Getting a Hand By Cutting Them Off: Drug Violence, Incumbency, and Uncertainty over Corruption in Mexico*

Paul Zachary† William Spaniel‡

December 21, 2014

Abstract

What role do politicians have in bargaining with violent non-state actors to determine the level of violence in their districts? Although some studies address this question in the context of civil war, it is unclear whether their findings generalize to organizations that do not want to overthrow the state. Unlike political actors, criminal groups monopolize markets by using violence to eliminate rivals. In the context of the Mexican Drug War, we argue that increased time in office increases cartels' knowledge about local political elites' willingness to accept bribes. With bribes accepted and levels of police enforcement low, cartels endogenously ratchet up levels of violence because its marginal value is greater under these conditions. We formalize our claims with a model and then test its implications with a novel dataset on violent incidents and political tenure in Mexico. We find that each additional year after an official initially takes office countrywide is associated with an additional 2,300 homicides.

*Previous version of this paper were presented at the Midwest Political Science Association, International Studies Association, and Western Political Science Association conferences. We thank Sergio Ascencio, Regina Bateson, Peter Bils, Kevin Clarke, James Fowler, Matthew Gichohi, David Lake, Alexander Lee, Benjamin Laughlin, Ben Lessing, John Patty, MaryClare Roche, Phil Roeder, Arturas Rozenas, Brad Smith, and Barbara Walter for their feedback. We are indebted to Andrea de Barros and Jacob Boeri for excellent research assistance. All errors remain the authors' own.

†PhD Student, Department of Political Science, University of California, San Diego, 9500 Gilman Drive, #0521, La Jolla, CA 92093 (pzachary@ucsd.edu)

‡PhD Candidate, Department of Political Science, University of Rochester, Harkness Hall 333, Rochester, NY 14627 (williamspaniel@gmail.com, http://williamspaniel.com).

1 Introduction

The use of violence as a means of establishing territorial control and generating rents is an increasingly prominent topic in the study of civil war and political violence. Beyond coercing states into making concessions and providing information about resolve, violence has a number of strategic benefits for armed groups (Pape 2003; Cohen 2014; Weinstein 2007). One strategic benefit is preventing civilian defection and maintaining territorial control (Kalyvas 1999). Harming civilians can serve as a means of forcing them to contribute rents to armed groups (Weinstien 2007; Humphreys and Weinstein 2006), improve unit cohesion and *esprit de corps* (Cohen 2014), and help armies maintain control of conquered territory (Downes 2007, 2008).

While these prior studies convincingly demonstrate the strategic logic behind the micro-dynamics of violence, they generally focus on periods of armed contestation between armed groups over control of the state (Sambanis 2004; Singer and Small 1982). With few exceptions in the literature, civil wars are coded as occurring when the government and one or more additional actor use violence to contest the legitimate control of the state. This definition excludes a number of ongoing conflicts, where armed groups are not currently attempting to assume control of the state or its institutions (Staniland 2010, 2012; Bateson 2012; Lessing 2014; Acemoglu, et al. 2009; Collier 2000). Examples of such violence include the "drug wars" in Mexico and Colombia and gang violence in Brazil, El Salvador, and Guatemala (Rios 2012; Reno 2000). In these cases, cartels and criminal organizations use violence to seize territory from rivals and force civilians to provide information about rivals' activities (Langton 2012; Kan 2012). It is currently unknown whether and how institutions affect the level and geographic dispersion of violence during these conflicts. To begin to address this question, we focus on the role of political leadership in cooperating with cartels and other criminal insurgents. We ask what role politicians have in bargaining with cartels and how they determine the level of violence in their district.

To answer our question, we develop a formal model of drug violence and police enforcement. In it, a local cartel attempts to reduce that enforcement by offering a bribe the local political party. The party then weighs its desire to minimize violence against bribery's economic benefit. Finally, the local cartel uses violence—endogenously determined—to maintain control of valuable territory against a rival cartel.

Uncertainty plays a critical role in determining the outcome of the interaction. When the cartel knows the politician's level of corruption, it can choose a precise bribe and ensure that enforcement will be lax. Consequently, levels of violence rise. In contrast, when the cartel faces great uncertainty about the corruptibility of the politician, it may offer small bribes that risk having the politician reject. This time, we expect levels of violence to be comparatively lower because the politician is more likely to enforce the laws. Thus, counter to standard models of costly conflict, we expect uncertainty to *decrease* levels of violence.

Drawing from recent theoretical and empirical conceptualizations of uncertainty in international relations (Wolford 2007; Rider 2013; Spaniel and Smith forthcoming), we then argue that regions where local political machines that have more recently taken control will have lower levels of violence. This might seem counterintuitive because experience seemingly should increase skill and thus decrease violence. However, in general, this literature argues that parties know less about leaders' preferences earlier in their tenure. Applying this to the Mexican Drug War, cartels facing greater uncertainty are more likely to see their offers fail, leading to properly enforced laws and less violence. As tenure progresses, though, the cartels can better narrow their suppositions about leader preferences. Bribery is more likely to succeed here, leading to laxly enforced laws and more violence.

To test this argument, we draw evidence from the ongoing violence between Mexican cartels. Using data from the Office of the President from 2000 to 2011 on all extralegal deaths by Mexican police agencies, we show that political tenure is positively correlated with a district's murder rate.[1] Estimating the marginal effects of tenure shows that each additional year a local political party serves in Congress is associated with an additional death in the respective municipality. Although every unnecessary death is tragic, this might seem substantively insignificant. Yet the 2,371 municipalities in our dataset implies that a countrywide increase of one year in tenure results in an increase of approximately as the same number of murders each year. That number is roughly equivalent to the 2011 murder totals of France, Germany, the United Kingdom, the Netherlands, and Belgium *combined* (UNODC 2014). While it might be tempting for political scientists to consider violence between cartels as apolitical, this finding shows that there are clear links between politics and criminal violence.

This result has several important implications for our current understanding of political violence. First, it suggests that the tendency to bifurcate episodes of violence into "civil wars" and other conflicts is potentially misleading (Singer and Small 1982; Sambanis 2004). Even in this context where Mexican cartels are not directly attempting to assume control of state institutions, there are clear political dynamics that shape the timing, location, and nature of violence (Beardsley and McQuinn 2009). Second, one of the most frequently cited arguments about civil war onset focuses on state capacity to explain onset (Fearon and Laitin 2003).

This paper proceeds as follows. We begin by discussing some of the historical background

[1]This is striking: theories of retrospective voting would predict the opposite (Fiorina 1981; Kinder and Kiewiet 1979; Kinder and Kiewiet 1981). Further, Cummins (2009) finds that governors and their parties in the United States suffer at the polls for high crime rates.

unique to the Mexican case to lay the microfoundations of our theoretical discussion. Second, we introduce our formal model of an interaction between two cartels and a local politician. Third, we test the empirical implications of this model using a novel dataset of violence and voting patterns. Finally, we conclude with suggestions for future research.

2 Historical Background

The argument outlined above depends upon four key assumptions. First, we assume that cartels do not have access to alternate dispute resolution mechanism, making violence necessary to control territory. Second, we claim that there is real uncertainty over whether a bribe to political elites is likely to be successful. As a corollary, this suggests that some politicians are more likely to take bribes than others. Finally, our argument requires that politicians and political elites can influence the behavior of police agencies. In this section, we offer qualitative evidence in support of these assumptions.

Our first assumption, namely that cartels use violence as a means of dispute resolution, has two principle observable implications. First, when disputes arise, cartel members should avoid formal dispute-resolution mechanisms like the courts; second, cartel leaders should harm rivals to acquire their territory and resources. The possession, sale, and/or distribution of narcotics is illegal under Mexican law. As Miron (1999) demonstrates, courts cannot enforce property rights when the good in question is illegal. This does not necessarily foreswear informal agreements, such as truces and gentlemen's agreements. Indeed, Skarbek (2011) demonstrates that rival gang leaders have established informal property rights within the California prison system. The California prison system, however, does not readily generalize to individual's behavior on the outside. Once in prison, gang leaders have a reasonable expectation of repeated interaction and means of sanctioning rule breakers are readily available (Axelrod 1984). Leaving the gang, or contravening its rules, is punishable by death (Skarbek 2011, 3). In contrast, cartel leaders live and operate in their own zones of control and cannot be easily punished for double-dealing.[2]

Despite the institutional and legal barriers to cooperation, it might still be possible if cartels were sufficiently incentivized. As profit-maximizing firms, cooperation and collusion does not happen regularly. Cartels do not have incentives to cooperate because their most important task is to establish control over territory, infrastructure, and smuggling routes that help it traffic and manufacture narcotics. Although Mexican cartels mostly transshipped

[2]Cartel leaders do sometimes form temporary alliances and cooperate with one another. Many of these, however, are relatively short lived and epiphenomenal. The story of Juárez Cartel leader Vincente Carrillo Fuentes is emblematic. Carrillo formed an alliance with the Sinaloa cartel early in the 2000s. When the head of the Sinaloa cartel killed Carrillo's borther in 2004, the alliance ended (Associated Press 2014).

South American narcotics in the 1980s, the country has since become one of the world's largest exporters of drugs. According to USAID (2014, 2), Mexico is the world's second largest cultivator of opium and the most important source of marijuana and methamphetamine in the United States. A 2014 report by the California Attorney General claims that Mexican cartels are "suspected of trafficking 70 percent of the U.S. supply of methamphetamine through the San Diego port of entry alone" (Harris 2014, ii). Controlling bases of operation along the border is therefore essential for a cartel's success in smuggling contraband. Cartels' preference for profit-maximization leaves cooperation over these facilities improbable, which makes conflict over valuable territory likely.

Second, we claim that cartels do not know *ex ante* corruptibility of local political elites. In other words, a politician's corruptibility is an innate quality that is difficult to suss out. For example, Andrés Granier was governor of Tabasco until 2012. During his time in office, Granier was not generally seen as corrupt. In 2013, a recording of a private conversation was leaked to a local radio station where he claimed to own enormous quantities of designer clothing (Zabludovsky 2013a). Later arraigned on tax evasion, Granier is alleged to have diverted $156 million in federal funds from the state budget (Castillo 2013). This example suggests that cartels have better information about a politician's receptiveness to bribes than private citizens.

After taking kickbacks, examples of politicians willing to directly collaborate with cartels are numerous. Tamaulipas, for instance, is the home state of the Gulf Cartel and Los Zetas. National daily newspaper *Excélsior* claims that cartels have "taken control of local security" and made the governor a "local employee" (Moreno 2011). In 2012, the U.S. Drug Enforcement Agency (DEA) accused a former governor of accepting millions of dollars in bribes from cartels (Fox News 2012). As part of an ongoing investigation into high level corruption in the state, the Federal Attorney General forbid the state's past three governors from leaving the country (El Universal 2012). The commander of the Tamaulipas State Police was arrested in 2013 when he crossed into the United States on a Washington, D.C. arrest warrant (Taylor 2013). Although this is a brief summary of the corrupt relationship between the political establishment and cartels in one state, similar relationships exist throughout the country.

