Supplemental Appendix For

Why Insurgents Kill Civilians in Capital Cities: A Disaggregated Analysis of Mechanisms and Trends

Contents

This appendix proceeds in six parts. In the first part, I discuss in detail the data and methods used in the main paper, which – due to space concerns – I was forced to reported in this appendix. This section also reports different summary statistics and plots and tables illustrating the distribution of violence within states. In the second part, a large number of different alternative specifications and sensitivity analyses are reported to highlight the robustness of the findings derived in the main paper. This part includes alternative samples used for CEM exercises, and analysis of urban and developed grid cells only. In the third part, I report tables listing negative binomial (NB) models and Vuong test results comparing these NB models to the ZINB models from the main paper, as well as the urban models reported below. In the fourth part, I report analyses illustrating the significant effect of *capital* as a predictive indicator of atrocities. In the fifth part, I provide a formal analysis of my argument using the Colonel Blotto game model. In the third part. Finally, in the last part different country level analyses are employed to illustrate that insurgent atrocities in capital cities are associated with both increased probability of regime failure, and lower levels of regime durability.

Data and Methods

In this section, a detailed discussion of the data and methods used for all analyses are reported, in addition to summary statistics and sample characteristics, and relevant figures and plots.

Variable Operationalization an Methodology

I test this hypothesis on a sample encompassing 14 years of data (1996-2009), the total temporal range for which information on both my independent and dependent variables was available. While many of these variables end in 2008, all independent variables are lagged by one year, which allows me to include atrocities occurring in the year 2009 for my dependent variable. Analyses conducted at the subnational level are useful for evaluating spatial patterns of violence. Therefore, these 14 years of data are structured into a cell-year level dataset, where cells are the cross-sectional unit of interest, and are measured at the 0.5 x 0.5 decimal degree cell resolution,[1] for the entire terrestrial globe (Tollefsen et al., 2012), with an average year in this 1996-2009 sample containing 64,818 cells. Importantly, because these grid cells are an arbitrary unit of analysis, a robustness model that focuses on the district/province level as a more theoretically meaningful unit of analysis is reported in Table A5 below.

As mentioned in the main paper, the dependent variable, *insurgent atrocities*, is operationalized as the yearly (t) count of intentional atrocities – or deliberate attacks done for political purposes – committed against civilians within a given cell by insurgent organizations unsanctioned by the regime. This measure was coded from the PITF Worldwide Atrocities Dataset, which defines atrocities as "implicitly or explicitly political, direct, and deliberate violent action resulting in the death of noncombatant civilians" (PITF, 2009, 3). The PITF uses international news sources[2] to collect and code a reasonably systematic sample of atroc-

[1]I.e., cells of approximately 3,025 square kilometers area, which become slightly smaller as one moves toward the Poles.

[2]Specifically, Agence France Presse, Associated Press, New York Times, Reuters, CNN, BBC World Monitor, All-Africa, and `http://syrianshuhada.com/`. Local and NGO/IGO sources can appear as primary or secondary sources in the PITF data when quoted by the aforementioned sources.

ities occurring worldwide between 1995 and 2014.[3] A subset of these atrocities, in which only incidents perpetrated by a non-sanctioned insurgents are coded, is then utilized to create the dependent variable. The PITF dataset also includes campaigns, which corresponds to a residual category for identified atrocities that lack sufficient information for the identification of incidents as defined above. I focus only on incidents here to ensure comparability across cases, and to facilitate temporal aggregation. After merging the PITF's atrocity incidents into my cell-year dataset based upon their recorded latitude-longitude coordinates, each cell's identified insurgent atrocity incidents are summed to the yearly level. For summary purposes, the frequencies of insurgent atrocities by location are presented in Figure A1.

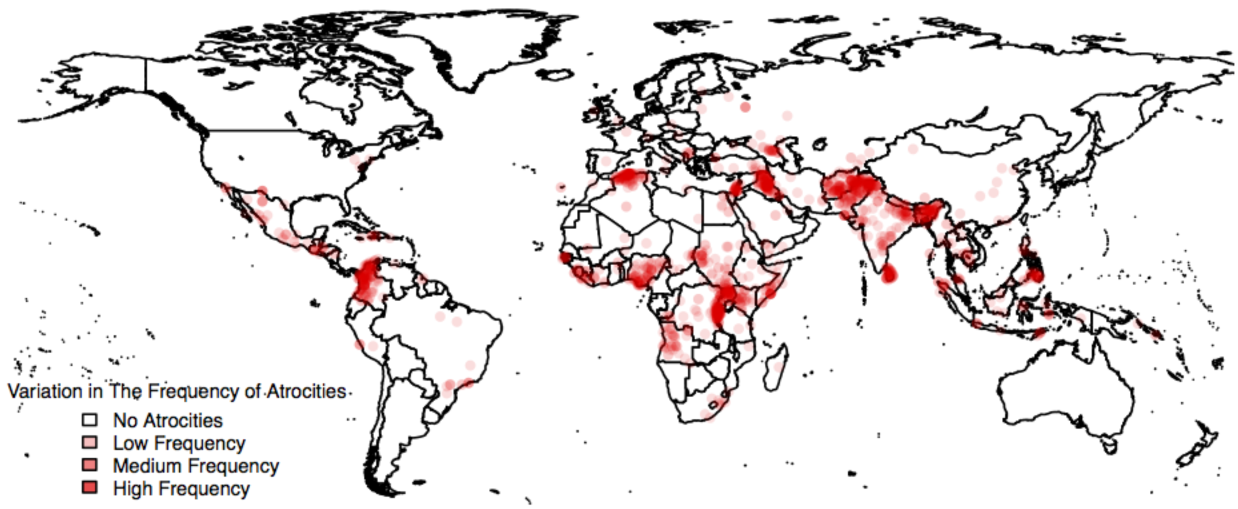
The PITF Worldwide Atrocities Dataset offers two notable advantages over other extant datasets for evaluating my hypothesis. First, it identifies all incidents "perpetrated by members of a single organization or communal group, or by members of multiple organizations or groups reportedly acting in concert, in a single locality" (PITF, 2009, 6). Human coders record each attack's geolocation, and do not report an event if no information on location is available (PITF, 2009, 8). This means that deaths coded as occurring within specific regions (e.g. capital cities) did in fact occur there, and not simply reported as such due to the lack of available information about attack location. Second, the PITF Worldwide Atrocities Dataset provides a *global* coverage of violence over the entire temporal period of analysis, which means that the linkage between capital cities and insurgent violence can be evaluated not only in certain regions but across the entire terrestrial globe for the period of concern. Nevertheless, to verify that the results are not driven by data choices, a robustness model that relies on one-sided violence by non-state groups data from the Georeferenced Event Dataset (GED) (Sundberg and Melander, 2013) is reported in Table A5 below.

Note, however, that the availability of this variable and temporal limitations with other controls (discussed below) limit my sample to the 1996-2009 period. Thus, this study's

[3] With the exception of the U.S, which is not coded by the PITF per the requirements of this data project's primary funding source (the CIA). Data for the U.S. has been collected independently, and includes three atrocities corresponding to the September 11 attacks, and one to the Oklahoma City bombing. My results are robust to this decision, as demonstrated in Table A5 below.

conclusions pertain specifically to insurgents operating after the cold war period and to campaigns occurring after the Rwandan Genocide and Balkan wars. Nevertheless, these conclusions are valid for research into wars occurring in the latter part of the 1990s, such as the civil war in Algeria, and insurgencies occurring at the post-9/11 period. As such, while these analyses are perhaps less valid for scholars studying political violence during the Cold War, these findings are valid for contemporary scholars and policymakers attempting to understand and forestall attacks by modern-day insurgent groups and the tactical variety they employ.

Figure A1: The Frequency of Insurgent Atrocities Worldwide, 1996-2009 (PITF)[1]



[1] Natural log

For *insurgent atrocities*, there were 2,708 atrocities by insurgents against civilians that affected 901 cells within my 1996-2009 sample, with an average, standard deviation, and range of 0.003, 0.142, and 0⇔60, respectively (frequency histograms are presented in Figure A2). Note that my sample contains a highly disproportionate (over 99% of recorded observations) number of zero values on *insurgent atrocities*. This extreme number of zeros suggests that in many of these cells atrocities were highly improbable for a variety of reasons, such as stringent rule of law or even an absence of human presence. To avoid these biases and account for the excess of structural zeros within the *insurgent atrocities* variable as well as this variable's count nature, I use a zero-inflated negative binomial (ZINB) model. The ZINB model adjusts the count model coefficient estimates to account for the excess of zero values by utilizing two equations. In the first – or inflation – equation, a binary logit equation is used to test for whether a zero observation is likely to have been produced by the zero-only data generating process. The covariates used in this stage account for an absence of atrocity-prone social and geographical characteristics. In the second – or count – equation, a negative binomial model is used to test the effect of covariates on the expected frequency of *insurgent atrocities*, conditional on a case being non-zero inflated based on estimates provided by the inflation equation (see, e.g., Bagozzi, 2015; Fjelde and Hultman, 2014).

As mentioned in the main paper, for atrocity-prone countries and regions, the hypothesis expects that capitals will experience higher frequencies of atrocities by insurgents compared with other regions. I accordingly construct an independent variable, *capital*, measured at the same 0.5 x 0.5 decimal degree cell resolution as my dependent variable. This variable is specifically operationalized to include cells located within a distance that is less than 55 kilometers of the exact coordinates of the nation's capital. This distance was chosen because it corresponds to the size of an edge of a square grid cell. However, to assess robustness, this indicator is operationalized as the capital province/district in Model A-X reported in the Table A5 below. Without missing information, four cells in each country are designated as capital for each given year, which makes it more likely that acts of violence aimed at the capital would be treated as such. For my 1996-2009 cell-year sample, a total of 620 cells were located in capital city confines, with a mean and standard deviation of 0.009 and 0.097, respectively, for the variable *capital*.

As additionally mentioned in the main paper, in the different stages, several lagged cell-year level controls are added to the count and inflation stages of my ZINB model specifications. I begin by accounting for economic development within a particular cell, *GCP* (gross cell product measured in billion USD) (Tollefsen et al., 2012). This measure minimizes the probability that observed effects of *capital* result from economic factors such as productivity and wealth, which are unrelated to the intrinsic value of capital cities. I additionally employ two different controls to account for the difference between capitals and other urban areas: *population*, to account for population densities within a given cell in a given year (Tollefsen et al., 2012); and the percentage of a cell denoted as urbanized, *urban* (Bontemps, Defourny and Van Bogaert, 2009). These controls ensure that it is the value inherent to capital cities, specifically, and not the density of available targets that are evaluated in relation to *insurgent atrocities*. To specifically account for the possibility that rural areas might experience higher frequencies of atrocities, I use two additional controls: *border distance*, or the distance from a given cell to the nearest border; and *travel time*, or the distance from a given cell to the nearest city with more than 50,000 inhabitants. These variables also serve as controls for the diffusion of violence across country borders.

