

Objectively Better, Subjectively Worse? Evaluating Criminal Justice Reform’s Impact on Security and Trust by Matching Municipalities in Mexico

Matthew C. Ingram

University at Albany, SUNY

mingram@albany.edu

July 30, 2020

Draft; comments welcome; please do not cite without permission.

Abstract

This paper employs a matching analysis to examine the impact of criminal procedure reform on security and public trust at the municipal level in Mexico from 2007-2012. The quasi-experimental design pairs communities that implemented the reform (“treated” units) to communities that did not (“control” units) to maximize analytic leverage. Results show the reform had a helpful impact on objective measures of security (reducing crime and violence across five outcomes), but had a harmful impact on subjective measures of security and public trust, especially on perceived security trend and trust in prosecutors. These mixed results highlight a tension in that the reform (a) succeeds in improving objective security conditions, yet (b) fails according to subjective perceptions of deteriorating security and justice institutions. The findings have broader implications for the politics of the reform process, and for the challenges of building and maintaining democratic institutions.

Word count: 9768 (excludes title page; includes all text, tables, references)

Key words: politics of reform, matching, criminal justice, Mexico, security, trust

Introduction

This paper employs a matching analysis at the municipal level in Mexico to examine the impact of criminal procedure reform on security and public trust in key justice institutions. Criminal procedure reform constituted a major transformation of justice systems across Latin America from the 1990s to 2010s (Langer 2007), shifting from a more opaque, inquisitorial system historically associated with the civil law tradition to a more open, adversarial system associated with the common law tradition.

The reform in Mexico was supposed to deliver improvements across a wide range of outcomes, including greater protection of the rights of both suspects and victims, increased efficiency, independence and accountability in the form of increased checks and balances among actors within and across institutions, oral proceedings and more transparency, and an overall increase in the quality of criminal justice across the entire system – from police to prisons. These gains promised reduced criminality, impunity, and corruption, improved public safety (security), and increased public trust and confidence in justice institutions (Blanco 2016; Ingram 2016b). However, throughout the reform process, advocates encountered backlash. The reform was widely regarded as a progressive achievement in the protection of due process rights. Conservative opponents argued the reform was soft on crime, and they pressed for more heavy-handed and repressive policies and practices consistent with trends towards “punitive populism” in region (Bailey and Dammert 2006; Dammert 2019). Efforts to reverse the reform have lobbied for harsher punishments (Ingram 2013), and more authoritarian measures, including extended periods of preventive custody (more than 2 years), preventive custody for all crimes, and less independence among judges (Cataño González 2020b; González Gómez 2020a; Novoa 2020). Despite this major transformation, ongoing debate, and high stakes, there is little research evaluating the reform’s impact in Latin America (c.f., Tiede 2012), including in Mexico (c.f., Blanco 2016; Huebert 2019).

This study examines nine outcomes of interest, including five objective measures of

crime and violence (homicide rates overall, homicide rates for men, homicide rates for women, total crime rate, robbery rate), and four subjective measures of security and trust (perceived security, perceived security trend, trust in prosecutors, and trust in judges). The analysis matches implementing municipalities with similar municipalities that had not yet implemented the reform to strengthen causal inferences (e.g., [King and Nielsen 2019](#)). The former, “treated” municipalities are then compared to the latter, “control” municipalities. Differences in outcomes across paired treated and control units reveal the reform’s impact.

Results show the reform had a helpful, beneficial impact on objective measures of security (reducing rates of crime and violence), but had a harmful impact on subjective measures of security and public trust, especially on trust in prosecutors. These mixed results highlight a tension between (a) the objective improvement in security due to the reform, and (b) the subjective perception that security and justice institutions are deteriorating. That is, the reform generated results that were objectively better, yet the subjective perception in these communities was that communities were worse off after the reform. The objective outcomes support consolidating the reform, while subjective perceptions provide ammunition for more conservative efforts aligned with punitive populism. Given persistently high levels of violence and insecurity accompanied by persistently low levels of public trust in public institutions, as well as the association between insecurity and declining support for democracy across Latin America ([Zechmeister and Lupu 2019](#)), a better understanding of the reform’s effects on security and public trust promises valuable insights into the politics of the reform process, long-term stability of criminal justice reform, and the challenges of maintaining democracy and rule of law in the region.

The paper proceeds as follows. I first explain the focus on criminal procedure in Mexico. I then turn to the theoretical framework and working hypotheses. The subsequent sections describe data and methods, including the use of coarsened exact matching (CEM, [Iacus, King and Porro 2009](#)) in favor of the commonly used propensity score matching ([King and Nielsen 2019](#)). The following section reports core results,

followed by robustness tests and a discussion of key findings and implications. The conclusion summarizes findings, limitations, and future directions.

Why Criminal Procedure in Mexico?

A massive, revolutionary transformation has been spreading across Latin America since the 1990s, consisting of movement away from the inquisitorial style of criminal procedure traditionally associated with the civil law tradition of continental Europe and toward a more accusatorial style historically associated with the common law tradition of England and the United States. Between 1992 and 2006, 14 countries and several subnational units in the region were part of this transformation, a change that constituted the “deepest transformation that Latin American criminal procedures have undergone in nearly two centuries” ([Langer 2007](#), 617).

Mexico was emblematic of this change starting in mid-2000s, so examining the impact of the reform here promises a better understanding of the phenomenon both in Mexico and elsewhere. Additionally, Mexico’s federal structure and the temporal variation of reform across states, across judicial districts within states, and across municipalities within judicial districts, provides a laboratory of diverse local experiences in order to assess the impact of reform.

Criminal procedure reform reached Mexico in the early 2000s, first in individual states, followed by a failed federal effort in 2004. The most important moment in decades was the successful federal constitutional reform in 2008 that mandated reform across all state and federal jurisdictions by June 2016. Most states then proceeded to adopt new codes of criminal procedure and other relevant laws, make new investments in training personnel to operate the new system, and built new courthouses and other facilities to accommodate new practices, e.g., oral proceedings. Despite the constitutional mandate and major effort in some states, the reform was implemented unevenly across the country, stalling around 2011-2012 as uncertainty increased regarding ongoing support for reform due to unclear positions of leading candidates in

2012 presidential election. Federal institutions were especially slow to reform, and as of 2013, five years after 2008 reform, there was no new federal legislation to guide reform, much less any implementation at federal level. Finally, a new, unified National Code of Criminal Procedure was approved in 2014 that applied equally to both state and federal jurisdictions. As of 2020, this code was still in place, but it was not fully implemented.

Figure 1 visualizes variation in the implementation of the reform. Panel 1a in Figure 1 shows the level of implementation in the 32 states as of 2011, and panel 1b shows the level of implementation in the federal justice system as of February 2018, where 38 federal jurisdictions overlap with state boundaries. In both panels, darker shades identify more advanced jurisdictions. As Figure 1 shows, this layered process of reform generated rules and practices in the criminal justice sector that were highly uneven over time, across geographic jurisdictions, and across levels of government.

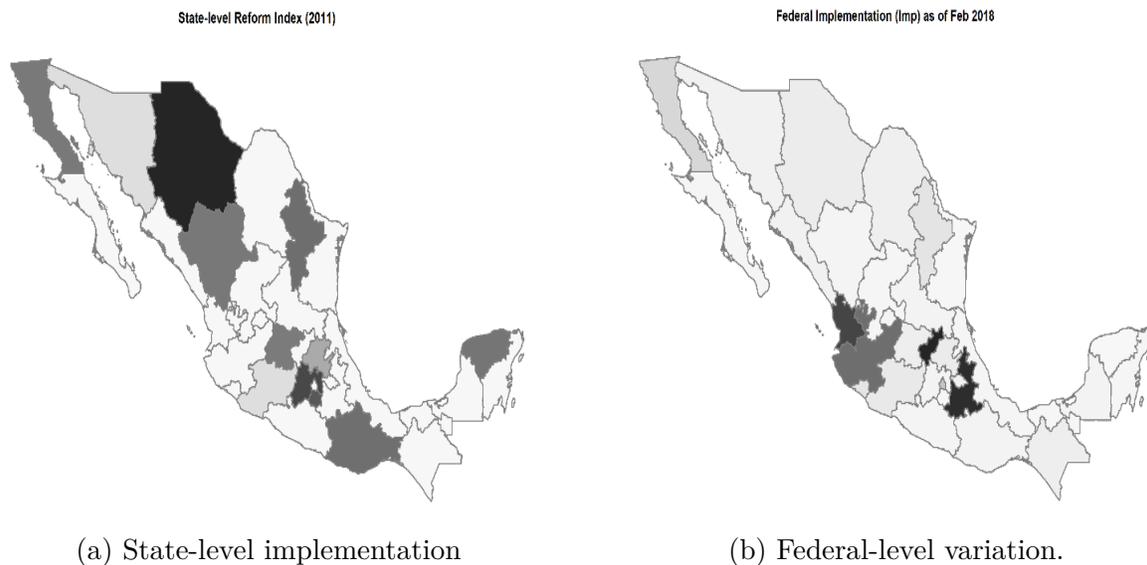


Figure 1: Geographic variation in implementation.