Finally, we assume that political elites can influence the deployment and enforcement priorities of police and security forces. There are three principle police forces in Mexico: the Policía Federal (PF), state, and local police forces. Political elites can influence policing decisions at all levels of government through encouraging corruption at the Attorney General's office, among police chiefs, and even by directly ordering police officers to ignore drug trafficking (GAO 1996, 9; Sullivan and Elkus 2008). For example, the former governor of Quintana Roo state, Mario Villanueva Madrid, was sentenced to almost eleven years in American prison for conspiracy to launder millions of dollars in bribes (Zabludovsky 2013b). According to prosecutors, "Mr. Villanueva had agreed to let the Juárez cartel... transport cocaine from Colombia through Quintana Roo and on to the United States in exchange for up to $500,000 per shipment. Traffickers were free to unload drug shipments at a state government hangar of a local airport" (Zabludovsky 2013b). As a clandestine activity, we cannot directly prove collusion between politicians and cartels. However, the number of arrests of high-ranking politicians suggests that these are not isolated incidents.

3 The Model

The game consists of three players: two cartels (denoted 1 and 2) and a local party.[3] Cartel 1 has status quo control over the local district (standardized to value 1) and needs to use violence to keep Cartel 2 from encroaching on its territory. Cartel 2, meanwhile, can use violence to challenge Cartel 1's control. Meanwhile, the party wishes to keep the level of violence down while, though it is willing to permit violence at the right price.

Play begins with Cartel 1 taking advantage of its regional ties and familiarity to offer a bribe $b > 0$ the party to limit the enforcement of anti-violence laws. If the party accepts the bribe, the party implements a "no enforcement" policy of $\lambda = 1$, where λ reflects the portion of violence that will succeed. In exchange, Cartel 1 pays b to the party.[4] To analyze how the outcome varies as a function of the party's level of corruption, the party internalizes bc from the bribe payment, where $c > 0$. Thus, higher levels of c reflect higher levels of corruption and a greater willingness to solicit a bribe.

If the party rejects, it selects a level of enforcement $\lambda \in [0, 1]$.[5] However, exerting such effort is costly. To reflect the enforcement cost, the party pays $k(\lambda)$, a function that is differentiable everywhere on the unit interval and where $-k'(\lambda) < 0$ and $-k''(\lambda) \leq 0$. This intuitively implies that effort is always helpful but costly.

Both cartels see the level of enforcement and simultaneously choose respective levels of

[3] Although we ultimately care about police enforcement, we focus on party-level bribery because such large-scale corrupt behavior requires political consent, and these party leaders ultimately have control over police policies.

[4] We are therefore analyzing a bargaining game with *quid-pro-quo* offers. This might seem strange given that the very nature of bribery means that such deals are not enforceable through traditional legal mechanisms. However, we could instead think of this game as the reduced form of a longer-horizon exchange. Rather than paying the entire bribe upfront, the cartel could make a number of smaller payments over time. Given this repetition, the party would not have incentive to defect on the deal when doing so would cancel the long-term gains from cooperation (Axelrod 1984). As such, another interpretation for the bribe value b is the total value of a large string of small bribes.

[5] While explicitly this means that the politician sets no enforcement, we might alternatively think of this value as the bare minimum level of enforcement that both the politician and cartel would find preferable to no enforcement at all.

violence $v_1 \geq 0$ and $v_2 \geq 0$. A contest success function uses the violence levels to determine the distribution of the district at the end of the game. Specifically, Cartel 1 takes $\lambda \frac{v_1}{v_1+v_2}$ portion and Cartel 2 takes the remainder, or $1 - \lambda \frac{v_1}{v_1+v_2}$. Thus, higher levels of enforcement (and in turn lower values of λ) permit Cartel 2 to encroach further on Cartel 1's territory, all other things equal.[6] Each pays a cost for its effort. We therefore subtract v_2 from Cartel 2's payoff and αv_1 from Cartel 1's payoff. Because it is easier to simpler previously controlled territory, we set $\alpha \in (0, 1)$; put differently, violence is comparatively cheaper for Cartel 1.[7]

Recapping, the timing is as follows:

1. Cartel 1 offers a bribe b to the party
2. The party accepts or rejects the bribe
3. If the party rejects the bribe, it sets a level of enforcement
4. The cartels simultaneously set violence levels v_1 and v_2
5. Payoffs are realized

Note that we make very few restrictions on the players' choices. The bribe, level of police enforcement, and levels of cartel violence are all endogenously selected; only the accept/reject choice is a binary decision. This helps ensure that the theoretical results we obtain are not a consequence of restrictive modeling decisions but instead the optimal strategies of the players.

Overall, those payoffs are as follows. If the bribe fails, the party suffers the total amount of violence minus its effort to reduce violence, or $-\lambda(v_1 + v_2) - k(\lambda)$. If the bribe succeeds, the party still suffers the total amount of violence but gains the value of the bribe multiplied by $c > 0$. Formally, this is $-(v_1 - v_2) + bc$. Thus, another way to interpret c is how much the party weighs self-enrichment to good policy. Cartel 1 receives $\lambda \frac{v_1}{v_1+v_2} - \alpha v_1$, while Cartel 2 earns $1 - \lambda \frac{v_1}{v_1+v_2} - v_2$.

3.1 Complete Information Equilibria

Since this is an extensive form game with complete information, we solve for its subgame perfect equilibria. SPE require that all strategy choices are sequentially rational, ensuring that players can only carry out threats that they have incentive to follow through on.

[6] An alternative interpretation is that our game is a close approximation of a game in which the status quo actor has closer ties to the local political party, which appears generally true across Mexican municipalities.

[7] Although Cartel 2 plays a secondary role in the interaction, this step is nevertheless critical to ensuring that our comparative statics on violence are correct. Indeed, our equilibrium analysis shows that Cartel 2's production of violence depends on the outcome of the bargaining stage. We could not pick up on this if Cartel 2's action was exogenous.

Table 1: Notation of the bribery game

Notation	Description
v_i	Cartel i's weakly positive level of violence
α	Cartel 1's relative advantage in producing violence
λ	Party's percentage of violence allowed
$k(\lambda)$	Party's strictly decreasing cost of enforcement function
c	Party's strictly positive level of corruption
b	Cartel 1's weakly positive bribe to the party

Proposition 1. *If the party's level of corruption is sufficiently high, Cartel 1 and the party reach an agreement. In the unique SPE for these parameters, violence levels are high. If the party's level of corruption is sufficiently low, no mutually acceptable bribe exists. In all SPE for these parameters, violence levels are low.*

See the appendix for a complete proof. There are two phases to analyze: bribery and violence. First, consider the violence subgame with varying levels of enforcement. As is standard with contest success functions, the parties choose an amount of violence that best responds to the opponent's level. Because the marginal value for increasing violence near 0 is great, the overall equilibrium level of violence is non-zero. Moreover, because Cartel 1 has a cheaper marginal value for violence on its territory (which α reflects), Cartel 1 produces more violence than Cartel 2, ensuring that it expects to win more of the prize at the end. Based on these predicted levels of violence, the party chooses a level of enforcement that optimizes its tradeoff between reducing the effectiveness of violence and exerting effort, which we call λ^*.

That optimal level of enforcement lingers throughout the game. Critically, enforcement erodes Cartel 1's status quo advantage over Cartel 2. Although one might think that Cartel 1 would respond to this limitation by producing more violence—and perhaps more than in a world without enforcement—this intuition is wrong. The additional enforcement effectively makes the per unit cost of additional violence greater. Consequently, higher levels of enforcement lead to lower levels of violence due to both the enforcement and the anticipation of that enforcement, which endogenously reduces the total amount of potential violence. Nevertheless, additional enforcement works to the detriment of Cartel 1, giving it incentive to bribe the party and secure a greater share of the good through the contest.

That concern impacts the other phase: bargaining between Cartel 1 and the party. Anticipating how various levels of enforcement affect its ability to capture drug rents, Cartel 1 can calculate its marginal gain for buying the party's compliance. Meanwhile, knowing the party's level of corruption and desire to reduce violence, Cartel 1 can calculate the party's minimally

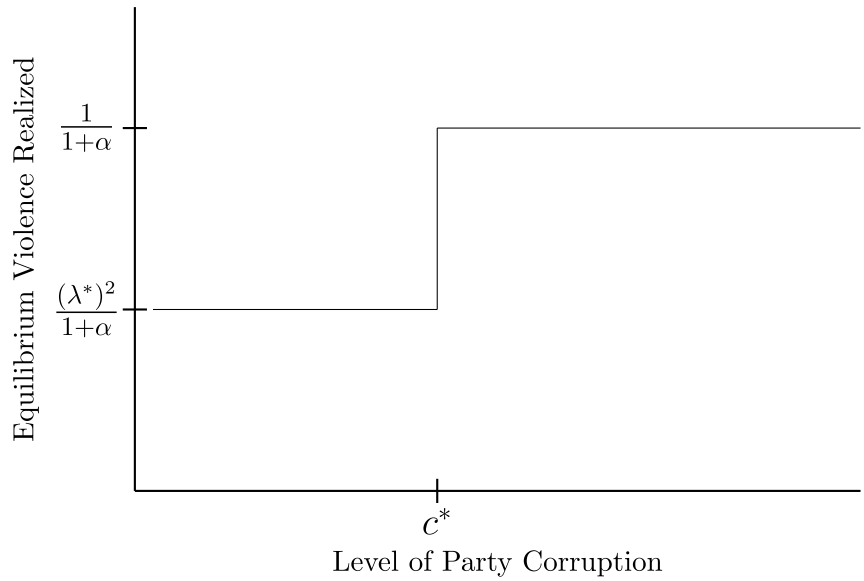

Figure 1: Equilibrium levels of realized violence by the level of party corruption. When corruption is below the critical threshold c^*, the cartel finds bribery too expensive. Enforcement then deters the production of violence (by factor λ^*) and also mitigates whatever violence is still produced (also by factor λ^*). When corruption is high, the bribe succeeds, resulting in higher levels of violence.

acceptable bribe and convert it to a monetary value. If the value to Cartel 1 is greater than the party's minimally acceptable bribe—because the party's level of corruption is sufficiently high—then a bargaining range exists. Because Cartel 1 has all the proposal power, it chooses a bribe exactly equal to the minimally acceptable amount, and negotiations succeed. The party subsequently shirks on enforcement, increasing the level of realized violence. If the value to Cartel 1 is less than the party's minimally acceptable bribe—because the party's level of corruption is sufficiently low—then the bargaining range disappears. The game has multiple equilibria here because Cartel 1 can optimally offer *any* unacceptable bribe to the party. However, the outcome remains the same: unable to reach an agreement, the party chooses a level of enforcement unencumbered and thus reduces the level of realized violence.

While this complete information game generates baseline results, it makes a strong assumption about the bargaining phase of the game: Cartel 1 knows the party's exact level of corruption. This allows it to tailor the appropriate offer to the party and reap all of the surplus through bribery. In practice, though, it would be very difficult for a cartel to know the exact amount it needs to offer a party to buy its compliance. After all, although levels of corruption correlate with many observable factors like platform and reputation, the exact level is an internal attribute of party officials. Thus, a more plausible setup would make Cartel 1 uncertain about party's minimally acceptable bribe. We investigate this exact scenario below.

