6 Appendix: Formal Proofs

6.1 Proof of Proposition 1

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

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

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

$$\frac{v_1}{(v_1+v_2)^2} - 1 = 0 \tag{3}$$

Using Equations 2 and 3 as a system of equations, the unique solution pair is $v_1^* = \frac{1}{(1+\alpha)^2}, v_2^* = \frac{\alpha}{(1+\alpha)^2}$. Note that when the politician accepts the bribe, $\alpha = \underline{\alpha}$ and therefore the solution pair is $v_1^* = \frac{1}{(1+\underline{\alpha})^2}, v_2^* = \frac{\underline{\alpha}}{(1+\underline{\alpha})^2}$.[38]

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

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:

$$-\frac{1}{1+\underline{\alpha}} + bc \tag{4}$$

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

$$-\frac{1}{1+\alpha^*} - k(\alpha^*) \tag{5}$$

[38]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.

Using Equations 4 and 5, the politician is willing to accept any a bribe if:

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

$$b \geq \underline{b} \equiv \frac{\frac{1}{1+\underline{\alpha}} - \frac{1}{1+\alpha^*} - k(\alpha^*)}{c} \tag{6}$$

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 $v_1^* = \frac{1}{(1+\underline{\alpha})^2}, v_2^* = \frac{\underline{\alpha}}{(1+\underline{\alpha})^2}$. Plugging these into Cartel 1's utility function gives:

$$\frac{1}{(1+\underline{\alpha})^2} \tag{7}$$

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

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

Combining Equations 7 and 8, Cartel 1's utility differential between successful and unsuccessful negotiations equals:

$$\bar{b} \equiv \frac{1}{(1+\underline{\alpha})^2} - \frac{1}{(1+\alpha^*)^2} \tag{9}$$

This is also the maximum bribe Cartel 1 is willing to pay. Using Equations 6 and 9 as the constraints, a mutually acceptable bargain exists if:

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

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

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. Note that Proposition 1 therefore applies to all cases where $c' < c^*$. □

6.2 Proof of Proposition 2 and 3

To begin, let $b' = \frac{\frac{1}{1+\underline{\alpha}} - \frac{1}{1+\alpha^*} - k(\alpha^*)}{c'}$ and $b'' = \frac{\frac{1}{1+\underline{\alpha}} - \frac{1}{1+\alpha^*} - k(\alpha^*)}{c}$. 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 $\frac{1}{(1+\underline{\alpha})^2}$ 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 $\frac{1}{(1+\underline{\alpha})^2}$. 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 $\frac{1}{(1+\alpha^*)^2}$. 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 $\frac{1}{(1+\underline{\alpha})^2} - b''$. Offering b' leads to a probabilistic outcome: Cartel 1 receives $\frac{1}{(1+\underline{\alpha})^2} - b'$ with probability p and $\frac{1}{(1+\alpha^*)^2}$ with probability $1-p$. As such, making the safe offer is optimal if:

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

$$p < p^* \equiv \frac{\frac{1}{(1+\underline{\alpha})^2} - \frac{1}{(1+\alpha^*)^2} - b''}{\frac{1}{(1+\underline{\alpha})^2} - \frac{1}{(1+\alpha^*)^2} - b'} \tag{10}$$

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

6.3 Proof of Proposition 4

Rewriting b' and b'' explicitly from Equation 10 yields:

$$p < \frac{\frac{1}{(1+\underline{\alpha})^2} - \frac{1}{(1+\alpha^*)^2} - \frac{\frac{1}{1+\underline{\alpha}} - \frac{1}{1+\alpha^*} - k(\alpha^*)}{c}}{\frac{1}{(1+\underline{\alpha})^2} - \frac{1}{(1+\alpha^*)^2} - \frac{\frac{1}{1+\underline{\alpha}} - \frac{1}{1+\alpha^*} - k(\alpha^*)}{c'}} \tag{11}$$

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 because 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 difference in 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 difference. □

7 Appendix: Model of Bribery and Competitive Effectiveness

In the model presented in the paper, the politician could control Cartel 1's marginal cost of violence. Alternatively, politicians may affect cartel competition by directly intervening on behalf one side. Such support directly increases the relatively likelihood that the preferred cartel prevails. We now analyze a model with this type of intervention to show that the empirical implications are equivalent.

