

ONLINE APPENDIX

The incursion of Leviathan: wartime territorial control and post-conflict state capacity in Peru

WIDER Working Paper 2022/136

Guillermo Kreiman*

November 2022

* Carlos III University, Madrid, Spain, gkreiman@clio.uc3m.es

This file contains additional information and a variety of results that complement the findings of the paper. First, I describe the territorial administrative changes that have taken place in Peru during the last decades, detailing how I dealt with this for the construction of the main database. Second, I include a codebook of the variables included in all the analyses, and I provide detailed descriptions of the main dependent and independent variables. Third, I report the descriptive statistics. Fourth, I deal with the relation and potential differential impact of wartime violence and wartime territorial control. Fifth, I test the potential heterogeneous effects in urban and rural areas. Sixth, I run a wide variety of robustness tests. Seventh, I test the fulfilment of some of the core assumptions of difference-in-differences models and try to deal with some concerns over endogeneity. And finally, I briefly describe the process of collection of qualitative data. This document includes:

- Appendix A – Changes in districts
- Appendix B – Codebook and main variables
 - B1. Codebook
 - B2. State capacity
 - B3. Territorial control
- Appendix C – Descriptive statistics
- Appendix D – Violence
 - D1. Violence by territorial control
 - D2. Violence over time and conflict time frames
 - D3. Violence or territorial control?
- Appendix E – Unit heterogeneity (urban–rural)
- Appendix F – Robustness tests
 - F1. DV specifications
 - F2. IV specifications
 - F3. Robustness controls
 - F4. Conflict time frames
 - F5. State bureaucrats and state security forces
 - F6. Provincial & departmental fixed effects
 - F7. Spill over effects
 - F8. Lagged DV models
 - F9. Ceiling effects
- Appendix G – DiD diagnostics & endogeneity
 - G1. Conditional parallel trends
 - G2. Lack of pre-treatment effects
 - G3. Covariate balance
 - G4. Matched sample
- Appendix H – Qualitative data

A Changes in districts

The administrative structure of Peru has been in constant change since the early 20th century. Three distinct divisions can be found: Departments, Provinces, and Districts, the latter being the lowest administrative unit of the country. Table A1 below shows the distribution of Peruvian departments, provinces, and districts during the last 80 years.

Table A1: Administrative structure of Peru 1940–2017

Date	N Departments	N Provinces	N Districts
June 1940	22	122	1064
July 1961	23	144	1491
June 1972	23	150	1676
July 1981	24	153	1680
July 1993	24	188	1793
August 2005	24	195	1811
October 2007	24	196	1833
October 2017	24	196	1874

Source: 'Perú: Proyecciones de Población por Departamento, Provincia y Distrito 2018-2020', page 23, and Instituto Nacional de Estadística e Informática.

During these decades, several districts have split, and others have merged, leading to an increase in the number of districts from 1064 in the 1940s to the current 1874. As stated in the paper, I make use of data from the Peruvian censuses between 1961 and 2007. Therefore, I use the number of departments, provinces, and districts from 1961 as the core reference. In order to ensure cross-district comparison, I follow a strategy of *minimum list*, which 'allows any local-level data collected at any point in time during the period considered to be effectively merged with other datasets' (Villamil 2020: 189). This strategy implies that if a district splits after 1961, it will be considered as merged, and that mergers that take place after this year will be considered as separate districts. Fine-grained information on individual territorial changes comes from the *Peruvian Instituto Nacional de Estadística e Informática*.

B Variables

B1 Codebook

State Bureaucrats: Number of state bureaucrats in a district. Source: Peruvian censuses between 1961 and 2007. See Appendix B2 below.

State Security Forces: Number of state security forces in a district. Source: Peruvian censuses between 1993 and 2007. See Appendix B2 below.

Electricity: Percentage of households with access to public electricity. Number of households with electricity by total number of households in a concrete district. Source: Peruvian censuses between 1961 and 2007. See Appendix B2 below.

Territorial Control: Type of territorial control during the conflict. Distinction between territories under state and insurgent control, as well as contested areas. Source: De la Calle (2017). See Appendix B3 below.

Illiteracy: Percentage of illiterate adult population. Illiteracy is defined as knowing how to write and read. Source: Peruvian censuses between 1961 and 2007.

Population Density: Number of people living in a district divided by the size of the district. Concretely, number of inhabitants by square kilometer. Sources: Peruvian censuses between 1961 and 2007, De la Calle (2017), and RIMISP (*Centro Latinoamericano para el Desarrollo Rural*).

% Spanish: Percentage of people speaking Spanish as their main language by district. Source: Peruvian censuses between 1961 and 2007.

Political Competition: Distance between incumbent party and the core opposition party. Source: Jurado Naciones de Elecciones Peru; Sulmont and Bazán (2011).

Killings: Total number of killings occurring during the civil war by district. Source: *Comisión de la Verdad y Reconciliación* (CVR 2003).

SL Killings: Total number of civilian killings committed by Sendero Luminoso during the civil war by district. Source: *Comisión de la Verdad y Reconciliación* (CVR 2003).

State Killings: Total number of civilian killings committed by the state during the civil war by district. Source: *Comisión de la Verdad y Reconciliación* (CVR 2003).

Urban–Rural: Percentage of the population living in an area categorized as urban with respect to the total population of the district. Source: RIMISP (*Centro Latinoamericano para el Desarrollo Rural*).

No Mayor 1989: Proxy of Sendero Luminoso territorial control, no mayor in office in 1989. Source: Albertus (2019) – Pareja and Gatti (1990): Pareja, Piedad, and Aldo Gatti. 1990. *Evaluación de las elecciones municipales de 1989*. Lima: *Instituto Nacional de Planificación*.

3 Guerrilla Events: Placebo of Sendero Luminoso territorial control, active presence in a district with more than three attacks. Source: Albertus (2019) – Truth and Reconciliation Commission report database: *Comisión de la Verdad y Reconciliación* (CVR). 2004. *Informe Final*. Lima, Peru. <https://sites.google.com/a/pucp.pe/informe-final-de-la-cvr-peru/>.

Distance to Provincial Capital: Distance from the district centroid to the capital of the province. Source: De la Calle (2017).

Slope: Slope of the district measured in degrees. Source: Albertus (2019) – FAO’s Global Agro-Ecological Zones database: Food and Agriculture Organization of the United Nations. ‘Global Agro-Ecological Zones’. Version 3.0. <http://www.fao.org/nr/gaez/en>

Cultivable Land: Cultivable land as percentage of the district area. Source: Albertus (2019) – FAO’s Global Agro-Ecological Zones database: Food and Agriculture Organization of the United Nations. ‘Global Agro-Ecological Zones’. Version 3.0. <http://www.fao.org/nr/gaez/en>

Vote Marxism 1980: Marxist vote share, 1980 Presidential Election. Source: Albertus (2019) – *Jurado Nacional de Elecciones: Jurado Nacional de Elecciones*. 1984. *Resultados de las elecciones municipales de 1980*. Lima, Peru.

Size: District area in square kilometers. Source: De la Calle (2017).

Height: District average altitude. Source: De la Calle (2017).

Density Roads 1973: Road density in 1973. Source: Albertus (2019) – Touring and Automobile Club and the *Banco de Crédito del Perú*: Touring and Automobile Club and the *Banco de Crédito del Perú*. 1973. *Carreteras del Peru*. Lima, Peru.

Caudillismo: Historical caudillo presence. Source: Albertus (2019) – Kammann (1982): Kammann, Peter. 1982. *Movimientos campesinos en el Peru: 1900-1968*. Lima: Universidad Mayor de San Marcos.

Social Movements: Number of previous social movements. Source: Albertus (2019) – Kammann (1982): Kammann, Peter. 1982. *Movimientos campesinos en el Peru: 1900-1968*. Lima: Universidad Mayor de San Marcos.

B2 State capacity

In this paper I make use of two distinct proxies of state capacity, each of them capturing distinct dimensions of the concept. First, as a proxy of state control, I make use of a variable on the number of state bureaucrats per district. And second, as a proxy of state legitimacy and the delivery of public goods and services, I use a variable on the percentage of households per district with access to public electric services. In this subsection, I succinctly describe the measurement details for these two variables.

It is important to remember that the dates of pre-war measures of state capacity diverge due to data limitations. With regards to the number of state bureaucrats, I rely on the information provided by the census of 1961, as there is no disaggregated information on state bureaucrats at the district level in the 1972 census. And second, I make use of data on access to electricity with information coming from the census of 1972. Disaggregated information on state bureaucrats and state security forces is only available for the 1993 and 2007 censuses.

State bureaucrats

- 1961 census: ‘*Servidores del Estado: Empleados y Obreros en la República*’ (‘State Servants: Employees and Workers in the Republic’). I sum up both the figures for employees and workers.
- 1981 census: ‘*Rama de Actividad = 9100: Administración Pública y Defensa*’ (‘Branch of Activity = 9100: Public Administration and Defense’).
- 1993 census:
 - State Personnel: ‘*Miembros del poder ejecutivo, de los cuerpos legislativos y personal de la administración pública*’ (‘Members of the executive branch, of the legislative bodies and personnel of the public Administration’).
 - State Security Forces: ‘*Miembros de las fuerzas armadas y de seguridad*’ (‘Members of the armed and security forces’).
- 2007 census:
 - State Personnel: ‘*Miembros del poder ejecutivo, de los cuerpos legislativos y personal de la administración pública*’ (‘Members of the executive branch, of the legislative bodies and personnel of the public Administration’).
 - State Security Forces: ‘*Defensa y Seguridad*’ (‘Defense and Security’).

Electricity

- 1972 census: '*Viviendas Particulares con Alumbrado Eléctrico*' ('Private Homes with Electric Lighting').
- 1981 census: '*Tipo de Alumbrado de la Vivienda - Eléctrico*' ('Type of Home Lighting - Electric').
- 1993 census: '*La Vivienda Tiene Alumbrado Eléctrico*' ('The House Has Electric Lighting').
- 2007 census: '*La Vivienda Tiene Alumbrado Eléctrico*' ('The House Has Electric Lighting').

B3 Territorial control

The development of adequate measures of subnational levels of territorial control has been one of the most difficult tasks for conflict scholars during the last decades (Anders 2020). Despite the fundamental importance played by the segmentation and fragmentation of sovereignty within national territories during internal armed conflicts (Kalyvas 2006: 88-89), the underlying complexity of territorial dynamics hampers the capacity of capturing reliable subnational measures on the distribution of wartime territorial control. Several variables have been used as proxies for territorial control. Most prominently, the majority of prior studies have made a direct connection between levels of violence and territorial control (Murshed and Gates 2005; Ishiyama and Widmeier 2013), implicitly assuming that rebels would commit more violent attacks in those areas under their control. However, recent research has shown that insurgent liberated areas are characterized by the presence of order, stability, and a lower level of violence than other type of territories (Arjona 2016), more particularly, contested areas. Other proxies of territorial control include levels of civilian displacement, the quality and collection of census data, as well as reports from national intelligence units (Hatlebakk 2007; Rubin 2020).