3.2 Uncertainty about Corruption

Consider the following modification to the game. Nature now begins by drawing a level of corruption of the party as one of two types.[8] Specifically, the party is more corrupt with probability p while the party is less corrupt with probability $1 - p$. These varying levels of corruptibility influence the intrinsic value of the bribe to the party. Thus, holding a bribe level fixed at b, a more corrupt type values that bribe at bc' whereas a less corrupt type values it at bc, with $c' > c$. In words, less corrupt types find bribes to be less valuable peso for peso. Because corruptibility is an internal attribute, it is private information to the party. The cartels therefore only know the prior at the start of the game. This prior may be strong or weak based on observable factors that correlate with corruptibility, and we will eventually investigate how the game's equilibria change as a function of the strength of Cartel 1's prior beliefs.

Since this is now a sequential game with incomplete information, we search for its perfect Bayesian equilibria (PBE). A PBE is a set of strategies and beliefs such that the strategies are sequentially rational and players update their beliefs via Bayes' rule wherever possible. Although this type of incomplete information game often yields multiple equilibria depending on off-the-path beliefs that cannot be derived from Bayes' rule, the equilibria we present here are unique. This is because the cartels' uncertainty about the party's level of corruption only has payoff-relevant ramifications during the accept/reject phase of the game. However, the actor facing uncertainty (Cartel 1) makes an offer before the informed actor (the party) decides how to respond. Consequently, the cartels do not need to analyze any signal before moving.

We limit our analysis to cases in which both types would be willing to reach an agreement with complete information. The reasons are twofold. First, all other cases are theoretically trivial. If the bargaining range is empty for both types, then negotiations are moot. If the bargaining range is empty for the less corrupt type but not the more corrupt type, then Cartel 1 can simply focus on settling with the more corrupt type.[9] The game only has an interesting

[8] Similar results would follow in an interaction where the party's level of corruption from a continuum of types.

[9] Further, under such conditions, the different types could credibly separate in a cheap talk extension to the game. This is because the less corrupt type, even with complete separation, receives the same offer as it would with all information revealed. Thus, the parameters we focus on are the parameters where information problems matter the most.

strategic interaction in the remaining case. Second, it is the substantively most compelling case. Based on our above qualitative discussion above, local officials and cartels seem willing to try negotiating agreements with one another. Stories and criminal proceedings of corruption and collusion between cartels and officials are not limited to any particular geographic region, political party, or socioeconomic background. With birbery so prevalent, we focus on that particular parameter condition.

With that, we are ready to describe the game's outcome. We consolidate the discussion into two cases, depending on whether Cartel 1's belief that the party is the more corrupt type exceeds a critical threshold $p^* \equiv \frac{\frac{1-\lambda^*}{(1+\alpha)^2} - \frac{1+(\lambda^*)^2}{c(1+\alpha)} + \frac{k(\lambda^*)}{c}}{\frac{1-\lambda^*}{(1+\alpha)^2} - \frac{1+(\lambda^*)^2}{c'(1+\alpha)} + \frac{k(\lambda^*)}{c'}}$:

Proposition 2. *If the party is sufficiently likely to be the more corrupt type (i.e., $p > p^*$), Cartel 1 offers a small bribe to the party. The more corrupt type accepts with certainty while the less corrupt type rejects with certainty. Violence levels are high against the more corrupt type but lower against the less corrupt type.*

The appendix contains a full proof. For intuition, note that high values of p mean that Cartel 1 believes it is very likely facing the more corrupt type. Consequently, it prefers tailoring its bribe to that type even though it knows that this smaller offer induces the less corrupt type to reject; it just is not worth paying more to cover the rare event that the party is not so easily corruptible. Because the less corrupt type proceeds to enforce the laws, both cartels select a lower level of violence. In contrast, when the bribe succeeds versus the high type, Cartel 1 chooses a higher level of violence because its marginal value is greater. Anticipating this, Cartel 2 increases its level of violence to compensate. Because both outcomes occur with positive probability in this case, we expect to see a middling level of violence here.[10]

Violence is more prominent in the next case, however:

Proposition 3. *If the party is sufficiently likely to be the less corrupt type (i.e., $p < p^*$), Cartel 1 offers a large bribe to the party. Both types accept with certainty. Without enforcement, violence levels are high.*

Again, the appendix contains a full proof. The intuition here is that Cartel 1 ought to tailor its bribe to the less corrupt type because that type is more prominent in this case. Unfortunately for Cartel 1, this requires offering a large amount. Because the more corrupt

[10]One may wonder if cheap talk signaling can resolve the bargaining breakdown here. It cannot. The key is that the more corrupt type always has incentive to mimic the less corrupt type; if believed, the more corrupt type receives a larger bribe then it would if Cartel 1 knew it was a more corrupt type. This incentive to misrepresent therefore prohibits meaningful communication under these circumstances.

type is receptive to small bribes, it is also willing to accept larger bribes. As a result, both types accept and do not enforce the laws. In turn, both Cartels choose high levels of violence for the reasons described above. As such, the expected level of violence for these parameters are greater when compared to Proposition 2's outcome.

3.3 Comparative Statics

Below, we empirically investigate the sources of violence in Mexican municipalities. To do this effectively, we first need to draw a comparative static from the model that we can then use to construct a testable hypothesis. Our qualitative overview at the beginning of this paper pointed to the ease of successful bribery as a critical driver of drug violence in Mexico. With incomplete information, such ease is a function of the informational environment. We thus focus on the "bandwidth" of potential types Cartel 1 might be facing:

Proposition 4. *If mutually acceptable bribes exist for both types of party, violence is weakly increases as uncertainty about the party (i.e., $c' - c$) decreases.*

Once more, the appendix contains the full proof. The basic intuition is as follows. Without uncertainty, per Proposition 1, Cartel 1 can appropriately tailor the bribe and reach a mutually preferable settlement with the party. In the incomplete information case, the bandwidth of types ($c' - c$, or how different the types are compared to one another) is one measurement of uncertainty. As that bandwidth diminishes, the potential types the cartel could be facing become increasingly similar. This helps Cartel 1 find an offer that both would prefer to bargaining breakdown.

Essentially, Cartel 1 faces a risk-return tradeoff. Broadly, it has two options. First, it can offer small amount, hope that it is actually facing the more corrupt type, and suffer through full enforcement against the less corrupt type. Second, it can offer a large amount and induce both types to accept. This second case is expensive because it requires paying the large bribe to both types, effectively costing Cartel 1 some fixed amount whenever the party is the more corrupt type. However, as $c' - c$ goes to 0, the risk premium Cartel 1 pays becomes vanishingly small. As such, the amount "wasted" on the bribe to the corrupt type becomes increasingly insignificant. In turn, Cartel 1 prefers offering the amount necessary to induce both types to accept.[11]

Although reducing uncertainty leads to an increase in the likelihood of settlement, note that it leads to an *increase* in the level of violence. This should be striking to researchers

[11] Note that Proposition 4 is a conditional statement. If the bargaining range is empty for one type but not the other, decreasing uncertainty could push the bribable type across c^* threshold, which in turn decreases violence. Due to the prevalence of corruption, we choose to focus on the case where both levels of corruption are greater than c^*.

familiar with bargaining and conflict. Normally such models show that reducing uncertainty reduces conflict. On a technical level, this remains true here: the level of observed conflict (i.e., bargaining breakdown) between Cartel 1 and the party decreases as uncertainty decreases. However, the purpose of an agreement between the two is to increase the effectiveness of violence for Cartel 1. As such, decreasing uncertainty has a negative externality on outsiders (i.e., private citizens) who want a decrease in the level of violence.

4 Empirics

This section introduces our novel dataset on violent events and political tenure in Mexico. We discuss our model of the effect of tenure on violence and conclude by presenting our results.

4.1 Hypothesis

Before delving into the data, we must first derive a testable implication from the model. The formal analysis demonstrates that high-quality information is critical for the parties to reach an agreement. This presents a major problem for empirical inquiry, however. Perfectly predicting bargaining failure would require the analyst to know more than the parties in the interaction. After all, if breakdown were perfectly predictable for the actors involved, the cartel would simply increase its offer to an acceptable level and eliminate any inefficiency. Thus, inevitably, bargaining breakdown (and thus variation in violence) is in the error term (Gartzke 1999).

Fortunately, despite this hurdle, fruitful inquiry is still possible. Rather than assume that researchers can better understand the information asymmetry than the players involved, we can instead investigate environments that correlate with uncertainty in general. Since Proposition 4 says that environments with greater uncertainty (as defined by the "bandwidth" of possible types), we would expect that the probability of bargaining breakdown to increase with these correlates of uncertainty.

Specifically, we draw on theoretical innovations from the international relations literature. Wolford (2007) argues that new leadership in the international system creates a shock to the informational structure. Opposing states must throw out their estimates of the old leader's resolve and begin the intelligence process anew. However, as a leader's tenure increases, those estimates become progressively better and therefore the bandwidth of possible types decreases. Bargaining is more likely to succeed under these circumstances.

This same logic ought to apply to local Mexican political machines and drug cartels. When a machine first takes local control, the dealers will be unfamiliar with the key political elite. As time progresses, though, observable information about these leaders will accumulate. Thus, although a level of corruption is an innate trait, cartels can update and narrow their expectations by seeing how these leaders behave over time. Per Proposition 4, this accumulation of knowledge decreases the probability of bargaining breakdown, which in turn decreases levels of law enforcement and increases violence. We can thus summarize our hypothesis as follows:

Hypothesis 1. *Violence levels are increasing in leader tenure.*

Below, we show that the data support this hypothesis.

4.2 Data and Model

There have been several different attempts to measure the ongoing violence in Mexico released within the past few years. In this paper, we use the official dataset released by the Office of the President in 2011. It contains the reported number of murders, by municipality, from 1990 to 2011.[12] For reasons discussed above, we drop observations before 2000 because electoral results were not free and fair (Magaloni 2008). Not all political parties faced the same potential consequences for collusion with cartels. Even if they were dissatisfied, voters faced enormous difficulties removing the PRI from power.[13] After dropping those observations, the resulting dataset has 28,368 municipality-year observations. Our electoral data comes from the Instituto Nacional Electoral (INE), which is the national agency responsible for conducting elections and tallying votes. Through their online portal, the INE releases data at the municipal level for both legislative and general elections. As voters elect new legislators every three years, we have electoral results from the elections in 2000, 2003, 2006, and 2009.

From among several potential measures of violence in Mexico, we use the dataset generated by the Office of the President for several reasons. First, its data acquisition process is the least likely to be geographically or temporally biased. While newspapers such as *Reforma* and *Milenio* also attempt to record all murders, they have particular regional focuses that might cause upwards bias in estimates from their home region.[14] Second, although the Department of Justice released some datasets with more detailed information the nature of the crime, they

[12]The Office of the President stopped updating this dataset in 2012 without explanation.

[13]This leaves us with a question of how to code tenure at the beginning of the time period. We choose to reset all party tenures to 0 for two reasons. First, from a theoretical standpoint, the lack of party competition and electoral responsiveness made the bargaining environment substantially different between these time periods. And second, resetting the data acts as a "hardest case" test—because we believe that long periods of tenure assist in the learning process and leads to more violence, this coding rule only makes it more difficult to obtain statistically and substantively significant results.