From a methodological perspective, this temporal and geographic variation offers unusually rich and valuable opportunities for comparative study of the causes and consequences of this major reform. These opportunities are not present to the same degree in other countries in the region since most countries in Latin America are

unitary, and, even among the region’s three large federal systems – Argentina, Brazil, and Mexico – Argentina and Brazil have moved slowly. Thus, Mexico provides the most information for this research, offering a rich laboratory of experiences.

Notably, the vast majority, more than 80%, of criminal cases in Mexico are matters of state jurisdiction. Within states, courts are organized into judicial districts, and these districts contain several municipalities. State authorities are in charge of implementation, but within each state the reform is implemented at level of judicial district, in a geographically gradual manner. In the time period studied here (2007-2012), sets of municipalities (i.e., a judicial district) within some states were implementing reform, while other municipalities in the same state were not implementing reform, and in other states all of the municipalities were not implementing reform. Among municipalities that were all implementing the reform, they might vary in a wide range of social, economic, or demographic features. This study leverages this high amount of information and variation across municipalities in Mexico to examine the impact of reform.

In this context, explaining the consequences of this reform is important because of the reform’s potential to reduce violence and insecurity and reshape public trust in institutions. Further, uneven institutions mean legal rights are distributed unequally (O’Donnell 2001), and people’s experience of justice and democratic life is very different from place to place.

Additionally, a wealth of resources has been expended promoting justice reforms in Latin America, and a large body of research has examined the patterns and causes of these reforms (Domingo and Sieder 2001; Hilbink 2007; Finkel 2008; Staton 2010; Ingram 2016a) . Between 2008-2018, Mexico invested no less than one billion USD to implement the current criminal procedure reform, a figure that does not include substantial additional material resources from states, as well as foreign and international sources (Novoa 2020), and existing research examines the patterns and sources of reform (Zwier and Barney 2012; Ingram 2016b). However, scholarship has generally neglected the assessment of the impact of justice reforms. This project helps to fill this

gap by examining two broader societal outcomes of reform: (a) public security, and (b) public trust in justice institutions.

Despite its scale and implications for a more democratic criminal justice system, criminal procedure no longer receives the attention, much less the funding, it once did. Further, throughout the reform process, advocates have encountered opponents of the reform, and opposition has remained consistent and increasingly relevant given the persistently high levels of violence and insecurity in Mexico. Opponents argue the reform is soft on crime, and advocate for more heavy-handed and repressive laws and practices associated with “punitive populism”, which has been a dominant current across the region ([Bailey and Dammert 2006](#); [Dammert 2019](#)). Opposition to the reform is also present within justice institutions, especially among prosecutors and older judges ([Ingram, Rodriguez Ferreira and Shirk 2011](#)).

The backlash to the reform generated several formal counterreform efforts. In late 2014, soon after approval of national code of criminal procedure and before federal courts had even begun implementation, legislative efforts already sought to reverse parts of the reform. This effort stalled after pro-reform mobilization, but another counterreform effort took its place in 2016. This was also thwarted by pro-reform advocates ([Negrete 2016](#)). However, in early 2020, the backlash was again underway. The 2020 effort was broader than prior efforts, covering multiple aspects of the reform, and seeking extended periods of preventive custody (including periods in excess of two years), application of preventive custody to all crimes, and even less judicial independence ([González Gómez 2020a](#); [Jimenez Padilla 2020](#); [Novoa 2020](#)).

Given the magnitude of the reform in Mexico and the resources and expectations invested in it, as well as the ongoing tension between proponents and opponents of the reform, one key area of research is on the impact of the reform, both in Mexico and elsewhere (e.g. [Tiede 2012](#); [Blanco 2016](#); [Huebert 2019](#)). If the impact is positive, then the effort was worthwhile and the opposition quieted; if the impact is mixed or negative, then the situation is more fraught.

Theoretical Framework and Working Hypotheses

The key anticipated effect is reform implementation had a helpful, beneficial impact on security and public trust. This expectation is examined while accounting for alternative explanations that emphasize the socio-economic and demographic drivers of insecurity and distrust.

The core expectation about the beneficial impact of reform implementation is derived from the motivations and design of the reform, as well as from existing research. As noted earlier, the reform was designed to address weaknesses in the criminal justice system, and was expected to increase competence and capacity across the board, from police to prisons. In doing so, the reform would generate improvements across several areas of institutional performance, including accessibility, efficiency, independence, and the overall quality of justice. These effects would have direct impact on institutional outcomes like case clearance rates, times to disposition, time in pretrial or preventive custody, etc. Further, if these improvements hold, then they lead to the logical-deductive expectation that there should be other, indirect effects in other areas, including security (Blanco 2016; Novoa 2020; Huebert 2019; Shirk 2011; Zwier and Barney 2012). To be more specific, if the overall performance of the justice system is increasing, including the detection, investigation, and prosecution of crime, then there should be reductions in crime, generating improved security. Also, the reform created new supports for witnesses, victims, and victims' families, and several features of the reform, prominently the new oral proceedings, added new levels of transparency (Ingram 2013; 2016b; Blanco 2016). These changes enhance accountability and reduce impunity, and should generate improved trust and legitimacy. Overall, reform implementation should reduce crime and violence, and increase public trust and confidence.

A limited amount of existing research on the impact of criminal procedure reform informs expectations. In Chile, Tiede (2012) found that reform reduced pretrial detention and also improved perceived security. In Mexico, Blanco (2016) – examining some of the same sources used here but focusing on the early reform period prior to 2009

– found that reform implementation both (a) reduced reported victimization and also (b) reduced perceived security. That is, actual improvement in lived security (lower victimization) was accompanied by more subjective insecurity, not less. Further, she found that reform lowered trust in local and federal police, but occasionally increased trust or confidence in prosecutors. Additionally, in a study focusing on areas in Mexico where organized crime was active, Huebert (2019) found that reform did not reduce homicides. Overall, this limited amount of empirical evidence on the impact of criminal procedure reform offers mixed results: real security improves (though not in areas of organized crime), but perceived security both improves (in Chile) and erodes (in Mexico). In sum, current research finds mixed effects of the reform on security and trust.

Taking these propositions and evidence together, the baseline theoretical expectation is that the reform will have beneficial effects. That is, communities that have implemented the reform will see improved security and trust in key justice institutions.

More concretely, one set of empirical implications is that communities that have implemented the reform will see reductions in objective measures of crime and violence compared to similar communities that have not implemented the reform. Five measures of objective security include homicide rates overall, homicide rates of men, homicide rates of women, total crime rate, and robbery rate. The section on data reports measurement details.

Another set of empirical implications anticipates that communities that have implemented the reform will see improvements in the subjective perception of security and trust compared to similar places that did not implement the reform. Four perception measures include perceived security, perceived trend in security, trust in prosecutors, and trust in judges. The expectations associated with security and trust are expressed in hypotheses 1a-1e and 2a-2d.

H1a: Reform implementation will vary negatively with aggregate homicide rate.

H1b: Reform implementation will vary negatively with homicide rate for men.

H1c: Reform implementation will vary negatively with aggregate homicide rate for women.

H1d: Reform implementation will vary negatively with total crime rate.

H1e: Reform implementation will vary negatively with robbery rate.

H2a: Reform implementation will vary positively with perceived security.

H2a: Reform implementation will vary positively with perceived security trend.

H2a: Reform implementation will vary positively with trust in prosecutors.

H2a: Reform implementation will vary positively with trust in judges.

Alternative explanations of crime and violence emphasize socio-economic and demographic drivers. Existing research, including research on Mexico, identifies income, inequality, family instability, education, and employment stability as key socio-economic factors, and population density and age distribution as key demographic factors (e.g., [Baller, Anselin, Messner, Deane and Hawkins 2001](#); [Rivera 2016](#); [Vilalta, Muggah and Fondevila 2020](#); [Vilalta, Lopez, Fondevila and Siordia 2020](#)). The same factors can influence public perception of security and trust in justice institutions, since security is associated with the legitimacy of public institutions and democracy ([Zechmeister and Lupu 2019](#)). For these reasons, the analysis accounts for per capita income, gini index, divorce rates, average years of education, proportion of the population not economically active, population density, and median age. Details on measurement and methods follow in next section.