Altogether, these five variables should adequately account for the difference between capital cells, urban cells, and rural cells. However, to further ensure robustness, I employ numerous additional controls. First, two cell-level variables were used to account for proximate levels of conflict: a binary indicator of civil conflict presence, *civil conflict*; and a one-year lagged measure of the number of atrocities perpetrated by both state of non-state actors, *atrocities*. I also control for a specific cell's geographic characteristics by adding *cell area*.

In order to further ensure that it is indeed the political value of the capital itself that affects *insurgent atrocities* and not different political and economic conditions associated with it, Models 1-4 also include two additional lagged country-year level controls. Because regime type has been shown to be related to atrocities against civilians (Koren, Forthcoming), I account for a country's political regime using the ordinal *Polity2* indicator (Marshall, Jaggers and Gurr, 2013). In line with argument presented above regarding the importance of the transferability of some resources, I also account for a given country's total oil exports, *oil* (Ross, 2011). To control for time dependencies unaccounted for by these lagged cell level variables, I also include yearly dummies (i.e., fixed effects) in each model. Note that clustered or robust standard errors were not used in the main models because such clustering can produce inferential biases (see King and Roberts, 2014). Nevertheless, to show that the results were robust to this concern, models with clustered standard errors by cell and by country are included Table A4 below.

I include several of the control variables listed above within the inflation stage of my ZINB models. Recall that this stage of analysis accounts for factors that may systematically predispose some cells and regions to be structurally atrocity-prone. Population presence is a necessary condition for a cell to have at least some opportunity to experience atrocities, and hence the variable *population* is included in this stage of analysis. In addition, because capital cities are urban areas, I also include the variable *urban* in this stage. I also add the variable *cell area* to my inflation stage, because cell area is decreasing as we approach the North and South Poles, which suggests that this measure can account for some geographic factors that limit the opportunity for atrocities (e.g., extreme temperatures, ice cover). My justifications for including these population-oriented variables in my inflation stages are consistent with other uses of zero-inflation modeling in civil conflict research (Fjelde and Hultman, 2014; Bagozzi, 2015).

Cells with no economic activity are also unlikely to involve interactions that can lead to atrocities, and I therefore include *GCP* in this stage of analysis. Building on studies that textitasize the importance of rural areas and border areas on conflict (Raleigh and Hegre, 2009), I also included the variables *border distance* and *travel time* in this stage of analysis. Accounting for the argument that oil exporting countries or autocracies are more likely to experience violence (Ross, 2011), this stage also includes the variables *oil* and *Polity2*. Lastly, stable socio-political environments and an absence of violence each likely limit the opportunities for atrocities to arise within some cells and regions. Hence, I additionally include the cell-level indicators for *civil conflict* and *atrocities* measures in my inflation stages. Again, this approach is consistent with previous studies employing ZINB models, which demonstrated that previous civil conflict levels are robust predictors of zero-inflation in these contexts (Fjelde and Hultman, 2014). All independent variables in my sample were lagged by one year to account for potential endogeneity issues. Summary statistics for all variables are listed in Table A1.

Table A1: Summary Statistics for Dependent and Independent Variables, 1996-2009

	Median	**Mean**	**Std. Dev.**	**Min**	**Max**
Nonstate atrocities	0	0.003	0.142	0	60
Capital	0	0.009	0.097	0	1
Civil conflict	0	0.063	0.243	0	1
Atrocities	0	0.005	0.341	0	231
Urban	0	0.207	1.218	0	51.550
Oil[1]	18.450	15.890	5.993	0	19.980
GCP[1]	0.087	0.520	0.869	0	6.966
Border distance[1]	5.740	5.600	1.461	0	9.305
Polity2	6.000	4.485	6.143	-10.000	10.000
Population[1]	8.313	7.812	3.708	0	16.690
Travel time[1]	6.230	6.339	1.250	0	10.310
Cell area[1]	7.696	7.392	1.022	-12.880	8.039
Mil. ex.[1,2]	16.030	15.680	2.624	0	20.130
Mountains[3]	0.2	0.4	0.352	0	1
Large urban area[4]	0	0.233	0.423	0	1
Nighttime light[1]	0	2.593	3.011	0	8.189
GED atrocities	0	0.019	0.594	0	121
Press	36	42.939	27.130	5	100

All independent variables are lagged by one year.
[1] Natural log.
[2] This variable is only available for the years 1996-2008.
[3] This variable is only available for the years 2001-2009.
[4] This variable is only coded for urbanized areas.

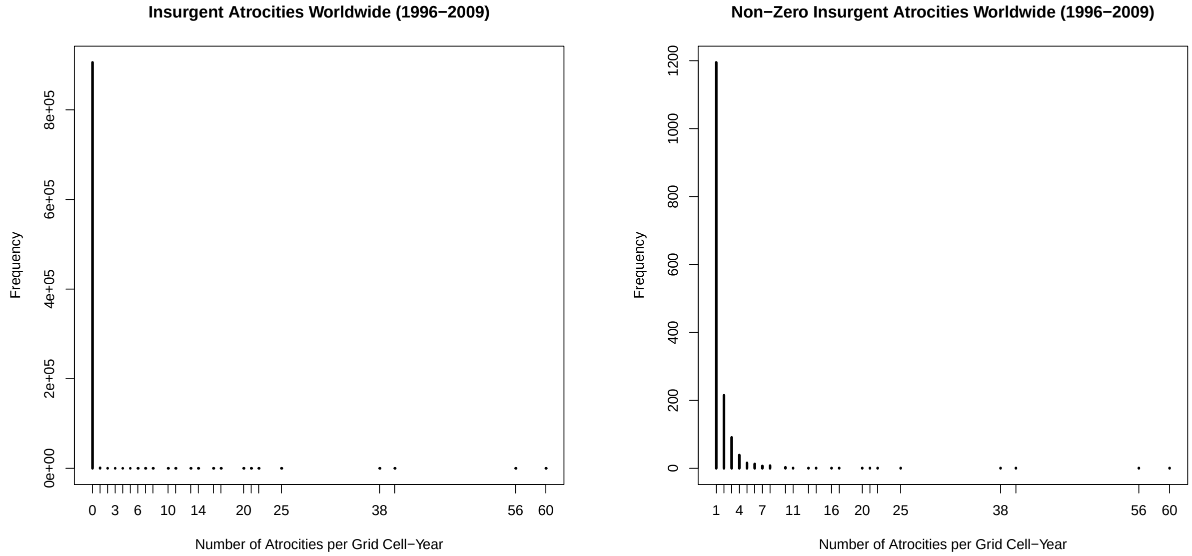


Figure A2: Counts of Insurgent Atrocity Incidents by Grid Cell Worldwide, 1996-2009

Clustering of Atrocity Incidents in National Capitals

In addition to the data, methods, and summary statistics reported above, Tables A2 highlights the concentration of insurgent atrocities in different countries, and atrocities in the capital as a percentage of all atrocities in any countries that experienced atrocities.

Table A2: Insurgent Atrocities As Percent of All Atrocities, 1996-2009

Country	Capital	All	Ratio	Country	Capital	All	Ratio
United States	1	18	5.56%	Iran	1	7	14.29%
Haiti	2	8	25%	Iraq	195	370	52.7%
Dom. Rep.	0	1	0%	Nicaragua	0	2	0%
Jamaica	1	1	100%	Egypt	2	8	25%
Mexico	4	30	13.33%	Syrian Arab Republic	1	1	100%
Guatemala	8	11	72.73%	Lebanon	2	4	50%
Honduras	1	5	20%	Jordan	1	1	100%
El Salvador	1	1	100%	Israel	27	55	49.09%
Colombia	0	213	0%	Ecuador	0	2	0%
Venezuela	1	3	33.33%	Saudi Arabia	2	4	50%
Guyana	1	2	50%	Afghanistan	43	150	28.67%
Peru	6	7	85.71%	Kyrgyzstan	1	1	100%
Tajikistan	0	1	0%	China	0	12	0%
Brazil	0	18	0%	Portugal	0	2	0%
United Kingdom	1	2	50%	Uzbekistan	2	3	66.67%
Spain	1	1	100%	India	5	317	1.58%
Poland	0	1	0%	Hungary	0	1	0%
Italy	0	2	0%	Yugoslavia	0	5	0%
Albania	1	1	100%	Madagascar	1	1	100%
Russian Federation	8	43	18.61%	Pakistan	7	49	14.29%
Lithuania	0	1	0%	Ukraine	0	1	0%
Finland	1	1	100%	Bangladesh	5	44	11.36%
Guinea-Bissau	1	1	100%	Myanmar	1	5	20%
Mali	0	3	0%	Senegal	0	13	0%
Benin	1	1	100%	Sri Lanka	13	65	20%
Niger	0	2	0%	Cote d'Ivoire	0	8	0%
Guinea	0	4	0%	Burkina Faso	0	2	0%
Liberia	2	9	22.22%	Nepal	5	11	45.46%
Sierra Leone	3	10	30%	Thailand	1	9	11.11%
Ghana	0	1	0%	Cameroon	0	3	0%
Nigeria	1	84	1.19%	Cambodia	1	9	11.11%
CAR	0	1	0%	Chad	0	6	0%
Congo	3	5	60%	Philippines	3	39	7.69%
Laos	0	2	0%	Malaysia	0	2	0%
DRC	0	95	0%	Tanzania	0	10	0%
Uganda	9	88	10.23%	Indonesia	7	50	14%
Kenya	2	32	6.25%	Solomon Islands	2	2	100&
Rwanda	24	44	54.55%	Burundi	49	91	53.85%
Somalia	56	82	68.29%	Sudan	3	66	4.55%
Ethiopia	0	2	0%	South Africa	0	8	0%
Eritrea	1	2	50%	Angola	1	40	2.5%
Mozambique	1	2	50%	Algeria	107	239	44.77%
Morocco	0	1	0%	Libya	0	2	0%
Turkey	0	14	0%	Yemen[1]	0	5	0%
Papua New Guinea	0	1	0%				

[1] Data for both Yemen (North) and Yemen (after 1990)

CEM Plots

This section reports all the CEM matching plots for the analyses conducted in the main paper (Figures A3 – A6), as well as the robustness CEM models analyzed in Table A3 (Figures A7 and A8).