[14]The newspapers generate statistics by compiling police reports, social media accounts, and other sources for information about the style of execution. This is incredibly detailed work, but requires reporters to have local contacts throughout the entire country. To cross validate our measure, we check the correlation between our figures and those released by *Reforma* and *Milenio*. Our measure of violence is highly correlated with both, but has better temporal and geographic coverage.

all have extremely short time series. For example, the time series in a dataset exclusively of "organized-crime style homicides" lasts only from January to September 2011. As these datasets do not have data before and after an election, they are inappropriate to test our argument. Finally, Department of Justice-compiled figures are generally reported at the state level. While all states maintain their own police force, the Policía Federal (PF) and municipal police forces are the most important law enforcement branches of the Mexican state (Bailey and Dammert 2006). It is more likely that corrupt political elites can influence the PF as well as their local police forces than state police, which are commanded by the state's governor (Bailey and Taylor 2009).

Although it is an official count of the number of murders in Mexico, there are certain caveats that are necessary to keep in mind when using this data to study the behavior of cartels. First, this dataset only reports murders that were reported to police agencies. While there is no obvious reason for a police agency to intentionally underreport figures, it is still possible that its totals are an underestimate of all cartel-related murders.[15] This is because cartels frequently punish rivals by disposing of their bodies in such a way that they cannot be located. For example, a cartel head based in Tijuana frequently "[boiled] rivals in barrels of lye" to dispose of their bodies (Lacey 2010). Given the dissolution of the human remains, it is unlikely that authorities received reports of the murder of someone disposed of this way. Second, although it does not state so explicitly, it is possible that the dataset records deaths in the year they were discovered — not necessarily the year they occurred. While this might be a concern, it is unlikely that the late discovery of criminal incidents should be correlated with Congressional party incumbency.

Finally, and perhaps most importantly, our dataset contains all murders and non-negligent manslaughters reported to the Office of the President. While this includes both cartel and non-cartel related deaths, our argument does not explain murders unconnected to the drug trade. Although occasionally it might be possible to use the particulars of a murder to code whether it is cartel-related, systematics attempts to do so cover very limited time periods or have specific regional focuses. Even so, drawing on Kalyvas (2006), we argue that this does not impose a serious constraint in our data analysis for three reasons. First, the measurement error in our dependent variable should make it more difficult for us to find a statistically significant relationship with our explanatory variable. Second, as non-cartel related violence is unrelated to politics, the frequency and geographic distribution of such crime should be relatively randomly distributed. Finally, we use several statistical techniques to control for the unobserved mixture of cartel- and non-cartel related violence.

[15] Indeed, it is easier to imagine a situation wherein police departments overestimate the number of murders to receive additional funds and matériel.

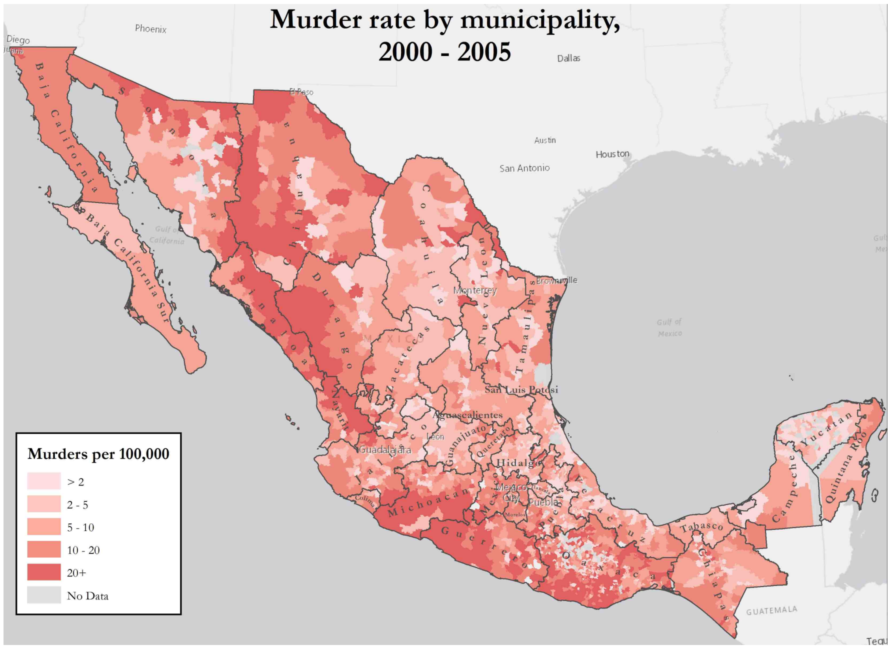

Figure 2: Murder rate by municipality, 2000 - 2005. During this period, this map shows that violence was primarily concentrated in the Altiplano around Oaxaca and along the Sierra Madre Occidental.

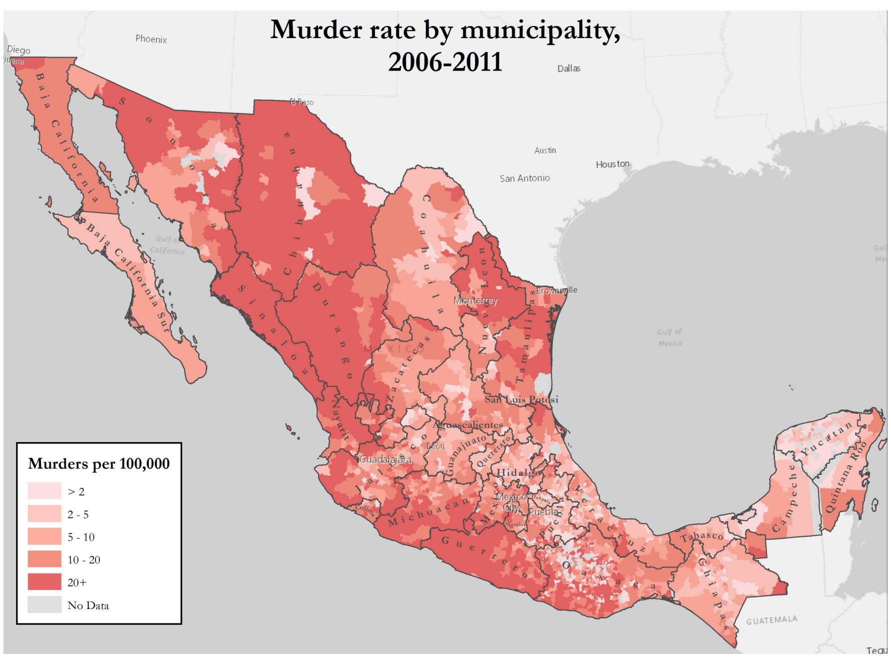

Figure 3: Murder rate by municipality, 2006 - 2011. During this period, this map shows that violence spread across Chihuahua, Sinaloa, Sonora, and Tamaulipas. The level of violence remains high along the Sierra Madre del Sur.

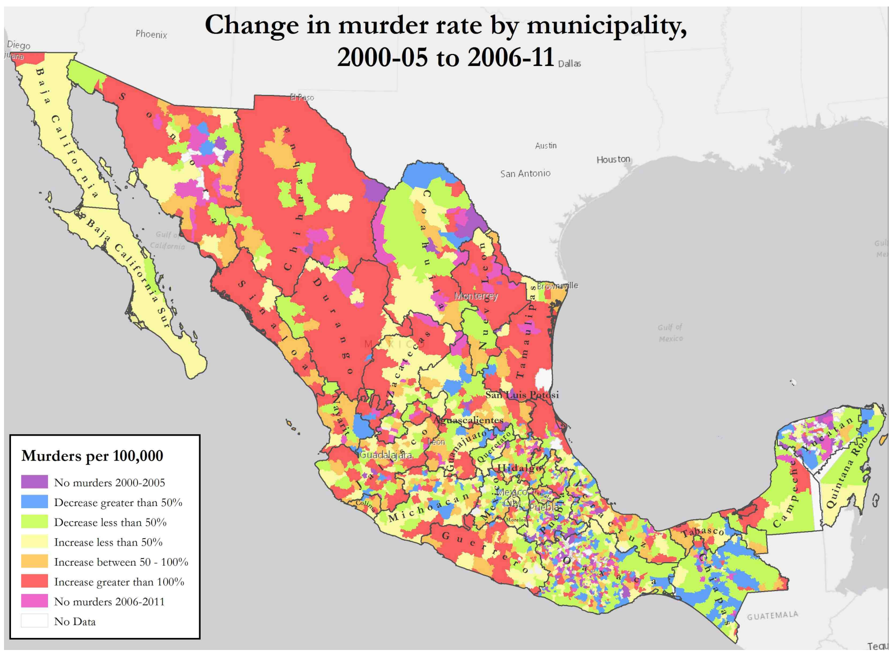

Figure 4: Change in murder rateby municipality, 2000-05 to 2006-11. This map shows that murder rates increased dramatically in the North and East of the country, but declined in the Altiplano and South. It also shows tremendous variation in the rate and location of these changes, which suggests that our results do not come from uniform increases throughout the country.

To test the empirical implications of our formal model, we run a linear ordinary least squares (OLS) model with municipal fixed effects and cubic restricted time splines (Green, Kim, and Yoon 2001). We use municipal fixed effects to control for unit-specific factors to reduce unobserved heterogeneity. For example, some municipalities might be more politically competitive, have their own media market, or be located more closely to international borders. While these are unit-specific, many such features are unobservable. We therefore include $\sum_{j=1}^{n} \theta_j$, where θ_j represents a set of unobserved fixed parameters for each of the n units in our sample. The short duration of our time series imposes certain restrictions on the number of additional factors for which we can control. Although scholars continue to debate the minimum number of observations per parameter necessary to avoid bias, simulations show that the bare minimum of observations in each group per parameter is around five (Harrell 1984, 2002; Vittinghoff and McCulloch 2007). Although we cannot completely eliminate the risk of omitted variable bias and autoregressive disturbance, our results are robust to a variety of model specifications.[16]

With this restriction in mind, we estimate the predicted murder level in municipality i in year t with Equation 1:

$$Murder_{it} = \beta_0 + \beta_1 Tenure_{it} + Murder_{(it-1)} + f(\gamma) + \sum_{j=1}^{n} \theta_j + \epsilon_{it} \qquad (1)$$

To account for the possibility of temporal dependence in our dependent variable and autoregressive disturbances, we control for temporal effects in two ways. First, we include a lagged dependent variable (Kiviet and Phillips 1993; Achen 2000). Second, we follow Beck, Katz, and Tucker (1998) and include restricted cubic time splines with knots at each quartile. A spline function is a "smoothly joined piecewise polynomial of degree n" (Durrleman and Simon 1989, 552). Splines control for nonlinear time effects, such as wars or new technology, which affect all municipalities in the panel differently.[17] Following Dupont and Plummer (2005), restricted time splines are defined as:

[16] As additional robustness checks, we present several alternate model specifications in the appendix. Across all models, including first differencing and standard OLS, Tenure remains significant and positive. It also remains significant and positive after controlling for whether the PRI controls a given municipality and election years.

