Online Appendix to “Competition and Civilian Victimization”

A Other Tables

Table A1. Summary Statistics

Variable	Observations	Mean	Std. Dev.	Minimum	Maximum
Coca area	611	0.16	0.96	0	9.74
Distance army base	611	92.09	107.28	0	884.99
Distance FARC’s place of origin	611	199.95	170.89	0	750.76
Distance AUC’s place of origin	611	293.52	200.68	0	918.67
Distance Magdalena river	611	107.02	74.88	0.2	409.96
Gini	611	0.43	0.12	0	0.53
Population	611	1.36	0.9	0.15	6.46
Poverty	611	45.17	19.75	8.17	100
Royalties (oil)	611	0.03	0.12	0	0.83
Variation Liberal party vote share	611	0.08	0.05	0.01	0.34
Vote share left	611	4.6	6.93	0	64.82

This table presents summary statistics of controls.

Table A2. Selective and Non-Selective Victimization

	FARC		AUC	
	Selective	Non-Selective	Selective	Non-Selective
Panel A. Strategic factors				
Rival's selective victimization probability	2.197**	0.507	2.867***	1.944*
	(0.856)	(1.244)	(0.928)	(1.075)
Rival's non-selective victimization probability	1.572*	1.392	2.109*	1.881
	(0.9)	(1.239)	(1.12)	(1.253)
Panel B. Controls				
Coca area	-0.126	0.451	0.048	0.119
	(0.173)	(0.684)	(0.269)	(0.275)
Distance army base	-0.002	-0.0003	0.0005	-0.0003
	(0.002)	(0.003)	(0.001)	(0.001)
Distance group's place of origin	-0.002**	-0.004**	0.002**	0.002*
	(0.001)	(0.002)	(0.001)	(0.001)
Distance Magdalena river	0.002	0.001	-0.006***	-0.0001
	(0.002)	(0.004)	(0.002)	(0.002)
Gini	-0.885	16.28**	1.865	0.444
	(1.883)	(7.570)	(1.536)	(1.803)
ln(Population)	0.024	0.264	0.182	0.528***
	(0.171)	(0.300)	(0.185)	(0.188)
Period 2002-2005	0.261	0.312	-0.073	-1.167***
	(0.334)	(0.502)	(0.256)	(0.271)
Poverty	0.03***	0.01	-0.02**	-0.01*
	(0.008)	(0.01)	(0.01)	(0.01)
Royalties (Oil)	0.461	-2.99	-1.09	-1.4
	(1.473)	(18.09)	(1.05)	(1.445)
Variation Liberal party vote share	-0.364	-8.474*	6.627**	8.303**
	(3.41)	(4.644)	(2.933)	(3.268)
Vote share left	0.018	0.058**	0.008	0.005
	(0.024)	(0.028)	(0.025)	(0.023)
Log likelihood				-1001.32
Observations				611

This table presents maximum likelihood estimates of the parameters of the civilian victimization model with three actions (non violence, selective victimization, and non-selective victimization). The model includes region intercepts. Bootstrapped standard errors are in parentheses. *** p<0.01, **p<0.05, *p<0.1.

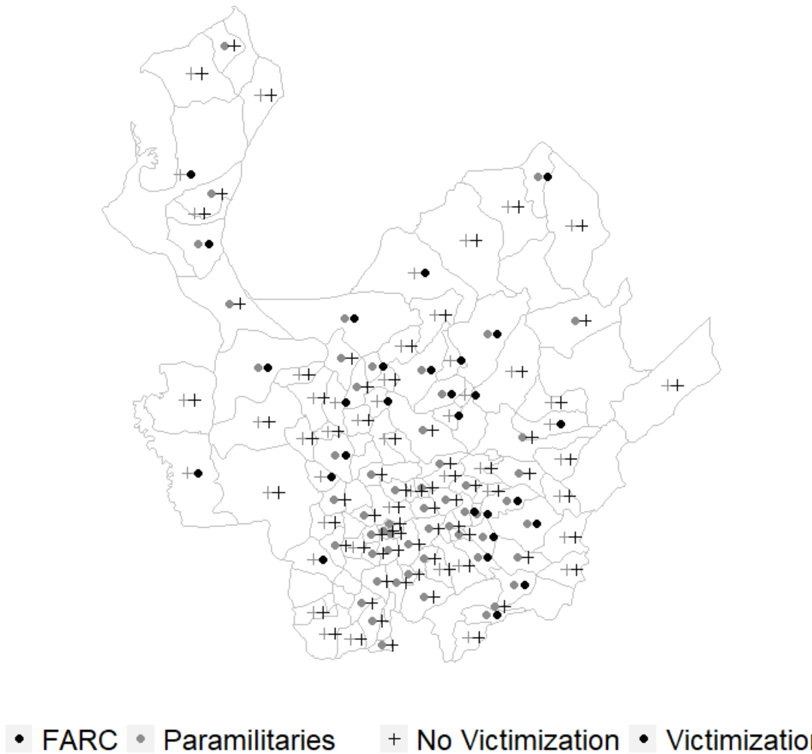
Table A3. Three-Player Game Estimates

	FARC	AUC	ELN
Panel A. Strategic factors			
AUC's Victimization Probability	2.387***		1.022
	(0.69)		(0.701)
ELN's Victimization Probability	1.956**	0.806	
	(0.852)	(0.787)	
FARC's Victimization Probability		3.026***	3.631***
		(1.01)	(1.134)
Panel B. Controls			
Coca area	1.196	-0.313	-0.19
	(0.936)	(0.697)	(1.34)
Distance army base	-0.002	0.001	0.001
	(0.002)	(0.002)	(0.002)
Distance group's place of origin	-0.002	0.001	-0.003**
	(0.002)	(0.001)	(0.002)
Distance Magdalena rive	0.004	-0.006*	-0.006
	(0.004)	(0.004)	(0.004)
Gini	13.958*	13.811**	5.159
	(7.449)	(6.999)	(8.537)
ln(Population)	0.061	0.58**	0.051
	(0.216)	(0.247)	(0.242)
Period 2002-2005	0.2	-0.593*	-0.64*
	(0.364)	(0.318)	(0.383)
Poverty	0.003	-0.019*	0.013
	(0.013)	(0.01)	(0.012)
Royalties (oil)	0.583	-0.271	-1.439
	(1.819)	(1.299)	(2.884)
Variation Liberal party vote share	-0.206	7.87	-3.049
	(4.735)	(5.615)	(4.859)
Vote share left	0.004	-0.017	-0.029
	(0.037)	(0.03)	(0.058)
Log likelihood			-443.45
Observations			320