In this paper, I use one of the most precise, albeit still imperfect, proxies used by conflict scholars to capture variation in wartime territorial dynamics: electoral boycotts (Allison 2010; Hatlebakk 2007; De la Calle 2017).² Non-state armed organizations relying on irregular war as its main technology of rebellion have been characterized by their territorial expansion, due to the set of benefits linked with the liberation of territory (e.g. resource extraction, civilian mobilization, safe havens, etc.). In direct connection with the assertions of prior authors, I argue that electoral boycott is an adequate proxy for capturing variation in types of territorial control during conflict, given the relevance of this tactic in liberated areas as well as in areas over which the insurgents aim to establish full territorial control (contested territories). Several insurgent groups rely on a tactic of boycotting local, regional, and national elections as a feature of its strategy against state structures. Indeed, the goal is to 'broadly delegitimize the political system' and it is 'often aimed at shaking the political foundations of a government' (Matanock and Staniland 2018: 713).

The use of electoral boycott has been widespread across insurgent groups. For example, the CPN-M boycotted several elections in Nepal since its inception (Hatlebakk 2007; Holtermann 2016);

² The boycott of local, regional, or national elections does not preclude the potential presence of participatory mechanisms in rebel-controlled areas, and more concretely, the implementation of so-called 'rebel elections'. These become another component of the multifaceted concept of rebel governance. As described by Cunningham et al. (2021: 7), insurgent elections serve 'as a strategic effort by the rebels to assert their authority and manage their control over a population while simultaneously increasing their legitimacy. Seen in this light, rebel elections represent not so much an incipient wartime democracy as a classic instance of the instrumental use of a political project by actors who seek to strengthen and legitimate their rule.'

the LTTE followed similar tactics in non-state-controlled municipalities in Sri Lanka (Schaffer 1995; Mampilly 2011: 107); while Steele and Schubigger (2018: 594) and Cunningham (2014: 179) also argue that elections were boycotted by insurgent groups in Bangladesh or India. Insurgent groups from Latin America were not an exception to this. Allison (2010) shows how the FMLN in El Salvador boycotted the municipal elections of 1984; the FARC similarly did so by the end of the 1990s in Colombia (Vargas Meza 1998); as well as the EZLN in Mexico (Paulson 2001). More relevantly for the purposes of this study, plenty of evidence shows how Sendero Luminoso aimed to boycott any type of election that took place either at the local, regional, or national level during its period of operation (Burt 2011: 127–28; Degregori et al. 1996: 61; McCormick 1992: 46, 97; CVR 2003; Montoya 2008: 69; McClintock 1998: 79–80).³ Indeed, this was part of the strategy of Sendero since its inception. In one of the initial documents of the group, they stated that: ‘The elections are, then, an instrument of the counterrevolutionary war [...] The boycott, therefore, is an incontrovertible reality and its indisputable success; and it blatantly shows how the policy of hindering elections, undermining them, and preventing them wherever possible is highly fruitful’ (CVR 2003: Parte I: Sección II, Capítulo 1: 35).

But what is the link between electoral boycott and territorial control? The main logic underlying this connection relates to the ‘degree of success of electoral boycotts promoted by the insurgents’ (De la Calle 2017: 432): insurgents prevent the running of elections in their areas of full control while actively try to avoid their adequate functioning in areas where they aim to expand their influence, in other words, areas of contested control. These assertions are supported by plenty of evidence from different cases and have been advocated by a variety of scholars. Schaffer (1995: 414) shows the direct connection between the lack of an election taking place in an area and insurgent territorial control in Sri Lanka: ‘The LTTE itself did not participate in the election, of course, and no voting took place in the territory it controlled in the North.’ Similarly, Steele and Schubigger (2018: 594) show how in Afghanistan and Cambodia local-level elections did not take place in insurgent controlled territories. Hatlebakk (2007: 3) argues that ‘the Maoists have boycotted and interrupted recent elections, and have been more successful in doing so within their base areas. [...] This is a potentially useful indicator that to some extent reflects the strength of the insurgents.’

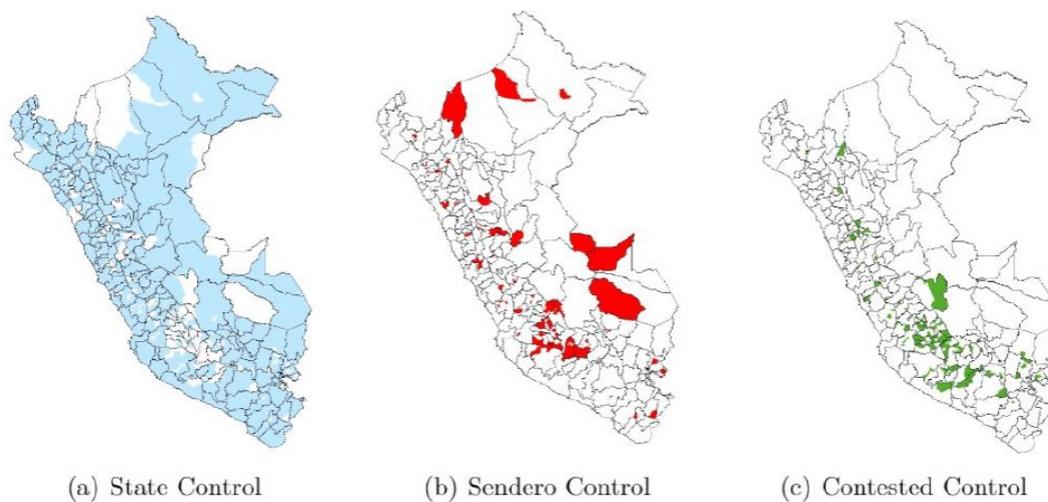
While in fully controlled territories elections do not take place, the core tactic of insurgent groups in territories under their scope of influence, but over which they do not have yet full control, is to spoil votes. Several studies on the Peruvian conflict confirm this. McClintock (1998: 79) argues that ‘invalid ballots in the departments where Sendero was considered to be strong were much higher than elsewhere’. Evidence from the Truth and Reconciliation Commission in Peru also provides support to the relation between spoiled votes and contested territories: ‘the number of spoiled votes in some rural districts of Ayacucho showed that the strength of the PCP-SL was greater than expected’ (CVR 2003: Parte I: Sección II, Capítulo 1: 35). In a similar vein, De la Calle (2017: 432–33) argues that: ‘Voting in Peru is mandatory, and the government used its full power to force people to vote [...] If the state is sufficiently strong to run the election, but not strong enough to prevent SP from forcing (or encouraging) local citizens to cast a spoiled ballot, the situation resembles one of contested power.’

³The boycotting actions of Sendero Luminoso are clearly described by Cárdenas (1992: 85–86), who shows that in the eve of the 1989 elections, ‘three members of the Electoral Board of the Huamanga district were assassinated in their respective homes. In this way, the Shining Path once again attacked the electoral process by ordering a boycott of the elections. The dead were three neighbours who were randomly drawn as members of the Electoral Board.’

In sum, and as stated by McClintock (1998: 80): ‘The best (although imperfect) indicator (of territorial control) [...] is the ability of the guerrillas to prevent official elections from being held in the area.’

This does not preclude the fact that this measurement has important limitations. Concretely, it is fundamental to acknowledge the fact that as the measurement of territorial control is directly linked to electoral boycott it could also be related to other sources of electoral behaviour, such as historically low levels of voting turnout. However, and more relevantly for the purposes of this paper, this could be related to the dependent variables, such as levels of state capacity. Another limitation relates to the fact that I am capturing a cross-sectional measure of territorial control, when territorial control is a feature that is in constant evolution during the conflict. In order to partially deal with some of these problems, I include a variety of robustness tests in Appendix F2 below.

Figure A1: Maps of territorial control in Peru (1983–93)



Source: author’s elaboration based on De la Calle (2017) data.

C Descriptive statistics

Table A2 below shows the descriptive statistics for the variables included in the main analyses as well as in the set of robustness tests included in this appendix.

Table A2: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
(Ln) State Bureaucrats	5775	2.379	1.644	0	9.883
% Electricity	5550	28.515	30.113	0	100
Territorial Control	5884	1.262	.62	1	3
Illiteracy	5687	29.983	18.972	.49	100
(Ln) Population Density	5732	2.625922	1.604059	0	11.25076
% Spanish	5869	59.584	38.725	0	100
Political Competition	4932	11.839	6.211	-52.3	78
Urban	5916	.396	.489	0	1
Distance Prov. Capital	5860	20.582	21.72	0	375.61
Slope	5576	5.698	3.661	0	19.526
Size	5880	1343.838	5011.087	2	119859.05
(Ln) Army	2900	1.99	1.877	0	9.396
No Mayor 1989	5576	.253	.434	0	1
3 Guerrilla Attacks	5576	.196	.397	0	1
Killings	5928	14.48036	99.04397	0	2997
SL Killings	5928	10.96417	84.16641	0	2704
State Killings	5928	3.516194	20.03352	0	440
Cultivable Land	5576	7	9.213	0	90
Vote Marxism 1980	5540	15.8	14.3	0	87.9
(Ln) Height	5864	7.491	1.286	0	8.541
Density Roads 1983	5576	35.09	66.859	0	992.939
Caudillismo	5576	.025	.174	0	1
Social Movements	5576	.1578	.7897	0	13

Source: author's calculations.

D Violence

D1 Violence by territorial control

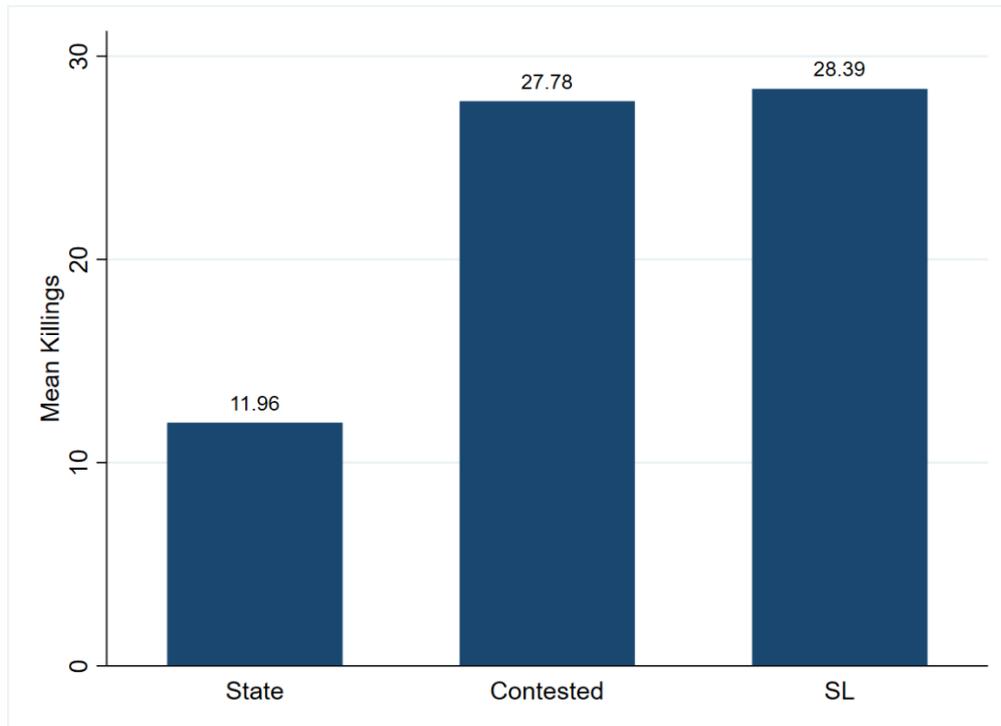
One of the core contributions of this paper is to trace the potential effect of other wartime dynamics beyond violence. However, it is also relevant to disentangle the potential effect that violence might also have on post-conflict state-building measures.