[17] Geospatial clustering is another concern. Clustering could bias our inferences by making coefficients inconsistant and inflating our model's R^2. To check for the presence of such clustering, we estimate a geographically weighted regression and present the results in the appendix. We then check for clustering in our residuals by estimating Moran's I (Moran 1950). Results from this analysis show that our data is randomly distributed geospatially.

$$
\begin{aligned}
\gamma_1 &= \gamma \\
\gamma_j &= (\gamma - t_{j-1})^3_+ - \left(\frac{(\gamma - t_{k-1})^3_+ (t_k - t_{j-1})}{(t_k - t_{k-1})} + \frac{(t - t_k)^3_+ (t_{k-1} - t_{j-1})}{(t_k - t_{k-1})} \right) \\
&\text{for } j = 2, \ldots, k-1 \\
(u)_+ &= \begin{cases} u: & u > 0 \\ 0 & u \leq 0 \end{cases}
\end{aligned}
\tag{2}
$$

where t_i, $i = 1, \ldots, n$ are the spline knot values; k is the number of knots. Our dependent variable Murder, is a count of the number of extralegal deaths recorded by the Office of the President.

We code our key dependent variable, Tenure, with data from the INE. As the INE releases data by political party, we assign Tenure a value of 0 in the year a district elected a political party to Congress.[18] It then increases by one every year a political party remains in office within a particular district. Should the party lose an election, Tenure resets to 0 in the year of the election.[19] We measure tenure at the party, rather than individual, level due to a quirk of Mexican law. Until electoral reforms passed in 2013, the Constitution strictly prohibited reelection in all political offices.[20] As a result, no Congressperson was reelected in the period of our study. Because politicians cannot build personal clientelist networks and support bases, Mexican political parties can exert substantial influence over their actions. Without independent bases of support, politicians who go against their local party officials' wishes encounter substantial difficulty in pursuing higher office or using their final year in office to seek alternate employment (Magaloni 2008; Mainwaring and Scully 1995; Morgenstern and Nacif 2002).

As we contend that cartels need time to learn whether it is possible to bribe political leaders, they likely work much more closely with local political party organizations and elites than individual legislators. Substantively, moreover, there is no variation in tenure at the individual level until 2018. For these reasons, we code tenure based on the number of years a party — and not a politician — remains in office.[21]

[18] As is common for studies using leader tenure as a key independent variable, one concern is how to assign transition years since there is split responsibility during that period. The appendix shows that the results are robust (and, indeed, slightly stronger) to dropping all transition years from the analysis.

[19] One concern might be that some parties are more professional and therefore less likely to lose an election. However, the correlation between political party and Tenure (0.33) is not significant.

[20] Coming into effect in 2018, mayors may now serve two consecutive terms, while legislators may serve for up to 12 years. Once elected, they are forbidden from switching political parties.

[21] In a few instances, political parties campaign in coalition with a junior party. For example, the *Alianza por el Cambio* was an alliance in the 2000 elections between the PAN and Green Ecological Party of Mexico.

4.3 Results

We report the results of our statistical model in Table 2. In line with our theoretical predictions, it shows that additional years of political tenure are associated with increased levels of violence. This result, moreover, is robust to municipality fixed effects, cubic restricted time splines, and a lagged dependent variable. While our model might appear sparse, it is worthwhile to note that including a lagged dependent variable controls for autocorrelation and the dynamics of the data generating process in $t - 1$ (Keele and Kelly 2006). Statistical research, moreover, suggests that lagged dependent variables can suppress the coefficients of the remaining independent variables; as such, the inclusion of a lagged dependent variable is a highly conservative model (Achen 2000; Durbin 1970). Together, this suggests the effect is quite robust to alternate model specifications and is not the result of autocorrelation. In Table 2, Model 1 includes country fixed effects, while Model 2 has both country fixed effects and cubic restricted time splines.

Table 2: Fixed Effects OLS of Incumbency's Effect on Violence with Lagged DV

	Dependent variable:	
	Murder	
	(1)	(2)
Tenure	0.39^{***}	0.16^{**}
	(0.06)	(0.07)
Murder_{-1}	0.84^{***}	0.84^{***}
	(0.01)	(0.01)
Municipal FE	Yes	Yes
Time Splines	No	Yes
Observations	25,538	25,538
R^2	0.59	0.59
Adjusted R^2	0.54	0.54

*p<0.1; **p<0.05; ***p<0.01
Standard errors (clustered on municipality) reported in parentheses.
Estimates for cubic restricted time splines not reported.

To uncover the substantive effect of an additional year of tenure on murder rates, Figure 5

In the following election, the alliance ended and the PAN competed separately. In such cases, we do not code these party renamings as a break in incumbency.

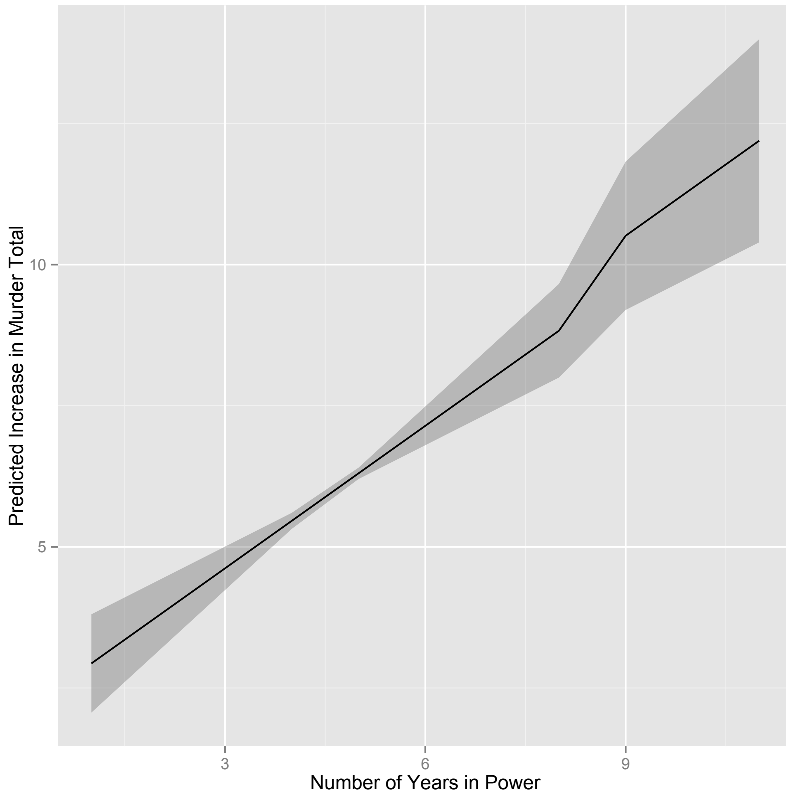

Figure 5: Predicted number of murders in municipalities conditioned on the duration of political tenure. This shows that each additional year of tenure increases a municipality's murder rate approximately by one additional death.

plots the marginal effects of each additional year of tenure on murder rates. After setting all other explanatory variables to their median values in Equation 1, the predicted murder rate across all Mexican municipalities increases from 2.93 ($\sigma = 0.87$) when leaders have been in power for only one year to 12.20 ($\sigma = 1.80$) after 11 years of tenure. This suggests that each additional year of tenure is associated with an additional murder within a given municipality.

While one additional murder per year might not sound substantively meaningful, it is important to remember that this effect accrues across all 2,371 municipalities in our study. As such, we predict that an additional 4,742 people die during the period between elections that would not otherwise. Across the eleven years in our study, this finding suggests that approximately 26,000 people have died in Mexico as a result of collusion between politicians and cartels.

As many political scientists are not familiar with crime statistics, it might not be clear how to interpret an additional 2,371 deaths per year in a comparative context. According to United Nations Office of Drugs and Crime estimates, 2,371 additional deaths per year is equivalent to the combined 2011 murder totals of France, Germany, the United Kingdom, the Netherlands, and Belgium (UNODC 2014). While our estimate might seem unrealistically high, Mexico is one of the world's most violent countries. With 27,213 violent deaths reported in 2011, our estimated treatment effect only represents approximately eight percent of all murders reported to Mexican authorities. This suggests that we can, with relative confidence, eliminate the possibility that the effects our model has captured may be due to modeling error.

5 Conclusion

This paper investigated Mexican drug cartels' strategic incentives to bribe local officials. Although bribes are costly, effective law enforcement hinders a cartel's ability to maintain control over its territory. As such, cartels may wish to buy off politicians to increase their stake in lucrative territorial possessions. Nevertheless, bargaining only succeeds with certainty when a cartel has sufficiently adequate information about an official's level of corruption. In turn, uncertainty is a key driver of enforcement. Using party tenure as a proxy for information and murder rates to reflect drug violence, we estimate that uncertainty is indeed statistically and substantively connected to fewer murders.

Our finding that every additional year countrywide of tenure leads increases homicides by 2,300 highlights the link between violence and politics. While there are good reasons to subset Mexico from standard datasets of civil war, this increase in intentional homicides is more than many civil wars' total casualty count (Sambanis 2004). Beyond the study of political violence, we focus on how cartels and political parties interact because understanding the links between incumbency and politicians behavior is not trivial for Mexicans. Despite forbidding reelection until recently, political parties exert great control over their home regions. To protest this, activists in Baja California launched a new political party in 2014 aimed at ending the leadership of the two main parties and increasing political turnover in the state. This strongly suggests that our theoretical mechanism is not far-fetched to Mexican voters.

Our results have important implications about institutions. If increases to party tenure decrease uncertainty, then electoral laws that favor status quo parties — such as minimum vote thresholds and public financial support for parties — have an unintended side effect of promoting increased violence. The theoretical and empirical analysis therefore indicate that groups fighting Mexican drug cartels could improve their chances by lobbying against such laws. Of course, implementing such reforms might require herculean effort—those in power have strong incentives to maintain and strengthen the systems that put them in office.

Moreover, artificially reducing party tenure could have unintended negative consequences, as experience and professionalization are desirable in other policy areas. Put simply, there is no easy fix to the cartel problems in Mexico.

The implications of our paper also call into question the recommendation to strengthen democratic institutions as a means of reducing violence. The literature on civil war suggests that increasing politicians' accountability to their voters via competitive elections should decrease rent-seeking behavior and their incentives to resort to violence (Fearon and Laitin 2003; Walter 1999). Indeed, the international community and donors tend to support elections when bargaining to end civil wars (Vreeland 2008; Walter 1999). While normatively appealing, many of these elections are rigged in favor of incumbents, leading to few power transitions. Using data from Mexico, we show that electoral incumbency is associated with increases in violence at the local level. Rather than encourage accountability, this suggests that numerous countries have higher levels of violence than they would in the absence of rigged elections. Further research is important to uncover the conditions in which post-conflict elections reduce violence — and when they might increase it.

Finally, the scope of our project is limited to understanding how cartels and local officials conspire with each other for mutual benefit at the expense of rival cartels and the municipality's citizens. However, this is only one interesting strategic aspect of the Mexican drug wars. Future research ought to consider how cartels negotiate with each other and how national intervention in local affairs complicates the larger bargaining and enforcement process.

References

Acemoglu, Daron, James A Robinson and Rafael J Santos. 2013. "The monopoly of violence: Evidence from colombia." *Journal of the European Economic Association* 11(s1):5–44.