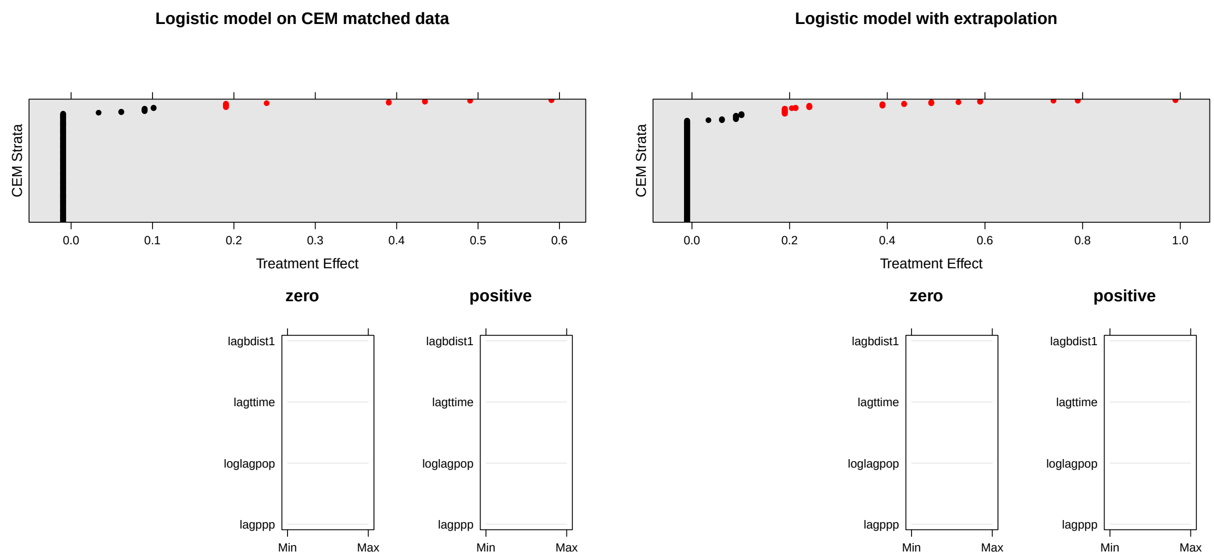


Figure A3: CEM Plots for Urbanized Regions in Atrocity Affected Countries, Grid Cell Years, 1996-2009

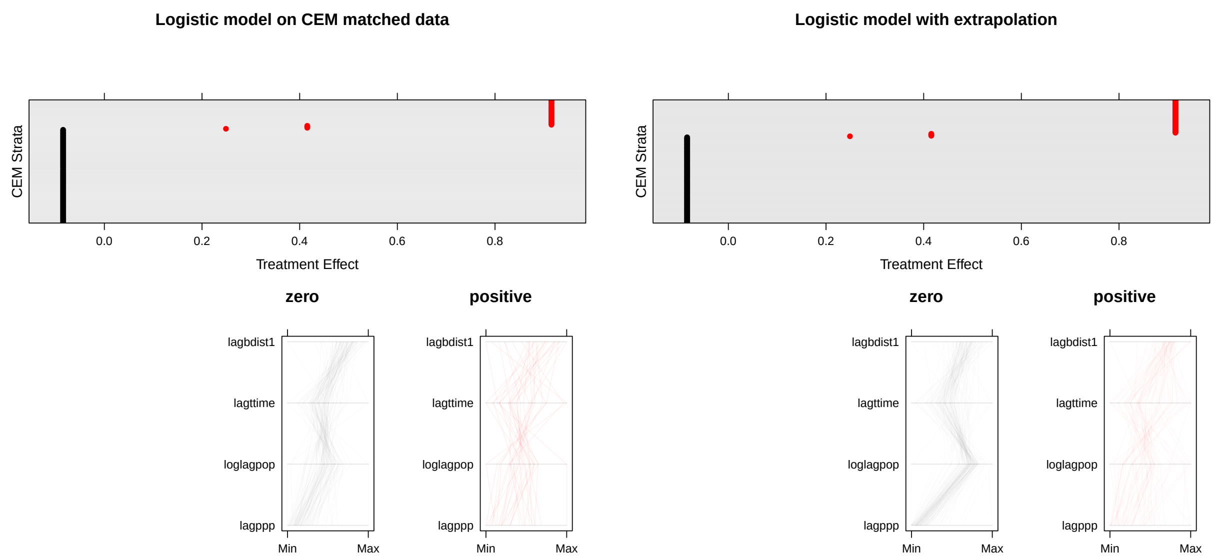


Figure A4: CEM Plots for Urbanized Regions in Atrocity Affected Countries, Average Cell Values, 1996-2009

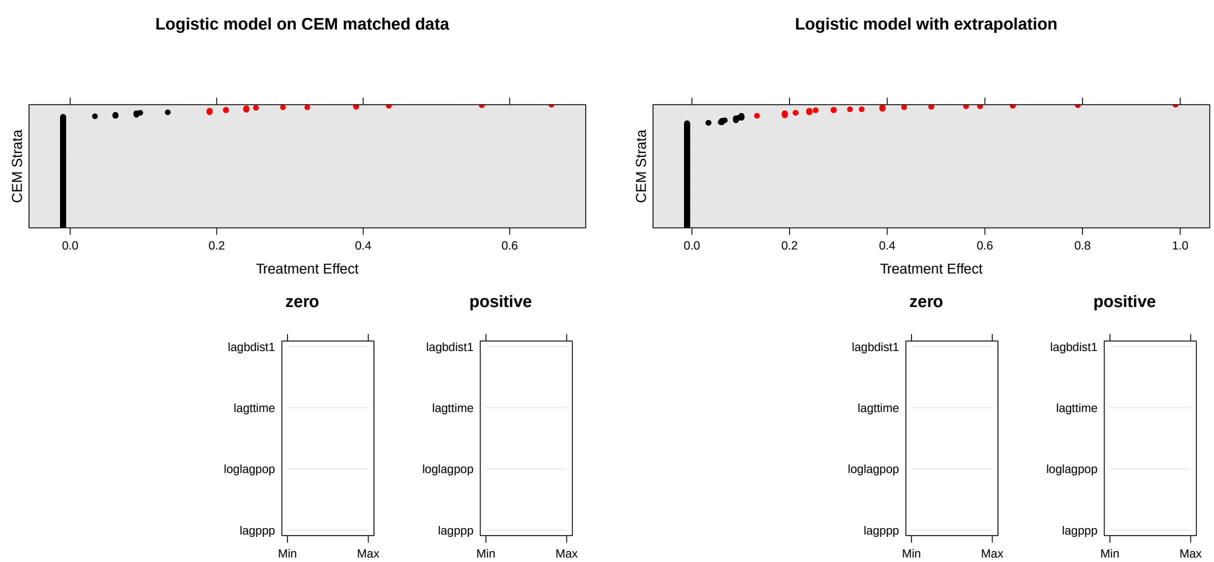


Figure A5: CEM Plots for Nighttime Light Regions in Atrocity Affected Countries, Grid Cell Years, 1996-2009

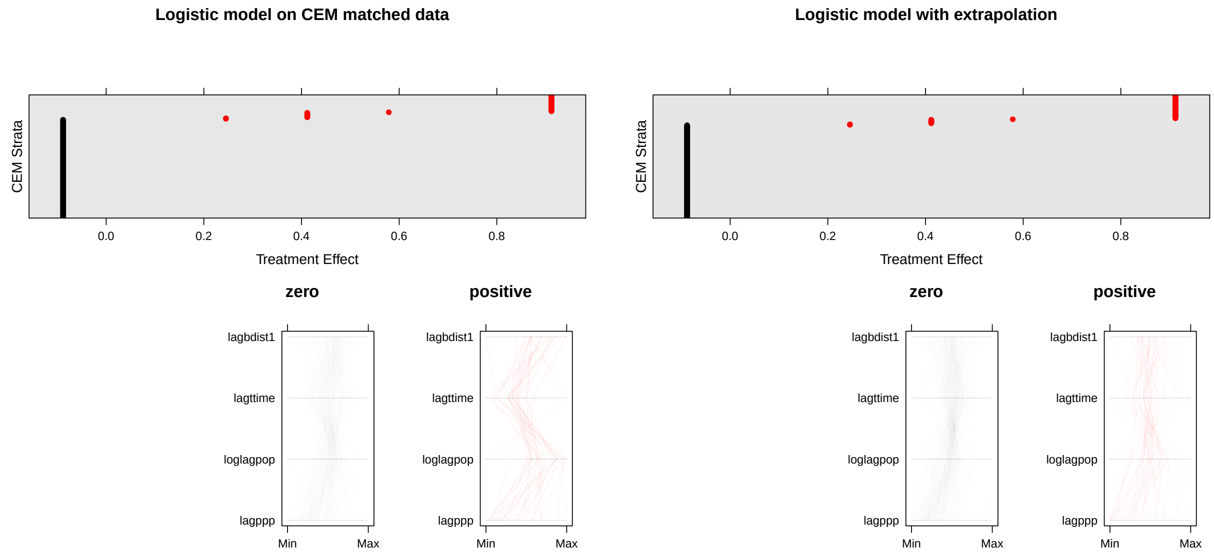


Figure A6: CEM Plots for Nighttime Light Regions in Atrocity Affected Countries, Average Cell Values, 1996-2009

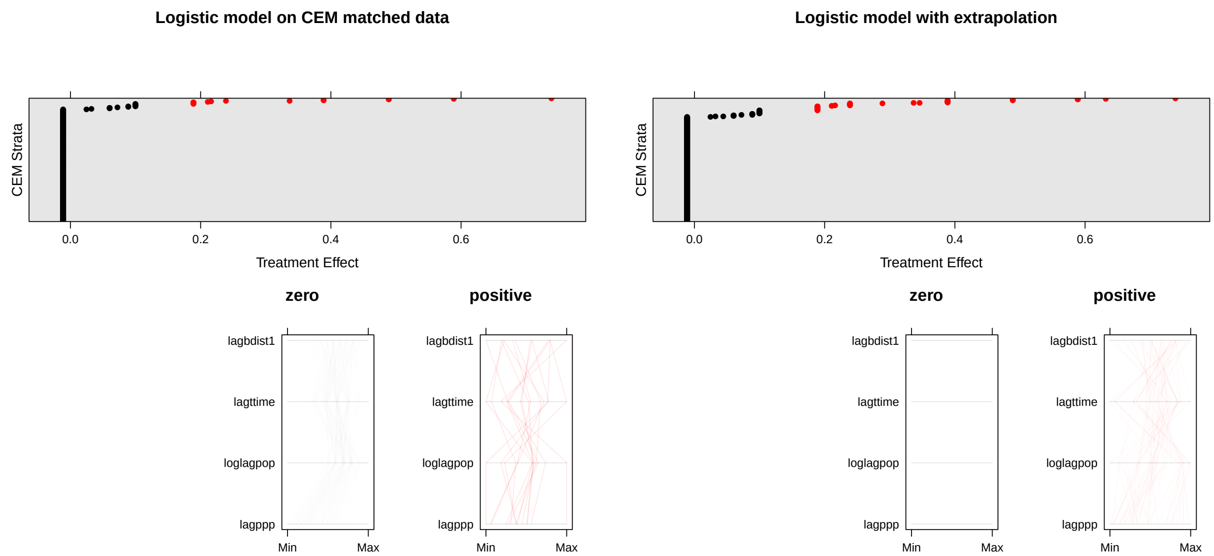


Figure A7: CEM Plots for All Grid Cells in Atrocity Affected Countries, Grid Cell Years, 1996-2009