Data and Methods

Data

The unit of analysis is the municipality and the focal time span is 2007-2012, though some explanatory variables are from 2005 due to data availability. There is a total

of 2455 municipalities in Mexico, but due to sample design, data availability, and matching process, the sample of matched observations across analyses varies from 14 to 1065 (see results).

The treatment variable, reform implementation, captures de facto implementation at the municipal level in the two-year time span of 2009-2010. Criminal procedure reform was implemented at the level of judicial districts within states in Mexico. Each of Mexico's 32 states contains multiple judicial districts, and each district consists of several municipalities. The treatment variable is coded according to whether the municipality was within a judicial district implementing the reform at any point in 2009-2010 (1=implemented, 0 otherwise).

The analysis examines nine outcomes of interest: (1) homicide rate (aggregate), (2) homicide rate for men (male victims only), (3) homicide rate for women (female victims only), (4) total crime rate (all reported crimes), (5) robbery rate (a high-frequency crime), (6) perceived security, (7) perceived trend in security, (8) trust in prosecutors, and (9) trust in judges. Each outcome reflects the two-year average in 2011-2012. Capturing outcomes over a two-year span after reform implementation gives the reform time to take effect and also reduces the risk of extreme or unusual values in any single year. Readers might ask whether measuring outcomes only two years after reform gives enough time for the reform to work. Setting a small window like this reduces the risk of confounders, similar to setting a small bandwidth in other design-based approaches (e.g., regression discontinuity). Further, setting a small window is a difficult test for the general expectation that reform has a meaningful effect; if any results are detectable in this small window, then I expect stronger results would appear over a larger window of time once the reform is more consolidated. I also generate the same two-year averages of each outcome for 2007-2008 (prior to reform) as baseline measures and controls (perception controls are measured in 2008-2009 due to data availability since same survey was not administered in 2007). Crime rates are expressed as rates per 100,000 people, a standard comparative measure.

All data on homicides, total crimes, robberies, and population for rate transfor-

mations are from the national statistics office, Instituto Nacional de Estadística, Geografía, e Información (INEGI), and their State and Municipal Database System (Sistema Estatal y Municipal de Bases de Datos, SIMBAD) (INEGI 2020). All other control variables are also from INEGI. Data on perceived security and trust come from two surveys: National Survey of Insecurity (Encuesta Nacional sobre Inseguridad, ENSI) and National Survey of Victimization and Perception of Public Safety (Encuesta Nacional de Victimization y Percepcion sobre Seguridad Publica, ENVIPE) (INEGI 2012-2013).

The data were divided into two main samples. The first sample (sample 1) comes from five of Mexico’s 32 states: (1) Baja California (BCN), (2) Durango (DUR), (3) Mexico State (MEX), (4) Morelos (MOR), and (5) Zacatecas (ZAC). These states were not the earliest reformers (like Nuevo León, Chihuahua, or Oaxaca), but they all completed the process of reform implementation in at least one judicial district in the state around the same time – in the two-year span of 2009-2010. The second sample (sample 2) adds municipalities from all other states (23) that had not yet implemented the reform. The core analysis is based on these two samples, matching municipalities first in sample 1 and the in sample 2, and comparing treated to control units within each sample. By focusing on reform implementation in 2009-2010 and its effects in 2011-2012, the analysis isolates a temporal window of reform and impact, maximizing control over temporal variation in the reform process and reducing influence of other potential confounders.

Thus, early reform periods from 2003-2008 in two states (Nuevo Leon and Chihuahua) are excluded. A third state, Oaxaca, saw reform implementation in two of its regions from 2007-2008, closer to the focal period of 2009-2010 examined here; the section on robustness checks adds municipalities from this state to test stability of results. Sample 3 consists of sample 1 plus municipalities from Oaxaca; sample 4 consists of sample 3 plus municipalities from the 23 non-reformed states.

Methods

Matching is “an increasingly popular method for preprocessing data to improve causal inferences in observational data” (King and Nielsen 2019, 435). Put briefly, matching is a “search for a data set that might have resulted from a randomized experiment but is hidden in an observational data set”; if this search is successful, many of the inferential problems of observational data “vanish” (King and Nielsen 2019, 435). There are several available matching techniques. Propensity score matching (PSM) is the most common. However, due to problems with PSM, King and Nielsen (2019) recommend coarsened exact matching (CEM).

CEM is preferable to unadjusted PSM because CEM “approximate[s] a fully blocked experimental design”, leading to better balance across treatment and control groups, lower model dependence, and lower bias (King and Nielsen 2019, 442). Also, given the mixed scales of measurement and the relatively small data set here – at least in first, smaller sample – CEM is advisable (King and Nielsen 2019, 449-450). Further, CEM is a more intuitive approach. Exact matching is the simplest approach; units are matched only if they share exactly the same value. However, this is rarely the case. CEM provides the next best solution, roughly approximating or coarsening the exact values into broader categories to approximate exact matching. All analyses were conducted in R (R Core Team 2019) using package *cem* (Iacus, M., King, Gary, Porro and Giuseppe 2018) and following Iacus, King and Porro (2009). All code is in replication materials.

Two issues were addressed prior to estimating causal effects: cutpoints and post-match imbalance. Cutpoints define how coarsely observations are matched on any given variable. A larger number of cutpoints can provide finer precision but can also limit the ability to match observations, especially in small or medium-size data sets. Rather than use default cutpoints (9, generating 8 bins for each variable), I defined my own (generally 6 cutpoints, establishing 5 bins). In some cases, redefining cutpoints made compelling substantive sense. For instance, with educational attainment, multiple cutpoints only 1-2 years apart does not make as much sense as cutpoints that

identify grade or schooling levels. In Mexico, these levels are primary, middle school (secundaria), and high school (preparatoria). No municipality had an average educational attainment higher than 12 years. Therefore, I defined only four cutpoints (0, 6.01, 9.01, and 12) to break data into three bins capturing primary, middle school, and high school. Aside from substantive reasons to redefine cutpoints, diagnostics showed that having a large number of cutpoints for some variables (fine coarsening) was limiting the ability to match observations (Iacus, King and Porro 2009). Population density, divorce rate, inequality, and baseline measures of outcomes of interest were frequent culprits that limited matching. I defined my own cutpoints for these variables. Cutpoints that are uniform across all analyses include education, age, population density, divorce rate, inequality, and unemployment. However, cutpoints for the baseline measure of the outcome of interest are specific to each outcome; in all of these cases, cutpoints created five bins. For objective crime rates, cutpoints differentiated primarily between units that had no or very little incidence of crime, several moderate amounts of crime, and very high levels. For subjective perception data, all measures were normalized between 0-100 to facilitate interpretation and equally spaced cutpoints created five bins. Details on cutpoints are in replication materials.

Turning to imbalance, there was still some imbalance among a subset of variables after matching, mainly population density, divorce rate, and the baseline measure of the outcome of interest. For this reason, I controlled for these variables in the post-matching estimation of causal effect.

Regarding estimation, the main results report the estimated average treatment effect (ATE) based on linear mixed-effects models that controls for imbalanced covariates. Generally, ATE after matching is the result of a bivariate linear regression of the outcome of interest (y) on the treatment variable, reform implementation (Tr [0,1]), i.e., $y \sim Tr$. In this simple estimation, the ATE is the coefficient on Tr , “which is the same as the difference in means, weighted by CEM stratum size” (Iacus et al. 2018, 3). However, if imbalance persists after matching, then this simple estimation is not sufficient. In the presence of such imbalance, the imbalanced covariates, e.g., x_1 and

x_2 , can be included in the model formula, i.e., $y \sim Tr + x_1 + x_2$. The main results report treatment effects controlling for imbalanced covariates.

Further, a basic linear estimation would assume homogeneous effects across all units. It is unrealistic to assume reform implementation has the same effect across all municipalities; therefore, the main models include fixed effects for all covariates plus both random intercepts and random coefficients for Tr (see replication materials and code in [Iacus et al. \(2018\)](#)). Mixed-effects models like these are appropriate for clustered or repeated measures across observations or over time; in this case, the matching stratum is the level at which observations are repeating. Each stratum consists of at least one treated and one control unit, but there are frequently more than one treated and/or more than one control unit per stratum. This approach captures a heterogeneous ATE rather than assuming a uniform effect across all units.