Achen, Christopher H. 2000. "Why lagged dependent variables can suppress the explanatory power of other independent variables." Working Paper.

Associated Press. 2014. "Vincente Carrillo Fuentes Arrested: Alleged Leader Of Juarez Cartel Nabbed In Mexico's Torreon.".

Axelrod, Robert. 1984. *The Evolution of Cooperation*. Basic Books.

Bailey, John and Lucía Dammert. 2006. *Public security and police reform in the Americas*. University of Pittsburgh Pre.

Bailey, John and Matthew M Taylor. 2009. "Evade, corrupt, or confront? Organized crime and the state in Brazil and Mexico." *Journal of Politics in Latin America* 1(2):3–29.

Bateson, Regina. 2012. "Crime victimization and political participation." *American Political Science Review* 106(03):570–587.

Beardsley, Kyle and Brian McQuinn. 2009. "Rebel Groups as Predatory Organizations: The Political Effects of the 2004 Tsunami in Indonesia and Sri Lanka." *Journal of Conflict Resolution* 53(4):624–645.

Beck, Nathaniel, Jonathan N. Katz and Richard Tucker. 1998. "Taking time seriously: Time-series-cross-section analysis with a binary dependent variable." *American Journal of Political Science* 42:1260–1288.

Castillo, E. Eduardo. 2013. "Mexico Corruption: State Government Scandals Reveal Lack Of Disclosure, Enforcement.".

Collier, Paul. 2000. "Rebellion as a Quasi-Criminal Activity." *Journal of Conflict Resolution* 44(6):839–853.

Downes, Alexander. 2007. "Draining the Sea by Filling the Graves: Investigating the Effectiveness of Indiscriminate Violence as a Counterinsurgency Strategy." *Civil Wars* 9(4):420–444.

Downes, Alexander. 2008. *Targeting Civilians in War*. Ithaca, NY: Cornell University Press.

Dupont, William and Dale Plummer. 2005. Using Stata v9 to model complex non-linear relationships with restricted cubic splines. In *North American Stata Users' Group Meetings*. Vol. 5.

Durbin, James. 1970. "Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables." *Econometrica: Journal of the Econometric Society* pp. 410–421.

Durrleman, Sylvain and Richard Simon. 1989. "Flexible regression models with cubic splines." *Statistics in Medicine* 8(5):551–561.

El Universal. 2012. "Busca PGR contra 3 ex gobernadores de Tamaulipas.".

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(1):75–90.

Federal Bureau of Investigation. 2014. "Uniform Crime Reports." Online Report.

Fotheringham, A Stewart, Martin E Charlton and Chris Brunsdon. 1998. "Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis." *Environment and planning A* 30(11):1905–1927.

Fox News. 2012. "Mexican Governor got Millions in Drug Cash, DEA Says.".

Green, Donald P., Soo Yeon Kim and David H. Yoon. 2001. "Dirty Pool." *International Organization* 55(02):441–468.

Harrell, Frank E. 2002. *Regression Modeling Strategies*. Springer.

Harris, Kamala. 2014. Gangs Beyond Borders: California and the Fight Against Transnational Organized Crime. Technical report California Attorney General Sacramento, CA: .

Humphreys, Macartan and Jeremy M. Weinstein. 2006. "Handling and Manhandling Civilians in Civil War." *American Political Science Review* 100(3):429–447.

Kalyvas, Stathis. 1999. "Wanton and Senseless? The Logic of Massacres in Algeria." *Rationality and Society* 11(3):243–285.

Kalyvas, Stathis. 2006. *The Logic of Violence in Civil Wars*. New York, NY: Cambridge University Press.

Kan, Paul Rexton. 2012. *Cartels at War: Mexico's Drug-Fueled Violence and the Threat to U.S. National Security*. Dulles, VA: Potomac Books.

Keele, Luke and Nathan J Kelly. 2006. "Dynamic models for dynamic theories: The ins and outs of lagged dependent variables." *Political Analysis* 14(2):186–205.

Kiviet, Jan F and Garry DA Phillips. 1993. "Alternative bias approximations in regressions with a lagged-dependent variable." *Econometric Theory* 9(01):62–80.

Lacey, Marc. 2010. "Mexico Holds Drug Suspect Accused of Grisly Tactics.".

Langton, Jerry. 2012. *Gangland: The Rise of Mexican Drug Cartels from El Paso to Vancouver*. Mississauga, ON: John Wiley and Sons Canada, Ltd.

Lessing, Benjamin. 2014. "How to Build a Criminal Empire From Behind Bars: Prison Gangs and Projection of Power.".

Liker, Jeffrey K, Sue Augustyniak and Greg J Duncan. 1985. "Panel data and models of change: A comparison of first difference and conventional two-wave models." *Social Science Research* 14(1):80–101.

Magaloni, Beatriz. 2008. *Voting for Autocracy: Hegemonic Party Survival and Its Demise in Mexico*. New York, NY: Cambridge University Press.

Mainwaring, Scott, Timothy R Scully et al. 1995. *Building democratic institutions: Party systems in Latin America*. Cambridge Univ Press.

Miron, Jeffrey A. 1999. "Violence and the U.S. Prohibitions of Drugs and Alcohol." *American Law and Economics Review* 1(1):78–114.

Moran, Patrick AP. 1950. "Notes on continuous stochastic phenomena." *Biometrika* pp. 17–23.

Moreno, Martín. 2011. "Tamaulipas: gobierno fallido.".

Morgenstern, Scott and Benito Nacif. 2002. *Legislative Politics in Latin America*. Cambridge University Press.

Pape, Robert A. 2003. "The Strategic Logic of Suicide Terrorism." *American Political Science Review* 97(3):343–361.

Paradis, Emmanuel. 2014. "Moran's Autocorrelation Coecient in Comparative Methods." R Vignette.

Reno, William. 2000. "Clandestine economies, violence, and states in Africa." *Journal of International Affairs* 53(2):433–459.

Rider, Toby J. 2013. "Uncertainty, salient stakes, and the causes of conventional arms races." *International Studies Quarterly* 57(3):580–591.

Rios, Viridiana. 2012. "Why did Mexico become so violent? A self-reinforcing violent equilibrium caused by competition and enforcement." *Trends in Organized Crime* pp. 1–18.

Sambanis, Nicholas. 2004. "What Is Civil War? Conceptual and Empirical Complexities of an Operational Definition." *Journal of Conflict Resolution* 48(6):814–858.

Singer, J. David and Melvin Small. 1982. *Resort to Arms: International and Civil Wars, 1816-1980*. 2nd ed. Beverly Hills, CA: Sage Publications.

Skarbek, David. 2011. "Governance and prison gangs." *American Political Science Review* 105(04):702–716.

Spaniel, William and Bradley C. Smith. Forthcoming. "Sanctions, Uncertainty, and Leader Tenure." *International Studies Quarterly* .

Staniland, Paul. 2010. "Cities on fire: social mobilization, state policy, and urban insurgency." *Comparative Political Studies* 43(12):1623–1649.

Staniland, Paul. 2012. "States, insurgents, and wartime political orders." *Perspectives on Politics* 10(02):243–264.

Sullivan, John P and Adam Elkus. 2008. "State of Siege: Mexico's Criminal Insurgency." *Small Wars Journal* 12.

Taylor, Jared. 2013. "Tamps. police boss, cousin of ex- governor, arrested in U.S. drug probe.".

UNODC. 2014. UNODC crime and criminal justice statistics. Technical report United Nations Office on Drugs and Crime.

U.S. Government Accountability Office. 1996. *Drug control counternarcotics efforts in Mexico: report to congressional requesters*. Washington, DC: GAO.

USAID. 2014. "Mexico." Country Profile.

Varese, Federico. 2011. *Mafias On the Move: How Organized Crime Conquers New Territories*. Princeton, NJ: Princeton University Press.

Vittinghoff, Eric and Charles E. McCulloch. 2007. "Relaxing the rule of ten events per variable in logistic and Cox regression." *American Journal of Epidemiology* 165(6):710–718.

Vreeland, James Raymond. 2008. "The effect of political regime on civil war unpacking anocracy." *Journal of Conflict Resolution* 52(3):401–425.

Walter, Barbara F. 1999. "Designing transitions from civil war: Demobilization, democratization, and commitments to peace." *International Security* 24(1):127–155.

Wolford, Scott. 2007. "The turnover trap: new leaders, reputation, and international conflict." *American Journal of Political Science* 51(4):772–788.

Zabludovsky, Karla. 2013*a*. "Former Mexican Governor Gets 11 Years for Taking Drug Bribes.".

Zabludovsky, Karla. 2013*b*. "Official Corruption in Mexico, Once Rarely Exposed, Is Starting to Come to Light.".

6 Appendix

6.1 Proof of Proposition 1

We proceed with backward induction. Fix a level of enforcement λ. Cartel 1's objective function is $\lambda\frac{v_1}{v_1+v_2} - \alpha v_1$, with its choice a value for v_1. Its first order condition is therefore:

$$\frac{\lambda v_2}{(v_1 + v_2)^2} - \alpha = 0 \tag{3}$$

Meanwhile, Cartel 2's objective function is $1 - \lambda\frac{v_1}{v_1+v_2} - v_2$, with its choice a value for v_2. Its first order condition is therefore:

$$\frac{\lambda v_1}{(v_1 + v_2)^2} - 1 = 0 \tag{4}$$

Using Equations 3 and 4 as a system of equations, the unique solution pair is $v_1^* = \frac{\lambda}{(1+\alpha)^2}, v_2^* = \frac{\alpha\lambda}{(1+\alpha)^2}$. Note that when the politician accepts the bribe, $\lambda = 1$ and therefore the solution pair is $v_1^* = \frac{1}{(1+\alpha)^2}, v_2^* = \frac{\alpha}{(1+\alpha)^2}$.[22]

Now consider the politician's enforcement level, conditional on its rejection of the bribe. The police's objective function is $-\lambda(v_1 + v_2) - k(\lambda)$, with its choice a value for λ. Because $v_1^* = \frac{\lambda}{(1+\alpha)^2}$ and $v_2^* = \frac{\alpha\lambda}{(1+\alpha)^2}$, we can rewrite this as $-\frac{\lambda^2}{(1+\alpha)^2} - k(\lambda)$. The first portion is

[22]Note that the objective functions are undefined for $v_1 = v_2 = 0$. Regardless of the rule we use to define each objective function's value in that instance, $v_1 = v_2 = 0$ cannot be part of any equilibrium—as is standard for contest success functions, the marginal value for investing a slight amount overwhelms the cost to do so and is therefore a profitable deviation for at least one player.

strictly concave, while the second is weakly concave. Therefore, the addition of the two is strictly concave. This implies that the objective function has a unique solution. Call that solution λ^*.