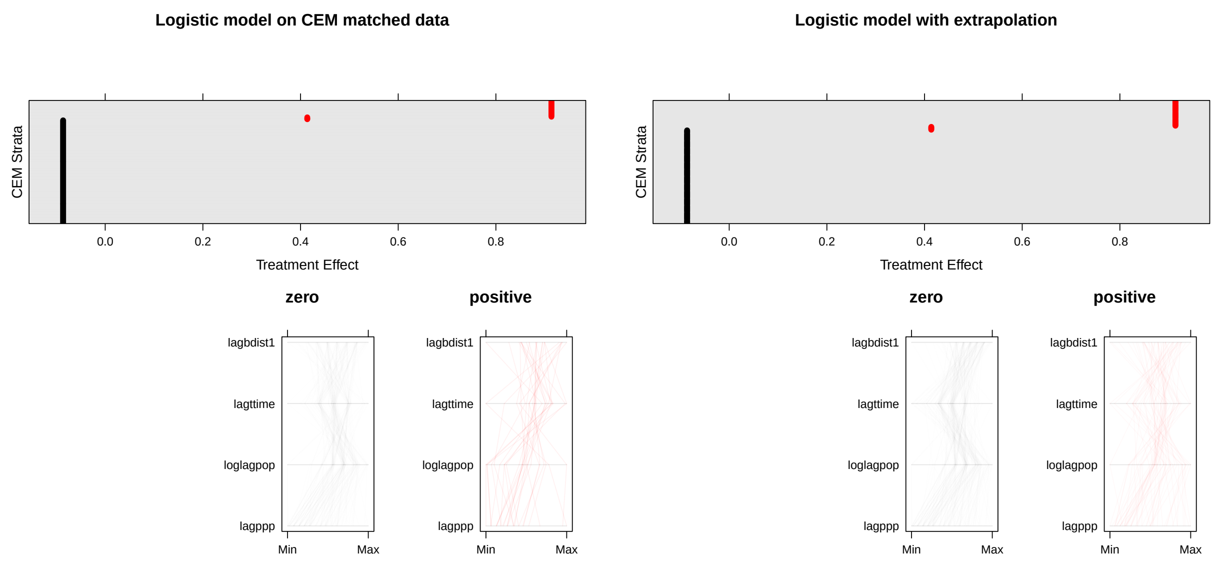


Figure A8: CEM Plots for Urbanized Regions in Atrocity Affected Countries, Average Cell Values, 1996-2009

Robustness Analyses

In this section I report several stages of robustness analyses corresponding to the Global Analysis section of the main paper. These robustness analyses include: (i) CEM exercises estimated on a sample of all grid cells in atrocity affected countries, not only urban or nighttime light cells; (ii) 12 ZINB models with different specifications and confounders corresponding to Table 5 in the main paper; (iii) a set of ZINB analyses that account for the effect of large urban areas conducted on a sample of only urban cells globally, and again, only for African urban cells; and (iv) a set of models that includes nighttime light to account for reporting biases estimated on all global grid cells and then again on solely urban cells.

Coarsened Exact Score Matching Sensitivity Analysis

I begin this section with an analysis of a sample matched using coarsened exact score matching on all grid cells in countries that experienced at least one atrocity event, not just urban or nighttime light cells as used in the main paper. To this extent, Table A3 replicates of the four logit models used in the main paper on a sample containing all grid cells within atrocity prone countries. As Table A3 shows, the findings presented in the first stage of analysis in the main paper are robust to the expansion of the sample to all affected countries. It is also important to stress that these findings hold when I run the same analyses on a sample consisting of *all* grid cells and grid cell years, although these analysis were not reported here.

Table A3: Coarsened Exact Matching Sample Analysis (Treatment=*Capital*)

	Cell Years		Averaged Cell Values	
	Matched Sample Only	**Entire Sample**	**Matched Sample Only**	**Entire Sample**
Capital_{t-1}	0.676***	494***	0.524**	0.476***
	(0.146)	(0.103)	(0.219)	(0.162)
Constant	-4.493 ***	–	-2.364***	–
	(0.058)		(0.070)	
N	595,486		42,130	
Matched (m_t/m_c)	2,650/27,256		197/2,566	
Unmatched (u_t/u_c)	1,123/564,457		70/39,297	

Note: *p<0.1; **p<0.05; ***p<0.01. m_t – matched treatment group; m_c – matched control group; u_t – unmatched treatment group; u_c – unmatched control group

Global Sample Sensitivity Analysis

The robustness of the findings presented in the second stage of analysis in the main paper is illustrated in two phases. I begin by reassessing my empirical models under nine alternative specifications. For each sensitivity analysis, the "large" model specifications presented in Table 2 in the main paper (i.e., Model 4) is estimated as a threshold model to assess the stability of each *capital* coefficient therein. First, to provide additional validation of my findings in respect to reporting biases as well as for other potential factors that affect both the grid cell and the country and are not captured by my independent variables, I employ a Bayesian framework that accommodates random effects for each grid cell and – separately – for each country worldwide in Model A-I. These random effects assume a hierarchical structure in which some country- and cell-specific factors might still independently influence the incidence of insurgent atrocities. The use of these random effects therefore allows taking into consideration the possibility that some cells might be me more likely to experience insurgent atrocities, but that these incidents might remain underreported due to the lack of journalistic presence in the region. Model A-I was estimated using 100,000 MCMC simulations, the first 90,000 of which were discarded as burn-in.

Second, to further assess that the effect of *capital* is not the result of factors related to the size of a given country's military during a given year, national military expenditure at a given year, *mil. exp.*, was included in Model A-II. Third, to ensure that the significance of the findings did not arise by a lax account of rural conditions that might accommodate conflict, I add to my model the percent of a given cell denotes as mountainous, *mountains*, in Model A-III. Fourth, to account for the potential heterogeneity of the error term (i.e., that errors are time dependent between one point and time and the next within specific locations) in respect to both the country and the cell, corresponding zero inflated models with the standard errors clustered by country and – alternatively – by grid cell are estimated in Models A-IV and A-V, respectively. Fifth, the effect of country specific factors is taken more throughly into account in Model A-VI, which includes binary indicators (i.e., fixed effects) for each of the countries analyzed to verify that the significant effect of *capital* does not result from a concentration of a high number of atrocities in the capitals of one or two violence-prone countries. Sixth, recall that the U.S. was not coded in the PITF atrocities data per the requirements of its funding sources. Drawing on independent research into historical records and the PITF code book, three atrocity incidents were coded during my years of interest, with three separate incidents corresponding to each location of the 9/11 attacks and one to the Oklahoma City Bombing were added to the sample in the main analyses. Nevertheless, to illustrate the robustness of my findings, Model A-VII estimates a sample where the United States are removed from analysis.

Seventh, note that some studies argue that the distance to capital is associated with an increased likelihood of conflict and political violence (Raleigh and Hegre, 2009). To show that it is capital cities, in-and-of themselves, that are more likely to experience insurgent atrocities, thus further ensuring that the significant association with insurgent atrocities is not the result of proximity – i.e. that attacks become more frequent as we move closer to the capital, but rather of the specific importance of capitals – the distance from each cell to the capital is included as a robustness measure in Model A-VIII. Next, recall that the grid cell framework, while highly effective in capturing the effects of spatial variations, is not necessarily a meaningful unit of analysis from a *theoretical* perspective. To address this concern, Model A-IX re-estimates the threshold model on a sample that has been aggregated to the district/province administrative unit level. Tenth, to account for the possibility that the strong associations between *capital* and *insurgent atrocities* is not the result of particular data choices, I reestimate Model 4 on a sample where the *insurgent atrocities* and the one-year lag *atrocities* variables were coded based on the Georeferenced Event Dataset (GED) (Sundberg and Melander, 2013) in Model A-X. The *insurgent atrocities* variables was operationalized based on the number of one-sided violence incidents in the GED dataset that were perpetrated by insurgent actors, while the *atrocities* variable was operationalized as the one-year lag of *insurgent atrocities*.[4] Note that GED data cover the entire terrestrial globe only starting 2005, which provides one reason as to why I chose to rely on PITF data in my main analyses. Eleventh,

Eleventh, it is possible that the event data used to compile the dependent variable is affected by media transparency and the ability of the press to report certain incidents in certain countries. To address this potential for reporting bias I incorporate a country level indicator coding freedom of the press levels, *press* coded by Freedom House (2017), at both the count and inflation stages in Model A-XI. Note that *higher* scores on this variable correspond to *less freedom* of the press. Finally, recall that, as Table A2 illustrates, a few countries in my shape experienced the lion's share of all insurgent atrocities, which might produce inferential biases. Accordingly, Model A-XII re-estimates Model 4 on a sample where 18 outlier countries that experienced the highest number of atrocities over the 1996-2009 period – i.e., countries in the 90th percentile of all countries that experienced at least

[4]This is different than with the PITF model, where the atrocities lag includes all atrocities by both state and nonstate actors. The reason for this difference was the result of the fact that the GED codes all one-sided violence events (Sundberg and Melander, 2013), while the PITF codes several different types of violence by different actors (PITF, 2009).

one insurgent atrocity – were removed to show that the findings are not driven by these outliers.[5] Critically, in all these robustness models, the *capital* coefficient maintains its sign and significance.

Urban Sample Analysis

To further illustrate that insurgents prefer to target civilians in capitals because these cities have a specific political value, the third stage of these robustness analyses examines the effect of *capital* on *insurgent atrocities* within urbanized areas, i.e. cells that included some degree of urban development, specifically (Bontemps, Defourny and Van Bogaert, 2009). To compare the effect of capitals to that of large urban centers, a new indicator, *large city*, is created to denote whether a given cell was included within the 95th percentile of all cells with any urban coverage, a value corresponding to a cell that is more than 4.38% urbanized. The effect of *capital*, taking into account large cities, is estimated on a *global sample of all urbanized cells*, and then again on a sample of *urbanized cells located solely in African countries*. In doing so, I further account for the possibility that in less urbanized regions where cities are likely to be rare – and as a result the accessibility to large concentrations of potential targets – atrocities by insurgents in capitals are nevertheless the result of the specific value of these cities, and not of their relatively large urban densities. These models also better account for potential reporting biases, because large cities in African countries are more likely to have, among other relevant factors, good cell-phone coverage compared with the rural countryside, suggesting that if bias exists it would affect all large urban areas equally, not just the capital (Weidmann, 2016). The global distribution and density of urbanized cells are presented in Figure A9.

Each of the ZINB models in Table A6 corresponds to an alternative specification of Model 4. In Model A-XIII-A, the variable *urban* is omitted from both stages, and the variable

[5] These countries were Colombia, Russia, Nigeria, the DRC, Uganda, Burundi, Rwanda, Somalia, Algeria, Sudan, Iraq, Israel, Afghanistan, India, Pakistan, Sri Lanka, and Indonesia.