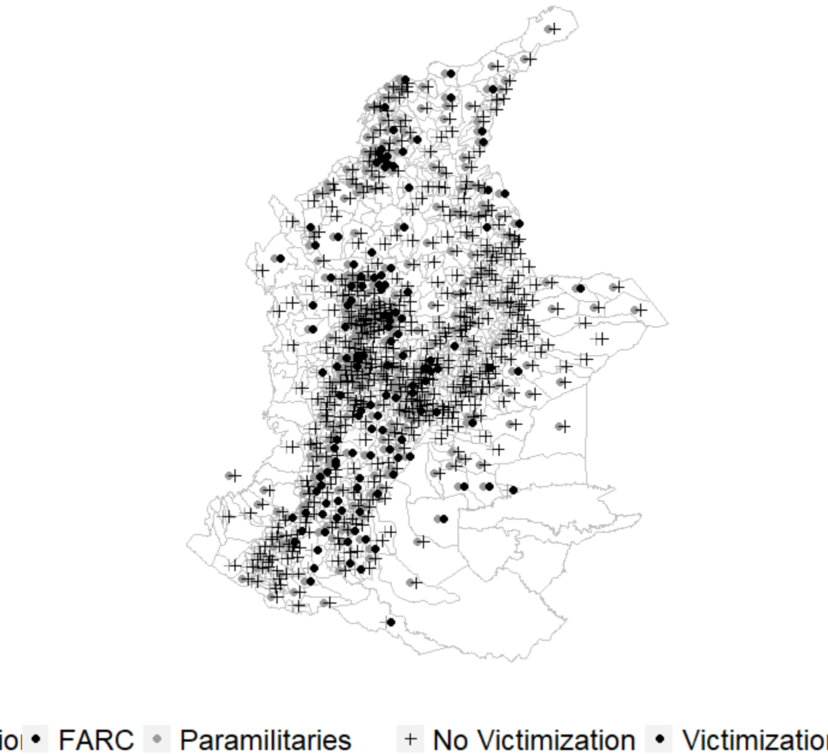
This table presents maximum likelihood estimates of the parameters of the three-player civilian victimization model. The model includes region intercepts. Bootstrapped standard errors are in parentheses. *** p<0.01, **p<0.05, *p<0.1.

B Other Figures

Areas of early influence are identified according to a variety of sources (Ugarriza and Ayala 2017; Alonso 1997; Bejarano 1997; Carlos Medina Gallego 2009; Centro Nacional de Memoria Histórica 2014; González 1991; Molano 2015; *Verdad Abierta* 2019).

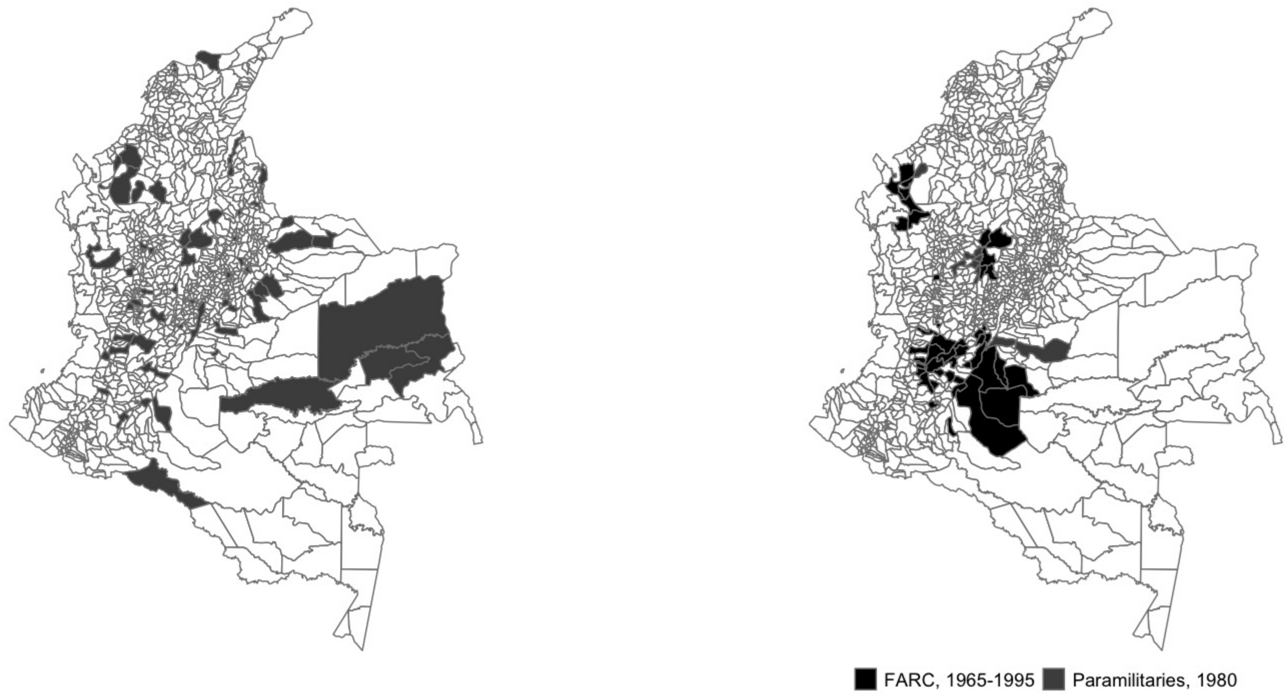


(a) Antioquia, 2002–2005.