Results

Tables 1-4 report results. First, tables 1-2 report the results from sample 1 consisting of municipalities drawn from five core states (BCN, DUR, MEX, MOR, and ZAC). This is the sample with the greatest control, but has the smallest number of observations. Second, tables 3-4 report the results from sample 2 (sample 1 plus municipalities drawn from the 23 non-reformed states). Each table reports the estimated average treatment effect (ATE), standard error, t-statistic, and p-value, along with total N in sample and number of matched observations from this sample. The columns in the tables correspond with the nine outcomes of interest, objective first (hr=homicide rate, hrm=homicide rate for men, hrf=homicide rate for women, tcr=total crime rate, and rr=robbery rate) and subjective second (ps=perceived security, pst=perceived security trend, tp=trust in prosecutors, tj=trust in judges). The number of matched observations varies slightly across columns because the matching process relies on a baseline measure of each outcome of interest, which in turn depends on data availability for these measures.

Results from Main Sample (sample 1)

Table 1: ATE for CEM analysis, objective outcomes, sample 1

	hr	hr(male)	hr(fem)	tcr	rr
estimate	-9.060	-7.730	-3.810	-22.600	-0.258
se	1.280	2.100	0.678	2.820	0.183
t	-7.090	-3.670	-5.620	-8.030	-1.410
p	0	0.0002	0.00000	0	0.159
total obs	260	260	260	260	260
matched obs	63	62	71	74	67

Table 2: ATE for CEM analysis, subjective outcomes, sample 1

	ps	pst	tp	tj
estimate	16.316	-10.919	-0.190	-7.811
se	3.024	3.004	1.527	2.476
t	5.396	-3.634	-0.125	-3.155
p	0	0	0.901	0.002
total obs	150	150	148	140
matched obs	24	24	14	17

Table 1 shows that effect of reform is negative for all objective outcomes, and this effect is statistically significant for all outcomes except robbery rate (rr). Indeed, for first four outcomes, the effect is highly significant ($p < 0.001$). That is, reform implementation has a helpful, dampening impact on homicide rates, homicide rates of men, homicide rates of women, and total crime rate. The effect size is also substantial. Reform generates a reduction of nine homicides overall per 100,000 per year, and generates almost eight fewer homicides of men and almost four fewer homicides of women per 100,000 people per year. Reform also generates a reduction of more than 22 crimes per 100,000 people per year.

Table 2 reports more varied results for subjective outcomes. The effect is positive for perceived security, and this effect is statistically significant ($p < .001$). That is, reform has a helpful effect on perceived security. Further, the effect size is substantial; all of the perception variables are normalized on a 100-point scale (0-100), so the estimate of 16.32 means the reform generates more than a 16% increase in perceived security. In contrast, the effect is negative for perceived security trend, trust in prosecutors, and trust in judges. However, this effect is statistically significant only for perceived security trend ($p < 0.001$) and trust in judges ($p < .01$). Effect size is again meaningful. Reform generates almost an 11% drop in perceived security trend, and almost an 8% drop in trust in judges.

Results from Main Sample Plus Non-Reformed States (sample 2)

Table 3: ATE for CEM analysis, objective outcomes, sample 2

	hr	hrm	hrf	tcr	rr
estimate	-2.61	-2.036	0.509	-56.854	-0.546
se	0.543	0.827	0.188	1.995	0.142
t	-4.81	-2.462	2.706	-28.497	-3.842
p	0	0.014	0.007	0	0
total obs	1404	1404	1404	1,404	1,404
matched obs	362	299	373	373	334

Table 4: ATE for CEM analysis, subjective outcomes, sample 2

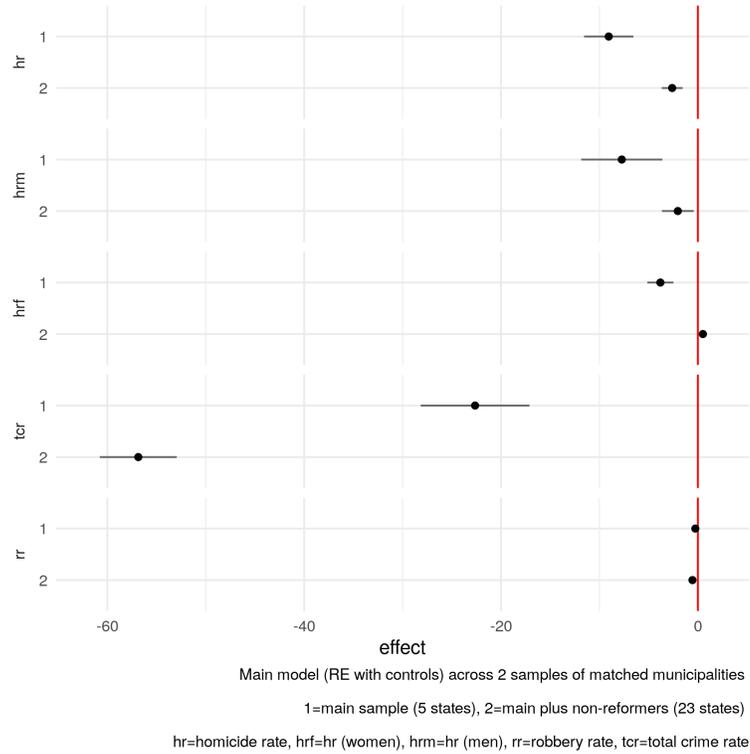
	ps	pst	tp	tj
estimate	-2.979	-8.995	-2.375	-1.095
se	1.192	1.136	0.761	0.832
t	-2.500	-7.918	-3.122	-1.317
p	0.012	0	0.002	0.188
total obs	604	604	600	581
matched obs	87	77	69	81

Turning to sample 2, table 3 shows results that are similar to table 1 with two exceptions: (1) the effect on homicides of women is now positive, and (2) the effect on robbery rate is now statistically significant. The effect on homicide rate overall and homicide rate for men is still negative and statistically significant, though the effect size is smaller (-2.61 and -2.04, respectively). The effect on total crime rate is still negative and highly significant, and the effect size is now greater; reform causes a reduction of almost 57 crimes per 100,000 people per year. Returning to the results that are different compared with table 1, the effect on homicides of women is positive and statistically significant, though the effect size is smaller (0.51). The effect on robbery rate is negative and statistically significant, and the magnitude of this effect is also relatively small (-0.55) compared to other effects on objective outcomes.

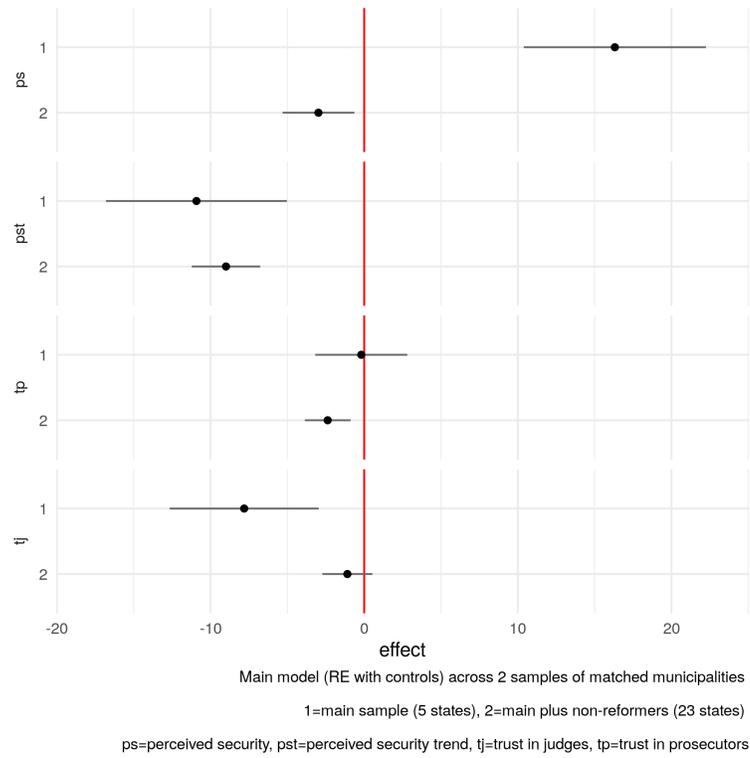
Table 4 reports mostly different results compared with analogous analysis in table 2. The only result that remains the same is the one for perceived security trend; the effect of reform is still negative and statistically significant, and the effect size is similar (-10.9 in table 2 and -9 in table 4). That is, reform generates a reduction of about 10% in perceived security trend. Otherwise, the effect on perceived security is now negative and statistically significant ($p < 0.05$), the effect on trust in prosecutors is still negative and now statistically significant ($p < 0.01$), and the effect on trust in judges is still negative but no longer significant.