Table A4: Zero-Inflated Negative Binomial Estimates of Atrocities, Robustness Analyses

	Model A-I (Hier.)[2]	**Model A-II** (Mil. Ex.)	**Model A-III** (Mnts.)	**Model A-IV** (CSEs)	**Model A-V** (GSEs)	**Model A-VI** (GFEs)[3]
Count Stage						
Capital	1.182***	0.526***	0.714***	0.687***	0.700***	0.680***
	(0.921 ⇔ 1.400)	(0.136)	(0.188)	(0.140)	(0.135)	(0.149)
Civil conflict	-0.166***	0.531***	1.746***	0.803***	0.759***	0.581***
	(-0.229 ⇔ -0.099)	(0.122)	(0.118)	(0.112)	(0.117)	(0.145)
Atrocities	0.062***	0.396***	1.404***	0.474***	0.416***	0.431***
	(0.040 ⇔ 0.086)	(0.039)	(0.114)	(0.043)	(0.037)	(0.046)
Urban	−0.056***	−0.013	0.052***	0.005	−0.008	0.035***
	(-0.079 ⇔ -0.026)	(0.011)	(0.020)	(0.011)	(0.011)	(0.012)
Oil[1]	0.025***	−0.029**	0.026***	0.009	0.021**	0.149***
	(0.012 ⇔ 0.029)	(0.013)	(0.009)	(0.010)	(0.010)	(0.043)
GCP[1]	0.240***	0.018	−0.413***	−0.030	0.127	−0.274***
	(0.138 ⇔ 0.331)	(0.090)	(0.097)	(0.086)	(0.088)	(0.101)
Border distance[1]	−0.061***	−0.015	−0.180***	−0.122**	−0.118**	−0.235***
	(-0.080 ⇔ -0.036)	(0.051)	(0.044)	(0.051)	(0.054)	(0.062)
Polity2	-0.048***	−0.022*	0.119***	−0.005	0.001	0.041**
	(-0.056 ⇔ -0.041)	(0.013)	(0.014)	(0.012)	(0.012)	(0.020)
Population[1]	-0.310***	0.116**	0.467***	0.190***	0.102*	0.215***
	(-0.323 ⇔ -0.295)	(0.054)	(0.065)	(0.055)	(0.054)	(0.065)
Travel time[1]	-0.268***	0.055	−0.454***	0.133	0.061	0.030
	(-0.300 ⇔ -0.242)	(0.087)	(0.101)	(0.087)	(0.085)	(0.098)
Cell area[1]	0.125***	0.168	0.296**	−0.051	0.020	−0.214
	(0.097 ⇔ 0.159)	(0.131)	(0.137)	(0.125)	(0.132)	(0.171)
Mil. ex.[1]	–	0.185***	–	–	–	–
		(0.053)				
Mountains	–	–	−0.175	–	–	–
			(0.164)			
Constant	2.543***	−7.092***	−9.926***	−4.614***	−2.426**	−6.700***
	(2.320 ⇔ 2.796)	(1.343)	(1.265)	(1.058)	(1.092)	(1.584)
Inflation Stage						
Civil conflict	-2.399***	−1.648***	−0.608**	−1.508***	−1.420***	−1.005***
	(-2.523 ⇔ -2.290)	(0.134)	(0.293)	(0.131)	(0.130)	(0.150)
Atrocities	-2.600***	−4.269***	−18.679	−10.201	−4.213***	−3.723***
	(-2.829 ⇔ -2.404)	(0.425)	(1,239.329)		(0.412)	(0.525)
Urban	-0.184***	−0.109***	−0.089**	−0.107***	−0.109***	−0.136***
	(-0.216 ⇔ -0.135)	(0.016)	(0.039)	(0.017)	(0.017)	(0.035)
Oil[1]	0.024***	−0.087***	0.090**	0.010	0.012	0.068***
	(0.015 ⇔ 0.031)	(0.015)	(0.035)	(0.011)	(0.011)	(0.013)
GCP[1]	0.630***	0.250**	−0.074	0.519***	0.463***	0.080
	(0.499 ⇔ 0.733)	(0.105)	(0.203)	(0.103)	(0.102)	(0.119)
Border distance[1]	0.192***	0.212***	0.013	0.133**	0.199***	0.043
	(0.163 ⇔ 0.219)	(0.054)	(0.085)	(0.056)	(0.056)	(0.066)
Polity2	-0.046***	−0.041***	0.564***	−0.006	0.0002	0.033**
	(-0.055 ⇔ -0.036)	(0.014)	(0.123)	(0.013)	(0.012)	(0.015)
Population[1]	-0.874***	−0.638***	−0.607***	−0.661***	−0.593***	−0.481***
	(0.910 ⇔ -0.817)	(0.060)	(0.118)	(0.061)	(0.060)	(0.069)
Travel time[1]	-0.356***	−0.134	−1.238***	−0.114	−0.013	−0.087
	(-0.418 ⇔ -0.305)	(0.103)	(0.211)	(0.103)	(0.100)	(0.114)
Cell area[1]	0.181***	0.371**	0.965***	0.131	0.227	0.076
	(0.072 ⇔ 0.293)	(0.155)	(0.293)	(0.145)	(0.144)	
Mil. ex.[1]	–	0.489***	–	–	–	–
		(0.056)				
Mountains	–	–	−0.736**	–	–	
			(0.375)			
Constant	15.523***	2.691*	2.526	9.931***	7.751***	7.209***
	(15.037 ⇔ 16.025)	(1.493)	(2.989)	(1.272)	(1.231)	(1.378)
N	800,634	744,941	515,599	800,634	803,795	800,634
AIC/DIC	12,362.33	12,827.67	8,696.18	13,588.86	13946.58	12,970.23

Note: *p<0.1; **p<0.05; ***p<0.01; year fixed effects included in each regression, though not reported here. All independent variables are lagged by one year.

[1] Natural log

[2] Based on 100,000 iterations, 90,000 burn-ins, thinning of 50. Values in parentheses for this model are 95% credible intervals

[3] Fixed effects by country not reported here

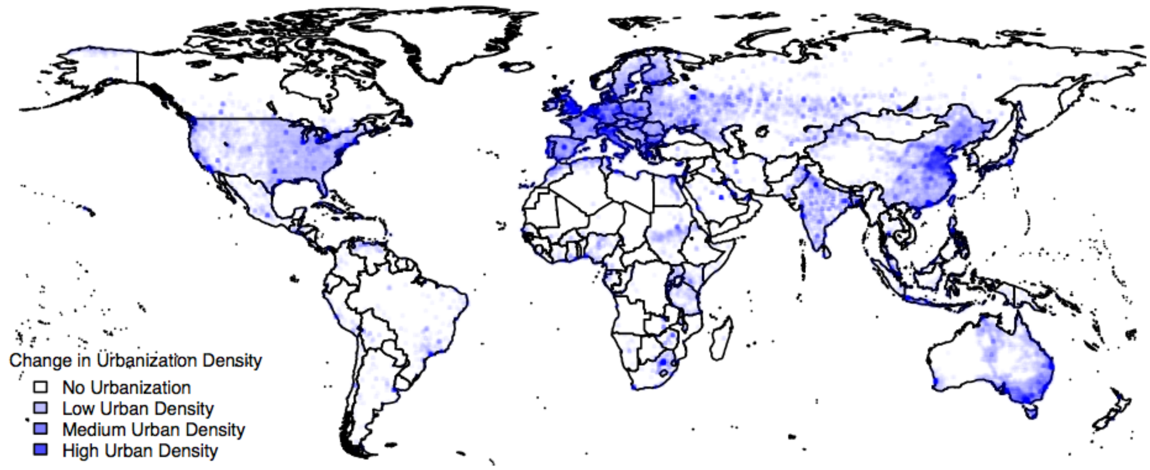
Table A5: Zero-Inflated Negative Binomial Estimates of Atrocities, Robustness Analyses (Continued)

	Model A-VII (No U.S.)	**Model A-VIII** (Cap. Dist.)	**Model A-IX** (Province)	**Model A-X** (GED)	**Model A-XI** (FP)	**Model A-XII** (Outliers)
Count Stage						
Capital	0.759***	0.622***	1.118***	0.499***	0.761***	1.067***
	(0.137)	(0.139)	(0.103)	(0.127)	(0.139)	(0.212)
Civil conflict	0.652***	0.620***	1.075***	0.416***	0.628***	0.446
	(0.120)	(0.119)	(0.117)	(0.087)	(0.119)	(0.296)
Atrocities	0.406***	0.424***	0.596***	0.137***	0.423***	0.629***
	(0.041)	(0.041)	(0.082)	(0.009)	(0.041)	(0.172)
Urban	−0.007	−0.006	−0.007	0.019	−0.023*	−0.032
	(0.011)	(0.011)	(0.021)	(0.013)	(0.014)	(0.028)
Oil[1]	−0.001	0.017*	0.064***	−0.017***	0.013	0.028
	(0.010)	(0.010)	(0.009)	(0.005)	(0.010)	(0.019)
GCP[1]	−0.023	−0.025	−0.337***	−0.361***	0.044	0.083
	(0.089)	(0.090)	(0.106)	(0.053)	(0.087)	(0.209)
Border distance[1]	0.011	−0.092*	0.011	−0.135***	−0.072	−0.124
	(0.055)	(0.053)	(0.055)	(0.027)	(0.052)	(0.094)
Polity2	0.007	0.005	0.063***	0.033***	−0.006	0.036
	(0.012)	(0.012)	(0.010)	(0.007)	(0.014)	(0.030)
Population[1]	0.101*	0.093*	0.072	0.130***	0.041	−0.222
	(0.055)	(0.054)	(0.055)	(0.033)	(0.054)	(0.135)
Travel time[1]	0.080	0.007	0.337***	0.094	−0.053	−0.474**
	(0.087)	(0.087)	(0.109)	(0.061)	(0.090)	(0.226)
Cell area[1]	0.103	0.093	0.383***	−0.152	0.129	0.135
	(0.125)	(0.126)	(0.127)	(0.101)	(0.125)	(0.264)
Cap. dist.	–	−0.0002***	–	–	–	–
		(0.0001)				
Press	–	–	–	–	−0.007*	–
					(0.004)	
Constant	−5.192***	−4.072***	−7.651***	0.204	−3.491***	0.222
	(1.082)	(1.073)	(1.266)	(0.861)	(1.151)	(2.410)
Inflation Stage						
Civil conflict	−1.501***	−1.502***	−1.256***	−2.403***	−1.315***	−2.084***
	(0.134)	(0.134)	(0.157)	(0.082)	(0.135)	(0.354)
Atrocities	−4.207***	−4.334***	−610.170***	−3.938***	−4.611***	−6.166***
	(0.411)	(0.462)	(158.320)	(0.225)	(0.519)	(1.597)
Urban	−0.117***	−0.118***	−0.201***	−0.066***	−0.132***	−0.193***
	(0.017)	(0.017)	(0.064)	(0.016)	(0.022)	(0.053)
Oil[1]	0.005	0.010	0.003	0.009*	0.028**	0.061***
	(0.011)	(0.011)	(0.012)	(0.005)	(0.011)	(0.022)
GCP[1]	0.413***	0.452***	0.460***	0.208***	0.373***	0.923***
	(0.106)	(0.104)	(0.143)	(0.065)	(0.103)	(0.234)
Border distance[1]	0.264***	0.177***	−0.061	0.248***	0.178***	−0.035
	(0.058)	(0.056)	(0.067)	(0.025)	(0.055)	(0.109)
Polity2	−0.006	−0.007	0.043***	0.001	−0.095***	0.009
	(0.013)	(0.012)	(0.013)	(0.007)	(0.017)	(0.033)
Population[1]	−0.593***	−0.628***	−0.509***	−0.476***	−0.611***	−0.966***
	(0.061)	(0.061)	(0.071)	(0.034)	(0.061)	(0.145)
Travel time[1]	−0.067	−0.154	−0.250*	−0.022	−0.179*	−0.312
	(0.102)	(0.104)	(0.135)	(0.066)	(0.107)	(0.271)
Cell area[1]	0.198	0.225	0.230	−0.122	0.276*	0.165
	(0.145)	(0.147)	(0.154)	(0.105)	(0.145)	(0.294)
Press	–	–	–	–	−0.035***	–
					(0.005)	
Constant	8.297***	9.208***	7.513***	10.256***	10.799***	13.970***
	(1.245)	(1.239)	(1.571)	(0.912)	(1.319)	(2.639)
N	732,552	803,795	24,026	436,186	622,787	547,675
AIC/DIC	13,505.80	13,998.32	8,301.37	25,035.33	12,125.99	4,502.99

Note: *p<0.1; **p<0.05; ***p<0.01; year fixed effects included in each regression, though not reported here. All independent variables are lagged by one year.