(b) Colombia, 2002–2005.

Figure A1. Geographical variation in victimization strategies.



(a) Army bases, 2000.

(b) Armed groups' early areas of influence.

Figure A2. Locations used in covariate specifications.

C Pseudo-likelihood and Identification

As this is a game of incomplete information with simultaneous moves, our solution concept is Bayesian Nash equilibrium. Because each player's utility is stochastic, in equilibrium i has a probability $p_i \in (0,1)$ of choosing victimization. In fact, per McKelvey and Palfrey (1995), an equilibrium is characterized by a pair of probabilities satisfying a rational expectations condition:

$$\begin{aligned} p_1 &= \Pr(\mathbf{x}_1 \cdot \beta_1 + p_2 \cdot \alpha_1 > \epsilon_1(0) - \epsilon_1(1)), \\ p_2 &= \Pr(\mathbf{x}_2 \cdot \beta_2 + p_1 \cdot \alpha_2 > \epsilon_2(0) - \epsilon_2(1)). \end{aligned} \tag{A1}$$

Substantively, this means each actor's strategy maximizes its own utility given the probability with which it expects the other group to victimize civilians. The equilibrium condition requires that neither player systematically over- or under-estimates the other's likelihood of engaging in violence given the local environment.

Collect the parameters of the model in $\theta = (\alpha_i, \beta_i)_{i=1,2}$, and let $\Psi_i(p_{-i}; \mathbf{x}_i, \theta)$ denote the corresponding best-response probabilities:

$$\Psi_i(p_{-i}; \mathbf{x}_i, \theta) = \int \mathbb{I}[\mathbf{x}_i \cdot \beta_i + p_{-i} \cdot \alpha_i > \epsilon_i(0) - \epsilon_i(1)]\, dF(\epsilon_i), \tag{A2}$$

where $\mathbb{I}$ is the indicator function and F is the prior distribution of the stochastic shocks. The equilibrium condition of Equation A1 is equivalent to $p_i = \Psi_i(p_{-i}; \mathbf{x}_i, \theta)$ for each $i = 1, 2$.

We obtain estimates of equilibrium beliefs, $\hat{p}_i^{mt}$, using the Nadaraya-Watson kernel estimator and covariates x_i^{mt} for each group $i \in \{\text{FARC}, \text{AUC}\}$, in each municipality m and time period $t \in \{1998-2001, 2002-2005\}$. Collect these estimates in vector $\hat{\mathbf{p}}$. Then we write the (conditional) pseudo-likelihood function from Hotz and Miller (1993) as

$$\mathcal{L}(\theta \mid \hat{\mathbf{p}}, \mathbf{X}) = \prod_t \prod_m \prod_i \Psi_i(\hat{p}_{-i}^{mt}; \mathbf{x}_i^{mt}, \theta). \tag{A3}$$

Given the definition of Ψ_i in Equation A2, Equation A3 has a natural interpretation: it is the likelihood assuming that each actor best responds to the equilibrium beliefs estimated in the first step. If $\hat{p}_i^{mt} = p_i^{mt}$ for all i, m and t, then it is the true likelihood in the data generating process. Furthermore, because ϵ_i^{mt} are drawn from the type one extreme value distribution and are independent across actions, the integral in Equation A2 takes the standard logistic form:

$$\Psi_i(\hat{p}_{-i}^{mt}; \mathbf{x}_i^{mt}, \theta) = [1 + \text{Exp}\{-\mathbf{x}_i^{mt} \cdot \beta_i - \hat{p}_{-i}^{mt}\alpha_i\}]^{-1}. \tag{A4}$$

Equation A4 helps to illustrate an identification problem that can arise when estimating this game. The first-stage estimate $\hat{p}_{-i}^{mt}$ is a function of x_{-i}^{mt} (estimated via the kernel estimator), and i's local municipality payoff of using violence ($x_i^{mt} \cdot \beta_i$) is a function of x_i^{mt}. If these covaraites are the same ($x_i^{mt} = x_{-i}^{mt}$), the analysis might suffer a collinearity problem when trying to separately identify the effects of β_i and α_i on the observed choices. This problem would be particularly accute if we used a linear probability model in the first stage, in which case both the local payoffs and the first-stage choice probabilities would be linear combinations of covariates. As described in Bajari et al. (2010), a sufficient (but not necessary) condition to separate the effects of β_i and α_i is an exclusion restriction. That is, find some variable that affects group i's local payoffs but that do not enter group $-i$'s local payoff. In our analysis, distance from a group's early area of control serves as our restricted variable, and this is inspired from previous work. Gibilisco and Montero (N.d.), for example, estimate a game of major-power interventions into civil wars and use a major power's distance from a war to specify its costs of intervention. In economics, Ellickson and Misra (2011) describe how distance from regional headquarters or regions of early openings are used to specify costs of retailers in market-entry game. That is, Kmart (founded in Michigan) and Walmart (founded in Arkansas) might be more profitable in the Midwest and South, respectively.