The timing of the game is identical to before. Cartel 1 begins by offering a bribe b to the politician. The politician accepts or rejects. If the politician rejects, it sets a level of enforcement $\lambda \in [0, 1)$, while an accepted bribe yields $\lambda = 1$. Higher levels of λ decrease the probability that Cartel 1 succeeds in the contest. Enforcement comes at cost $k(\lambda)$, where k is differentiable everywhere and $k'(\lambda) < 0$ and $k''(\lambda) \geq 0$. After, the cartels simultaneously choose v_1 and v_2, with Cartel 1's marginal cost now exogenously set at $\alpha > 0$. Cartel 1's final payoff is $\lambda \frac{v_1}{v_1+v_2} - \alpha v_1$, while Cartel 2's final payoff is $1 - \frac{v_1}{v_1+v_2} - v_2$.

All propositions as written in the paper remain true for this alternative model. We sketch these out below.

7.1 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:

$$\frac{\lambda v_2}{(v_1 + v_2)^2} - \alpha = 0 \tag{12}$$

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:

$$\frac{\lambda v_1}{(v_1 + v_2)^2} - 1 = 0 \tag{13}$$

Using Equations 12 and 13 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 $v_1^* = \frac{1}{(1+\alpha)^2}, v_2^* = \frac{\alpha}{(1+\alpha)^2}$.

Now consider the politician's enforcement level, conditional on its rejection of the bribe. The politician'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 $-\frac{\lambda^2}{(1+\alpha)} - k(\lambda)$. The first portion is 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:

$$-\frac{1}{(1+\alpha)} + bc \tag{14}$$

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

$$-\frac{(\lambda^*)^2}{(1+\alpha)} - k(\lambda^*) \tag{15}$$

Using Equations 14 and 15, the politician is willing to accept any a bribe if:

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

$$b \geq \underline{b} \equiv \frac{\frac{1-(\lambda^*)^2}{(1+\alpha)} - k(\lambda^*)}{c} \tag{16}$$

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:

$$\frac{1}{(1+\alpha)^2} \tag{17}$$

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:

$$\frac{\lambda^*}{(1+\alpha)^2} \tag{18}$$

Combining Equations 17 and 18, Cartel 1's utility differential between successful and unsuccessful negotiations equals:

$$\bar{b} \equiv \frac{1-\lambda^*}{(1+\alpha)^2} \tag{19}$$

This is also the maximum bribe Cartel 1 is willing to pay. Using Equations 16 and 19 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. In turn, if $c' < c^*$, Cartel 1 is not willing to offer an acceptable bribe to either type. It is then free to offer any bribe less than $\underline{b}$, guaranteeing the politician's rejection.

7.2 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:

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

$$p < p^* \equiv \frac{\frac{1-\lambda^*}{(1+\alpha)^2} - b''}{\frac{1-\lambda^*}{(1+\alpha)^2} - b'} \tag{20}$$

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

7.3 Proof of Proposition 4

Rewriting b' and b'' from Equation 20 in its unreduced form yields:

$$p < \frac{\frac{1-\lambda^*}{(1+\alpha)^2} - \frac{1-(\lambda^*)^2}{c(1+\alpha)} - \frac{k(\lambda^*)}{c}}{\frac{1-\lambda^*}{(1+\alpha)^2} - \frac{1-(\lambda^*)^2}{c'(1+\alpha)} - \frac{k(\lambda^*)}{c'}} \tag{21}$$

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 because 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. □

8 Appendix: Empirical Robustness Checks

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

8.1 Reverse Causality

One potential concern with interpreting the above results is reverse causality; states with greater corruption are more likely to reelect mayors and Congresspeople. The logic in this case is intuitive. By accepting bribes from cartels, political parties have additional resources to spend on their campaign and buy votes. To assess this possibility, we create an indicator variable that takes a value of one whenever a political party is reelected within the same district. We then estimate whether corruption, as measured by Transparencía Mexico, affects the probability of reelection with a Probit. The results, presented in Table A1, show no such effect for corruption.