[1] Natural log

Figure A9: The Distribution and Density of Urbanization by Cell Worldwide, 1996-2009



large city is included in the count stage only. In Model A-XIII-B *large city* is included in both inflation and count stages to account for the possibility that – due to their relatively wide global spread – large cities might systematically dispose a cell to experience insurgent atrocities. Model A-XIII-C is similar to Models A-XIII-A and A-XIII-B, only this time the percentage of a given cell's degree of urbanization, *urban* is added to the inflation stage, while *large city* is added to the count stage. Models A-XIV-A – C then reestimate the same specifications on a sample of cells located only in Africa. The coefficient of *capital* is positive and significant ($p < 0.01$) across these different samples and specifications, which confirms the results of the main analyses presented in Table 5 in the main paper.

Models A-XIII-A – C also suggest that large cities are significantly more disposed to experiencing insurgent atrocities. Yet, the direction and significance of *capital* strongly supports the hypothesis that the effect of capital cities is distinct from that of large cities. Moreover, Models A-XIII-A – C show that even in African countries, where (as Figure A9 illustrates) large cities are relatively rare, the effect of *capital* on *insurgent atrocities* remains significant, but not so in the case of *large city*. Large cities offer a large number of targets and easy access to insurgent perpetrators. Yet, as Table A6 shows – and the CEM models

Table A6: Zero-Inflated Negative Binomial Estimates of Insurgent Atrocities in Urban Areas, 1996-2009

	Model A-XIII (Global Sample)			A-XIV (Africa Sample)		
	A-XIII-A	**A-XIII-B**	**A-XIII-C**	**A-XVI-A**	**A-XVI-B**	**A-XVI-C**
Count Stage						
Capital	0.764***	0.765***	0.764***	0.656***	0.649***	0.660***
	(0.168)	(0.168)	(0.168)	(0.207)	(0.206)	(0.208)
Large city	0.510***	0.500**	0.321*	0.121	−0.264	0.155
	(0.178)	(0.226)	(0.193)	(0.287)	(0.435)	(0.348)
Civil conflict	2.446***	2.446***	2.419***	0.831***	0.885***	0.825***
	(0.134)	(0.134)	(0.134)	(0.242)	(0.264)	(0.242)
Atrocities(lag)	0.816***	0.816***	0.809***	0.175***	0.187***	0.174***
	(0.085)	(0.085)	(0.085)	(0.042)	(0.045)	(0.043)
Oil[1]	−0.035***	−0.035***	−0.036***	−0.007	−0.005	−0.008
	(0.012)	(0.012)	(0.011)	(0.018)	(0.018)	(0.018)
GCP[1]	−0.056	−0.055	−0.023	−0.048	−0.022	−0.046
	(0.102)	(0.103)	(0.101)	(0.208)	(0.208)	(0.208)
Border distance[1]	−0.125**	−0.125**	−0.121**	0.073	0.055	0.074
	(0.055)	(0.055)	(0.055)	(0.089)	(0.091)	(0.089)
Polity2	0.131***	0.131***	0.130***	−0.093***	−0.111***	−0.091**
	(0.016)	(0.016)	(0.016)	(0.035)	(0.042)	(0.037)
Population[1]	0.372***	0.373***	0.377***	0.146	0.170	0.142
	(0.079)	(0.080)	(0.078)	(0.127)	(0.132)	(0.129)
Travel time[1]	−0.004	−0.005	−0.060	0.431**	0.414**	0.430**
	(0.115)	(0.116)	(0.117)	(0.195)	(0.196)	(0.195)
Cell area[1]	0.160	0.159	0.174	0.130	−0.050	0.153
	(0.162)	(0.162)	(0.162)	(0.315)	(0.342)	(0.340)
Constant	−10.210***	−10.205***	−10.124***	−8.172***	−7.009**	−8.287***
	(1.553)	(1.554)	(1.548)	(3.047)	(3.150)	(3.125)
Inflation Stage						
Large city	–	−0.027	–	–	−0.796	–
		(0.365)			(0.749)	
Civil conflict	0.353	0.350	0.180	−1.542***	−1.498***	−1.547***
	(0.264)	(0.266)	(0.268)	(0.286)	(0.305)	(0.286)
Atrocities	−2.791***	−2.788***	−2.775***	−4.160***	−4.335***	−4.142***
	(0.716)	(0.719)	(0.744)	(0.967)	(1.116)	(0.957)
Urban	–	–	−0.064***	–	–	0.009
			(0.023)			(0.052)
Oil[1]	−0.021	−0.021	−0.020	−0.036*	−0.035	−0.036*
	(0.026)	(0.026)	(0.026)	(0.022)	(0.023)	(0.022)
GCP[1]	0.254*	0.256*	0.351**	0.142	0.187	0.140
	(0.152)	(0.155)	(0.152)	(0.252)	(0.260)	(0.252)
Border distance[1]	0.185**	0.185**	0.199**	0.248**	0.230*	0.250**
	(0.093)	(0.093)	(0.093)	(0.117)	(0.120)	(0.117)
Polity2	0.509***	0.509***	0.494***	−0.152***	−0.178***	−0.149***
	(0.060)	(0.060)	(0.061)	(0.044)	(0.060)	(0.045)
Population[1]	−0.840***	−0.838***	−0.775***	−0.649***	−0.631***	−0.654***
	(0.133)	(0.136)	(0.133)	(0.155)	(0.162)	(0.156)
Travel time[1]	−0.311	−0.315	−0.541**	−0.052	−0.082	−0.046
	(0.209)	(0.216)	(0.227)	(0.252)	(0.258)	(0.254)
Cell area[1]	0.484	0.482	0.527*	0.358	0.131	0.383
	(0.306)	(0.308)	(0.307)	(0.399)	(0.426)	(0.425)
Constant	5.508**	5.515**	5.493**	7.260*	8.900**	7.091*
	(2.602)	(2.605)	(2.593)	(3.770)	(3.835)	(3.924)
N		164,259			16,155	
AIC	6,328.72	6,330.71	6,322.16	2,797.19	2,797.85	2,799.16

Note: *p<0.1; **p<0.05; ***p<0.01; year fixed effects included in each regression, though not reported here. All independent variables are lagged by one year.

[1] Natural log

presented above and in the main paper illustrate – capital cities *systematically* experience a higher frequency of atrocities perpetrated by insurgents compared with other (large) urban areas. This suggests – again – that the main paper's findings are not driven by reporting bias, at least in relation to the difference between capital cities and other urban or developed areas.

Nighttime Light Sample Analysis

In this final robustness phase, two sets of models that include the amount of nighttime light for a given cell at both the inflation and count stages and are estimated on a global sample consisting of all cells, and then again on a sample composed of urban cells only, to specifically account for the effect of reporting bias. Here, higher levels nighttime light approximate cells where infrastructure conducive of reporting is more likely to exist (Weidmann, 2016; Koren and Sarbahi, Forthcoming), and hence that the propensity of these cells to (i) be inflated, and (ii) include reports of higher numbers of atrocities due to a higher probability of reporting.

In the absence of *global* data on newspaper reporting bias at the highly disaggregated level, nighttime light – similar to other proxies such as cellular phone cover (Weidmann, 2016) – provides good approximation for potential reporting bias by highlighting areas where information on atrocities is more likely to be obtained and from which it can be transmitted. The distribution of electricity is more likely in areas where the government can regulate and provide public goods (Koren and Sarbahi, Forthcoming). From this perspective, nighttime light can be used to identify peripheral regions where reporting is more or less likely. Nighttime light also characterizes regions and times where economic development is pursued, a process that involves the expansion of infrastructure including roads, telecommunication and electricity. Indeed, nighttime light has shown to be closely correlated with cell phone coverage in Western Africa (Martinez-Cesena et al., 2015).

I therefore chose to use nighttime light data as an approximation of development and hence of reporting biases both in the first stage of analysis in the main paper (which employs coarsened exact score matching) and in this robustness analysis. The nighttime light data used in all stages of analysis are from the Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS) Nighttime Lights Time Series dataset (Koren and Sarbahi, Forthcoming). This indicator, *nighttime light*, measured in five year intervals starting in 1992 and ending in 2007 to alleviate simultaneity concerns, measures the total number of pixels – or squares of approximately 1km x 1km around the equator – within a given cell that had any nighttime light. Everything else equal, I assume that a higher number of luminous pixels correspond to a higher degree of development within a given cell.

Table A7 reports two different variations of Model 4 estimated on a global sample of all grid cell years and again on a separate sample composed of only urban cells. The count stage of Model A-XV includes the same coefficients as Model 4, but only a small number of controls at the inflation stage, whereas Model A-XVI exactly replicated Model 4 from the main paper.In both models, the the variable *nighttime light* is additionally included in both equations. Models Model A-XVII and Model A-XVIII estimate the same specifications on a sample subset consisting solely of grid cells that had some urbanization according to Bontemps, Defourny and Van Bogaert (2009). Crucially, in all stages, the effect of *capital* is again positive and significant (to a $p < 0.01$ level), which provides strong confirmation that the main analysis' findings are not the result of reporting biases.