D Victimization Data

Our information on civilian victimization comes from information provided by the Historic Memory Group (*Grupo de Memória Histórica*, GMH). This group was created by the Colombian government (Law 1448 of 2011) as part of the National Center for Historical Memory with the explicit aim of gathering and disseminating accurate information about the recent history of Colombia's internal conflict. We use 6 datasets from GMH on the following forms of violence: terrorist attacks, infrastructure and property attacks, incursions in population centers, clashes between groups, massacres, and selective killings. The datasets contain detailed information about violent incidents including geographic location, dates, groups involved, and the number of victims. Unlike fatalities, in some of these datasets, the total number of injured is included but it is not possible to dissagregate it into civilians and combatants. The GMH also publishes information on victims of mines and kidnappings. Although these forms of violence greatly affect civilians in the Colombian conflict, we focus on fatalities. This is because with mines the perpetrator is unknown, and with kidnappings the dataset is missing information on the confirmed perpetrator for most cases. Along with datasets from Restrepo, Spagat and Vargas (2004) and Palao-Mendizabal et al. (2019), the GMH datasets uses the periodical Noche y Niebla published by the NGO *Centro de Investigacíon y Educacíon Popular* (CINEP) as one of its main sources of information. The CINEP data have been used a source for documenting the conflict by the U.S. State Department, Human Rights Watch, and Amnesty International (Palao-Mendizabal et al. 2019). A key difference with other datasets is that the GMH complemented the CINEP data, with multiple sources including the Interamerican Commission of Human Rights, the Permanent Committee for the Defense of Human Rights, and importantly, official confessions from paramilitaries given to prosecutors as part as their demobilization process, among other publications and NGOs.[29] The information was compared across the sources by the GMH to avoid duplication. A more extensive number of sources and in particular the fact that members of one of the armed groups provided information regarding their crimes could reduce underreporting. A separate difference with the dataset of Restrepo, Spagat and Vargas (2004) is that the GMH datasets allows us to separate victims by the type of attack, a feature that is used in our analysis of selective and non-selective victimization. For example, a bomb that explodes and kills civilians will be classified as a terrorist attack while civilian fatalities in a clash between two groups would be included in the clashes dataset. The original CINEP data as well as the dataset in Restrepo, Spagat and Vargas (2004) are both more comprehensive in other aspects of the conflict like combatants' fatalities, and number of captured and injured combatants. The CINEP data also includes information on other forms of victimization like threats.

[29]For the full list of sources see (GMH 2013).

E Robustness Checks

E.1 Presence Criteria

We first consider the robustness of our results to alternative sample inclusion criteria. As described in the text, we use the CEDE dataset to identify municipalities in which armed groups operate but do not engage in systematic violence against civilians according to the Grupo de Memória Histórica information. Although our temporal aggregation should reduce the noise in the data, it is still possible that we have excluded municipalities where groups operated undetected or included those where groups tended not to operate but were detected in a one-off incident.[30] This can potentially bias our estimates. In this section, we examine these issues with two robustness checks. In the first one, we address potential under-inclusion by adding municipalities with only one group present to the baseline sample. To estimate our model in this case, we impute the beliefs of the group that is present about whether its rival would victimize civilians if it were present as well. In the second exercise, we vary our threshold on the number of incidents in the CEDE data that determine whether a group is considered to be present or not.

Recall that in our main analysis, we code an armed group as present in a municipality-period if the average number of CEDE incidents involving the group is above the pooled sample median or if the group engaged in civilian victimization there. With this coding, we identify 402 municipality-periods where only the FARC entered and 193 municipality-periods where only the AUC did. Our baseline sample excludes these observations, but in the first robustness exercise, we add them by imputing first-stage beliefs. If group i enters a municipality-period but its rival $j \neq i$ does not, we first assume i believes its rival j uses violence with probability $\frac{1}{2}$. That is, the group has maximum uncertainty about what its rival does. We also explore the scenario where i believes j uses violence with probability equal to j's pooled propensity to use violence, as measured in Table 1. These exercises add considerable noise to the estimation of the groups' best responses because, in more than half of the new sample, we are estimating the entering group's beliefs about a rival that does not in fact enter.

With the expanded sample and the two methods of imputed beliefs, we re-estimate the model, and the first four rows of Table A4 report the strategic interdependence parameters. When compared to Table 2, the expanded samples indeed attenuate the strategic complements in violence decisions, but the coefficients are still positive and significant at conventional levels. Consequently, we are confident saying that our main finding—that strategic incentives drive violence that would not otherwise occur—is not an artifact of our sample selection criteria.

In the second robustness check, we explore stricter or looser criteria of presence. Specifically, we examine requiring the average number of CEDE incidents to be above the 75th or 25th percentile in the sample, in contrast with the 50th percentile threshold in our main analysis. In loosening the threshold, we account for the possibility that low levels of reported group activity actually reflect a group's dominance (Kalyvas 2006). As noted by Ch et al. (2018), presence without reports of illegal activities is unlikely to occur in long periods of time because challenges by other groups eventually arise as well as opportunities to exploit that dominance by breaking the

[30] Using yearly data and survey based information in a subsample of municipalities, Arjona and Otálora (2011) find that indicators of presence based on the CEDE information underestimate presence of both FARC and AUC. Ch et al. (2018), however, show significant correlations of CEDE indicators of presence with measures based on areas where the groups demobilized for the periods 2007–2010 for the FARC and 1997–2002 for the AUC.

	Exercise	α_i	SE	t	p-value
Including single-entrant municipality-periods					
FARC	Random beliefs	1.20	0.50	2.40	0.02
AUC	Random beliefs	1.47	0.49	3.02	<0.01
FARC	Mean beliefs	1.16	0.60	1.95	0.05
AUC	Mean beliefs	1.85	0.63	2.93	<0.01
Alternative entry criteria					
FARC	75% percentile	2.39	0.87	2.76	<0.01
AUC	75% percentile	2.45	0.87	2.81	<0.01
FARC	25% percentile	1.44	0.48	3.01	<0.01
AUC	25% percentile	1.96	0.65	3.03	<0.01
Alternative threshold of victimization					
FARC	25% percentile	1.46	0.66	2.21	0.03
AUC	25% percentile	1.94	0.71	2.74	<0.01
Alternative first-stage model					
FARC	Leave-one-out	1.89	0.93	2.04	0.04
AUC	Leave-one-out	2.34	1.01	2.32	0.02

Table A4. Robustness of the strategic spillover parameter estimates across four alternative sample selection rules, first-stage estimates, and alternative definition of victimization. Standard errors are estimated from the outer product of gradients.

law. Nevertheless, our main results persist under these alternative criteria—see the middle four rows of Table A4. Even when municipality-periods where both groups are choosing not to victimize are eliminated by adopting a much more stringent requirement to classify groups as entering the municipality (CEDE incidents above the 75th percentile), we still find significant strategic complements. The strategic complements are also maintained if we include municipalities where one group's dominance might induce the other not to engage in many illegal activities captured by the CEDE indicators. In general, the coefficients reported in Table 2 are on the more conservative side.

