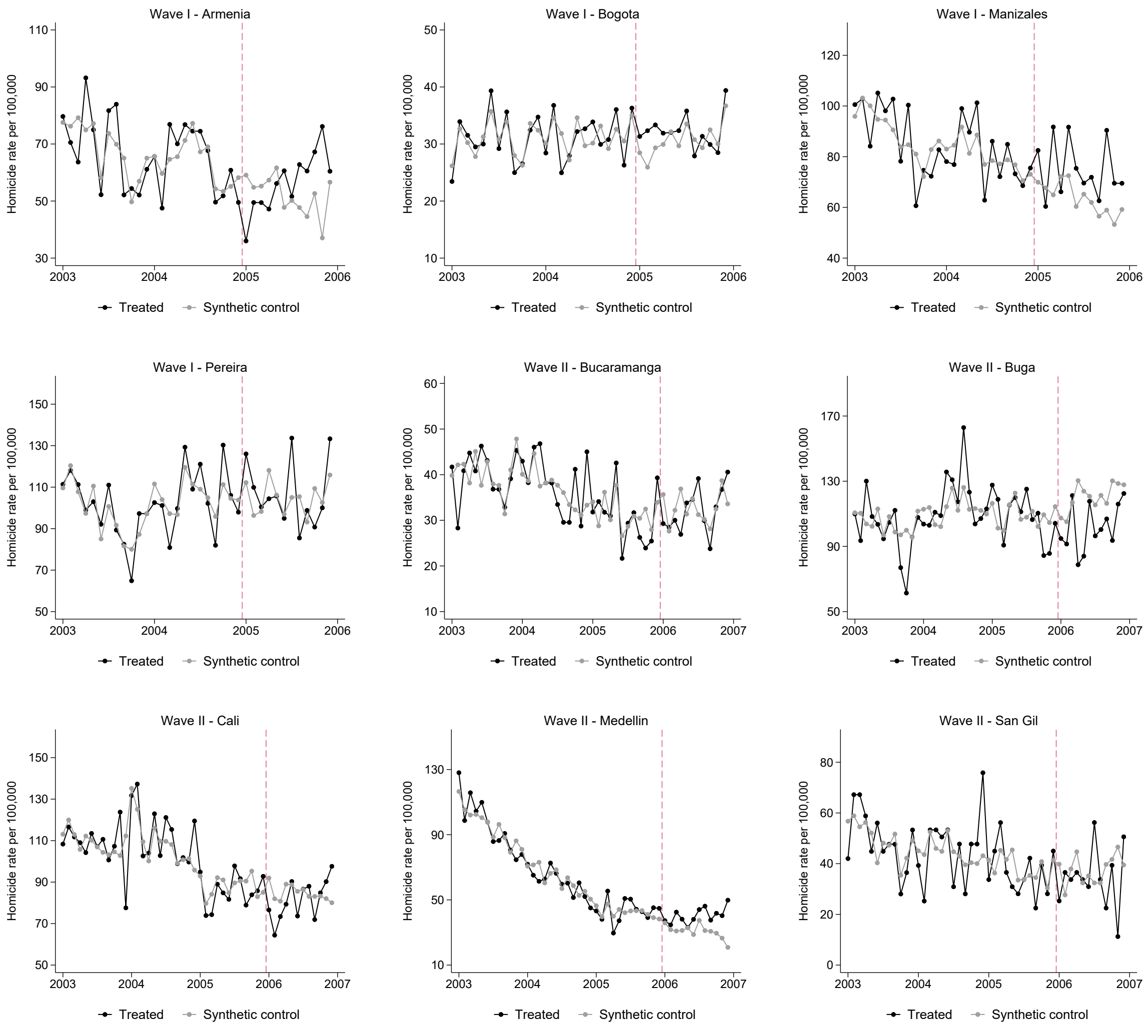


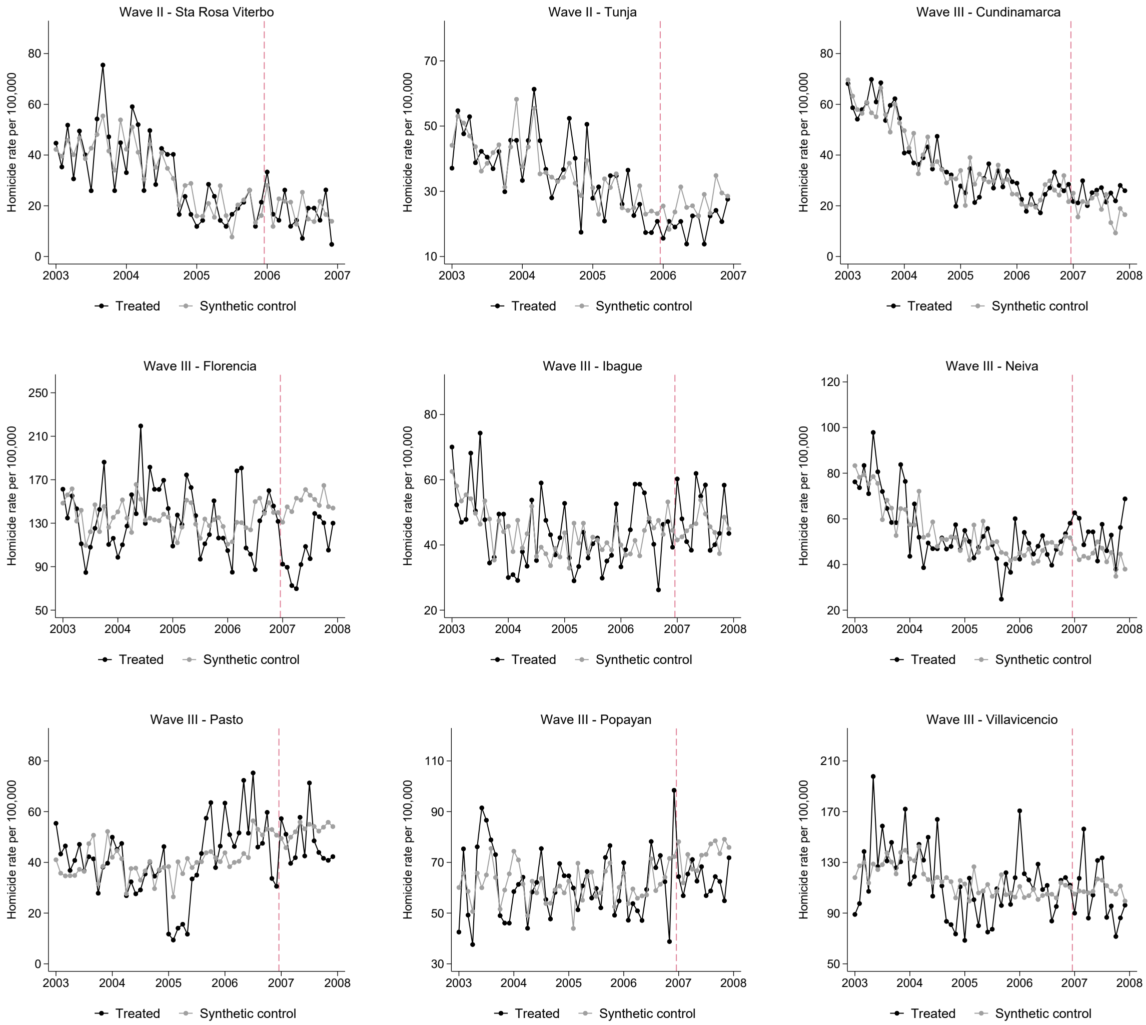
(c) Ratio smaller than 0.15 (−344 municipalities)



(d) Ratio smaller than 0.25 (−266 municipalities)

E Synthetic Control by Judicial District




Wave II - Sta Rosa Viterbo
Wave II - Tunja
Wave III - Cundinamarca
Wave III - Florencia
Wave III - Ibague
Wave III - Neiva
Wave III - Pasto
Wave III - Popayan
Wave III - Villavicencio
Homicide rate per 100,000
Treated
Synthetic control

F Vehicle Theft

Because the data on vehicle theft (from insurers) is so sparse, we attempt to discuss the effect of the new code on vehicle theft merely by examining time trends in vehicle theft rates in the three cities that contain 78% of insured vehicles: Bogotá, Cali, and Medellín.