Firstly, Figure A2 below traces the average level of killings by territorial control during the Peruvian civil war. Using data from the Truth and Reconciliation Commission (CVR 2003), results show that areas controlled by Sendero Luminoso suffered the highest level of violence during the war with an average of 28.39 killings throughout the entire conflict. Contested areas had the second highest level of mean killings, with a mean of 27.78. Nevertheless, a t-test statistic comparing the means of areas controlled by insurgent and areas that remained contested shows that there is no statistical significance between the mean of the two types of territories (p -value = 0.13). The difference in levels of violence widens to over 15 points when compared to areas controlled by the government.

This evidence further strengthens the main argument developed in the paper, which argues that in decisions over where to allocate state-building measures in the aftermath of conflict, the distribution of territorial control, and not the level of violence, will play a fundamental role. However, it is important to acknowledge that violence is a multifaceted concept and killings are just one of the potential measures of violent actions, but other types of violence might take place

in a district, such as kidnappings or civilian displacement, among many others. The lack of available data on these measures hampers the capacity to test the distinctive role of different types of violent actions.

Figure A2: Violence by territorial control



Source: author's elaboration based on data from the Truth and Reconciliation Commission (CVR2003).

D2 Violence or territorial control?

Tables A3–A4 below test the potential effect of the variables on violence on the main difference-in-differences models included in the paper. Table A3 includes a variable capturing the total number of killings at the district level during the conflict lifespan, while Table A4 distinguishes according to the identity of the perpetrator.

In both cases, the inclusion of the variables on violent actions does not take away the effect of territorial control on post-conflict state-building measures. The core coefficients remain similar to the ones presented in the core models. However, the variable on the total number of killings is positive and statistically significant for the deployment of state bureaucrats ($p < 0.01$) and the effects vanishes when focusing on the provision of electricity ($p < 0.1$). Table A4 shows that these effects hold when unpacking the identity of the perpetrator of violent actions, as the coefficients remain positive and significant both for violence perpetrated by the state as well as by insurgents.

Table A3: Difference-in-differences violence vs territorial control

	<i>(ln) State Bureaucrats</i>		% <i>Electricity</i>	
	Model (1)	Model (2)	Model (3)	Model (4)
SL * Post	0.435*** (0.0752)		6.556*** (1.969)	
CT * Post		0.219** (0.0905)		0.131 (2.298)
(ln) Killings	0.360*** (0.0356)	0.359*** (0.0363)	0.799* (0.480)	0.344 (0.507)
Illiteracy	-0.00621*** (0.00184)	-0.00579*** (0.00191)	-0.108*** (0.0301)	-0.0873*** (0.0292)
(ln) Population Density	0.0330 (0.0241)	0.0297 (0.0242)	1.115*** (0.338)	1.059*** (0.336)
% Spanish	-0.00120*** (0.000447)	-0.00148*** (0.000463)	0.0137* (0.00803)	0.0153* (0.00814)
Political Competition	-0.00198** (0.000896)	-0.00201** (0.000932)	0.138*** (0.0196)	0.151*** (0.0202)
Distance Prov. Capital	-0.0223*** (0.00223)	-0.0241*** (0.00233)	-0.364*** (0.0331)	-0.379*** (0.0328)
Slope	-0.0393*** (0.00959)	-0.0359*** (0.00974)	-0.986*** (0.151)	-1.051*** (0.153)
Cultivable Land	0.000291 (0.00371)	-0.000325 (0.00368)	0.0699 (0.0497)	0.0660 (0.0492)
Vote Marxism 1980	1.189*** (0.291)	1.043*** (0.282)	11.89*** (4.612)	12.77*** (4.251)
Constant	4.263*** (0.469)	3.429*** (0.464)	21.02* (10.90)	11.80 (7.178)
Observations	4097	4015	4033	3977

Note: dependent variables: Models (1) and (2): *(ln) State bureaucrats* by district; Models (3) and (4): *Proportion access to electricity* by district. State controlled districts as reference category for the independent variables. All models include province and year fixed effects. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: author's calculations.

Table A4: Difference-in-differences violence by perpetrator vs territorial control

	<i>(ln) State Bureaucrats</i>		% <i>Electricity</i>	
	Model (1)	Model (2)	Model (3)	Model (4)
SL * Post	0.432*** (0.0751)		6.545*** (1.968)	
CT * Post		0.212** (0.0909)		0.147 (2.297)
(ln) State Killings	0.297*** (0.0563)	0.301*** (0.0607)	-0.247 (0.793)	-0.621 (0.842)
(ln) SL Killings	0.251*** (0.0352)	0.251*** (0.0361)	0.892* (0.509)	0.572 (0.526)
Illiteracy	-0.00633*** (0.00183)	-0.00592*** (0.00190)	-0.108*** (0.0301)	-0.0872*** (0.0292)
(ln) Population Density	0.0346 (0.0239)	0.0320 (0.0240)	1.121*** (0.338)	1.059*** (0.335)
% Spanish	-0.00120*** (0.000447)	-0.00147*** (0.000463)	0.0136* (0.00803)	0.0152* (0.00815)
Political Competition	-0.00201** (0.000898)	-0.00205** (0.000933)	0.138*** (0.0196)	0.151*** (0.0202)
Distance Prov. Capital	-0.0223*** (0.00224)	-0.0240*** (0.00233)	-0.365*** (0.0332)	-0.380*** (0.0329)
Slope	-0.0387*** (0.00953)	-0.0359*** (0.00974)	-0.988*** (0.151)	-1.054*** (0.153)
Cultivable Land	0.000279 (0.00371)	-0.000135 (0.00368)	0.0698 (0.0496)	0.0655 (0.0491)
Vote Marxism 1980	1.206*** (0.292)	1.026*** (0.283)	11.87** (4.626)	12.82*** (4.256)
Constant	4.256*** (0.422)	3.371*** (0.478)	21.18* (11.05)	12.27* (7.205)
Observations	4097	4015	4033	3977

Note: dependent variables: Models (1) and (2): *(ln) State bureaucrats* by district; Models (3) and (4): *Proportion access to electricity* by district. State controlled districts as reference category for the independent variables. All models include province and year fixed effects. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

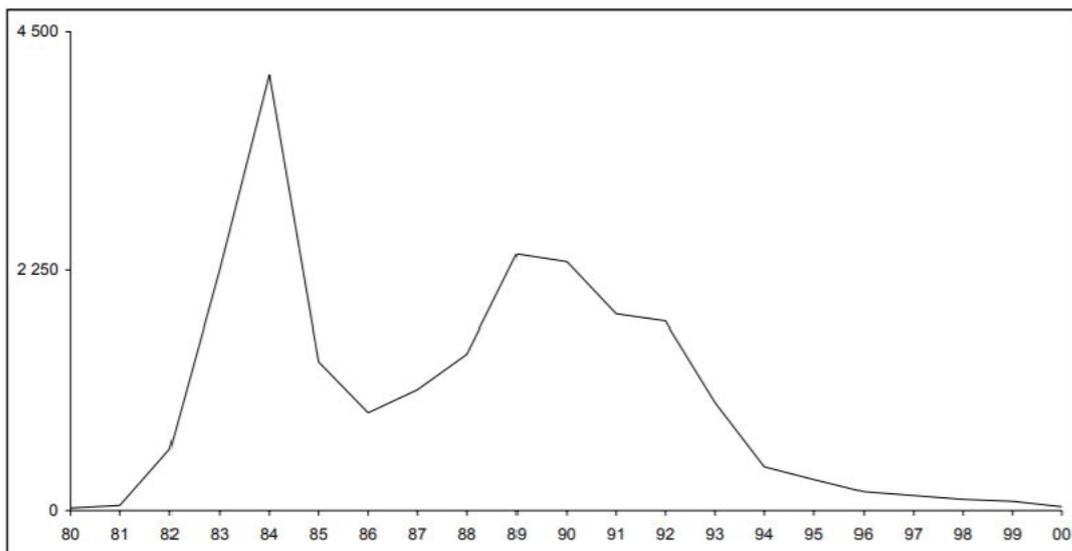
Source: author's calculations.

D3 Violence over time and conflict time frames

Finally, Figure A3 below portrays the evolution of violence related to the conflict between 1980 and 2000. Two elements deserve to be highlighted. First, the almost flat line between 1980 and 1981, showing that the use of measures of state capacity in 1981 is effectively capturing pre-war levels of state capacity, as until late 1982 there is not a considerable increase in armed actions, nor in territorial expansion (Degregori et al. 2012). And second, the profound decline in violence after 1993. Despite there was certain level of violence after 1994, it was mostly state violence (CVR 2003), showing how Sendero stopped being an effective organization by the end of 1993.

As stated in the paper, a clear definition of when the conflict starts and ends has been widely debated in the civil war literature. I follow the usual categorization of intrastate conflict developed by the Correlates of War Project, which argues that ‘an intra-state war [...] must involve sustained combat, involving organized armed forces, resulting in a minimum of 1,000 battle-related combatant fatalities within a twelve month period’ (Singer and Small 1994). Based on this, Figure A3 provides further justification to the use of the conflict time frame covering the years 1983 and 1993. Results remain robust to the use of different conflict frames, as shown in Appendix F4 below.

Figure A3: Evolution of violence over time

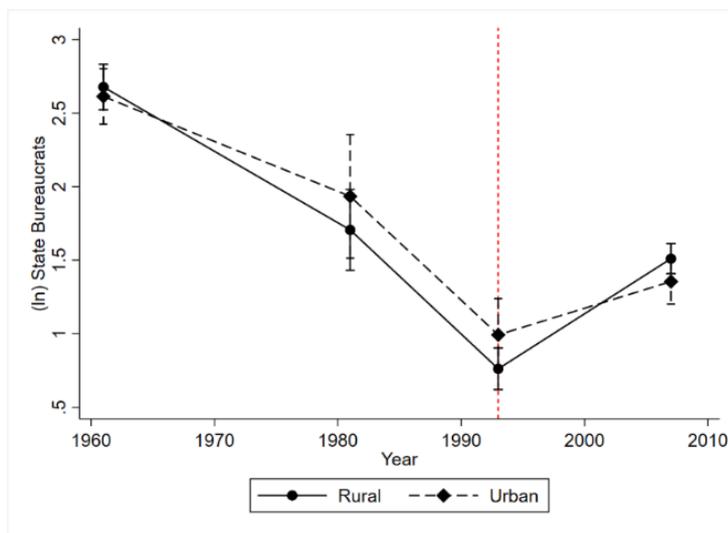


Source: CVR (2003).