E.2 Threshold of Victimization

Recall that in the baseline results we code victimization as occurring if the fraction of civilians killed intentionally by the group (victims of massacres, selective killings, or terrorist attacks) out of the total of the group's victims is greater or equal than the sample median in a given period. Formally, the indicator is coded as 1 if $\frac{\sum_{a\in C} K^i_{m,t,a}}{\sum_a K^i_{m,t,a}}$ is above or at the median of these fractions across municipalities and groups, where $K^i_{m,t,a}$ is the number of civilians killed by group i in the period t in municipality m in a type of attack a and C is the subset of types of attacks in which the intention was to kill civilians (massacre, selective, terrorist). Our conclusions still hold, however, if we define victimization as occurring when the fraction of civilians killed intentionally by the group in a given period is greater or equal than the 25th percentile. The results are reported in Table A4. Given that the median of the sample is already 1 (all killings are intentional) in both periods, the results will be identical as those reported in the paper if one was to use a higher percentile than the median to define the victimization threshold.

E.3 First Stage Estimates

We also would like to ensure that our finding of strategic complements is not driven by revenge dynamics or other non-strategic behavior within a municipality. Specifically, we want to ensure that our first-stage estimates of victimization probabilities only capture *ex ante* strategic expectations, not *ex post* observed violence within the given municipality. To this end, we recalculate our choice probability estimates using a leave-one-out procedure: to estimate each municipality's victimization probabilities, we take out-of-sample predictions from a model trained using only the data from outside that municipality. This rules out the possibility that our estimated *ex ante* probability of victimization is mistakenly picking up realized victimization within the given municipality. The estimated strategic complements in this robustness check are even stronger than those in the baseline model, as shown in the bottom rows of Table A4. This reinforces our claim that the correlation between FARC and AUC victimization is driven by strategic expectations in a competitive process.

The procedure for generating the first-stage estimates of choice probabilities is as follows. Remember that our goal in the first stage is to consistently estimate p_i^{mt}, the probability of victimization by group i in municipality m during period t. For each municipality m' in our data (441 total), we extract the subset of observations from other municipalities, i.e., in which $m \neq m'$. With this subset of data excluding observations from m', we train random forest models (one for FARC, one for AUC) to predict victimization as a function of the same set of covariates as in the baseline model.[31] Finally, for each time period t at which municipality m' enters our data, we let our first-stage estimate $\hat{p}_i^{m't}$ equal the out-of-sample prediction from our model of group i. Therefore, our first-stage estimate $\hat{p}_i^{m't}$ is not even partially a function of realized violence $v_i^{m't}$, as the model from which it is computed does not "see" data from within municipality m'.

Given the computational intensity of this procedure, resampling-based inference like the bootstrap is infeasible. We thus report nominal standard errors in Table A4. In order to conclude that the strategic spillover parameters were statistically insignificant at the 0.05 level, the true standard errors would have to be 1.41 times the nominal ones in case of the FARC, or 1.30 times in case of the AUC.

E.4 Control Specification

Finally, we explore the robustness of our findings to changes in the controls used in the players' utility functions. Our ability to identify the strength of strategic interdependence resides in isolating characteristics of the municipalities that could make different non-state armed actors treat civilians in the same way. Although our baseline specifications covers key demographic, economic, and political determinants of violence, as well as time-invariant region characteristics, it is possible there are still sources of unobserved heterogeneity driving our findings. To address this, we estimate a model in which we replace the region intercepts with state intercepts, a finer-grained geographical grouping. A drawback of this specification is that by focusing on within-state comparisons, the

[31] Our baseline specification uses kernel regressions rather than random forests. We use a random forest here due to technical problems extracting out-of-sample predictions from the kernel regression model.

explanatory power of the distance to an armed group's early place of influence is necessarily reduced. With this caveat in mind, Table A5 below presents the parameters of interest, showing that there are still strategic complements in the use of violence. We also estimate a model that includes interaction terms of the period 2002–2005 dummy with all time-invariant municipality characteristics, finding substantively similar results. Finally, to rule out the possibility that overall intensity of conflict is generating the observed patterns, we estimate a model that controls for such intensity at the beginning of the period. Because overall conflict intensity is a function of victimization, the specification implicitly assumes that the groups do not take into account how their actions in one period affects their future actions. With this important caveat in mind, we still find very similar estimates. We also see that initial intensity of conflict is not significant for any of the groups, indicating that the baseline specification captures the main determinants of victimization.

Table A5. Strategic victimization (alternative specifications)

	FARC	AUC
Model 1. State effects		
Rival's victimization probability	1.244***	1.223***
	(0.337)	(0.416)
Distance group's place of origin	-0.002	0.002
	(0.002)	(0.002)
Log likelihood		-600.37
Model 2. Time invariant controls interacted with period 2002-2005 dummy		
Rival's victimization probability	1.544***	1.937***
	(0.357)	(0.506)
Distance group's place of origin	-0.004**	0.0002
	(0.002)	(0.001)
Log likelihood		-644.57
Model 3. Initial intensity of violence as control		
Rival's victimization probability	1.447***	2.025***
	(0.353)	(0.488)
Distance group's place of origin	-0.002**	0.002**
	(0.001)	(0.001)
Initial violent incidents	0.015	0.008
	(0.01)	(0.012)
Log likelihood		-654.26

This table presents maximum likelihood estimates of the parameters of the civilian victimization model. Model 1 includes state intercepts and baseline controls. Model 2 includes region intercepts, baseline controls, and interactions of the period 2002-2005 dummy with all the time invariant controls. Model 3 includes the sum of all violent incidents involving the FARC, paramilitaries, and ELN at the beginning of the period. All models use 611 observations. Bootstrapped standard errors are in parentheses. *** p<0.01, **p<0.05, *p<0.1.