Figure 2 visualizes these results. Effect size is on horizontal x-axis, and a vertical line marks 0 (no effect). Effects are organized by outcome, listed on vertical axis, and then by sample (1=sample 1, 2=sample 2). Filled circles identify the point estimate for effect size, and horizontal segments identify the 95% confidence interval. Panel 2a shows that all effects are negative (left of 0 line) for objective outcomes with only one exception: homicides of women (hrf) in second sample. Also, we can quickly see that the confidence interval does not include 0 for all estimates of homicide rate, homicides rate of men, and total crime rate. Panel 2b offers visual summary of results for subjective outcomes. Results are most stable for perceived security trend, with both effect size and confidence interval to left of 0 line. For other outcomes, direction

of effect switches for perceived security and confidence interval varies in whether it includes or excludes 0 for both trust in prosecutors and trust in judges.



(a) Effect of reform across 5 objective outcomes.



(b) Effect of reform across 4 subjective outcomes.

Figure 2: Effects of reform across two main samples.

Robustness Checks

I report two sets of robustness checks. First, I analyze two expanded samples: (1) sample 3, which is sample 1 plus municipalities drawn from the state of Oaxaca (municipalities from total of six states), and (2) sample 4, which is sample 3 plus all non-reformed states (drawing municipalities from total of 24 states). Second, I re-analyze all four samples (samples 1-4) with all controls as above, plus a control for state of origin. This specification accounts for unobserved state-level factors that might influence results.

Expanded Samples with Oaxaca

Table 5: ATE for CEM analysis, objective outcomes, sample 3

	hr	hrm	hrf	tcr	rr
estimate	-2.144	-9.962	-2.146	-6.162	-0.516
se	0.325	0.742	0.172	1.302	0.068
t	-6.596	-13.432	-12.482	-4.733	-7.582
p	0	0	0	0	0
total obs	830	830	830	830	830
matched obs	390	466	558	468	542

Table 6: ATE for CEM analysis, subjective outcomes, sample 3

	ps	pst	tp	tj
estimate	7.618	-8.242	-5.459	-5.844
se	2.073	1.922	1.247	1.153
t	3.674	-4.288	-4.378	-5.068
p	0.0002	0.00002	0.00001	0.00000
total obs	210	210	208	194
matched obs	39	36	39	33

Table 7: ATE for CEM analysis, objective outcomes, sample 4

	hr	hrm	hrf	tcr	rr
estimate	-3.407	-7.319	-1.063	-16.720	-0.632
se	0.266	0.507	0.109	1.188	0.033
t	-12.809	-14.445	-9.794	-14.077	-19.120
p	0	0	0	0	0
total obs	1974	1974	1974	1,974	1,974
matched obs	984	923	1065	871	974

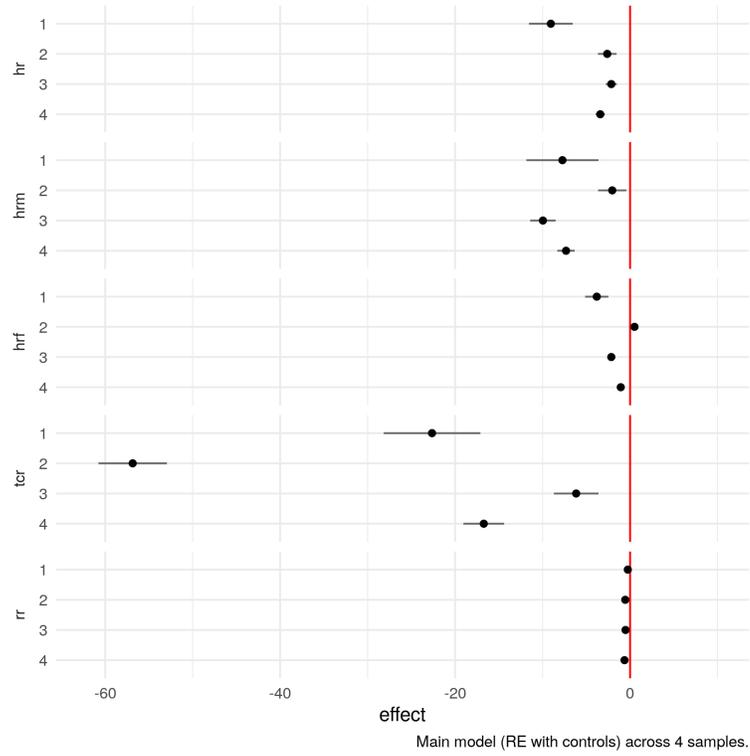
Table 8: ATE for CEM analysis, subjective outcomes, sample 4

	ps	pst	tp	tj
estimate	-1.561	-7.927	-5.335	-0.831
se	0.963	0.758	0.587	0.732
t	-1.622	-10.454	-9.094	-1.136
p	0.105	0	0	0.256
total obs	664	664	660	635
matched obs	155	200	192	178

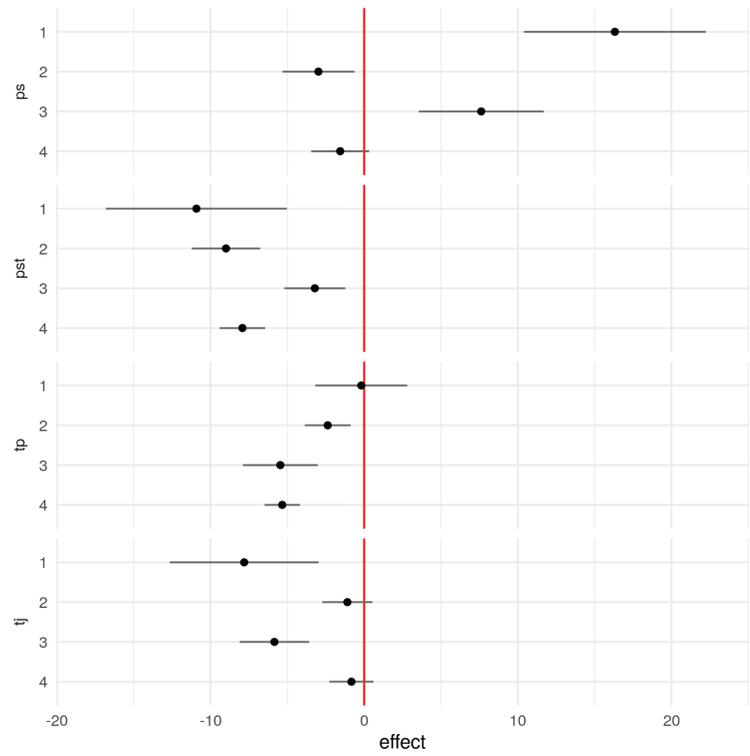
Tables 5-8 summarize CEM results from analysis of samples 3 and 4, adding state of Oaxaca and then adding all non-reformed states, respectively. Following same structure above, tables 5 and 7 report results for objective outcomes. Across the board for objective outcomes, effects are negative and highly statistically significant ($p < 0.001$). Effect sizes range from -2.1 to -3.4 for homicide rate overall, -7.3 to -10 for homicide rate for men, and -1 to -2.1 for homicide rate for women, -17 to -6 for total crime rate, and -0.5 to -0.6 for robbery rate. Put simply, reform has a uniformly helpful effect on objective security.

Tables 6 and 8 report results for subjective outcomes. Effects are again mixed. Among most consistent results, effect is uniformly negative and highly statistically significant for perceived security trend and trust in prosecutors. Effect size on perceived security trend varies from -7 to -8, and for trust in prosecutors is consistently around -5. Among more mixed results, effect is positive (7.6) and statistically significant for perceived security in sample 3, but is negative and not statistically significant in sample

4. For trust in judges, effect is negative (-5.8) and statistically significant in sample 3, and negative (-0.8) but not statistically significant in sample 4. Figure ?? visualizes results for samples 3-4.



(a) Effect of reform across 5 objective outcomes.

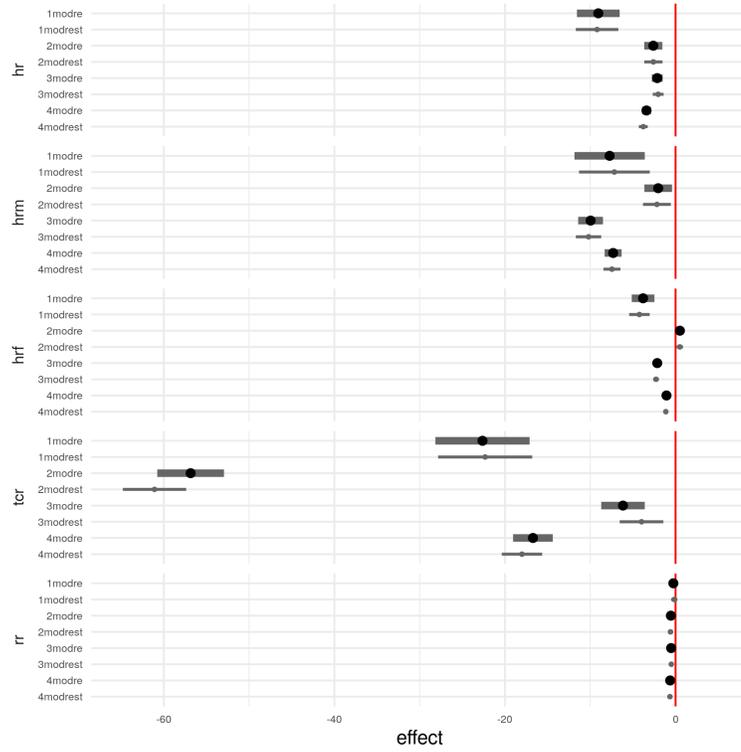


(b) Effect of reform across 4 subjective outcomes.