E Unit heterogeneity (urban–rural)

In this section I trace potential heterogeneity across the Peruvian districts on the effects of wartime territorial control over state-building measures. Concretely, I aim to capture whether investments in state capacity in the post-conflict period are substantially different in urban and rural areas of the country.⁴ My initial expectation is that, given that rural areas were the ones mostly affected by the conflict, results might remain robust in these areas while the effects in urban areas should be weaker. For testing this, I first show in Figures A4–A7 below the overall trajectory of state bureaucrats and access to electricity in urban and rural areas since the pre-war to the post-conflict period. As the parallel trends assumption seems to graphically hold, difference-in-differences models will be implemented.

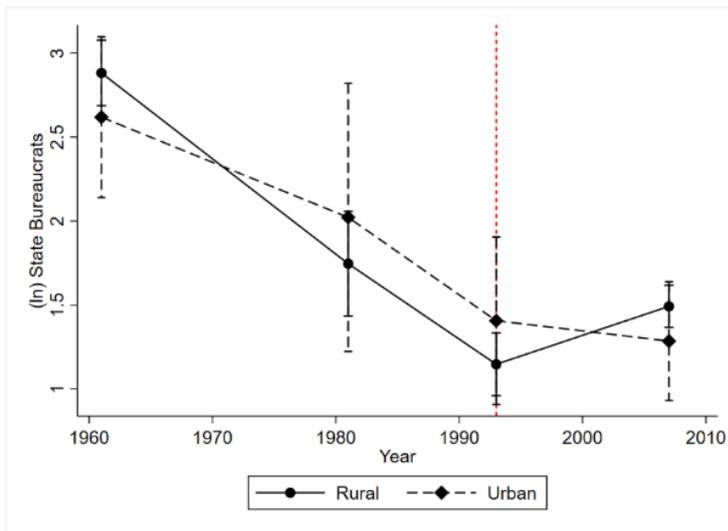
Figure A4: State bureaucrats in insurgent areas by urban–rural



Source: RIMISP (*Centro Latinoamericano para el Desarrollo Rural*).

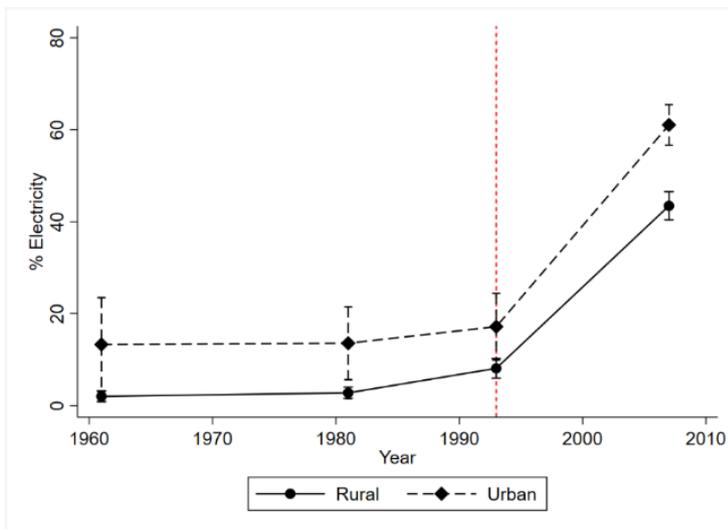
⁴ I measure urban and rural areas with a variable coming from RIMISP (*Centro Latinoamericano para el Desarrollo Rural*). This variable captures the percentage of the population living in an area categorized as urban with respect to the total population of the district. In order to compare urban and rural areas, I transform this into a dummy variable that takes the value of 1 if at least 50% of the population in a district lives in an area categorized as urban and 0 otherwise.

Figure A5: State bureaucrats in contested areas by urban–rural



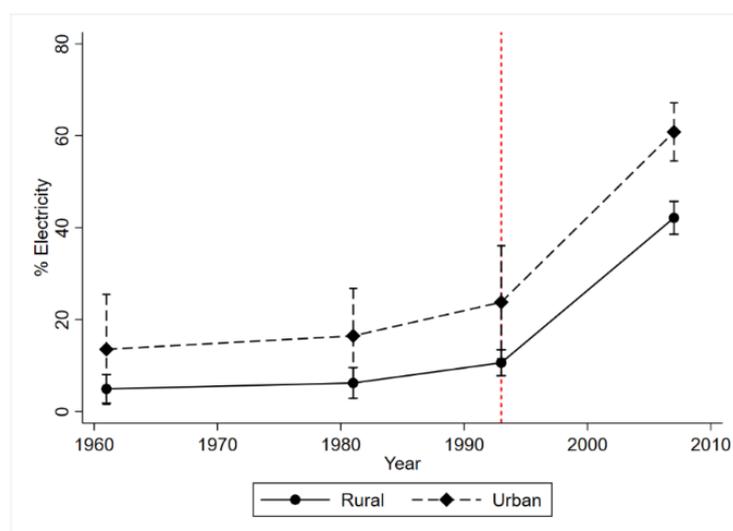
Source: RIMISP (*Centro Latinoamericano para el Desarrollo Rural*).

Figure A6: Electricity in insurgent areas by urban–rural



Source: RIMISP (*Centro Latinoamericano para el Desarrollo Rural*).

Figure A7: Electricity in contested areas by urban–rural



Source: RIMISP (*Centro Latinoamericano para el Desarrollo Rural*).

Tables A5–A6 below present the results. When subsetting the sample according to the urban–rural nature of the district, two findings should be highlighted. First, the Peruvian state invested in the deployment in state bureaucrats in insurgent and contested territories in rural areas. And second, in urban areas, it only deployed a higher level of state agents in territories that were under the control of Sendero Luminoso. This finding shows that the results of the core models seem to be driven by the effect of contested territories in rural areas, as most of urban territories were controlled by incumbents during the conflict.⁵

Regarding access to electricity, results for rural areas stay in line with the core models: the state only aimed to spur its legitimacy through the distribution of public goods and services to areas that remained under insurgent control, but not to areas that remained contested. Results remain positive and significant as well for insurgent-controlled areas in urban settings. However, contested districts in urban scenarios had a positive and significant effect on access to electricity.

Table A5: State bureaucrats by urban–rural

	<i>Rural</i>		<i>Urban</i>	
	Model (1)	Model (2)	Model (3)	Model (4)
SL * Post	0.519*** (0.0916)		0.513*** (0.149)	
CT * Post		0.287** (0.113)		-0.0355 (0.182)
Illiteracy	-0.00556** (0.00241)	-0.00553** (0.00256)	-0.00291 (0.00354)	-0.00254 (0.00369)

⁵ 89% of urban territories remained under state control during the conflict, while the number drops to 79% of rural areas. Similarly, the percentages of contested and insurgent controlled districts in rural areas are 9.35% and 11.26%, respectively. Urban areas had a proportion of contested areas of 3.44%, and 7.04% were districts controlled by Sendero Luminoso.

InPopDensity1	0.0000443 (0.0383)	0.00372 (0.0398)	0.0331 (0.0397)	0.0287 (0.0407)
% Spanish	-0.000725 (0.000506)	-0.000778 (0.000522)	-0.00190** (0.000960)	-0.00250** (0.00102)
Political Competition	-0.000252 (0.00114)	-0.000107 (0.00117)	-0.00284** (0.00142)	-0.00327** (0.00149)
Constant	3.357*** (0.196)	3.334*** (0.203)	3.984*** (0.178)	4.091*** (0.184)
Observations	2614	2595	1754	1687
R^2	0.423	0.408	0.498	0.504
Adjusted R^2	0.422	0.406	0.495	0.502

Note: dependent variable: (ln) *State bureaucrats* by district. State controlled districts as reference category for the independent variables. All models include district and year fixed effects. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

Table A6: Electricity by urban–rural

	<i>Rural</i>		<i>Urban</i>	
	Model (1)	Model (2)	Model (3)	Model (4)
SL * Post	4.534** (2.162)		14.43*** (4.693)	
CT * Post		-1.541 (2.846)		11.83** (5.197)
Illiteracy	0.0763** (0.0329)	0.0703** (0.0320)	-0.324*** (0.0649)	-0.308*** (0.0661)
(ln) Pop. Density	0.294 (0.595)	0.276 (0.607)	-1.925*** (0.656)	-2.108*** (0.666)
% Spanish	-0.0135 (0.00845)	-0.0145* (0.00865)	0.00832 (0.0163)	0.00839 (0.0175)
Political Competition	0.127*** (0.0238)	0.129*** (0.0252)	0.0863** (0.0344)	0.117*** (0.0340)
Constant	-11.66*** (3.077)	-10.81*** (3.161)	33.75*** (3.512)	32.44*** (3.447)
Observations	2574	2571	1729	1671
R^2	0.708	0.696	0.648	0.648
Adjusted R^2	0.707	0.695	0.647	0.646

Note: dependent variable: Proportion *access to electricity* by district. State controlled districts as reference category for the independent variables. All models include district and year fixed effects. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

F Robustness tests

I have conducted several robustness checks in order to test the sensitivity of the results. This includes the use of different variables specifications of the core dependent and independent variables (Appendices F1 and F2); the use of a battery of control variables (Appendix F3); the use of distinct conflict time frames (Appendix F4); the distinction between state agents and state security forces (Appendix F5); the use of provincial and departmental fixed effects (Appendix F6); the control of potential spatial effects (Appendix F7); the use of lagged dependent variable models (Appendix F8); and finally, controlling for potential ceiling effects (Appendix F9). The robustness of the results across these analyses ensures the validity of the findings presented in the core models.

F1 DV specifications

The first robustness test is towards the measurement of the core dependent variables, state bureaucrats and access to electricity. Concretely, I make use of three different specifications. First, I use the raw number of state bureaucrats and households with electricity, including the total population and the total population of households as control variables. Results from Table A7 below stay very similar to the ones of the core models of the paper. Secondly, I use inverse sine measures of the core dependent variables in Table A8. Results stay very similar, however, the coefficient on electricity for contested areas becomes positive and significant ($p < 0.05$). Finally, I make use of a variable capturing the percentage of state bureaucrats per district in Table A9, providing similar results to the main empirical models. Overall, it can be concluded that results remain robust to the use of different specifications of the two dependent variables.