Table A6. Strategic and Bivariate Normal Models Comparison

Voung test (p-value)	Clarke Test (p-value)	AIC		BIC	
		Strategic	Binormal	Strategic	Binormal
0.634	0.037	1383.91	1385.40	1534.02	1531.10

The null hypothesis in Voung and the Clarke test is that the models are equivalent, $H_0 : E_0[ln(f/g)] = 0$, where E_0 denotes expectations over the true data generating process, f the likelihood of the strategic model, and g that of the bivariate normal. The alternative hypothesis is $H_f : E_0[ln(f/g)] > 0$. The model with the smallest AIC and BIC is preferred.

E.5 Non-Nested Model Comparison

We execute a comparison of our model with a simple bivariate normal model that effectively assumes victimization choices are not strategic (i.e., do not depend on the expectations of the other group's actions), but that allows the unobserved determinants of victimization of each group to be correlated with each other. For the bivariate normal model we include contextual variables and the distance to the group's areas of influence as determinants of victimization choices. Table A6 reports the results of this comparison. While the Voung test of non-nested models does not discriminate between the two and the BIC favor the bivariate probit, the Clarke test, and the Akaike Information criterion favor the strategic model. The Clarke test has been shown to perform better than the Vuong test when the empirical distribution of individual log-likelihood ratios has a high kurtosis coefficient (Clarke 2007), as is the case here where we find a kurtosis of 25.02.

E.6 Municipality-Year Results

In this appendix we assess the robustness of the results to choosing as unit of observation the municipality-year rather than the municipality-period.[32] Before we present these results, we discuss some limitations of this approach. In addition to the potential attenuation caused by mismeasurement described in the main text, having temporally more dissagregated observations introduces stronger time dependence. For example, if a group uses violence against civilians in one municipality-month, this may alter its costs and benefits of using violence in the same municipality but in the next month. While a common solution to account for this dependence is to include lags of the dependent variables, in our structural framework, however, such an approach would treat past endogenous actions as exogenous covariates and would imply that the groups do not take into account how their actions in one month affect payoffs in future months. This is an unappealing assumption given our view on the strategic nature of the AUC and FARC.

With these caveats in mind, Table A7 replicates the baseline analysis with municipality-year observations. Here, we see that the strategic interdependence parameters are positive and significant at conventional levels. Even though we still find evidence of strategic complementarities, the strategic coefficients are smaller, a finding that is consistent with the having a noisier measure of victimization with the municipality-year data. Figure A3 replicates the analysis of different types of violence against civilians. It shows that the strategic complementarities in the use of civilian

[32] We include the same control variables as in our baseline specification. Population and oil royalties are available every year at the municipality level. For the liberal party vote share, we used the one from the most recent election. For other controls that vary over time but that are not available every year, we use interpolations.

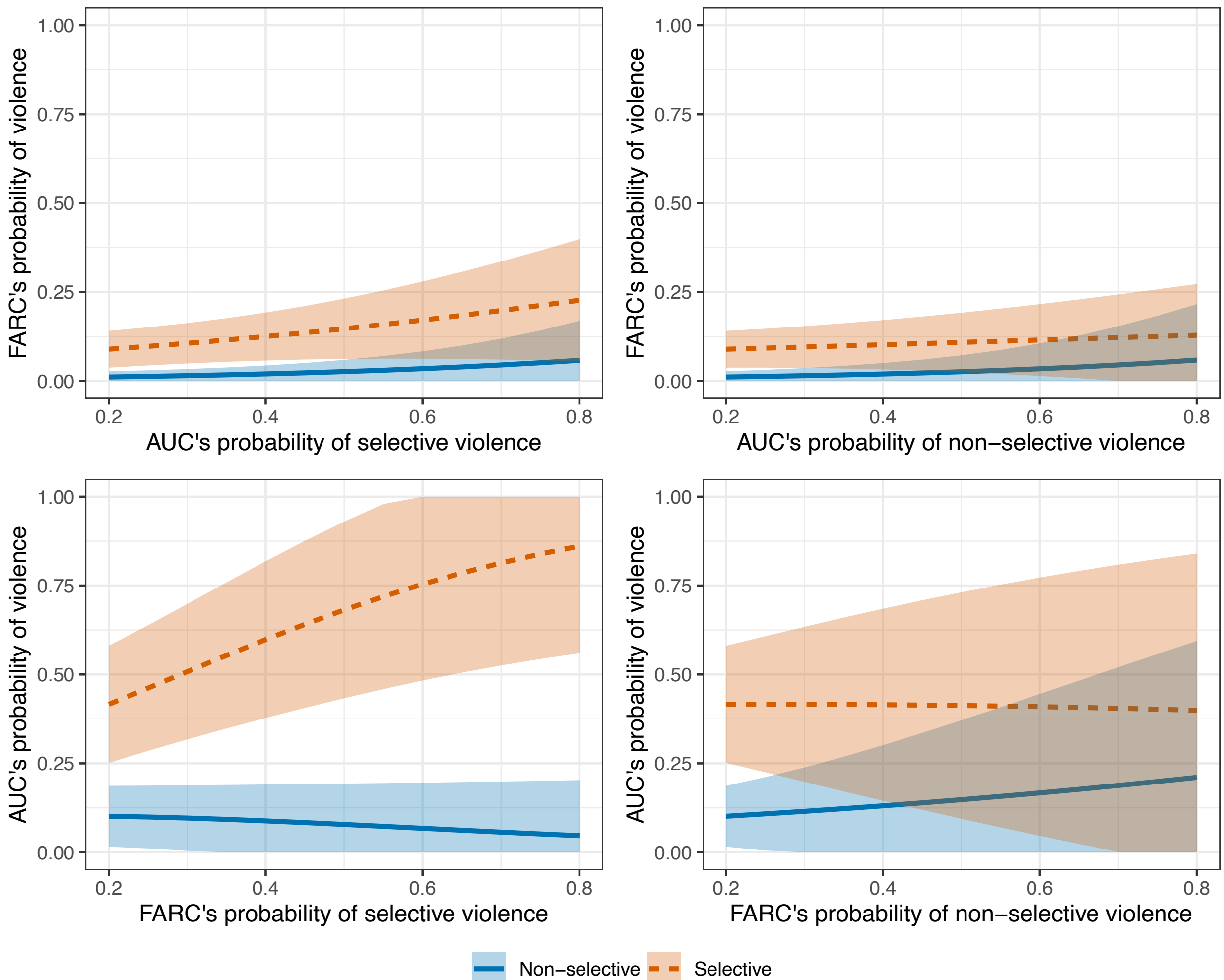
victimization are driven by selective victimization. When including the ELN as a third player, no strategic substitution between the ELN and the FARC violence is evident (Figure A4).