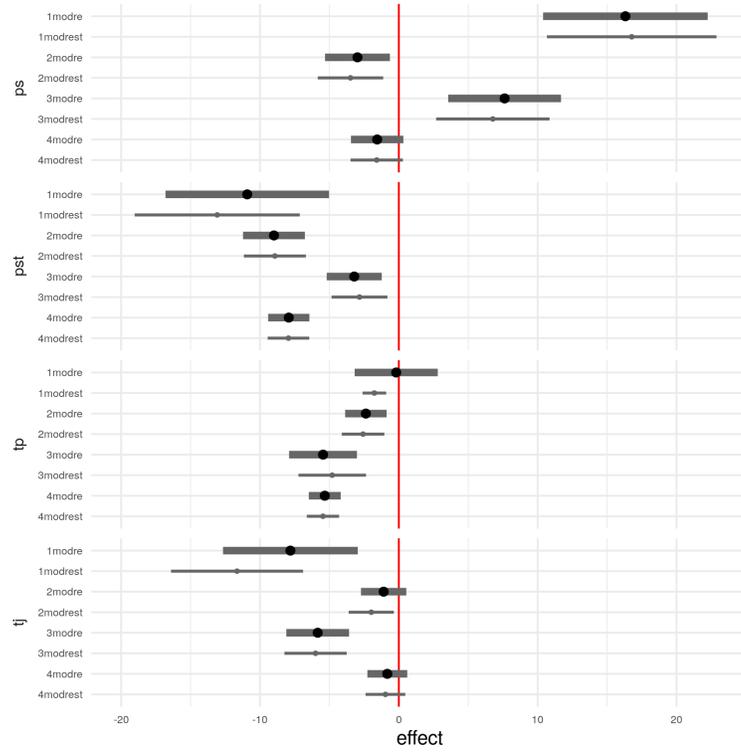
Figure 3: Effects of reform, expanded sample.

Controlling for State of Origin

The main results report treatment effects based on units matched across states, but without controlling for higher-level, state-specific influences. While matching process should account for most relevant factors, it is possible that state-level dynamics influence outcomes of interest. In this stability check, the analysis adds a control for state of origin to the estimation process. Figure 4 reports results. Along vertical axis, labels identify the sample and model grouped by outcome. For instance, “1modre” identifies the main results from sample 1, and “1modrest” identifies the same model but now controlling for state. The original models for each outcome have thicker dots and confidence bands, and models with state control have smaller dots and thinner bands. Overall, the results remain remarkably stable.



(a) Effect of reform across 5 objective outcomes.



(b) Effect of reform across 4 subjective outcomes.

Figure 4: Effects of reform, expanded sample.

Discussion

Table 9: Summary of range of effects based on 95% confidence intervals

hr	hrm	hrf	tcr	rr	ps	pst	tp	tj
-11.6, -6.6	-11.9, -3.6	-5.1, -2.5	-28.2, -17.1	-0.6, 0.1	10.4, 22.2	-16.8, -5	-3.2, 2.8	-12.7, -3
-11.7, -6.7	-11.3, -3	-5.4, -3	-27.8, -16.8	-0.5, 0.2	10.7, 22.9	-19, -7.1	-2.6, -0.9	-16.4, -6.9
-3.7, -1.5	-3.7, -0.4	0.1, 0.9	-60.8, -52.9	-0.8, -0.3	-5.3, -0.6	-11.2, -6.8	-3.9, -0.9	-2.7, 0.5
-3.7, -1.5	-3.8, -0.6	0.1, 0.9	-64.8, -57.4	-0.9, -0.3	-5.8, -1.1	-11.2, -6.7	-4.1, -1	-3.6, -0.4
-2.8, -1.5	-11.4, -8.5	-2.5, -1.8	-8.7, -3.6	-0.6, -0.4	3.6, 11.7	-5.2, -1.2	-7.9, -3	-8.1, -3.6
-2.7, -1.4	-11.7, -8.7	-2.6, -1.9	-6.5, -1.4	-0.6, -0.4	2.7, 10.9	-4.8, -0.8	-7.2, -2.4	-8.2, -3.8
-3.9, -2.9	-8.3, -6.3	-1.3, -0.8	-19, -14.4	-0.7, -0.6	-3.4, 0.3	-9.4, -6.4	-6.5, -4.2	-2.3, 0.6
-4.3, -3.3	-8.5, -6.5	-1.4, -0.9	-20.4, -15.6	-0.7, -0.6	-3.5, 0.3	-9.4, -6.5	-6.6, -4.3	-2.4, 0.5

Recapping analysis, figure 4 shows which results are most stable and also visualizes the range of effect size for each outcome of interest. Table 9 complements that figure by summarizing range of effect sizes for each outcomes across all models (this range is based on lowest and highest bounds of 95% confidence intervals, not on point estimates). Overall, reform implementation has a nearly uniform, consistently helpful, beneficial effect on objective measures of security. This relationship holds across all but two models in the analysis (the two exceptions are the results for homicide rate for women in sample 2 and for robbery rate in sample 1). Reform has a more mixed but generally negative, harmful effect on subjective perception of security and trust.

Focusing first on objective security, the magnitude of the helpful effect on homicide rate overall ranges from -1.4 to -11.7 (based on upper and lower bounds of confidence intervals, not point estimates), i.e., the reform generates about 1-12 fewer homicides per 100,000 people per year. The effect range for homicide rates of men was similar (-0.4 to -11.9), and the effect range for homicides of women was 0.9 to -5.4 (the range includes 0 so no effect is also possible). These effects are substantial. Mexico has more than 100 cities each with more than 100,000 people. For every 100,000 people, reform has the identified effect per year, so a city with 200,000 people would have two times the estimated effect per year, a city with 400,000 people would have four times the estimated effect per year, and so on. Mexico also has more than 10 cities each

with more than one million people, and the nation's capital, Mexico City, has more than 10 million. The most guarded interpretation is that if each of these cities with at least one million people implemented the reform, they could on average expect at least 14 (1.4 x 10) fewer homicides per year, 4 fewer homicides of men, and possibly an increase of 9 homicides of women. On the more optimistic end, the results mean that if these cities implemented the reform they would see around 117 fewer homicides overall, disaggregated into around 119 fewer homicides of men and about 54 fewer homicides of women. Although the worst scenario includes the possibility of an increase in lethal violence against women, this possibility is offset by the larger decrease in homicides overall, for a net reduction in lethal violence. To be sure, this does not mean that even one homicide is tolerable or acceptable. This is especially true in a country like Mexico that has an existing problem with femicide. One potential policy solution to address possible increase in homicides of women is to pair this reform with additional investments in women-centered programs, including efforts to combat violence against women.

For other objective outcomes, reform effects for total crime rate are uniformly helpful and range from -3.6 to -64.8, and for robbery rate range from 0.2 to -0.6 (this includes 0, or no effect). Recalling Mexico's 10 cities with more than one million people, this means that if these cities implemented the reform they should conservatively expect at least 36 fewer crimes per year, though perhaps an increase of 20 in specific category of robberies. On more optimistic end, these cities could expect to see more than 600 fewer crimes per year, including 60 fewer robberies. Although the worst scenario includes the possibility of an increase in robberies, this potential outcome is offset by the greater decrease in crime overall, for a net reduction in crime.

Given persistently high and increasing violence and insecurity in Mexico over last decade, these results offer mostly good news. To put these numbers in broader context, the national homicide rate has been rising from around 10 throughout most of 2000s, to 22 in 2010, to 28 in 2019 ([WorldBank 2020](#); [Rivera 2020](#)). This increase has not been monotonic, but 2019 was the fifth consecutive year with an increase and set a

record with around 35,000 intentional homicides. Any reduction would be welcome, and a reduction of 6 per 100,000 (only half the maximum of 12 in ranges discussed earlier) would put homicide rate back at what it was around 2010. In short, the results show the reform had a helpful, dampening effect on lethal violence and crime generally in municipalities within judicial districts that implemented the reform. This evidence supports deepening and consolidating the reform.