The remaining task is to solve for the bargaining game. We first look at the politician's accept or reject decision. Accepting yields $-(v_1^* + v_2^*) + bc$. Substituting for the equilibrium levels of violence, we have:

$$-\frac{1}{(1+\alpha)} + bc \tag{5}$$

Meanwhile, the politician receives $-(v_1^* + v_2^*) - k(\lambda^*)$ if it rejects. Again substituting for the equilibrium levels of violence, we have:

$$-\frac{(\lambda^*)^2}{(1+\alpha)} - k(\lambda^*) \tag{6}$$

Using Equations 5 and 6, the politician is willing to accept any a bribe if:

$$-\frac{1}{(1+\alpha)} + bc \geq -\frac{(\lambda^*)^2(1)}{(1+\alpha)} - k(\lambda^*)$$

$$b \geq \underline{b} \equiv \frac{[1+(\lambda^*)^2]\left(\frac{1}{(1+\alpha)}\right) - k(\lambda^*)}{c} \tag{7}$$

That leaves Cartel 1's bribe decision. To analyze this, we first need to find 1's payoffs in the violence decision subgames with and without enforcement. Without enforcement, recall that the equilibrium levels of violence are $v_1^* = \frac{1}{(1+\alpha)^2}, v_2^* = \frac{\alpha}{(1+\alpha)^2}$. Plugging these into Cartel 1's utility function gives:

$$\frac{1}{(1+\alpha)^2} \tag{8}$$

In contrast, with enforcement, the equilibrium levels of violence are $v_1^* = \frac{\lambda}{(1+\alpha)^2}, v_2^* = \frac{\alpha\lambda}{(1+\alpha)^2}$. Thus, Cartel 2's utility function for an unsuccessful bribe is:

$$\frac{\lambda^*}{(1+\alpha)^2} \tag{9}$$

Combining Equations 8 and 9, Cartel 1's utility differential between successful and unsuccessful negotiations equals:

$$\bar{b} \equiv \frac{1-\lambda^*}{(1+\alpha)^2} \tag{10}$$

This is also the maximum bribe Cartel 1 is willing to pay. Using Equations 7 and 10 as the constraints, a mutually acceptable bargain exists if:

$$\underline{b} < \bar{b}$$

$$c > c^* \equiv \frac{[1 + (\lambda^*)^2](1 + \alpha) - k(\lambda)(1 + \alpha)^2}{1 - \lambda^*}$$

So if $c > c^*$, Cartel 1 offers the politician's minimally acceptable amount ($\underline{b}$), and the politician accepts. If $c < c^*$, no bribe is mutually acceptable. Cartel 1 is then free to offer any bribe less than $\underline{b}$, guaranteeing the politician's rejection.

6.2 Proof of Proposition 2 and 3

To begin, let $b' = \frac{1+(\lambda^*)^2}{c'(1+\alpha)} - \frac{k(\lambda^*)}{c'}$ and $b'' = \frac{1+(\lambda^*)^2}{c(1+\alpha)} - \frac{k(\lambda^*)}{c}$. These values represent the minimally acceptable bribe to the more corrupt and the less corrupt types. Note that $b'' > b'$, so it costs more to bribe the less corrupt type.

No equilibria exist in which Cartel 1 offers a value not equal to b'' or b'. To see why, consider proof by cases. If Cartel 1 offers $b > b''$, both types accept. Cartel 1 receives $\frac{1}{(1+\alpha)^2}$ for the remainder of the game. However, Cartel 1 could alternatively offer the midpoint between that offered bribe and b''. Because that value is still strictly greater than b'', both types still accept. Cartel 1 in turn receives $\frac{1}{(1+\alpha)^2}$. But note that it receives this same payoff but pays a strictly smaller bribe. This is a profitable deviation. Therefore, offering $b > b''$ is never optimal.

Next, offering $b < b'$ is not optimal either. Such an offer induces both types to reject. Cartel 1's payoff therefore equals $\frac{\lambda^*}{(1+\alpha)^2}$. In contrast, consider an offer $b \in (b', b'')$ instead. That amount induces the more corrupt type to accept and the less corrupt type to reject. In turn, Cartel 1's payoff is equivalent if it is facing the less corrupt type. However, with positive probability, it is facing the more corrupt type. Because that offer is in the bargaining range for the more corrupt type, Cartel 1 earns strictly more than in this case than if bargaining fails. This is a profitable deviation. Therefore, offering $b < b'$ is not optimal.

Finally, consider $b \in (b', b'')$. As discussed above, such an offer induces the more corrupt type to accept and the less corrupt type to reject. Now consider a deviation to the midpoint between that offer and b'. This amount is still strictly greater than b' and strictly less than b''. Consequently, the more corrupt type still accepts and the less corrupt type still rejects. Cartel 1's payoff for the contest portion of the game remains the same. However, it pays a strictly smaller bribe to the more corrupt type. This is a profitable deviation. Therefore, offering $b \in (b', b'')$ is not optimal.

That information means that strategies can only satisfy equilibrium conditions if Cartel 1 offers b' or b''. In the first case, note that the weak type is indifferent between accepting and rejecting; in the second case, the strong type is indifferent. For reasons standard to ultimatum games like this one, no equilibria exist when one of those types rejects with positive probability when indifferent. This leaves two possibilities: Cartel 1 offers b'' and both types accept with certainty and Cartel 1 offers b', the more corrupt type accepts with certainty, and the less corrupt type rejects.

To see which offer prevails under equilibrium conditions, note that offering b'' yields Cartel 1 a flat payoff of $\frac{1}{(1+\alpha)^2} - b''$. Offering b' leads to a probabilistic outcome: Cartel 1 receives $\frac{1}{(1+\alpha)^2} - b'$ with probability p and $\lambda\left(\frac{1}{(1+\alpha)^2}\right)$ with probability $1-p$. As such, making the safe offer is optimal if:

$$\frac{1}{(1+\alpha)^2} - b'' > p\left(\frac{1}{(1+\alpha)^2} - b'\right) + (1-p)\left(\frac{\lambda^*}{(1+\alpha)^2}\right)$$

$$p < p^* \equiv \frac{\frac{1-\lambda^*}{(1+\alpha)^2} - b''}{\frac{1-\lambda^*}{(1+\alpha)^2} - b'} \tag{11}$$

By analogous argument, Cartel 1 offers b' if $p > p^*$. □

6.3 Proof of Proposition 4

Rewriting b' and b'' from Equation 11 in its unreduced form yields:

$$p < \frac{\frac{1-\lambda^*}{(1+\alpha)^2} - \frac{1+(\lambda^*)^2}{c(1+\alpha)} - \frac{k(\lambda^*)}{c}}{\frac{1-\lambda^*}{(1+\alpha)^2} - \frac{1+(\lambda^*)^2}{c'(1+\alpha)} - \frac{k(\lambda^*)}{c'}} \tag{12}$$

Because we care about how this function behaves as $c' - c$ decreases, we implicitly need to know how the cutpoint behaves as c' decreases and as c increases. This is easy to show since both the numerator and denominator must be positive for the parameter space. As c' decreases, the size of the optimal bribe against the more corrupt type increases. That in turn decreases the value of the denominator, increasing the size of the fraction overall. Meanwhile, as c increases, the size of the optimal bribe against the less corrupt type decreases. That in turn increases the value of the numerator, again increasing the size of the fraction overall. Both of these effects make it easier to fulfill the inequality overall.

In relating this to the equilibrium level of violence, decreasing the bandwidth of possible types ($c' - c$) either has no effect because it does not change whether p^* is greater or less than p or it changes p from being greater than p^* to less than. Therefore, the level of violence is weakly decreasing in the bandwidth. □

7 Robustness Checks

When presenting quantitative models, it is important to test the robustness of the conclusions to alternate model specifications. Here, we present a first differences model and a naïve OLS model. The results in both cases are highly similar to those presented in the paper, suggesting that our results are quite robust to alternate specifications.

7.1 First Differences

Our first robustness check is a first differences time series pooled ordinary least squares (OLS) regression. Unlike standard OLS, a first difference model subtracts the observed values of the dependent and independent variables in $t = 1$ from $t = 2$. In the process of subtracting, taking the first difference removes all invariant, unit-specific factors, denoted by θ.[23] This is because all of the factors contained within θ do not change between time periods, meaning they reduce to zero.

We estimate the predicted murder level in municipality i in year t with Equation 13:

$$\Delta Murder_{it} = \Delta\beta_0 + \beta_1 \Delta Tenure_{it} + \Delta f(\gamma) + \Delta\epsilon_{it} \tag{13}$$

The results from this model are presented in Table 3:

Table 3: Fixed Effects OLS of Incumbency's Effect on Violence with First Differences

	Dependent variable:	
	Murder	
	(1)	(2)
D.Congress Tenure	0.05**	0.05**
	(0.03)	(0.03)
District FE	Yes	Yes
Time Splines	No	Yes
Observations	25,517	25,517

*p<0.1; **p<0.05; ***p<0.01
Standard errors (clustered on municipality) reported in parentheses.
Estimates for cubic restricted time splines not reported.

[23] First differences are not the only estimation technique to control for unobserved heterogeneity. Many scholars use fixed effects to do so. Although we present results using district fixed effects in the online appendix, we believe fixed effects' assumption that the error term is serially independent to be harder to justify. Murder rates in $t = 0$ likely are likely highly predictive of violence in $t = 1$. As serial correlation incorrectly decreases the coefficients' standard errors, this is a serious specification issue. First differences, in contrast, are more robust to violations of this assumption (Liker, et al. 1985).

7.2 Fixed Effects OLS

In this section, we estimate use the following equation to estimate a fixed effects OLS regression:

$$Murder_{it} = \beta_0 + \beta_1 Tenure_{it} + f(\gamma) + \sum_{j=1}^{n} \theta_j + \epsilon_{it} \tag{14}$$

The results, presented in the table below, are also statistically significant and in the right direction.

Table 4: Fixed Effects OLS of Incumbency's Effect on Violence

	Dependent variable:		
	Murder		
	(1)	(2)	(3)
Tenure	0.84***	0.47***	0.26**
	(0.07)	(0.11)	(0.11)
PRI			4.21***
			(0.68)
ElectionYear			0.93**
			(0.41)
Observations	27,861	27,861	27,861
Municipal FE	Yes	Yes	Yes
Time Splines	No	Yes	Yes
R^2	0.01	0.01	0.01
Adjusted R^2	0.01	0.01	0.01

*p<0.1; **p<0.05; ***p<0.01

Standard errors (clustered on municipality) reported in parentheses.

Estimates for cubic restricted time splines not reported.

7.3 Power Transition Years

Empirical models that use tenure as a key independent variable inevitably face problems coding transition periods. With elections midway through the election year, it is difficult for the researcher to know exactly whom to assign the murders to. In the interest of completeness, the main model included the transition years. However, there are two relevant alternative coding schemes. The first assigns all the murders to the party in power at the beginning of the year. Some may find this coding scheme preferable because fresh leaders may not have held office long enough to sway policy in a meaningful way. Running the main model specification with the alternative coding scheme generates the following results:

Table 5: Fixed Effects OLS with Lagged DV Subsetting Transition Years

	Dependent variable:	
	Murder	
	(1)	(2)
Tenure	0.17^{**}	0.17^{**}
	(0.07)	(0.07)
Murder_{-1}	0.84^{***}	0.84^{***}
	(0.01)	(0.01)
Observations	25,517	25,517
R^2	0.59	0.59
Adjusted R^2	0.54	0.54

*p<0.1; **p<0.05; ***p<0.01
Standard errors (clustered on municipality) reported in parentheses.
Estimates for cubic restricted time splines not reported.

As Table 5 shows, our Tenure variable remains positive and significant with this alternate specification.