Table A7: Zero-Inflated Negative Binomial Estimates of Insurgent Atrocities in Areas with Nighttime Light, 1996-2009

	Global Sample		Urban Sample	
	Model A-XV	**Model A-XVI**	**Model A-XVII**	**Model A-XVIII**
Count Stage				
Capital	0.947***	0.768***	0.892***	0.626***
	(0.146)	(0.135)	(0.163)	(0.173)
Nighttime light[1]	−0.021	−0.089***	−0.020	0.071*
	(0.020)	(0.033)	(0.034)	(0.040)
Civil conflict	2.052***	0.617***	2.501***	2.443***
	(0.070)	(0.116)	(0.109)	(0.141)
Atrocities	2.003***	0.405***	1.363***	0.786***
	(0.092)	(0.039)	(0.108)	(0.088)
Urban	0.059***	−0.020*	0.075***	0.065***
	(0.011)	(0.012)	(0.012)	(0.016)
Oil[1]	0.006	0.019*	−0.029***	−0.054***
	(0.006)	(0.011)	(0.010)	(0.015)
GCP[1]	−0.295***	0.141	−0.298***	−0.168
	(0.056)	(0.093)	(0.073)	(0.112)
Border distance[1]	−0.252***	−0.063	−0.277***	−0.105*
	(0.024)	(0.052)	(0.037)	(0.060)
Polity2	0.009	0.002	−0.028***	0.136***
	(0.006)	(0.012)	(0.009)	(0.017)
Population[1]	0.659***	0.122**	0.638***	0.332***
	(0.031)	(0.056)	(0.050)	(0.089)
Travel time[1]	0.085	−0.027	0.165	0.105
	(0.060)	(0.089)	(0.104)	(0.128)
Cell area[1]	−0.045	0.127	−0.307**	0.094
	(0.083)	(0.124)	(0.146)	(0.162)
Constant	−12.910***	−4.418***	−10.017***	−9.248***
	(0.713)	(1.063)	(1.129)	(1.558)
Inflation Stage				
Nighttime light[1]	4.348***	−0.057	−395.212	0.060
	(1.229)	(0.037)		(0.097)
Civil conflict	–	−1.506***	–	0.281
		(0.131)		(0.272)
Atrocity	–	−4.289***	–	−2.609***
		(0.413)		(0.612)
Oil[1]	–	0.017	–	−0.044
		(0.012)		(0.029)
GCP[1]	–	0.559***	–	0.214
		(0.108)		(0.164)
Border distance[1]	–	0.213***	–	0.218**
		(0.055)		(0.098)
Polity2	–	−0.008	–	0.468***
		(0.012)		(0.058)
Population[1]	–	−0.608***	–	−0.835***
		(0.062)		(0.135)
Urban	−5,091.327	−0.122***	8.765	−0.018
	(48,806.470)	(0.017)	–	(0.022)
Travel time[1]	−0.797	−0.130	199.692	−0.337
	(0.815)	(0.106)	–	(0.228)
Cell area	1.138	0.244*	−1,348.758	0.445
	(1.135)	(0.146)	–	(0.302)
Constant	−38.514***	8.628***	9,039.268	6.202**
	(10.866)	(1.247)	–	(2.572)
N	786,449		164,259	
AIC	14,660.42	14,006.87	6,540.01	6,321.46

Note: *p<0.1; **p<0.05; ***p<0.01; year fixed effects included in each regression, though not reported here. All independent variables are lagged by one year.

[1] Natural log

Negative Binomial Models

In this section, estimates obtained from non-inflated Negative Binomial (NB) models as well as the results of Vuong's non-nested hypothesis tests are provided in Tables A8 and A10, and Tables A9 and A11, respectively, to additionally highlight the robustness of the hypothesized relationship between capital cities and the frequency of atrocities perpetrated by insurgents by relaxing the argument concerning the effect of excessive zero observations in my sample.

Table A8: Negative Binomial Estimates of Insurgent Atrocities, 1996-2009

	Model A-NB1	**Model A-NB2**	**Model A-NB3**
Capital	1.174***	0.916***	0.933***
	(0.131)	(0.139)	(0.144)
Civil conflict	2.007***	2.042***	2.042***
	(0.070)	(0.071)	(0.071)
Atrocities (lag)	2.047***	2.022***	2.023***
	(0.040)	(0.039)	(0.039)
Urban	–	0.072***	0.072***
		(0.010)	(0.010)
Oil[1]	–	–	0.003
			(0.006)
GCP[1]	−0.227***	−0.339***	−0.346***
	(0.049)	(0.053)	(0.055)
Border distance[1]	−0.251***	−0.253***	−0.256***
	(0.022)	(0.022)	(0.022)
Polity2	0.002	0.007	0.007
	(0.006)	(0.006)	(0.006)
Population	0.638***	0.649***	0.651***
	(0.030)	(0.030)	(0.030)
Travel time[1]	0.082	0.126**	0.126**
	(0.058)	(0.058)	(0.058)
Cell area[1]	−0.058	−0.074	−0.073
	(0.084)	(0.083)	(0.084)
Constant	−12.779***	−12.952***	−12.998***
	(0.721)	(0.724)	(0.733)
N		803,795	
AIC	14,731.990	14,687.370	14,689.130

Note: *p<0.1; **p<0.05; ***p<0.01; year fixed effects included in each regression, though not reported here.
[1] Natural log

Table A9: Vuong Non-Nested Hypothesis Tests Results, NB and ZINB Models Comparison

	Model 1	**Model 2**[1]	**Model 3**	**Model 4**
Raw	-3.670***	−9.913***	−10.140***	−10.145***
AIC-Corr.	-3.378***	−9.648***	−9.850***	−9.826***
BIC-Corr.	-1.526*	−8.111***	−8.168***	−7.964***

Note: *p<0.1; **p<0.05; ***p<0.01. Z values for each test are reported, with negative values denoting ZINB is preferred.

[1] ZINB model estimates compared to Model A-NB1.

Table A10: Negative Binomial Estimates of Insurgent Atrocities, Urban Sample 1996-2009

	Model A-NB4	**Model A-NB5**
Capital	1.048***	0.396*
	(0.156)	(0.213)
Large urban area	0.536***	0.045
	(0.153)	(0.293)
Civil conflict	2.483***	2.080***
	(0.108)	(0.151)
Atrocities (lag)	1.398***	0.910***
	(0.042)	(0.040)
Oil[1]	−0.027***	0.012
	(0.009)	(0.011)
GCP[1]	−0.242***	−0.127
	(0.069)	(0.121)
Border distance[1]	−0.279***	−0.089
	(0.035)	(0.061)
Polity2	−0.035***	0.032*
	(0.008)	(0.017)
Population[1]	0.636***	0.712***
	(0.049)	(0.081)
Travel time[1]	0.008	0.497***
	(0.102)	(0.142)
Cell area[1]	−0.022	−0.174
	(0.132)	(0.179)
Constant	−12.117***	−15.006***
	(1.085)	(1.853)
N	164,259	16,155
AIC	6,562.664	2,952.189

Note: *p<0.1; **p<0.05; ***p<0.01; year fixed effects included in each regression, though not reported here.
[1] Natural log

Table A11: Vuong Non-Nested Hypothesis Tests Results, NB and ZINB Models Comparison, Models 5A-C and 6A-C

	Model 5A	**Model 5B**[1]	**Model 5C**[1]	**Model 6A**	**Model 6B**[2]	**Model 6C**[2]
Raw	-5.716***	−5.716***	−5.802***	−5.850***	−5.988***	−5.842***
AIC-Corr.	-5.266***	−5.221***	−5.316***	−5.181***	−5.241***	−5.108***
BIC-Corr.	-3.013***	−2.742***	−2.882***	−-2.611***	−2.369***	−2.284**

Note: *p<0.1; **p<0.05; ***p<0.01. Z values for each test are reported, with negative values denoting ZINB is preferred.

[1] ZINB model estimates compared to Model A-NB4.

[2] ZINB model estimates compared to Model A-NB5.

Predictive Analysis

Given the importance of forecasting to the study of atrocities and mass killing (Koren, Forthcoming), I also evaluate the utility of *capital* as a *predictive* indicator of atrocities perpetrated by insurgents. To do so, I first calculate the predicted probability of a non-zero atrocity count for each cell-year in my sample based on the coefficient estimates obtained from Model 4 by subtracting an observation's full ZINB zero-atrocity predicted probability from one. I then repeat this exercise for a similar model that does not include the variable capital[6] – but otherwise remain unchanged – and evaluate each set of predictions against a binary indicator of whether or not a non-zero atrocity count occurred in each global cell-year (1996-2009). I then repeat this process using k-fold cross validation to estimate change in forecast error. Specifically, all of the country-year observations used in Model 4 were randomly divided into ten segments. Nine of these segments were combined to create a "training set,"' which was used to reestimate the model. The tenth segment, or "test set," was then utilized to assess the predictive power of the coefficients estimated using the training set (Ward, Greenhill and Bakke, 2010). Forecast error is defined as the ratio of times a given model failed to correctly predict the outcome in the tenth segment using a dataset composed of the other nine segments to all prediction attempts.[7]

The predictive power of Models 4 with and without *capital* was measured by calculating the area under the receiver operator characteristic (ROC) curve (Ward, Greenhill and Bakke,

[6]Coefficient estimates for this model, Model 4B, are reported below.

[7]Trimming proportion of 0.1 was used to obtain the trimmed mean of forecast error for each model.

2010) for my entire sample, and the estimates of each model are reported in Table A13 below. I then test whether the difference between a given pair of areas under curve (AUCs) is statistically significant using a nonparametric significance test developed by DeLong, DeLong and Clarke-Pearson (1988). The difference in the AUCs, model fit, and forecasting error (by way of cross validation) of Models 4 and 4B are provided in Table A12. Note that the addition of *capital* produces a significant (to a $p < 0.01$ level) improvement in the predictive power of the model and reduces its forecasting error. AIC scores also confirm these findings by favoring Model 4 over Model 4B. This suggests that the designation of a given cell as capital makes it an effective *predictive* indicator of atrocities, and this highlights the substantive contribution of the analyses provided in the main paper and in this appendix.

Figure A10: In-sample ROC curve *with* capital$_{t-1}$

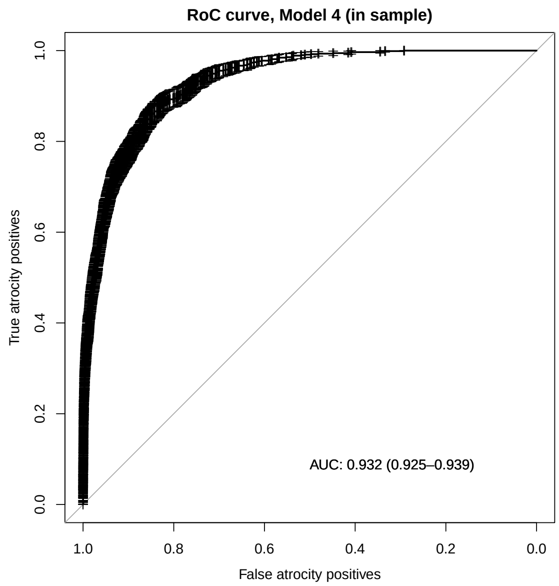


Figure A11: In-sample ROC curve *without* capital$_{t-1}$

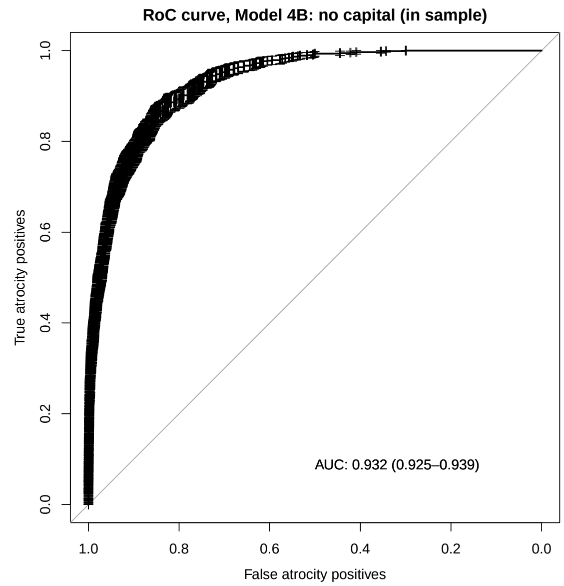


Figure A12: Comparison of In-Sample ROC Curves for Model 4 With and Without Capital$_{t-1}$, 1996-2009