Turning to subjective, perception-based measures of security and trust, the most stable result is the one for perceived security trend. The effect of reform on *pst* is uniformly harmful, and this effect ranges from -0.8 to -19. That is, reform causes anywhere from around a 1% to a 13% decrease in perceived security trend, i.e., that security is deteriorating over time. The second most stable result is that the reform has a harmful effect on trust in prosecutors. In all but the first model (which had no control for state), the reform has a harmful effect, so this result is consistent. Also, the magnitude of this effect increases from sample 1 to sample 4, and ranges from 2.8 to -7.9. That is, reform may possibly increase trust in prosecutors as much as 2.8%, or may have no effect, but the more common outcome suggested by analysis is that reform generates a reduction in trust in prosecutors – as much as an 8% reduction. Reform also has a mostly harmful effect on trust in judges. This result is not as stable since the result is not significant in three of eight models in figure 4 and summarized in table 9. The effect ranges from 0.6 to -16.4. That is, in best scenario reform has a minor helpful effect on trust in judges, raising trust by less than 1%, and reform might have no effect at all. But the likeliest outcome suggested by results is that reform has a harmful effect on trust in judges, reducing this trust by as much as 16%. Lastly, the least consistent result is for perceived security. The effect of reform on perceived security ranges from 22.9 to -5.8. Half the models suggest a large, helpful, positive effect, but half the models suggest no effect or a moderate harmful, negative effect. If the reform is helpful, this effect ranges from 2.7 - 22.9; that is, the reform might increase perceived security anywhere from 2% to 23%. If the reform is harmful, it decreases perceived security by as much as 6%. A guarded interpretation is that the

reform has no effect or perhaps even a harmful effect.

The results about subjective perception are puzzling and preoccupying because the previous evidence shows security is objectively improving after reform, yet the perception is that security and trust in key justice actors is deteriorating. That is, even if a key goal of improving security is objectively and empirically met – subjective community perception worsens. These results are striking in that they match the same objective-security-plus-subjective insecurity result also found in Mexico by (?) using a different approach, relying on different data, and examining a different time period.

Implications

The tension between objective security and subjective insecurity presents a major challenge to policymakers and supporters of the reform in Mexico. On one hand, the reform had the desired effect of reducing violence and criminality and objectively improving security. These results provide clear support for deepening and consolidating the reform. On the other hand, reform was accompanied by the subjective perception of worsening security and trust. This feeds punitive populism and other voices opposing the reform, and could give energy to efforts to reverse the reform and move towards more authoritarian practices outlined in earlier sections (e.g., lengthy periods of preventive custody, mandatory preventive custody for most or all crimes, fewer controls on probative value of evidence, fewer checks and balances among judges, broad discretion for investigatory detentions absent clear legal grounds similar to “reasonable suspicion” in U.S.). To avoid this reversal and continue to build and maintain more democratic institutions, and specifically a more democratic system of criminal justice institutions, this tension must be resolved.

One explanation for increased dissatisfaction after reform relates to the theory of relative deprivation [Gurr \(1970\)](#); [Gurney and Tierney \(1982\)](#). That is, even if someone is well off, if they expected to be more well off, then relative to those heightened expectations they may feel aggrieved. Applied to criminal justice reform in contexts of

persistently high violence and insecurity like Mexico, this expectation-achievement gap is likely largest in places where a high-profile reform is implemented, promising a wide range of improvements. Any implementation is likely to be imperfect and fall short, at least in near term, and then require ongoing effort and consolidation. However, falling short leads to an expectation-achievement gap, which then leads to the public being doubly disappointed, even if the community is objectively better off than comparable communities without the reform. This possibility raises compelling dynamics about the cognitive-psychological dimensions of the reform process and the long-term stability of reform. That is, even if the reform is objectively successful, cognitive processes might still shape negative perceptions that undermine the reform.

This explanation is likely since we know that the reform was (and continues to be) implemented incompletely and unevenly. That is, even in jurisdictions where the reform was functionally operational, some practices persisted from the pre-reform past. Even in early 2020, some of the problems identified by actors internal to justice institutions include: (1) police detain individuals for one crime but then charge a different, unrelated crime without sustaining legality of original detention ([González Gómez 2020b](#)); (2) judges allow investigatory steps to enter court record as proof of facts such that the investigation itself is evidence of guilt, rather than requiring an examination of relevant evidence to clarify what evidence is admissible ([González Gómez 2020a](#); [Cataño González 2020a](#)); (3) prosecutors tend to focus on punitive goals (conviction and punishment), and neglect victims and witnesses, including restitution ([Cataño González 2020b](#)); and (4) prosecutors close too many cases without taking any action (60% at state level and nearly 40% at federal level), contributing to impunity ([Novoa 2020](#)). Solutions that can help deepen and consolidate the reform include renewed attention to the reform, renewed investment in the reform, more coordination across federal and state levels of government, continued training and education, and more systematic data collection and evaluation ([Cataño González 2020b;a](#); [Novoa 2020](#)). These steps could help reduce the expectation-achievement gap. However, even if this gap exists, another solution is to prioritize outreach and public communication

strategies about the nature and challenges of the reform in order to keep expectations realistic and educate and inform the public about the reform process.

Another possible explanation for the contrasting results between objective security and subjective security and trust is that communities that have implemented the reform have also seen an increase in the discussion of the reform and related criminal justice issues. That is, implementing the reform makes criminal justice issues more salient, and this salience creates openings for new actors to influence public opinion and perception. In these windows where opinion is more open to influence, vocal critics of the reform might have an outsized negative impact on perception of the reform. Again, one potential solution is to strengthen outreach and public communication about the reform, as mentioned earlier.

Lastly, some readers might wonder about a more methodological explanation: might the results be influenced by non-reporting of crime? That is, low trust in justice institutions can dissuade victims from reporting crimes to authorities; therefore, seeing that reform is associated with both lower trust and lower crime, perhaps the finding that reform is associated with lower crime reports is a function of the fact that people do not trust the authorities and therefore report fewer crimes. The research design and the stability of the results across a range of measures strongly suggest this is not the case. First, municipalities are matched on a range of relevant covariates identified earlier, including baseline measures of the outcomes of interest. For instance, if the outcome is the two-year total crime rate in 2011-2012, units are matched on the two-year average of total crime rate in 2007-2008. If there was non-reporting in the matched municipalities, it should be accounted for in the baseline crime rates that motivate the matching. Second, robbery and total crime might be vulnerable to measurement error due to non-reporting; however, homicide is widely regarded as a measure that is not subject to non-reporting, in part because there is usually good evidence associated with the crime – a body. The stability of the helpful effect of reform across different measures of objective security therefore supports concluding the results are not driven by non-reporting. Further, an auxiliary model (not reported here) for total crime rate

in sample 2 that also matched units on trust in prosecutors generated results that were virtually indistinguishable from those in table 3 (i.e., treatment effect of -57 in that table and -59 in auxiliary analysis); if anything, adding the trust covariate increased the effect size. In sum, several reasons justify confidence in results.

Conclusion

In sum, the implementation of criminal procedure reform has a consistently helpful effect on measures of objective security, but has a generally harmful effect on subjective measures of security and trust in key justice institutions. The discussion raises several implications related to this tension between objective security and subjective insecurity and distrust generated by the reform, including the need to deepen and consolidate the reform to avoid incomplete or uneven institutional change, as well the need to prioritize outreach and public communication strategies to educate and inform the public about the reform process, and to keep expectations realistic.

It bears emphasizing that reform is normatively desirable beyond its positive impact on objective security. The main goal of reform was not to reduce criminality or improve security. Rather, these were downstream effects expected from broader, fundamental improvement in due process and the accessibility, efficiency, independence, and overall quality of the justice system. Even if the reform had temporary harmful effects on security, it would still be worth doing in order to secure broader, systemic improvements over time.

However, one of the main arguments of critics of the reform is that the reform is soft on crime and making people less safe. Based on that proposition, opponents advocated “punitive populism” (Bailey and Dammert 2006; Dammert 2019) – harsher, more heavy-handed punishments and repressive strategies. In Mexico, these positions advocate extended preventive custody (even beyond two years), preventive custody for all crimes, expansive powers of investigation and detention, less transparency, and less independence among judges. If adopted, these measures would push Mexico’s criminal

justice system backwards and in a decidedly authoritarian direction.