Table A7. Strategic Victimization (municipality-year)

	FARC	AUC
Panel A. Strategic factors: α_i		
Rival's victimization probability	**1.242 (0.213)**	**1.008 (0.181)**
Panel B. Controls		
Coca area	0.01 (0.084)	0.098 (0.081)
Distance army base	-0.001 (0.001)	-0.0004 (0.001)
Distance group's place of origin	**-0.002 (0.001)**	0.001 (0.001)
Distance Magdalena river	0.001 (0.001)	**-0.003 (0.001)**
Gini	1.19 (1.145)	0.951 (0.799)
ln(Population)	**0.186 (0.088)**	**0.44 (0.089)**
Poverty	**0.024 (0.005)**	-0.001 (0.004)
Oil royalties	0.888 (0.795)	**-1.174 (0.571)**
Liberal party vote share	0.011 (0.017)	0.023 (0.015)
Variation Liberal party vote share	0.942 (1.846)	**4.611 (1.517)**
Log-likelihood	-1363.05	
Observations	1318	

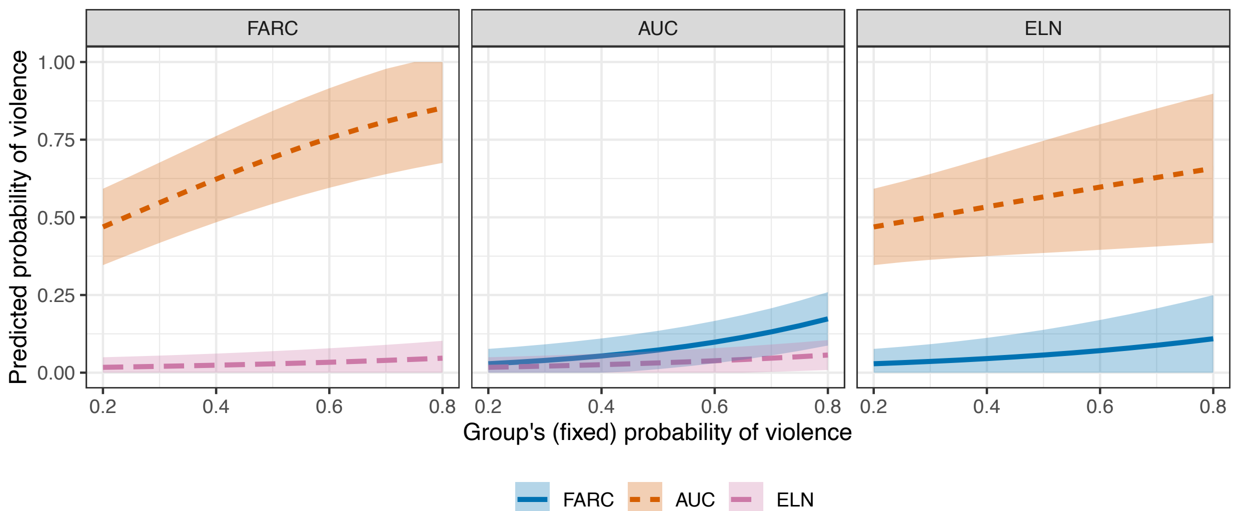
Estimates of the payoff parameters in the victimization model using municipality-year observations. The model includes region intercepts. Bootstrapped standard errors are in parentheses. Estimates in bold are statistically significant with $p < 0.05$. The results should be compared to Table 2 which uses municipality-period observations.

Figure A3. Strategic interdependence distinguishing selective and non-selective victimization (municipality-year).



Notes: Using municipality-year observations, we reestimate the model with two types of violence (selective and non-selective), and then use estimated model to predict how i's propensity of each violence type varies as it expects more or less violence from its rival $-i$. The results should be compared to Figure 1, which uses municipality-period observations.

Figure A4. Strategic interdependence adding the ELN as a third group (municipality-year).



Notes: Using municipality-year observations, we reestimate the model with three groups (FARC, AUC, and ELN), and then use estimated model to predict how i's propensity of each violence type varies as it expects more or less violence from its rival $-i$. The results should be compared to Figure 2, which uses municipality-period observations.

F Expanded Model

In this section, we expand our baseline model to incorporate an arbitrary number of groups and types of victimization. This version of the model covers the additional analyses in Section 6.

There are G groups, indexed by $i = 1, \ldots, G$. Each group simultaneously chooses a type of civilian victimization $v_i \in A_i = \{0, 1, \ldots, K\}$. Here, we interpret $v_i = 0$ as the choice not to victimize and $v_i > 0$ as a choice to employ victimization with type v_i. For example, in Section 6.1, we would set $K = 2$, where $v_i = 1$ represents selective victimization and $v_i = 2$ represents non-selective.