Table A1: Probit of Corruption's Effect on Probability of Reelection

	Dependent variable: Homicide	
	Reelection	
	(1)	(2)
Corruption	-0.00	-0.00
	(0.00)	(0.00)
Time Splines	No	Yes

*p<0.1; **p<0.05; ***p<0.01
Standard errors reported in parentheses.

8.2 Local capacity

One concern with our results is that they are biased by unobserved variation in local enforcement capacity. We explore this possibility in two ways. First, we generate state-year fixed effects.[39] These capture unobservable trends at the state-year level, such as policing levels or bureaucratic capacity. To explore whether there are non-linearities in these unobservable effects, we also created state-year^2 and state-year^3 effects. As shown in Table A2, controlling for these effects does not affect the sign or statistical significance of our main variable of interest, $Tenure$.

[39] In other words, we include a separate indicator variable that controls is unique for every state and year (i.e. Guerrero in 2001, 2002, etc.).

Table A2: Random Effects OLS of Federal Incumbency's Effect on Violence with State-Year FE

	Dependent variable: Homicide
	(1)
Tenure	0.15***
	(0.04)
L.Homicide	0.98***
	(0.02)
PAN	2.16***
	(0.52)
Corruption	-0.01
	(0.03)
Joint Rule	-0.30
	(0.32)
State Year FE	-0.00
	(0.00)
State Year FE^2	0.00
	(0.00)
State Year FE^3	-0.00
	(0.00)
N	21,079
R^2	0.64

*p<0.1; **p<0.05; ***p<0.01
Standard errors (clustered on municipality) reported in parentheses.

Second, we use data from Holland and Rios (2015) on changes in poverty and economic inequality. We expect that poorer areas will have lower bureaucratic and state capacity. Unfortunately, these measures are only available for years 2008-2010. Moreover, they are more likely to be available in PRI municipalities than PAN ones. While Tenure remains positive and significant, we are cautious in interpreting the results in Table A3:

Table A3: Fixed Effects OLS of Federal Incumbency's Effect on Violence with Economic Controls

	Dependent variable: Homicide
	(1)
Tenure	0.72***
	(0.20)
L.Homicide	1.16***
	(0.18)
Corruption	-0.03
	(0.06)
Joint Rule	-0.48
	(1.43)
Inequality	-2.25
	(5.67)
Poverty	2.28*
	(1.36)
N	7,720
R^2	0.64

*p<0.1; **p<0.05; ***p<0.01
Standard errors (clustered on municipality) reported in parentheses.

As before, *Tenure* remains positive and statistically significantly correlated with Homicide.

8.3 Border States

Exploiting the expiration of the 2004 Federal Assault Weapons Ban and California's state-level ban on the sale of assault weapons, Dube, Dube, and García-Ponce (2013) find that violence is higher in Mexican states that do not border California. They speculate that this because cartels and other violent actors can purchase assault weapons in Texas, Arizona, and New Mexico and carry them across the border. As a municipality's position along the US-Mexico border is invariant with time, all of our fixed effects regressions in the paper control for these potential cross-border spillovers. However, to separately control for the effect of border proximity on violence, we estimate two new random effects models. The first column includes an indicator variable for Mexican border states.[40]. While this new indicator variable is significant, the effect of tenure remains positive and significant. In the second column, we include a "spill over" indicator that omits Baja California — the state where Dube, Dube, and García-Ponce (2013) find that California's assault weapons sale ban decreases violence levels. As before, *Tenure* remains positive and statistically significant in Table A4:

[40] From west to east, these are Baja California, Sonora, Chihuahua, Coahuila, Nuevo León, and Tamaulipas.

Table A4: Random Effects OLS of Federal Incumbency's Effect on Violence with Border Effects

	Dependent variable: Homicide	
	(1)	(2)
Tenure	0.14***	0.14***
	(0.04)	(0.04)
L.Homicide	0.98***	0.98***
	(0.02)	(0.02)
PAN	1.97***	1.99***
	(0.46)	(0.47)
Corruption	0.02	0.01
	(0.02)	(0.02)
Joint Rule	-0.37	-0.36
	(0.31)	(0.32)
Border	1.91***	
	(0.67)	
Spill Over		1.73***
		(0.60)
N	21,079	21,079
R^2	0.64	0.64

*p<0.1; **p<0.05; ***p<0.01
Standard errors (clustered on municipality) reported in parentheses.