Table A7: Negative binomial regression raw DVs

	<i>State Bureaucrats</i>		<i>Electricity</i>	
	Model (1)	Model (2)	Model (3)	Model (4)
SL * Post	0.464*** (0.0935)		0.380*** (0.0479)	
CT * Post		0.233** (0.105)		0.0182 (0.108)
(ln) Population	0.145*** (0.0209)	0.143*** (0.0210)		
(ln) Household Pop.			1.140*** (0.0176)	1.131*** (0.0170)
Illiteracy	0.000557 (0.00123)	0.000830 (0.00125)	-0.0340*** (0.00256)	-0.0327*** (0.00281)
% Spanish	-0.00124*** (0.000383)	-0.00153*** (0.000386)	-0.00346*** (0.000408)	-0.00290*** (0.000412)
Political Competition	-0.00141 (0.000931)	-0.00161* (0.000944)	-0.00623*** (0.00105)	-0.00523*** (0.00108)
Constant	0.121	0.149	-1.356***	-1.399***

	(0.202)	(0.204)	(0.160)	(0.168)
Observations	4369	4289	4367	4307
Wald χ^2	2033.23	1962.51	7247.69	7228.25

Note: dependent variables: Models (1) and (2) number of *State bureaucrats* by district. Models (3) and (4) number of households with access to *public electricity*. State controlled districts as reference category for the independent variables. All models include district and year fixed effects. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

Table A8: Inverse sine DV

	<i>State Bureaucrats</i>		<i>Electricity</i>	
	Model (1)	Model (2)	Model (3)	Model (4)
SL * Post	0.658*** (0.0951)		1.216*** (0.162)	
CT * Post		0.340*** (0.120)		0.529** (0.210)
Illiteracy	-0.00357* (0.00214)	-0.00300 (0.00225)	-0.0217*** (0.00319)	-0.0219*** (0.00318)
(ln) Population Density	0.0574* (0.0324)	0.0513 (0.0331)	0.00658 (0.0494)	0.000764 (0.0501)
% Spanish	-0.00140*** (0.000502)	-0.00171*** (0.000517)	-0.00579*** (0.000835)	-0.00537*** (0.000851)
Political Competition	-0.00224** (0.00101)	-0.00240** (0.00105)	-0.00742*** (0.00178)	-0.00669*** (0.00182)
Constant	4.145*** (0.143)	4.176*** (0.148)	4.939*** (0.211)	4.971*** (0.213)
Observations	4368	4282	4303	4242
R^2	0.411	0.401	0.652	0.643
Adjusted R^2	0.409	0.400	0.651	0.643

Note dependent variables: Models (1) and (2) inverse sine of *State bureaucrats*. Models (3) and (4) inverse sine of access to *public electricity*. State controlled districts as reference category for the independent variables. All models include district and year fixed effects. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

Table A9: % state bureaucrats

	<i>State Bureaucrats</i>	
	Model (1)	Model (2)
SL * Post	0.0919*** (0.0210)	
CT * Post		0.0463** (0.0224)
Illiteracy	-0.00448*** (0.000794)	-0.00450*** (0.000808)
% Spanish	-0.000843*** (0.000159)	-0.000913*** (0.000163)
Political Competition	-0.00177*** (0.000305)	-0.00189*** (0.000317)
Constant	0.936*** (0.0454)	0.950*** (0.0462)
Observations	4407	4319
R^2	0.450	0.451
Adjusted R^2	0.450	0.450

Note: dependent variable: (ln) % *State bureaucrats* per district. State controlled districts as reference category for the independent variables. All models include district and year fixed effects. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

F2 IV specifications

One of the core limitations of the paper is the measurement of wartime distribution of territorial control. This is related to two different factors: (i) The use of electoral boycott as a measure of territorial control; and (ii) the reliance of a cross-sectional measure of territorial control that aims to capture variation during the entire conflict, when the conquering of territorial control during civil war is known to be a very dynamic process. In order to partially deal with these limitations, I make use of five distinct variables on territorial control in order to test its robustness (Tables A10–A14).

First, I use a measure coming from Albertus (2019) capturing the absence of local authorities (mayor) in 1989 as a proxy of insurgent control. Coefficients on SL control from Table A10 remain positive and with high levels of statistical significance ($p < 0.01$). Second, Table A11 includes a dummy measure of insurgent control using a variable of three guerrilla attacks as a placebo proxy. In line with the argument developed above, the coefficient lacks any statistical significance. Third, and in order to accurately capture the intensity of territorial control, Tables A12–A13 include the results when the main independent variables are measured in a continuous way, capturing the total number of years in which a district has been either insurgent control or that remained contested. Most of the coefficients remain in the expected direction and with considerable levels of statistical

significance.⁶ Finally, I create a variable on territorial control that captures the last two years of the conflict (1992–93). In this case, results remain robust, except for access to electricity in contested areas, where the coefficient becomes positive and significant (Table A14).

Table A10: DiD models with no mayor as a proxy of control

	Model (1)	Model (2)
SL No Mayor * Post	0.412*** (0.0571)	8.782*** (1.385)
Illiteracy	-0.00373** (0.00189)	0.0288 (0.0292)
(ln) Population Density	0.0257 (0.0292)	-0.674 (0.477)
% Spanish	-0.00128*** (0.000433)	0.00324 (0.00751)
Political Competition	-0.00216** (0.000859)	0.144*** (0.0190)
Constant	3.547*** (0.129)	-0.184 (2.234)
Observations	4452	4371
R^2	0.425	0.670
Adjusted R^2	0.424	0.670

Note: dependent variables: Model (1): (ln) *State bureaucrats* by district; Model (2): Proportion *access to electricity* by district. State controlled districts as reference category for the independent variables. All models include district and year fixed effects. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

⁶ There are two relevant exceptions. First, for state bureaucrats in districts that were controlled by SL over more than nine years. And second, for electricity in districts that remained contested between seven and eight years.

Table A11: DiD Models with 3 guerrilla events as placebo proxy of control

	Model (1)	Model (2)
SL * Post	-0.00742 (0.0661)	-2.326 (1.415)
Illiteracy	-0.00406** (0.00190)	0.0189 (0.0293)
(ln) Population Density	0.0161 (0.0293)	-0.852* (0.473)
% Spanish	-0.00154*** (0.000435)	-0.00197 (0.00754)
Political Competition	-0.00225*** (0.000863)	0.144*** (0.0194)
Constant	3.616*** (0.128)	1.325 (2.247)
Observations	4452	4371
R^2	0.417	0.664
Adjusted R^2	0.416	0.664

Note: dependent variables: Model (1): (ln) *State bureaucrats* by district; Model (2): Proportion *access to electricity* by district. State controlled districts as reference category for the independent variables. All models include district and year fixed effects. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

Table A12: DiD models with number of years under SL control as IV

	<i>(ln) Bureauc.</i>	<i>% Electricity</i>								
	Model									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SL * 5 Years	0.606*** (0.0999)	10.44*** (2.668)								
SL * 6 Years			0.630*** (0.116)	12.64*** (2.917)						
SL * 7 Years					0.450** (0.193)	9.495** (4.493)				
SL * 8 Years							0.445** (0.212)	7.851* (4.603)		
SL * 9 Years									0.0973 (0.257)	8.019 (5.042)
Constant	3.635*** (0.130)	2.654 (2.308)	3.648*** (0.130)	2.744 (2.303)	3.675*** (0.130)	3.238 (2.298)	3.676*** (0.130)	3.289 (2.297)	3.689*** (0.130)	3.384 (2.297)
Observations	4368	4303	4368	4303	4368	4303	4368	4303	4368	4303
R ²	0.429	0.660	0.428	0.661	0.425	0.658	0.425	0.658	0.424	0.658
Adjusted R ²	0.428	0.660	0.427	0.660	0.424	0.658	0.424	0.658	0.423	0.658

Note: dependent variables: Models (1), (3), (5), (7) and (9): *(ln) State bureaucrats* by district; Models (2), (4), (6), (8) and (10): *Proportion access to electricity* by district. All models include district and year fixed effects as well as the control variables included in the core models. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

Table A13: DiD models with number of years under contested control as IV

	<i>(ln) Bureauc</i>	% <i>Electricity</i>								
	Model	Model								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
CT * 5 Years	0.313*** (0.116)	4.330 (2.798)								
CT * 6 Years			0.244** (0.124)	4.672 (3.108)						
CT * 7 Years					0.243 (0.178)	-0.653 (5.421)				
CT * 8 Years							0.411 (0.291)	-11.27 (11.95)		
CT * 9 Years									0.695*** (0.267)	-0.00187 (0.00784)
Constant	3.649*** (0.134)	1.549 (2.309)	3.662*** (0.134)	1.628 (2.307)	3.674*** (0.134)	1.996 (2.301)	3.676*** (0.134)	2.110 (2.292)	3.678*** (0.134)	2.008 (2.300)
Observations	4282	4242	4282	4242	4282	4242	4282	4242	4282	4242
R ²	0.422	0.654	0.422	0.654	0.421	0.653	0.421	0.653	0.422	0.653
Adjusted R ²	0.421	0.653	0.421	0.653	0.420	0.652	0.420	0.653	0.421	0.653

Note: dependent variables: Models (1), (3), (5), (7) and (9): *(ln) State bureaucrats* by district; Models (2), (4), (6), (8) and (10): *Proportion access to electricity* by district. All models include district and year fixed effects as well as the control variables included in the core models. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

Table A14: Difference-in-differences control in the last years of the war (1992–93)

	<i>(ln) Bureaucrats</i>	<i>(ln) Bureaucrats</i>	% <i>Electricity</i>	% <i>Electricity</i>
	Model	Model	Model	Model
	(1)	(2)	(3)	(4)
SL * Post	0.524*** (0.0616)		7.074*** (1.467)	
CT * Post		0.270*** (0.0627)		3.536** (1.440)
Illiteracy	-0.00370* (0.00207)	-0.00176 (0.00208)	-0.00583 (0.0366)	0.0490 (0.0332)
(ln) Population Density	0.0159 (0.0316)	0.0612** (0.0310)	-0.677 (0.536)	-0.756 (0.557)
% Spanish	-0.00118** (0.000507)	-0.00182*** (0.000489)	-0.00943 (0.00948)	-0.00239 (0.00849)
Political Competition	-0.00235** (0.00101)	-0.00259*** (0.00100)	0.140*** (0.0231)	0.128*** (0.0223)
Constant	3.664*** (0.144)	3.554*** (0.143)	4.466* (2.654)	2.378 (2.557)
Observations	3532	3644	3473	3617
R ²	0.439	0.437	0.652	0.647
Adjusted R ²	0.437	0.436	0.651	0.646

Note: dependent variables: Models (1) and (2) : (ln) *State bureaucrats* by district; Models (3) and (4): Proportion *access to electricity* by district. State controlled districts as reference category for the independent variables. All models include district and year fixed effects. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

F3 Robustness controls

In Table A15 below, I get rid of the district fixed effects and the time-varying controls in order to include a variety of variables to check whether the main results of the paper remain similar after the inclusion of these measures. Concretely, I control by distance to provincial capital; district slope; percentage of cultivable land; vote for Marxist parties in 1980; municipality size and height; density of roads in 1973; number of previous social movements; presence of historical caudillos; and level of urbanization by 1993 (see Appendix A for further details on the sources and operationalization of these variables). In line with the expectations, results present scarce changes and the core variables on wartime territorial control remain very similar.