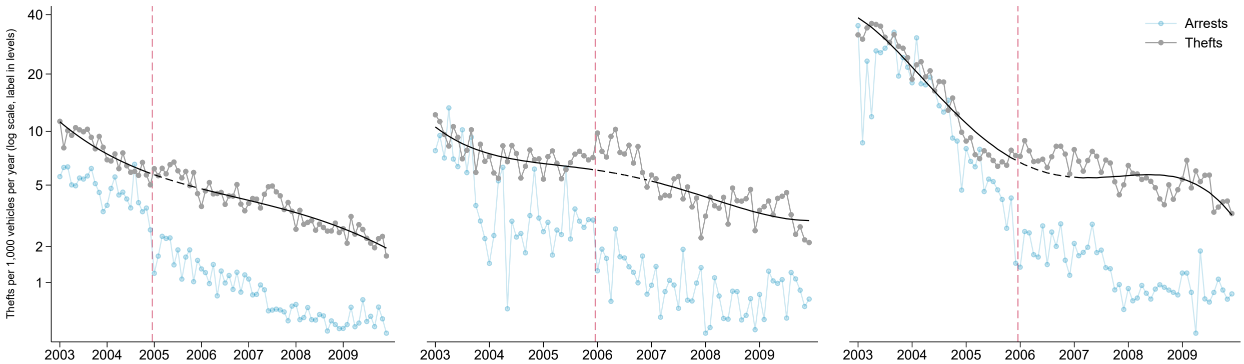
Figure F.21 plots the number of vehicle thefts per 1,000 insured vehicles per year; we use a log scale in order to facilitate comparison across the three cities (for readability, the y-axis labels are in levels). In Bogotá, the overall vehicle theft rate fell quickly and nearly monotonically throughout 2003–2009—*except* for a blip beginning in January, 2005, when the new code of criminal procedure came into effect in Bogotá. This time series alone would not be sufficient to attribute that blip to the new code (as opposed to other coincident changes), but the time series from Cali and Medellín are highly suggestive. In Cali, the vehicle theft rate fell by nearly 50% in 2003 (from 12 to 7 per 1,000 per year), flattened out in 2004–2005, and then jumped up to 8 in 2006, when the new code came into effect in Cali. After 2006, the vehicle theft rate in Cali resumed a downward trajectory. In Medellín, the vehicle theft rate fell even more faster in 2003–2004, from a peak of 35 per 1,000 down to 7 per 1,000; this dramatic improvement in citizen security parallels the 66% drop in the homicide rate over the same period (Figure 13). The vehicle theft rate in Medellín was remarkably stable in 2005, the year prior to the new code, and then increased slightly in 2006. These trends suggest that the new code did increase vehicle theft.

Though *arrests* for vehicle theft did not decline nearly as much as drug or vandalism arrests, they did drop when the new code of criminal procedure came into effect. The blue lines in Figure F.21 mark the number of vehicle-theft *arrests* per 1,000 insured vehicles per year (i.e., using the same denominator with which we normalize the number of vehicle thefts). In Bogotá and Medellín, the number of *arrests* for vehicle theft hovered close to the number of *thefts* of insured vehicles themselves in the years prior to the new code, suggesting a significant degree of enforcement; in Cali, the ratio (of vehicle-theft arrests to thefts of insured vehicles) was much lower than in the other two cities, but similarly stable in the two years leading up to the implementation of the new code. When the new code came into effect, however, the number of vehicle-theft arrests dropped sharply in all three cities—and especially in Cali, the city with the largest uptick in vehicle

Figure F.21: The Effect of the New Code on Vehicle Theft

The gray 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). Black 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.

(a) Bogotá (b) Cali (c) Medellín



theft. Moreover, vehicle-theft arrests appear to be falling for several months in anticipation of the arrival of the new code in Medellín, just as the vertiginous drop in vehicle theft begins to flatten out.