8.4 Simulated Confound

We do not find evidence that there are unobserved processes biasing our estimate of $Tenure's$ effect, we consider this possibility quantitatively by generate simulated data at various levels of correlation with $Tenure$ to estimate the robustness of our results to unobserved variable bias. As Table A5 shows, our results are robust up to a simulated confounder correlated with $Tenure$ at 0.50. Were this confounder to exist empirically, it would to predict reelection better than PAN (-0.05); PRI (-0.10); poverty (0.04); or inequality (-0.03). Given the extreme unlikelihood of such a variable existing, we are confident for the reasons above that our results are not biased by some alternate factor explaining which political parties succeed or fail in office.

Table A5: Fixed Effects OLS of Federal Incumbency's Effect on Violence with Simulated Confound

	Dependent variable: Homicide
	(1)
Tenure	0.002*
	(0.001)
L.Homicide	0.87***
	(0.03)
PAN	1.64**
	(0.71)
Corruption	-0.03
	(0.03)
Joint Rule	-0.11
	(0.42)
Simulated Confound (0.50)	-27.69
	(23.37)
Time Splines	Yes
Municipal FE	Yes
N	21,402
R^2	0.64

*p<0.1; **p<0.05; ***p<0.01
Standard errors (clustered on municipality) reported in parentheses.

8.5 Count Data

Our dependent variable, the number of homicides a municipality experiences in a year, cannot take negative values. As such, it is an example of count data. While OLS estimators are BLUE for count data, they could potentially be inefficient. Correcting for this inefficiency might affect estimated coefficient direction, significance, or show that the model does not appropriately fit the data. To address this possibility, we present the results here from a class of models specifically designed for count data — a cross sectional Poisson regression. These results, presented in Table A6, show that our main coefficient of interest is not affected by accounting for count data.

Table A6: Poisson Regression

	Dependent variable:	
	Homicide	
	(1)	(2)
Tenure	0.03***	0.03***
	(0.002)	(0.002)
L.Homicide	0.00***	0.00***
	(0.00)	(0.00)
Observations	25,538	24.224
Municipal FE	No	Yes
Time Splines	Yes	Yes

*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.

8.6 Lagged DV with Fixed Effects

The Nickell effect is a concern when including fixed effects in a lagged dependent variable model. By including fixed effects, one might bias results by artificially deflating the model's mean square error (MSE). This is particularly problematic with a small number of time periods in the data. Beck and Katz (2009) show that the bias produced by the Nickell effect is less than some of the proposed solutions, such as the Kiviet method. As such, this suggests that the results presented in the body of the paper are one modeling choice with TCSC data. It is still possible, however, that our results remain biased by the Nickell effect. To explore this possibility, we first reestimate our model with a lagged dependent variable without fixed effects. Second, we reestimate our results with year- and municipal-fixed effects. The results from these models are presented in Table A7:

Table A7: Lagged DV Without FE and Year FE

	Dependent variable:	
	Homicide	
	(1)	(2)
Congress Tenure	0.22***	0.47**
	(0.05)	(0.23)
L.Homicide Rate	0.99***	-
	(0.05)	-
Municipal FE	No	Yes
Year FE	No	Yes

*p<0.1; **p<0.05; ***p<0.01
Bootstrapped errors reported in parentheses.
Estimates for year fixed effects not reported.

As Table A7 shows, our results remain broadly consistent (and in fact the coefficient on Congress Tenure becomes larger) without municipal FE. This suggests that the results reported in the main body of the paper are not strongly biased by the Nickell effect. Second, we show that our choice to model time with splines is not necessary to obtain our results. After using year fixed effects, our variable of interest continues to have a positive and statistically significant effect on violence.

8.7 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 homicides 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 homicides 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. As Tables A8 and A9 show, our Tenure variable remains positive and significant with this alternate specification.

Table A8: Fixed Effects OLS with Lagged DV Subsetting Transition Years

	Dependent variable:
	Homicide
	(1)
Tenure	0.16**
	(0.07)
L.Homicide	0.93***
	(0.00)

*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.