Table A15: Difference-in-differences controls

	<i>(ln) Bureaucrats</i>		% <i>Electricity</i>	
	Model (1)	Model (2)	Model (3)	Model (4)
DiDSL	0.461*** (0.0745)		6.729*** (1.936)	
DiDCT		0.261*** (0.0913)		-0.495 (2.190)
Distance Prov. Capital	-0.0110*** (0.00188)	-0.0129*** (0.00197)	-0.205*** (0.0261)	-0.223*** (0.0257)
Slope	-0.0389*** (0.00933)	-0.0343*** (0.00944)	-0.627*** (0.140)	-0.691*** (0.143)
Cultivable Land	-0.000517 (0.00370)	-0.00157 (0.00370)	-0.00483 (0.0404)	-0.0105 (0.0400)
Vote Marxism 80	1.034*** (0.269)	0.874*** (0.255)	7.974** (4.011)	8.756** (3.754)
Size	0.000106*** (0.0000264)	0.000104*** (0.0000263)	0.000383 (0.000236)	0.000318 (0.000231)
(ln) Height	0.0367 (0.0573)	0.0352 (0.0577)	-0.547 (0.541)	-0.192 (0.544)
Density Roads 1973	0.00180*** (0.000544)	0.00181*** (0.000546)	0.0292*** (0.00663)	0.0270*** (0.00643)
Social Movements	-0.0234 (0.0631)	0.00413 (0.0633)	0.895 (0.793)	0.901 (0.784)
Caudillismo	0.486** (0.235)	0.499* (0.256)	0.311 (3.849)	4.263 (3.300)
Urban 93	0.0102*** (0.00137)	0.0101*** (0.00137)	0.328*** (0.0186)	0.331*** (0.0185)
Constant	3.722*** (0.728)	3.097*** (0.627)	8.968 (9.348)	-0.876 (6.856)
Observations	4081	3999	4017	3961

Note: dependent variables: Models (1) and (2) : *(ln) State bureaucrats* by district; Models (3) and (4): *Proportion access to electricity* by district. State controlled districts as reference category for the independent variables. All models include province and year fixed effects as well as the control variables of main models. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

F4 Conflict time frames

One of the potential criticisms to this paper is that the results might be driven by the specific time frame of the conflict chosen for the main analyses. While the justification of the choice of the conflict time frame is thoroughly developed in Appendix D3 above, I test the robustness of the results in light of the use of three distinct conflict time frames: (i) 1980–93; (ii) 1981–93; and (iii) 1980–94. Results remain very similar to the ones of the core models (Table A16). The only exception to this is present in Models (4) and (8), where the variable on access to electricity in contested areas becomes barely significant ($p < 0.1$)

Table A16: Differences-in-differences conflict frames

	1980–93				1981–93				1980–94			
	(1) State	(2) State	(3) Elec.	(4) Elec.	(5) State	(6) State	(7) Elec.	(8) Elec.	(9) State	(10) State	(11) Elec.	(12) Elec.
DiDSL	0.665*** (0.101)		10.99*** (2.717)		0.627*** (0.0911)		12.64*** (2.625)		0.599*** (0.103)		10.65*** (2.852)	
DiDCT		0.397*** (0.0909)		3.947* (2.233)		0.368*** (0.0947)		3.783* (2.200)		0.473*** (0.0913)		3.563 (2.228)
Constant	3.645*** (0.130)	3.601*** (0.130)	3.056 (2.317)	1.434 (2.275)	3.645*** (0.130)	3.618*** (0.131)	2.886 (2.323)	1.723 (2.273)	3.686*** (0.131)	3.611*** (0.130)	3.170 (2.316)	1.692 (2.273)
Observations	4328	4458	4264	4401	4331	4434	4267	4380	4359	4470	4295	4412
R^2	0.429	0.424	0.656	0.655	0.430	0.423	0.657	0.654	0.430	0.424	0.656	0.655
Adjusted R^2	0.428	0.423	0.656	0.654	0.429	0.422	0.657	0.654	0.429	0.423	0.655	0.654

Note: dependent variables: Models (1), (2), (5), (6), (9) and (10): (ln) *State bureaucrats* by district; Models (3), (4), (7), (8), (11) and (12): *Proportion access to electricity* by district. All models include district and year fixed effects as well as the control variables included in the core models. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

F5 State bureaucrats and state security forces

Table A17 below shows the results when bringing together as a dependent variable the number of state bureaucrats and state security forces per district. Due to the lack of information on state security forces before 1993, the only pre-treatment period is 1993, while the post-treatment period remains 2007. Results stay positive and statistically significant ($p < 0.01 - p < 0.1$)

Table A17: Difference-in-differences state bureaucrats and state security forces

	Model (1)	Model (2)
SL * Post	0.336*** (0.0905)	
CT * Post		0.203* (0.107)
Illiteracy	-0.00272 (0.00202)	-0.00159 (0.00208)
(ln) Population Density	0.182*** (0.0307)	0.176*** (0.0310)
% Spanish	0.0000269 (0.000452)	-0.000406 (0.000462)
Political Competition	-0.000591 (0.000944)	-0.000687 (0.000954)
Constant	2.955*** (0.137)	2.966*** (0.140)
Observations	4312	4230
R^2	0.134	0.123
Adjusted R^2	0.133	0.121

Note: dependent variable: (ln) *State bureaucrats and security forces* by district. State controlled districts as reference category for the independent variables. All models include district and year fixed effects. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

F6 Provincial & departmental fixed effects

Tables A18–A19 below test the robustness of the results when making use of provincial and departmental fixed effects. These tables include a variety of district-level control variables. Results remain highly similar to the ones of the core models.

Table A18: Difference-in-differences provincial fixed effects

	<i>(In) State Bureaucrats</i>		% <i>Electricity</i>	
	Model (1)	Model (2)	Model (3)	Model (4)
SL * Post	0.447*** (0.0747)		6.566*** (1.967)	
CT * Post		0.266*** (0.0905)		0.187 (2.302)
Illiteracy	-0.00559*** (0.00188)	-0.00496** (0.00195)	-0.107*** (0.0300)	-0.0867*** (0.0291)
(In) Population Density	0.0136 (0.0253)	0.00942 (0.0253)	1.061*** (0.337)	1.037*** (0.337)
% Spanish	-0.00123*** (0.000448)	-0.00150*** (0.000464)	0.0137* (0.00803)	0.0153* (0.00814)
Political Competition	-0.00187** (0.000898)	-0.00187** (0.000939)	0.139*** (0.0196)	0.151*** (0.0202)
Distance	-0.0228*** (0.00235)	-0.0245*** (0.00245)	-0.365*** (0.0333)	-0.380*** (0.0329)
Slope	-0.0510*** (0.00977)	-0.0479*** (0.00981)	-1.011*** (0.150)	-1.063*** (0.152)
Cultivable Land	-0.000377 (0.00404)	-0.000904 (0.00400)	0.0686 (0.0496)	0.0654 (0.0492)
Vote Marxist 80	1.288*** (0.308)	1.104*** (0.290)	12.09*** (4.594)	12.82*** (4.242)
Constant	4.954*** (0.661)	4.228*** (0.468)	22.59** (11.16)	12.58* (7.177)
Observations	4097	4016	4033	3978

Note: dependent variables: Models (1) and (2): *(In) State bureaucrats* by district; Models (3) and (4): *Proportion access to electricity* by district. State controlled districts as reference category for the independent variables. All models include province and year fixed effects. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

Table A19: Difference-in-differences department fixed effects

	<i>(ln) State Bureaucrats</i>		% <i>Electricity</i>	
	Model (1)	Model (2)	Model (3)	Model (4)
SL * Post	0.411*** (0.0741)		5.158*** (1.734)	
CT * Post		0.199** (0.0896)		-1.459 (2.218)
Illiteracy	-0.00851*** (0.00180)	-0.00857*** (0.00187)	-0.240*** (0.0302)	-0.231*** (0.0302)
(ln) Population Density	0.0996*** (0.0290)	0.0920*** (0.0289)	2.016*** (0.317)	1.905*** (0.320)
% Spanish	-0.00116*** (0.000433)	-0.00135*** (0.000449)	0.0222*** (0.00800)	0.0255*** (0.00816)
Political Competition	-0.00192** (0.000906)	-0.00188** (0.000943)	0.131*** (0.0192)	0.142*** (0.0197)
Distance	-0.0179*** (0.00217)	-0.0194*** (0.00224)	-0.306*** (0.0332)	-0.308*** (0.0339)
Slope	-0.0722*** (0.00930)	-0.0709*** (0.00943)	-1.252*** (0.147)	-1.341*** (0.152)
Cultivable Land	0.00385 (0.00394)	0.00259 (0.00395)	0.0896 (0.0547)	0.0819 (0.0545)
Vote Marxist 80	0.764*** (0.271)	0.634** (0.259)	4.880 (4.291)	3.709 (4.160)
Constant	3.434*** (0.173)	3.464*** (0.176)	7.153** (3.272)	6.631** (3.328)
Observations	4097	4016	4033	3978

Note: dependent variables: Models (1) and (2): *(ln) State bureaucrats* by district; Models (3) and (4): *Proportion access to electricity* by district. State controlled districts as reference category for the independent variables. All models include department and year fixed effects. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

F7 Spill over effects

One of the potential concerns from an empirical perspective is the potential bias given by the presence of spatial autocorrelation in the data, as the distribution of wartime dynamics, as well as of state-building measures, could potentially be spatially clustered in certain geographic regions. The presence of spatial autocorrelation in the dependent variables produces biased estimates under OLS models, as it violates the assumption of unit independence.

I test for the presence of spatial autocorrelation through the use of spatial lag and spatial error models, controlling for the values of the DV of the neighbouring districts and capturing in the error term the level of spatial correlation. I make use of a weights matrix constructed on Queen neighbours. Results from Tables A20–A21 support the evidence of the main models even after controlling for potential spatial effects.

Table A20: Spatial lag models

	<i>State Bureaucrats</i>		<i>Electricity</i>	
	Model	Model	Model	Model
	(1)	(2)	(3)	(4)
SL Control	0.468*** (0.136)	0.435*** (0.136)	7.057*** (1.745)	6.295*** (1.746)
Contested Control	0.390*** (0.148)	0.318** (0.148)	1.360 (1.886)	0.629 (1.896)
ρ	0.504*** (0.028)	0.453*** (0.029)	0.611*** (0.024)	0.596*** (0.026)
Constant	-0.315*** (0.044)	0.946*** (0.159)	11.226*** (0.891)	23.487*** (2.245)
Observations	1,783	1,754	1,769	1,742
Log Likelihood	-3,386.879	-3,300.876	-7,890.214	-7,753.620
σ^2	2.481	2.421	403.663	397.636
Akaike Inf. Crit.	6,789.757	6,625.752	15,796.430	15,531.240
Wald Test (df = 1)	323.555***	247.617***	623.104***	589.558***
LR Test (df = 1)	322.955***	246.082***	448.782***	425.912***

Note: Reference category: state control.
 DV: Difference state capacity between 2007 and 1993.
 Models (2) and (4) include the main control variables.
 Robust standard errors in parentheses. *p<0.10; **p<0.05; ***p<0.01.

Source: author's calculations.