The second alternative scheme is the most conservative option available. It removes all transition years from the data. In other words, if a party does not hold office for the entire year, we subset it out of the analysis. Note that this is not the same as removing all election years—for elections where the incumbent party wins, we know who is responsible for the murders in that year.

As Table 6 illustrates, the model is robust to this specification. Moreover, the effect is stronger here. While it is important not to overanalyze these differences, we would expect this

Table 6: Fixed Effects OLS of Incumbency's Effect on Violence with Lagged DV

	Dependent variable:	
	Murder	
	(1)	(2)
Tenure	0.41***	0.20**
	(0.06)	(0.08)
$Murder_{-1}$	0.84***	0.83***
	(0.005)	(0.005)
Observations	23,985	23,985
Municipal FE	Yes	Yes
Time Splines	No	Yes
R^2	0.59	0.59
Adjusted R^2	0.53	0.53

*p<0.1; **p<0.05; ***p<0.01

Standard errors (clustered on municipality) reported in parentheses.

Estimates for cubic restricted time splines not reported.

to be true given our mechanism. After all, if transition years mix the "correct" assignment of homicides, the model would systematically underreport the actual effect due to newer leaders being unfairly assigned murders from their predecessor.

7.4 Geospatial Dependence

In the previously discussed models, our estimation strategies depend upon two assumptions that may be violated in our data. First, we assume that the effect of congressional tenure is consistent across units. Second, we assume that each municipality is statistically independent, i.e. that there is no spatial autocorrelation. Examples of such spatial autocorrelation include community and spillover effects. With spatially dependent data, estimated coefficients can be unstable and estimated measures of model fit can be inflated. The independence assumption, moreover, might be especially hard to justify in the case of political violence. Due to the quality of their security institutions, their terrain, or social structures, certain regions might be more prone to experience violence than others (Fearon and Laitin 2003). Regions with these characteristics might be more susceptible to diffusion from neighboring units. As the inclusion of such units would violate ordinary least squares' (OLS) independence assumption, we perform spatial statistics and visually plot residuals to check for evidence of such spatial

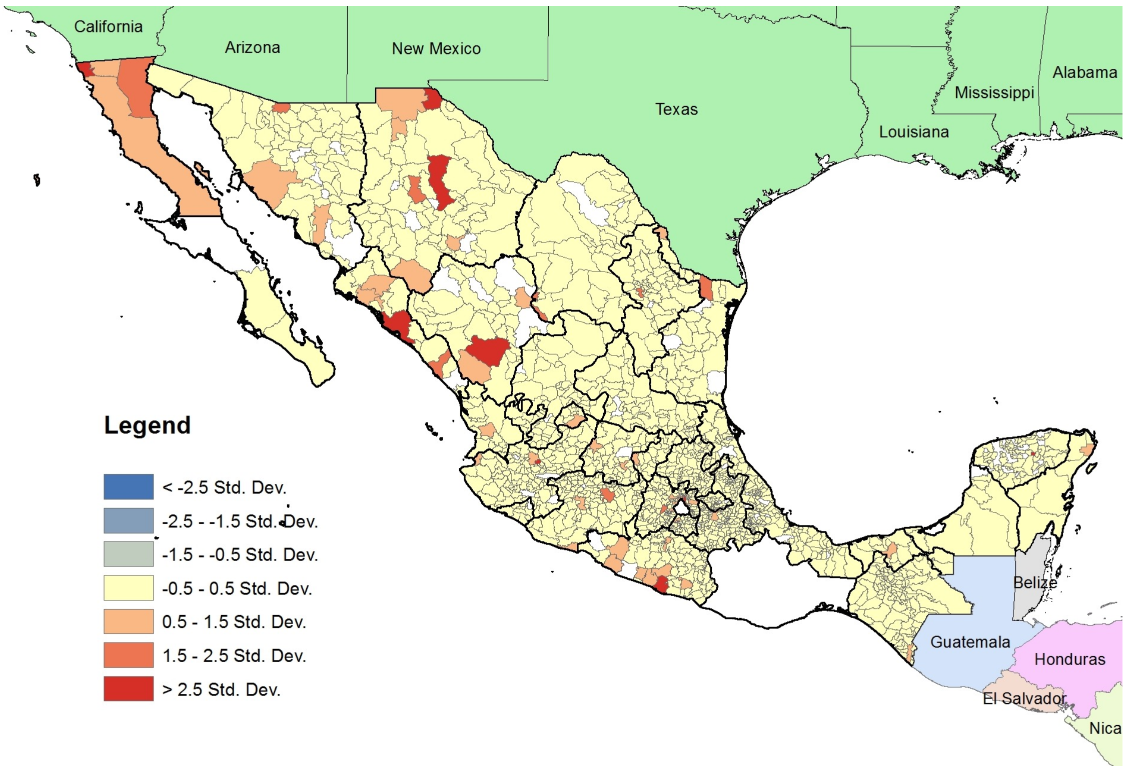

Figure 6: OLS residuals by municipality for 2008. Note that Oaxacan municipalities do not follow a consistent naming pattern and are excluded from this analysis.

autocorrelation. As a first cut, Figure 6 plots the residuals from a spatial bivariate OLS regression with data from 2008:

Although Figure 6 does not show unambiguous evidence of spatial dependence, it does appear that the effect of tenure on violence is greater in border regions and in the north of the country. It also shows that the center of the country is less influenced by tenure. For the reasons stated above, it is possible that our estimated coefficients are unstable due to our inclusion of data from border municipalities. To systematically test whether our results are robust to controlling for spatial autocorrelation, we take two steps. First, we estimate a geographically weighted regression (GWR). Second, we test for spatial dependence in the residuals with Moran's I, which is a measure of spatial autocorrelation. Following Fotheringham, et al. (1998), we estimate our GWR using the following equation:

$$Murder_{it} = a_0(u_i, v_i) + \sum_k a_k(u_i, v_i) Tenure_{ik} + \epsilon_{it} \tag{15}$$

where "(u_i, v_i) denotes the coordinates of the ith point in space and $a_k(u_i, v_i)$ is a realization of

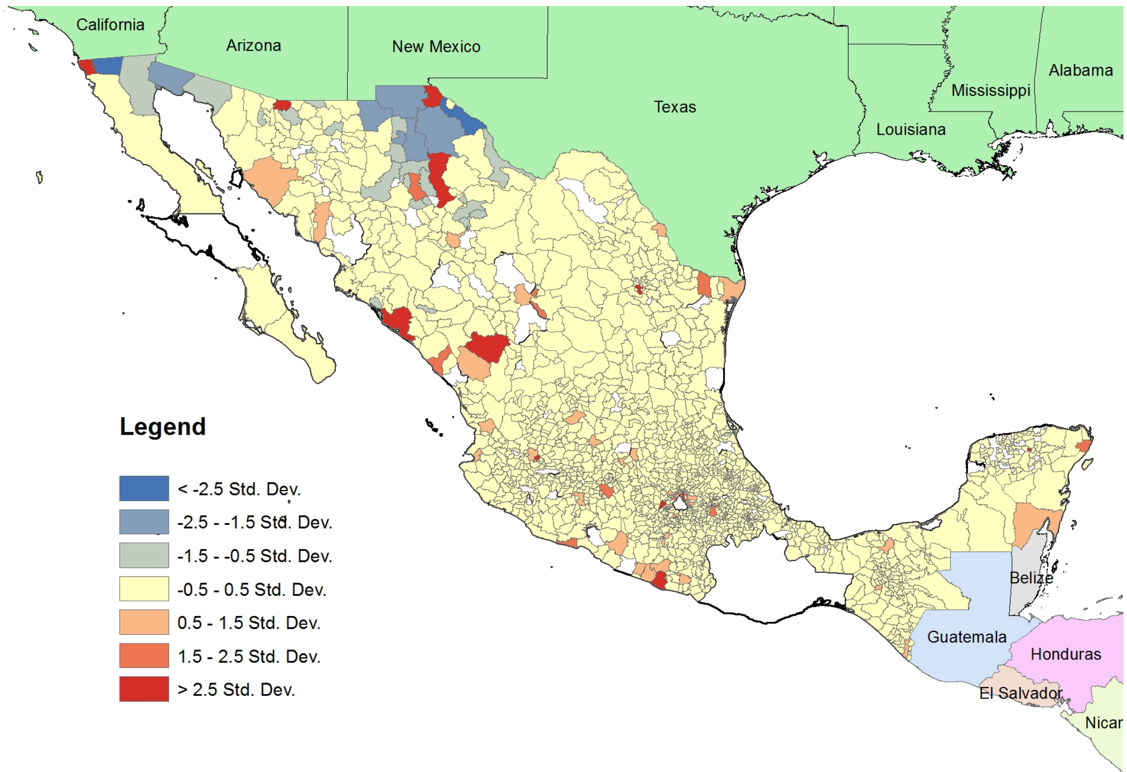

Figure 7: Coefficient estimated by GWR for all municipalities in 2008. Note that Oaxacan municipalities do not follow a consistent naming pattern and are excluded from this analysis.

the continuous function $a_k(u, v)$ at point i" (Fotheringham, et al. 1998, 1907). The estimated coefficients from this regression are displayed in Figure 7:

Unlike the potentially problematic clustering in the OLS model's residuals, evidence of clustering in the estimated coefficients from our GWR is far less obvious. As shown in Figure 7, municipalities with positive coefficients — such as Tijuana and Ciudad Juárez — appear to be surrounded by a randomly distributed mixture of positive and negative units. Beyond visual inspection, we can test whether these observed coefficients are geosnatially clustered with Moran's I statistic. Moran's I is a means of detecting the presence of multidimensional correlation in geospatial data (Paradis 2014). Moran's I, as defined by Moran (1950) is:

$$I = \frac{N}{\sum\limits_{i=1}\sum\limits_{j=1} w_{ij}} \frac{\sum\limits_{i=1}\sum\limits_{j=1} w_{ij}(X_i - \bar{X})(X_j - \bar{X})}{\sum\limits_{i=1}(X_i - \bar{X})^2} \tag{16}$$

where N is the number of units in the sample; X is our variable of interest (in this case, Tenure), $\bar{X}$ is the mean of X; and w_{ij} is an index of spatial weights. As the null hypothesis is no spatial autocorrelation, the expected value of I_0 is defined as $I_0 = -1/n - 1)$. The expected value of 1_0 is known, we can test for a statistically significant difference between the observed I ($\hat{I}$) and I_0. When $I_0 > \hat{I}$, it suggests evidence of positive spatial correlation. In contrast, when $1_0 < \hat{I}$, it is evidence of negative spatial correlation. Finally, when I_0 is not statistically distinguishable from $\hat{I}$, we cannot reject the null hypothesis that the data is randomly distributed spatially (Paradis 2014).

In this case, we use Equation 16 to estimate the Moran's I for our model. Our estimated I — i.e. $\hat{I}$ —is 0.002. We then test whether this estimated value is statistically distinguishable from the expected value of I under the null hypothesis, which is -0.0005. The p-value of the difference between these two values is 0.20, which means we cannot reject the null hypothesis. Although this does not definitely prove that there is no spatial dependence, it strongly suggests that there is no statistical evidence for it that is discernible in our data.