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 homicides in that year. As Table A8 illustrates, the model is robust to this specification.

Table A9: Fixed Effects OLS of Incumbency's Effect on Violence with Lagged DV

	Dependent variable:	
	Homicide	
	(1)	(2)
Tenure	0.41***	0.20**
	(0.06)	(0.08)
L.Homicide	0.84***	0.83***
	(0.01)	(0.01)
Observations	23,985	23,985
Municipal FE	Yes	Yes
Time Splines	No	Yes

*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.

8.8 First Differences

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.[41] 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 homicide level in municipality i in year t with Equation 22:

$$\Delta Homicide_{it} = \Delta\beta_0 + \beta_1 \Delta Tenure_{it} + \Delta f(\gamma) + \Delta\epsilon_{it} \tag{22}$$

The results from this model are presented in Table A10:

[41]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. Homicide 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).

Table A10: Fixed Effects OLS of Incumbency's Effect on Violence with First Differences

	Dependent variable:	
	Homicide	
	(1)	(2)
D.Tenure	0.05**	0.05*
	(0.03)	(0.02)
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.

8.9 Lagged Independent Variable

One concern with our results is that they might be driven by reverse causality, i.e. a party's electoral fate is decided by the current level of violence. This would introduce simultaneity bias wherein an unobserved variable that explains both levels of violence and electoral success biases the error term in our results. One common way to address this issue is through an instrumental variable, where an exogenous variable is used to perform a two-stage OLS. Unfortunately, we are unable to identify a good instrument that predicts which parties are reelected and is also satisfies the exclusion restriction that it is uncorrelated with violence. To address this potential problem, we follow Clemens, et al. (2012, 1) and "avoid poor quality instrumental variables and instead address potential biases from reverse and simultaneous causation by the more transparent methods of lagging and differencing." This is a relatively common econometric technique to overcome simultaneity bias. We then estimate the predicted homicide level in municipality i in year t with Equation 23:

$$\Delta Homicide_{it} = \Delta\beta_0 + \beta_1 \Delta Tenure_{it-1} + \beta_2 \Delta Homicide_{it-1} + \Delta f(\gamma) + \Delta\epsilon_{it} \tag{23}$$

The results from this model are presented in Table A11:

Table A11: Lagged Fixed Effects OLS of Incumbency's Effect on Violence with First Differences

	Dependent variable:	
	Homicide	
	(1)	(2)
LD.Tenure	0.22***	0.29***
	(0.04)	(0.04)
LD.Homicide	-0.13***	-0.13***
	(0.05)	(0.04)
Municipal 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.

As these results show, the finding presented in the body of the paper likely does not result from simultaneity bias and is robust to a lagged IV.

8.10 Coding of Independent Variable

In the main body of the paper, we code our main independent variable — $Tenure$ — as linearly increasing year-on-year. One concern with this might be that each additional year a political party remains in office might not in fact provide substantial additional information to cartels. Instead, cartels might update only after each election and then adopt similar bargaining strategies with politicians during their time in office. To address this possibility, we recode $Tenure$ so that it increases only when a political party secures reelection. This reduces the scale of our new $Tenure^*$ variable from 11 to 4 ($\sigma = 0.95$, $\mu = 1.83$). We then estimate the effect of an increase in this $Tenure$ variable with a cross sectional Poisson regression. As before, increases in $Tenure$ are positively and significantly associated with an increased number of homicides in a given municipality.

Table A12: Recoding Key Independent Variable

	Dependent variable:	
	Homicide	
	(1)	(2)
$Tenure^*$	0.06***	0.06***
	(0.01)	(0.01)
L.Homicide	0.00***	0.00***
	(0.00)	(0.00)
Observations	25,943	24,629
Municipal FE	No	Yes
Time Splines	Yes	Yes

*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.