Table A21: Spatial error models

	<i>State Bureaucrats</i>		<i>Electricity</i>	
	Model	Model	Model	Model
	(1)	(2)	(3)	(4)
SL Control	0.448*** (0.149)	0.411*** (0.149)	6.390*** (1.943)	5.439*** (1.943)
Contested Control	0.338** (0.156)	0.311** (0.156)	-1.861 (1.988)	-2.237 (1.988)
λ	0.507*** (0.028)	0.464*** (0.029)	0.618*** (0.024)	0.611*** (0.025)
Constant	-0.573*** (0.078)	0.694*** (0.192)	29.569*** (1.274)	42.671*** (2.650)
Observations	1,783	1,754	1,769	1,742
Log Likelihood	-3,389.165	-3,311.746	-7,891.656	-7,755.981
σ^2	2.486	2.446	403.412	396.867
Akaike Inf. Crit.	6,794.330	6,647.492	15,799.310	15,535.960
Wald Test (df = 1)	326.582***	248.294***	640.412***	614.957***
LR Test (df = 1)	318.383***	224.342***	445.898***	421.189***

Note: Reference category: state control.
 DV: Difference state capacity between 2007 and 1993.
 Models (2) and (4) include the main control variables.
 Robust standard errors in parentheses. *p<0.10; **p<0.05; ***p<0.01.

Source: author's calculations.

F8 Lagged DV models

In order to ensure that the results are robust to the use of different model specifications, I make use of lagged dependent variable models. As shown in Table A22 below, results do not significantly change.

Table A22: Lagged dependent variable models

	<i>(In) State Bureaucrats</i>		<i>% Electricity</i>	
	Model (1)	Model (2)	Model (3)	Model (4)
SL * Post	0.513*** (0.0994)		6.203** (2.411)	
CT * Post		0.295** (0.123)		0.314 (2.889)
(In) State Bureaucrats t-1	-0.0243 (0.0254)	-0.0212 (0.0255)		
% Electricity t-1			-0.176*** (0.0271)	-0.155*** (0.0268)
Illiteracy	-0.0193*** (0.00453)	-0.0175*** (0.00464)	-0.410*** (0.0840)	-0.441*** (0.0841)
(In) Population Density	-0.221*** (0.0760)	-0.243*** (0.0770)	-1.193 (1.374)	-0.948 (1.378)
% Spanish	-0.00255*** (0.000604)	-0.00278*** (0.000612)	0.0169* (0.00991)	0.0180* (0.00984)
Political Competition	-0.00227** (0.00107)	-0.00175 (0.00110)	0.119*** (0.0213)	0.124*** (0.0219)
Constant	4.184*** (0.193)	4.151*** (0.200)	24.34*** (3.639)	24.63*** (3.635)
Observations	1254	1231	1243	1228
R^2	0.350	0.339	0.651	0.642
Adjusted R^2	0.348	0.338	0.650	0.641

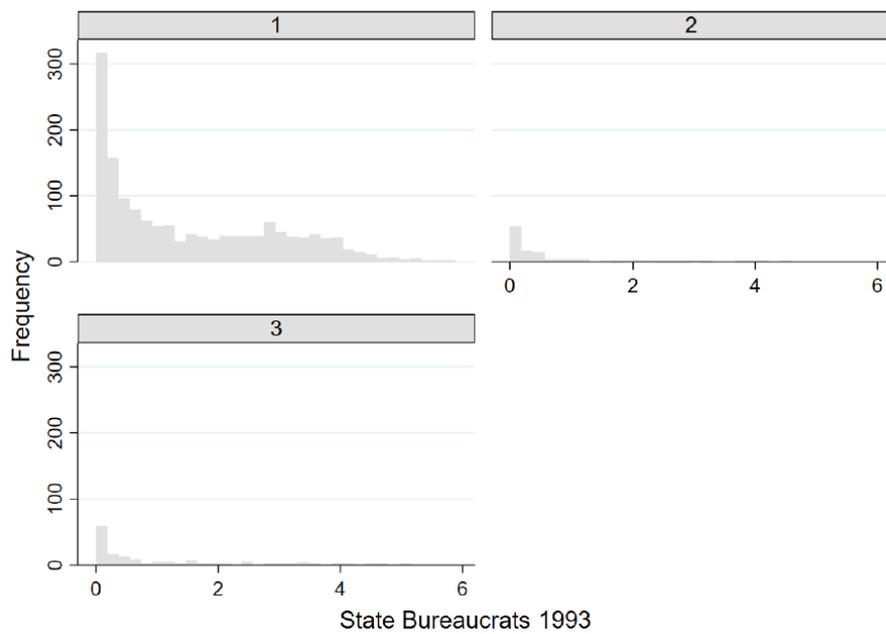
Note: dependent variables: Models (1) and (2): *(In) State bureaucrats* by district; Models (3) and (4): *Proportion access to electricity* by district. State controlled districts as reference category for the independent variables. All models include district fixed effects. Robust standard errors clustered by district in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

F9 Ceiling effects

In order to trace the potential existence of ceiling effects, I display the distribution of state bureaucrats by territorial control right upon conflict termination (1993) in Figure A8 below. Moreover, I test the robustness of the results when dropping the 5 per cent of state-controlled areas with the highest number of state bureaucrats in Table A23.

Figure A8: Distribution of state bureaucrats by type of territory



Source: author's elaboration.

Table A23: Ceiling effects dropping top 5%

	(1) (ln) State Bureaucrats	(2) (ln) State Bureaucrats	(3) % Electricity	(4) % Electricity
DiDSL	0.627*** (0.0838)		8.479*** (1.984)	
DiDCT		0.327*** (0.101)		2.573 (2.531)
Illiteracy	-0.00375* (0.00212)	-0.00362* (0.00217)	0.0762*** (0.0295)	0.0749** (0.0292)
(ln) Pop. Density	-0.0169 (0.0340)	-0.0162 (0.0352)	-0.197 (0.557)	-0.335 (0.566)
% Spanish	-0.00114** (0.000460)	-0.00120*** (0.000465)	-0.00189 (0.00772)	0.000309 (0.00782)
Political Competition	-0.00208** (0.000927)	-0.00166* (0.000950)	0.150*** (0.0196)	0.163*** (0.0199)
Constant	3.445*** (0.156)	3.437*** (0.160)	-7.108*** (2.642)	-7.211*** (2.730)
Observations	3908	3848	3908	3848
R ²	0.397	0.389	0.675	0.667
Adjusted R ²	0.396	0.388	0.674	0.666

Note: standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

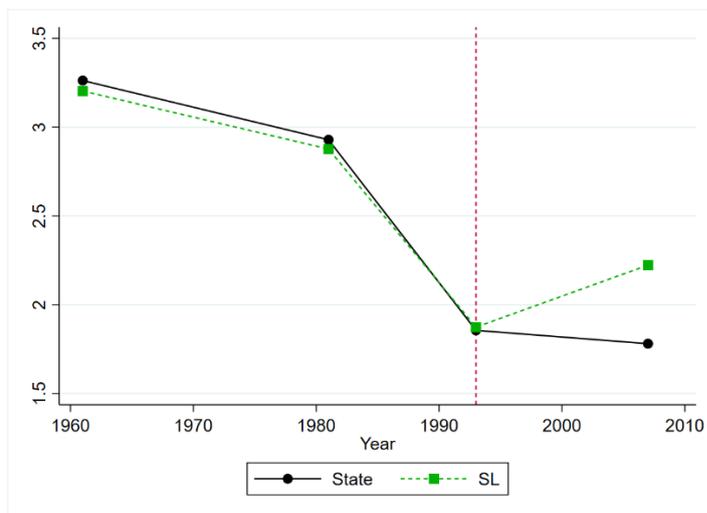
G DiD diagnostics & endogeneity

Another important source of robustness to the results is the fulfilment of some of the core diagnostics of difference-in-differences models as well as to the potential endogeneity issues present between the core variables. Subsections G1 to G4 aim to partially deal with some of these issues.

G1 Conditional parallel trends

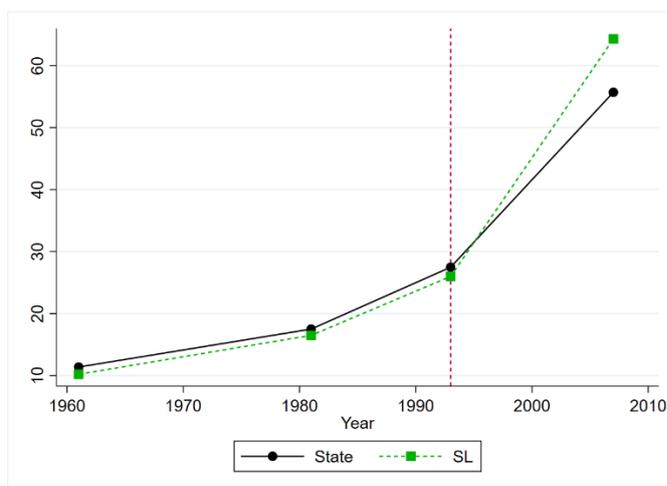
While Figure 3 of the paper already portrays sufficient graphical evidence on the core assumption of DiD models, that on parallel trends in the pre-treatment period, I provide further evidence of this with the use of conditional parallel trends figures. As shown in Figures A9 to A12 below, the assumption of parallel trends across categories in the pre-treatment periods holds in all cases.

Figure A9: SL executive



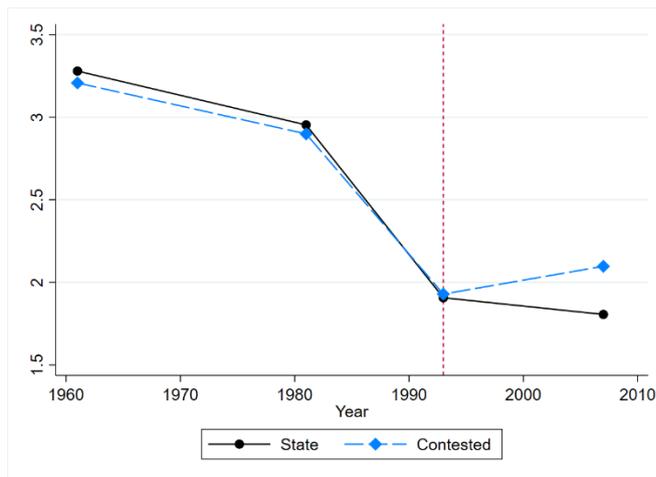
Source: author's elaboration.

Figure A10: SL electricity



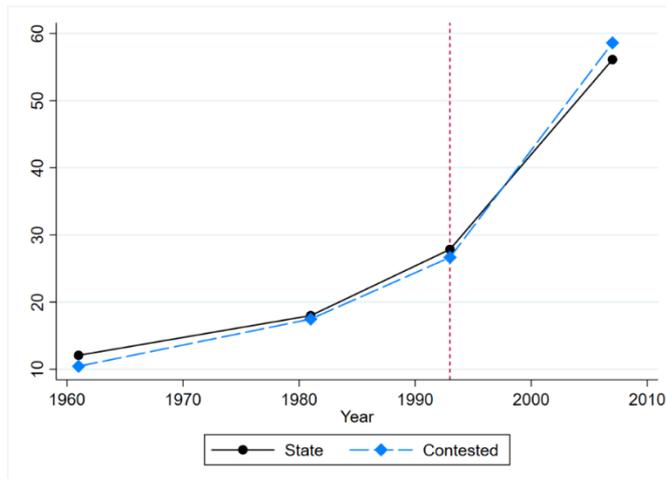
Source: author's elaboration.