Payoffs are as follows:

$$u_i(v_i, v_{-i}, \epsilon_i) = \mathbf{x}_i \cdot \beta_i^{v_i} + \sum_{j \neq i} \alpha_{i,j}^{v_i, v_j} + \epsilon_i(v_i). \tag{A5}$$

In the above equation, $\beta_i^{v_i}$ captures the impact of contextual variables $\mathbf{x}_i$ on group i's payoff from choosing action $v_i \in A_i$. In addition, $\alpha_{i,j}^{v_i, v_j}$ captures the impact of group j's victimization choice of $v_j \in A_j$ on i's payoff from choosing $v_i \in A_i$.

As in the baseline model, we normalize the *ex ante* expected utility (i.e., before the stochastic shock) for no victimization to zero. That is, $\beta_i^0 = 0$ and $\alpha_{i,j}^{0, v_j} = 0$ for all i, all $j \neq i$, and all $v_j \in A_j$. In addition, we normalize $\alpha_{i,j}^{v_i, 0}$ to zero for all i, all $j \neq i$ and all $v_i \in A_i$, which is also carried over from the baseline model in Equation 1. Essentially, we can only identify the effect of j's victimization choice on i's utility relative to a baseline action v_j, where we use not committing violence $v_j = 0$ as the relative baseline action.

Collect the to-be-estimated payoff parameter in θ. As above, equilibria can be represented as choice probabilities satisfying a rational expectations condition. Let $p_i(v_i)$ denote the probability that Group i chooses $v_i \in A_i$. Let $\Psi_i(v_i, p_{-i}; \mathbf{x}_i, \theta)$ denote the corresponding best-response probabilities:

$$\Psi_i(v_i, p_{-i}; \mathbf{x}_i, \theta) = \int \mathbb{I}\left[v_i = \arg\max_{a_i \in A_i} \left\{ \mathbf{x}_i \cdot \beta_i^{a_i} + \sum_{j \neq i} \sum_{v_j \in A_j} \alpha_{i,j}^{a_i, v_j} p_j(v_j) + \epsilon_i(a_i) \right\} \right] dF(\epsilon_i)$$

An equilibrium is a vector of choice probabilities $(p_1, \ldots, p_G)$ such that for all i and all $v_i \in A_i$ we have

$$\Psi_i(v_i, p_{-i}; \mathbf{x}_i, \theta) = p_i(v_i).$$

When the action-specific payoff shocks $\epsilon_i(v_i)$ are drawn i.i.d. from the type-one extreme value distrubtion, $\Psi_i(v_i, p_{-i}; \mathbf{x}_i, \theta)$ takes the form:

$$\Psi_i(v_i, p_{-i}; \mathbf{x}_i, \theta) = \frac{\exp\left\{ \mathbf{x}_i \cdot \beta_i^{v_i} + \sum_{j \neq i} \sum_{v_j \in A_j} \alpha_{i,j}^{v_i, v_j} p_j(v_j) \right\}}{\sum_{a_i \in A_i} \left[\exp\left\{ \mathbf{x}_i \cdot \beta_i^{a_i} + \sum_{j \neq i} \sum_{v_j \in A_j} \alpha_{i,j}^{a_i, v_j} p_j(v_j) \right\} \right]}.$$

Additional References

Alonso, Manuel A. 1997. *Conflicto armado y configuración regional: El caso del Magdalena Medio*. Medellin, Colombia: Universidad de Antioquia.

Arjona, Ana M. and Laura Otálora. 2011. "Presencia vs. violencia: problemas de medición de la presencia de actores armados en Colombia.". http://focoeconomico.org/2011/12/20/presencia-vs-violencia-problemas-de-medicion-de-la-presencia-de-actores-armados-en-colombia/.

Bejarano, Jesús A. 1997. *Colombia : inseguridad, violencia y desempenño econoómico en las áreas rurales*. Bogotá, Colombia: Universidad Externado de Colombia.

Carlos Medina Gallego. 2009. *FARC EP. Notas para una historia política 1958-2008*. Bogotá, Colombia: Universidad Nacional de Colombia facultad de derecho ciencias políticas y sociales.

Centro Nacional de Memoria Histórica. 2014. *Guerrilla y población civil. Trayectoria de las FARC 1949-2013*. Bogotá, Colombia: Imprenta Nacional de Colombia.

Ch, Rafael, Jacob Shapiro, Abbey Steele and Juan F. Vargas. 2018. "Endogenous Taxation in Ongoing Internal Conflict." *American Political Science Review* 112(4):996–1015.

Clarke, Kevin A. 2007. "A Simple Distribution-Free Test for Nonnested Model Selection." *Political Analysis* 15(3):347–363.

Ellickson, Paul B. and Sanjog Misra. 2011. "Estimating Discrete Games." *Marketing Science* 30(6):997–1010.

Gibilisco, Michael and Sergio Montero. N.d. "Do Major-Power Interventions Encourage the Onset of Civil Conflict? A Structural Analysis." *Journal of Politics*. Forthcoming.

González, José J. 1991. "Espacio, sociedad y conflicto en Colombia. Las repúblicas independientes en Colombia 1955-1965." *Revista UIS Humanidades* 20(1):67–75.

Hotz, V. Joseph and Robert A. Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *The Review of Economic Studies* 60(3):497–529.

Kalyvas, Stathis N. 2006. *The Logic of Violence in Civil War*. Cambridge University Press.

McKelvey, Richard D. and Thomas R. Palfrey. 1995. "Quantal Response Equilibria for Normal Form Games." *Games and Economic Behavior* 10(1):6–38.

Molano, Alfredo. 2015. Fragmentos de la Historia del Conflicto Armado 1920-2010. In *Conflicto social y rebelión armada en Colombia. Ensayos críticos*. Bogotá: Gentes del común.

Ugarriza, Juan Esteban and Nathalie Pabón Ayala. 2017. *Militares y guerrillas. La memoria histórica del conflicto armado en Colombia desde los archivos militares 1958-2016*. Bogotá, Colombia: Editorial Universidad del Rosario.

Verdad Abierta. 2019. https://verdadabierta.com/. Accessed: 2019-02-04.