8.11 Irregular leadership change

Are there distinct patterns from the aftermath or electoral (regular) changes as compared to changes that follow irregular (assassination, death, etc.) changes? We explore this question using data on assassinations of Mexican mayors from the Justice in Mexico project. Beginning in 2005, they record 46 separate assassinations. We create an indicator variable, $Assassinate$, that takes a value of 1 in the year a mayor was assassinated in a given municipality. Controlling for municipalities that experienced an assassination does not affect our results substantively. If anything, the coefficient for $Tenure$ slightly increases. We report these results in Table A13:

Table A13: Fixed Effects OLS of Federal Incumbency's Effect on Violence with Assassination Control

	Dependent variable: Homicide
	(1)
Tenure	0.18**
	(0.07)
L.Homicide	0.86***
	(0.02)
PAN	1.83**
	(0.75)
Corruption	-0.03
	(0.03)
Joint Rule	-0.27
	(0.44)
Assassination	65.34
	(48.21)
N	21,079
R^2	0.64

*p<0.1; **p<0.05; ***p<0.01
Standard errors (clustered on municipality) reported in parentheses.

8.12 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 autocorrelation. As a first cut, Figure A1 plots the residuals from a spatial bivariate OLS regression with data from 2008:

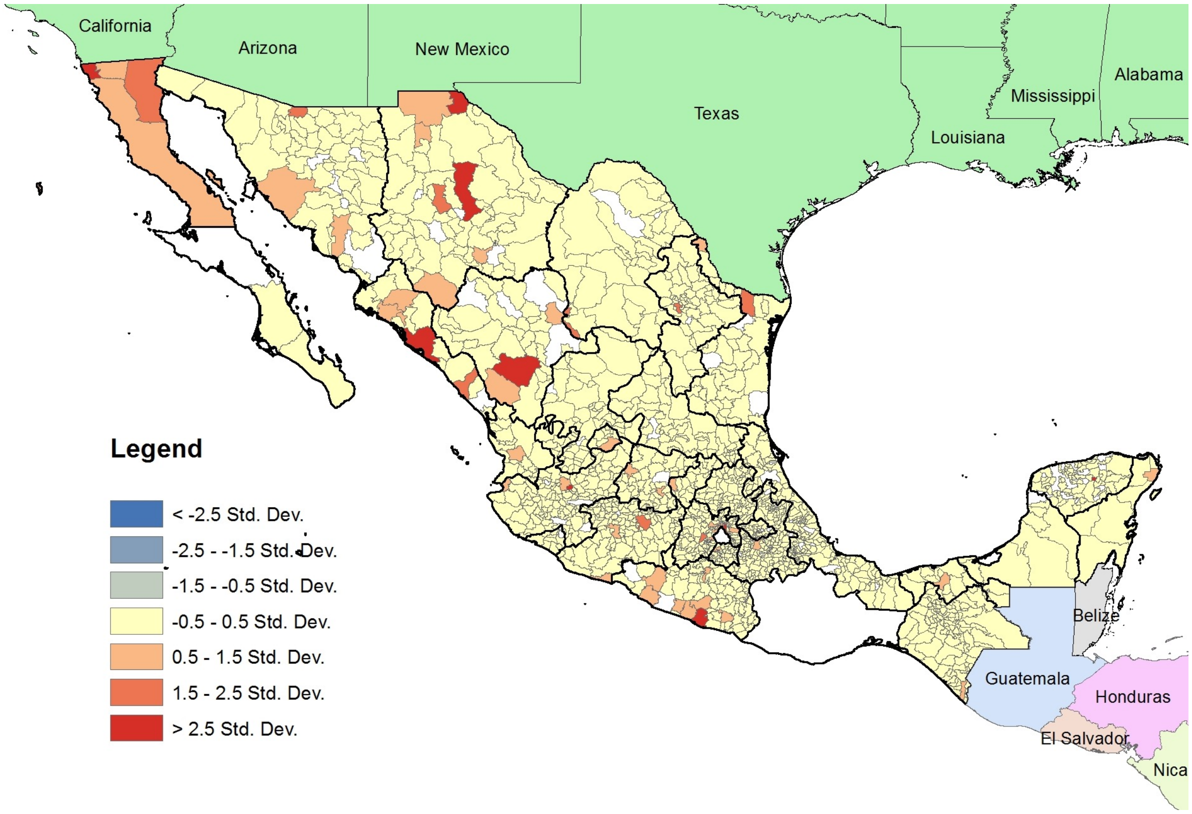

Figure A1: OLS residuals by municipality for 2008. Note that Oaxacan municipalities do not follow a consistent naming pattern and are excluded from this analysis.