Figure A11: Contested executive



Source: author's elaboration.

Figure A12: Contested electricity



Source: author's elaboration.

G2 Lack of pre-treatment effects

Another test is the analysis of the potential impact of pre-treatment effects. Concretely, I make use of a placebo treatment, using the year 1981 as the placebo unit for defining the beginning of the treatment period, which in the core model is captured by the year 2007. As shown in Table A24 below, most of the coefficients either go in the contrary direction of the main results or lack statistical significance.

Table A24: Placebo 1981 as treatment

	(1) (ln) State Bureaucrats	(2) (ln) State Bureaucrats	(3) % Electricity	(4) % Electricity
SL * Post	-0.207 (0.164)		-3.536** (1.706)	
CT * Post		-0.393** (0.171)		-2.330* (1.275)
Illiteracy	-0.000272 (0.00219)	-0.000559 (0.00230)	0.0967*** (0.0255)	0.0930*** (0.0247)
(ln) Population Density	0.104*** (0.0336)	0.104*** (0.0342)	0.504 (0.468)	0.640 (0.474)
% Spanish	-0.000824 (0.000539)	-0.00118** (0.000563)	0.0435*** (0.00831)	0.0475*** (0.00836)
Political Competition	0.000579 (0.00128)	-0.000634 (0.00134)	0.267*** (0.0245)	0.275*** (0.0244)
Constant	3.004*** (0.166)	3.142*** (0.171)	-17.51*** (2.447)	-18.20*** (2.438)
Observations	3031	2979	2966	2939
R^2	0.458	0.446	0.405	0.406
Adjusted R^2	0.457	0.445	0.403	0.404

Note: standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: author's calculations.

G3 Covariate balance

The final DiD test is related to the presence of covariate balance in the pre-treatment period. As shown in Tables A25 and A26 below, insurgent and contested areas are not significantly different from state areas with regards to the main time-varying covariates included in the core empirical models.⁷

⁷ As argued by García-Montoya et al. (2022: 9), who rely on the same covariate balance test used in this paper, covariates will be balanced as long as the variance ratio (the ratio of the variance of the propensity score in treated cases and the variance of the propensity score in control cases) lies between 0.5 and 2.

Table A25: Insurgent controlled areas

	Treated			Control			Balance	
	Mean	Variance	Skewness	Mean	Variance	Skewness	Std-diff	Var-ratio
<i>Illiteracy</i>	32.7014	298.9663	1.011357	33.36715	383.6986	.8028288	-.0360353	.7791697
<i>(ln) Pop Density</i>	3.358742	4.016074	.3134207	2.619496	2.587397	1.162411	.4068345	1.552168
<i>% Spanish</i>	59.0221	903.4183	.0997574	58.64517	1547.262	-.3448502	.0107679	.5838819
<i>Pol. Compet.</i>	60.53509	571.4059	.298869	63.54451	575.589	-.063216	-.1256656	.9927323

Source: author's calculations.

Table A26: Contested areas

	Treated			Control			Balance	
	Mean	Variance	Skewness	Mean	Variance	Skewness	Std-diff	Var-ratio
<i>Illiteracy</i>	32.46117	304.2442	1.057084	33.28581	382.5591	.8037118	-.0445	.7952869
<i>(ln) Pop Density</i>	3.079694	5.032444	.7088872	2.637514	2.619496	1.145789	.2260625	1.92115
<i>% Spanish</i>	59.96011	882.4624	-.0696884	58.48859	1552.47	-.3386305	.0421734	.5684247
<i>Pol. Compet.</i>	60.75324	540.8885	.2219404	63.7539	574.4255	-.0703584	-.127067	.9416164

Source: author's calculations.

G4 Matched sample

Finally, and in order to further increase the confidence in the findings presented so far, I make use of difference-in-differences models with the use of matched samples. Concretely, the goal of matching under this setting is to compare pairs of districts that are similar in a set of covariates, but that differed in their type of territorial control. Once a subset of the sample has been formed with these matched units, the DiD models will be applied. Relying on nearest neighbour matching methods, I distinguish across districts that are either 'treated' with territorial control by insurgents or by being contested. The main specifications are constructed with the pre-treatment values of the variables included in Tables 1 and 2, as well as with those from Appendix F6, Table A18 (provincial fixed effects).

Table A27: Difference-in-differences matched sample

	(1) (ln) State Bureaucrats	(2) (ln) State Bureaucrats	(3) % Electricity	(4) % Electricity
SL * Post	0.369*** (0.108)		5.302* (3.190)	
CT * Post		0.247** (0.124)		-3.086 (3.694)
Illiteracy	-0.0113*** (0.00386)	-0.0133*** (0.00449)	-0.478*** (0.0843)	-0.190** (0.0757)
lnPopDensity1	0.0921 (0.0776)	0.0773 (0.0612)	1.669 (1.031)	0.806 (1.043)
% Spanish	0.0000548 (0.000966)	0.000789 (0.00124)	0.0650*** (0.0250)	0.0623** (0.0275)
Political Competition	-0.00249	-0.000306	0.0306	0.104**

	(0.00205)	(0.00257)	(0.0528)	(0.0521)
Constant	3.544*** (0.337)	3.548*** (0.379)	23.23*** (8.083)	2.620 (8.000)
Observations	652	639	588	589
R^2	0.476	0.396	0.652	0.673
Adjusted R^2	0.470	0.389	0.647	0.669

Source: author's calculations.

H Qualitative data

Qualitative data collection for this project took place between July and October 2020. Due to the existent COVID-19 travelling restrictions, interviews were held online. These interviews have the correspondent research ethics approval from the institutional review board (IRB) of my institution. The set of qualitative evidence corresponds to five semi-structured interviews. Interviewees were selected based on their in-depth knowledge of the processes under study in the manuscript. The list of interviewees includes: one former combatant of Sendero Luminoso, three state representatives, and the director of one of the key development programmes in the country. All the participants required to preserve their anonymity.

References

- Albertus, M. (2019). Land Reform and Civil Conflict: Theory and Evidence from Peru. *American Journal of Political Science*, 64(2), 256-274.
- Allison, M. E. (2010). The legacy of violence on post-civil war elections: The case of El Salvador. *Studies in Comparative International Development*, 45(1), 104-124.
- Anders, T. (2020). Territorial control in civil wars: Theory and measurement using machine learning. *Journal of Peace Research*, 57(6), 701-714.
- Arjona, A. (2016). *Rebelocracy*. Cambridge University Press.
- Burt, J.M. (2011). *Violencia y autoritarismo en el Perú: bajo la sombra de Sendero y la dictadura de Fujimori*. Lima.
- Cárdenas, A. V. (1992). *Los crímenes de Sendero Luminoso en Ayacucho*. Editorial Impacto.
- CVR (2003). *Informe final*. Lima: Comisión para la Verdad y la Reconciliación (CVR).
- Cunningham, K. G., Huang, R., & Sawyer, K. M. (2021). Voting for militants: Rebel elections in civil war. *Journal of Conflict Resolution*, 65(1), 81-107.
- De la Calle, L. (2017). Compliance vs. constraints: A theory of rebel targeting in civil war. *Journal of Peace Research*, 54(3), 427-441.
- Degregori, C. I., Coronel, J., Pino, P. D., & Starn, O. (1996). *Las rondas campesinas y la derrota de Sendero Luminoso*. Instituto de Estudios Peruanos.
- Degregori, C. I., Stern, S. J., Appelbaum, N., Drzewieniecki, J., Flores, H., & Hershberg, E. (2012). *How difficult it is to be god: Shining Path's politics of war in Peru, 1980–1999*. University of Wisconsin Press.

- García-Montoya, L., Arjona, A., & Lacombe, M. (2022). Violence and Voting in the United States: How School Shootings Affect Elections. *American Political Science Review*, 116(3), 807-826.
- Hatlebakk, M. (2007). LSMS data quality in Maoist influenced areas of Nepal. *CMI Working Paper*.
- Holtermann, H. (2016). Relative capacity and the spread of rebellion: Insights from Nepal. *Journal of Conflict Resolution*, 60(3), 501-529.
- Ishiyama, J., & Widmeier, M. (2013). Territorial control, levels of violence, and the electoral performance of former rebel political parties after civil wars. *Civil Wars*, 15(4), 531-550.
- Kalyvas, S. N. (2006). *The logic of violence in civil war*. Cambridge University Press.
- Kammann, P. (1982). *Movimientos campesinos en el Perú: 1900-1968*. Lima: Universidad Mayor de San Marcos.
- Matanock, A. M., & Staniland, P. (2018). How and why armed groups participate in elections. *Perspectives on Politics*, 16(3), 710-727.
- McClintock, C. (1998). *Revolutionary movements in Latin America: El Salvador's FMLN & Perú's shining path*. US Institute of Peace Press.
- McCormick, G. H. (1992). *From the sierra to the cities: the urban campaign of the Shining Path*. Rand National Defense Research Inst Santa Monica Ca.
- Montoya, C. (2008). La Guerra de los Tenientes, Memorias de la Guerra con Sendero Luminoso.
- Murshed, S.M. and Gates, S. (2005). 'Spatial–Horizontal Inequality and the Maoist Insurgency in Nepal'. *Review of Development Economics*, 9(1): 121-134
- Paulson, J. (2001). Peasant struggles and international solidarity: the case of Chiapas. *Socialist Register*, 37.
- Rubin, M. A. (2020). Rebel territorial control and civilian collective action in civil war: Evidence from the communist insurgency in the Philippines. *Journal of Conflict Resolution*, 64(2-3), 459-489.
- Schaffer, H. B. (1995). The Sri Lankan elections of 1994: The chandrika factor. *Asian Survey*, 35(5), 409-425.
- Singer, J. David, & Small, M. (1994). 'Correlates of War Project: International and Civil War Data, 1816–1992' (ICPSR 9905). Codebook, Ann Arbor, MI: Interuniversity Consortium of Political Social Research.
- Steele, A., & Schubiger, L. I. (2018). Democracy and civil war: The case of Colombia. *Conflict Management and Peace Science*, 35(6), 587-600.
- Sulmont, D., & Bazán, J. L. (2011). *Base de datos de elecciones presidenciales en el Perú: 1980-2011*. Proyecto DGI 2011-0173. PUCP.
- Vargas Meza, R. (1998). The FARC, the war and the crisis of the state. *NACLA Report on the Americas*, 31(5), 22-27.
- Villamil, F. (2020). *Mobilizing memories: The long-term consequences of wartime violence against civilians*. Doctoral Dissertation, ETH Zurich.