The results here strongly counter the arguments of opponents to the reform. Even though security was not a primary, direct goal of the reform, reform implementation had a helpful impact across a range of objective measures of security, including lethal violence (homicides), total crime rate, and the high-frequency crime of robbery. This is strong support for consolidating the reform. There are harmful effects on subjective perceptions of security and trust; as noted, the discussion identified strategies for addressing these. In any case, the positive news about helpful effects should be good news for a country suffering from persistently high levels of violence and insecurity.

Limitations of this study relate to time span and level of analysis. First, this study focuses on a small period of implementation from 2009-2010. Reform implementation that took place later is excluded. Thus, work that expands the temporal scope of this work can address this limitation, updating the analysis and relying on more recent data. Second, this study examines outcomes of interest only within two years (2011-2012) of reform implementation. This might not allow sufficient time for full effects to manifest; however, there is a tradeoff between expanding the window of time until outcomes are measured and the quasi-experimental control of the matching design. Measuring outcomes soon after reform maximizes analytic leverage, but might miss the longer-term effects of the reform. Future studies can try to examine longer-term effects or repeat the study at different intervals. Third, this study emphasizes the municipal-level effects of state-level reform implemented in judicial districts, excluding the role of advancing harmonization and implementation at the federal level. Work that more fully examines the multi-level patterns of reform and its impact across municipalities, judicial districts, states, and federal districts promises valuable contributions and a better understanding of the system-wide process of reform and its consequences. These and related lines of research promise to enhance our understanding of how to build and maintain democratic institutions.

References

- Bailey, John and Lucia Dammert. 2006. *Public Security and Police Reform in the Americas*. Pittsburgh, PA: University of Pittsburgh.
- Baller, Robert D., Luc Anselin, Steven F. Messner, Glenn Deane and Darnell F. Hawkins. 2001. "Structural Covariates of U.S. County Homicide Rates: Incorporating Spatial Effects*." *Criminology* 39(3):561–588.
- Blanco, Luisa. 2016. "The impact of judicial reform on crime victimization and trust in institutions in Mexico." *Violence and victims* 31(1):27–50.
- Cataño González, Verónica. 2020a. "Panel de discusión: ¿Hacia adelante o hacia atrás? El futuro del sistema penal acusatorio." Symposium Internacional OASIS, Jun. 5.
- Cataño González, Verónica. 2020b. "¿Qué tan necesaria es una reforma?" Symposium Internacional OASIS, Jun. 4.
- Dammert, Lucia. 2019. Challenges of police reform in Latin America. In *Routledge Handbook of Law and Society in Latin America*, ed. Rachel Sieder, Karina Ansolabeher and Tatiana Alfonso. New York: Routledge.
- Domingo, Pilar and Rachel Sieder, eds. 2001. *Rule of Law in Latin America: The International Promotion of Judicial Reform*. London: Institute of Latin American Studies, University of London.
- Finkel, Jodi. 2008. *Judicial Reform as Political Insurance*. Notre Dame, IN: University of Notre Dame Press.
- González Gómez, Alejandro. 2020a. "Dando voz a los Estados." Symposium Internacional OASIS, Jun. 3.
- González Gómez, Alejandro. 2020b. "Panel de discusión: ¿Hacia adelante o hacia atrás? El futuro del sistema penal acusatorio." Symposium Internacional OASIS, Jun. 5.

- Gurney, Joan Neff and Kathleen J. Tierney. 1982. "Relative Deprivation and Social Movements: A Critical Look at Twenty Years of Theory and Research." *The Sociological Quarterly* 23(1):33–47.
- Gurr, Ted Robert. 1970. *Why Men Rebel?* Princeton, NJ: Princeton University Press.
- Hilbink, Lisa. 2007. *Judges beyond Politics in Democracy and Dictatorship: Lessons from Chile*. New York: Cambridge University Press.
- Huebert, Erin Terese. 2019. "Criminal procedure reform and the impact on homicide: evidence from Mexico." *Trends in Organized Crime* . online first.
- Iacus, Stefano, Gary King and Giuseppe Porro. 2009. "cem: Software for Coarsened Exact Matching." *Journal of Statistical Software, Articles* 30(9):1–27. version April 2018.
URL: <https://www.jstatsoft.org/v030/i09>
- Iacus, Stefano M., King, Gary, Porro and Giuseppe. 2018. *cem: Coarsened Exact Matching*. R package version 1.1.19.
URL: <https://CRAN.R-project.org/package=cem>
- INEGI. 2012-2013. "Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública (ENVIPE)". Instituto Nacional de Estadística, Geografía, e Información.
URL: <http://www.inegi.org.mx/programas/envipe/>
- INEGI. 2020. "Sistema Estatal y Municipal de Bases de Datos (SIMBAD)". Instituto Nacional de Estadística, Geografía, e Información.
URL: <https://sc.inegi.org.mx/cobdem/index.jsp>
- Ingram, Matthew C. 2013. "Criminal Procedure Reform in Mexico: Where Things Stand Now". Mexico Institute.
- Ingram, Matthew C. 2016a. *Crafting Courts in New Democracies: The Politics of Subnational Judicial Reform in Brazil and Mexico*. New York: Cambridge University Press.

- Ingram, Matthew C. 2016b. "Mandates, Geography, or Networks? Diffusion of Criminal Procedure Reform in Mexico." *Latin American Politics and Society* 58(1):121–145.
- Ingram, Matthew C., Octavio Rodriguez Ferreira and David A. Shirk. 2011. "Justi-ciabarómetro."
- Jimenez Padilla, Alejandro. 2020. "¿La reforma que viene? Buscar la paz con un proceso penal autoritario." *Nexos* . January 20 (accessed May 15, 2020).
URL: <https://eljuegodelacorte.nexos.com.mx/?p=10894>
- King, Gary and Richard Nielsen. 2019. "Why Propensity Scores Should Not Be Used for Matching." *Political Analysis* 27(4):435–454.
- Langer, Máximo. 2007. "Revolution in Latin American Criminal Procedure: Diffusion of Legal Ideas from the Periphery." *American Journal of Comparative Law* 55:617–676.
- Negrete, Layda. 2016. "Contrarreforma, muerta o en pausa?". June 6 (accessed May 15, 2020).
URL: <https://www.mexicoevalua.org/contrarreforma-penal-muerta-o-en-pausa/>
- Novoa, Maria. 2020. "Los vaivenes en el sistema de justicia penal en México." Simpo-sium Internacional OASIS, Jun. 2.
- O'Donnell, Guillermo. 2001. "Democracy, Law, and Comparative Politics." *Studies in Comparative and International Development* 36:5–36.
- R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
URL: <https://www.R-project.org/>
- Rivera, Mauricio. 2016. "The sources of social violence in Latin America: An empirical analysis of homicide rates, 1980–2010." *Journal of Peace Research* 53(1):84–99.

- Rivera, Ruben. 2020. "México registró 34,582 asesinatos durante 2019, la cifra más alta en 20 años." Univision (Jan. 21).
URL: <https://www.univision.com/noticias/america-latina/mexico-registro-34-582-asesinatos-durante-2019-la-cifra-mas-alta-en-20-anos>
- Shirk, David A. 2011. "Criminal justice reform in Mexico: An overview." *Mexican Law Review* 3:189–228.
- Staton, Jeffrey K. 2010. *Judicial Power and Strategic Communication in Mexico*. Cambridge, UK: Cambridge University Press.
- Tiede, Lydia Brashear. 2012. "Chile's Criminal Law Reform: Enhancing Defendants' Rights and Citizen Security." *Latin American Politics and Society* 54(3):65–93.
- Vilalta, Carlos J., Pablo Lopez, Gustavo Fondevila and Oscar Siordia. 2020. "Testing Broken Windows Theory in Mexico City." *Social Science Quarterly* 101(2):558–572.
- Vilalta, Carlos J., Robert Muggah and Gustavo Fondevila. 2020. "Homicide as a function of city block layout: Mexico City as case study." *Global Crime* 0(0):1–19.
- WorldBank. 2020. "Intentional homicides (per 100,000 people) - Mexico." Based on UN Office on Drugs and Crime's International Homicide Statistics database (accessed June 23, 2020).
URL: <https://data.worldbank.org/indicator/VC.IHR.PSRC.P5?locations=MX>
- Zechmeister, Elizabeth and Noam Lupu. 2019. "The Pulse of Democracy in the Americas: Results from the 2019 AmericasBarometer Study." accessed June 3, 2020.
URL: <https://www.vanderbilt.edu/lapop/ab2018/>
- Zwier, Paul J. and Alexander Barney. 2012. "Moving to an Oral Adversarial System in Mexico: Jurisprudential, Criminal Procedure, Evidence Law, and Trial Advocacy Implications." *Emory International Law Review* 26:189–225.