Although Figure A1 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:

$$Homicide_{it} = a_0(u_i, v_i) + \sum_k a_k(u_i, v_i)Tenure_{ik} + \epsilon_{it} \tag{24}$$

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 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 A2:

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 A2, municipalities

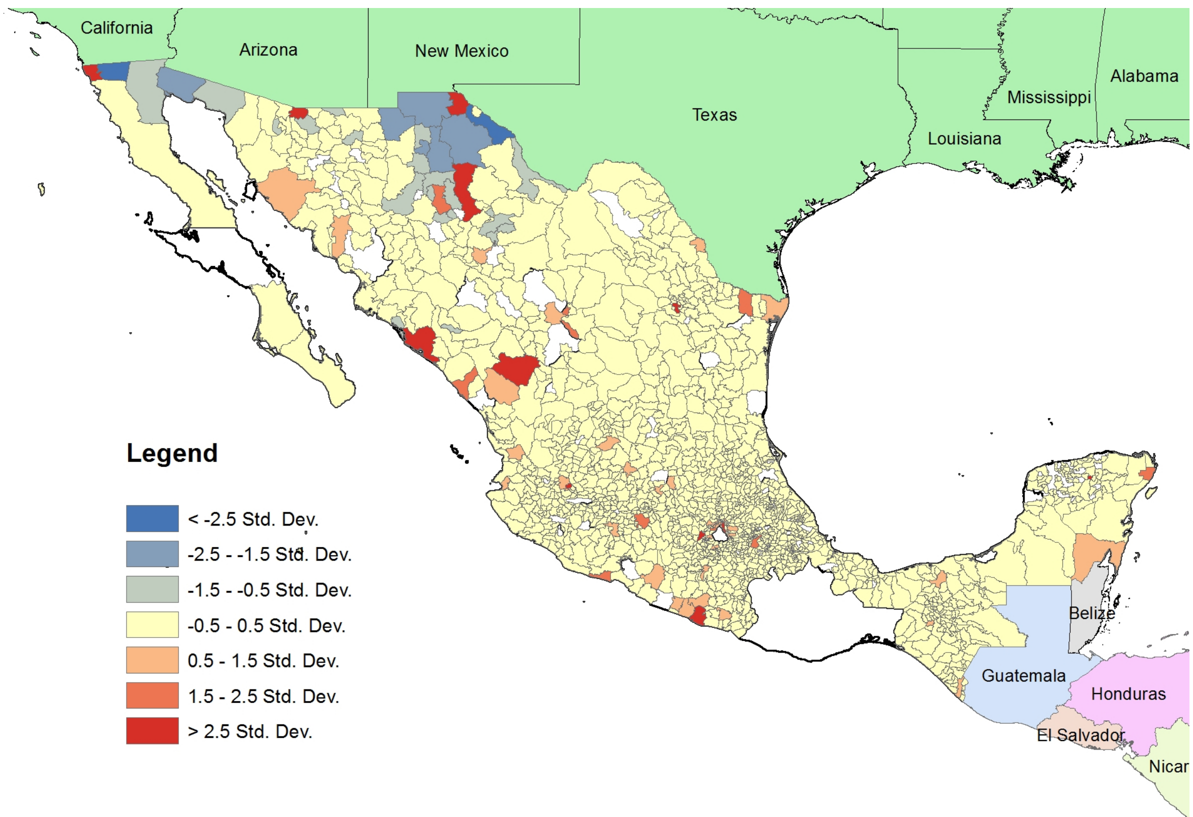

Figure A2: Coefficient estimated by GWR for all municipalities in 2008. Note that Oaxacan municipalities do not follow a consistent naming pattern and are excluded from this analysis.

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:

$$I = \frac{N}{\sum_{i=1}\sum_{j=1} w_{ij}} \frac{\sum_{i=1}\sum_{j=1} w_{ij}(X_i - \bar{X})(X_j - \bar{X})}{\sum_{i=1}(X_i - \bar{X})^2} \tag{25}$$

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 25 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.