Table A12: In Sample and Cross Validation Prediction Assessments of Insurgent Atrocities , Model 4 Estimates, 1996-2009

	AUC (in sample)		**AIC (in sample)**		**Forecast error (cross validation)**	
	(with capital$_{t-1}$)	(w/o capital$_{t-1}$)	(with capital$_{t-1}$)	(w/o capital$_{t-1}$)	(with capital$_{t-1}$)	(w/o capital$_{t-1}$)
Value	93.22%	93.16%	14,011.51	14,038.15	6.5098e-4	6.6326e-4
95% confidence interval	92.53% ⇔ 93.90%	92.48% ⇔ 93.85%	–	–	–	–
DeLong et al. test	$z = 2.802$ [$p < 0.01$]		–	–	–	
Favors:	AUC with capital$_{t-1}$		AIC with capital$_{t-1}$		Forecast error with capital$_{t-1}$	
N			803,795			

Note: Null hypothesis for Delong et al.'s Test for two correlated ROC curves: true difference in AUC's is equal to zero.

Formal Model

A Formal Illustration of Theoretical Implications

My argument can be illustrated formally using an extension of the Colonel Blotto game.[8] This relatively simple game model is helpful in illustrating how the calculations of insurgents are affected by the value the government attributes to the capital, even in primarily rural insurgencies.[9] Assume two actors, the insurgent group, I, and the government, G. Assume N locations for the game, $L \in \{1, 2, ..., N\}$, where $N > 2$. Building on the premise that the insurgents are more likely to operate in the periphery, assume a binary space where a location can be either rural or capital city, such that $L^{-C} \in \{1, 2, ..., N_{-C}\}$ locations are rural, and the remaining locations $L^C \in \{1, 2, ..., N_C\}$ are located within the confines of the capital. Formally, denote $L^{-C} = L^R$ and $N_{-C} = N_R$, and all locations as $L = L^C \cup L^R$, where $|L^C| = N_C$, $|L^R| = N_R$, and $N = N_C + N_R$. Last, assume that the model is static, i.e. that a given attack would not affect the likelihood of or the cost incurred by a later attack.

Both actors move simultaneously. Note that for the purpose of argument the insurgent group I wishes to harm the regime only by targeting civilians, and not, for example, by killing government troops or winning territory. The government must position its troops in anticipation of an attack in a given location, denoted L_G, while the insurgents must choose a given location in which to target civilians, denoted L_I. If government troops G defeat the enemy in a rural area, G receives a utility normalized to 1. Now, as I argued above that civilian targeting by insurgents produces a greater effect on the government if it occurs in the capital, assume that if I strikes in any location L^C, the government G always pays a given cost κ even if it thwarts the attack, the result of economic, political, or other types of repercussions. Therefore, its utility for thwarting an attack in the capital is $1 - \kappa$, where $1 > \kappa > 0$. Each actor's utility is now given by:

[8] Shubik and Weber, 1981.

[9] For all derivations, see below.

$$u_G(L_G, L_I) = \begin{cases} 1 \text{ if } L_G = L_I \text{ and } L_G \notin L^C \\ 1 - \kappa \text{ if } L_G = L_I \text{ and } L_G \in L^C \\ 0 \text{ otherwise.} \end{cases}$$

$$u_I(L_G, L_I) = \begin{cases} 0 \text{ if } L_G = L_I \\ 1 \text{ otherwise.} \end{cases}$$

where $1 > \kappa > 0$

Characterize a mixed strategy Nash equilibrium to this game. Firstly, note that for the insurgents I, and before the additional political value of the capital city – captured by the additional cost parameter κ – is taken into account, targeting civilians in all locations L is identical: each yields a utility of 1 if the government does not choose to defend it, and a utility of 0 if it does. Denote the probability with which the government stations troops at a particular location as p_L. The point of indifference for the insurgents I before capital city-specific costs are taken into account (i.e. $\kappa = 0$) is therefore:

$$p_L^* = \frac{1}{N} \forall L \in \{1, 2, ..., N\} \tag{1}$$

As Equation 1 suggests, for equilibrium conditions to hold, the insurgents I must target all capital city locations with equal probability, denoted q_C, and similarly all locations in rural areas with equal, but distinct, probability, denoted $q_{-C} = q_R$. Likewise, an equilibrium would require the government G to be indifferent to defending a particular location within the capital, and similarly it should be indifferent to defending a particular rural location. Hence, for the government G, the expected utility from defending any capital city location should be equal to the expected utility from defending any rural location:

$$E[u_G(L = L^C, q_C)] = E[u_G(L = L^R, q_R)] \; \forall \; L^C, L^R$$

$$q_C(1-\kappa) = q_R \tag{2}$$

Knowing these probabilities, the insurgents I can now take the additional political value of the capital, the cost parameter κ, into consideration. Hence, I must chose a location in which to target civilians, with the following probabilities:

$$q_C^* = \frac{1}{N_C + N_R(1-\kappa)} \tag{3}$$

$$q_R^* = \frac{1-\kappa}{N_C + N_R(1-\kappa)} \tag{4}$$

To see how the positioning of attacks and defenses changes as a function of the additional political value of atrocities in the capital, the derivative of each equilibrium probability in respect to κ is taken. Note that p_L^* is invariant in κ, as mentioned above, so the government's G positioning of troops does not change as a function of it:

$$\frac{\partial q_C^*}{\partial \kappa} = \frac{N_R}{[N_C + N_R(1-\kappa)]^2} > 0 \tag{5}$$

$$\frac{\partial q_R^*}{\partial \kappa} = -\frac{N_C}{[N_C + N_R(1-\kappa)]^2} < 0 \tag{6}$$

As equations 5 and 6 illustrate, as the additional political value of using violence against civilians (i.e. the cost incurred by G) κ increases, the insurgent group I is more likely to try and target locations within the capital, even if I is originally indifferent between targeting any of these locations. Building on the argument developed above and the results of my formal model, I expect capital cities to experience insurgent atrocities significantly more frequently than other localities.

Formal Model Proofs

Proof of Equation 3: We know based on the laws of probability that the sum of all probabilities and utilities should equal 1, which can be written as:

$$N_C \times q_C + N_R \times q_R = 1$$

If q_C^* denotes equilibrium probability, we need to find the equation according to which $q_C = q_C^*$. Similarly, we need to find $q_R = q_R^*$. Therefore, we only need to isolate q_C and q_R. This is done as follows, beginning with Equation 2:

$$q_R = q_C(1-\kappa)$$

And similarly we can isolate q_C:

$$q_C = \frac{q_R}{(1-\kappa)}$$

Now we can assign values for q_C and q_R to solve each equation for one unknown. To find $q_C = q_C^*$:

$$N_C q_C + N_R q_C(1-\kappa) = 1$$
$$q_C[N_C + N_R(1-\kappa)] = 1$$

Divide by $[N_C + N_R(1-\kappa)]$ to get:

$$q_C = q_C^* = \frac{1}{N_C+N_R(1-\kappa)}$$

Proof of Equation 4: As mentioned above, to find $q_R = q_R^*$ we begin with the equation:

$$N_C \times q_C + N_R \times q_R = 1$$

Assigning $q_C = \frac{q_R}{(1-\kappa)}$, we can solve for $q_R = q_R^*$:

$$N_C \frac{q_R}{(1-\kappa)} + N_R q_R = 1$$
$$q_R[N_C \frac{1}{(1-\kappa)} + N_R] = 1$$

Divide by $[N_C \frac{1}{(1-\kappa)} + N_R]$ to get:

$$q_R = \frac{1}{[N_C \frac{1}{(1-\kappa)} + N_R]}$$

Multiply N_R by $\frac{1-\kappa}{1-\kappa} = 1$ to rewrite this equation as:

$$q_R = \frac{1}{[\frac{N_C+N_R(1-\kappa)}{(1-\kappa)}]}$$

Which simplifies to:

$$q_R = q_R^* = \frac{(1-\kappa)}{N_C+N_R(1-\kappa)}$$

Proof of Equation 5: To show the comparative statics differentiation of capital cities, i.e. the change in the likelihood of attack in capital cities as the cost κ increases, the derivative of Equation 3 in respect to κ is taken. Beginning with Equation 3, and proceeding using the quotient rule:

$$q_C = q_C^* = \frac{1}{N_C+N_R(1-\kappa)}$$
$$q_C = q_C^* = \frac{1}{N_C+N_R-\kappa N_R}$$
$$\frac{\partial q_C^*}{\partial \kappa} = \frac{N_C+N_R(1-\kappa)\times 0-1\times(-N_R)}{[N_C+N_R(1-\kappa)]^2}$$

We know that the denominator is always positive because it is raised to an even power. We also know that there must be at least one location that is rural, especially if insurgencies are primarily a rural phenomenon, meaning that N_R is also always positive. Therefore, it follows that:

$$\frac{\partial q_C^*}{\partial \kappa} = \frac{N_R}{[N_C+N_R(1-\kappa)]^2} > 0$$

And the probability of the insurgents I targeting civilians in the capital increases as the cost parameter κ increases.

Proof of Equation 6: To show the comparative statics differentiation of rural locations, i.e. the change in likelihood of attack in rural areas as the cost κ increases, the derivative of Equation 4 in respect to κ is taken. Beginning with Equation 4, and proceeding using the quotient rule:

$$q_R = q_R^* = \frac{(1-\kappa)}{N_C+N_R(1-\kappa)}$$
$$q_C = q_C^* = \frac{1-\kappa}{N_C+N_R-\kappa N_R}$$
$$\frac{\partial q_R^*}{\partial \kappa} = \frac{[-1\times(N_C+N_R-\kappa N_R)]-[(1-\kappa)\times(-N_R)]}{[N_C+N_R(1-\kappa)]^2}$$
$$\frac{\partial q_R^*}{\partial \kappa} = \frac{-N_C-N_R+\kappa N_R+N_R-\kappa N_R}{[N_C+N_R(1-\kappa)]^2}$$
$$\frac{\partial q_R^*}{\partial \kappa} = \frac{-N_C}{[N_C+N_R(1-\kappa)]^2}$$

Note that that the denominator here is, again, raised to an even power and hence always positive. However, the nominator in this case is always negative, as there is always at least one location that is in capital city. Therefore, it follows that:

$$\frac{\partial q_R^*}{\partial \kappa} = -\frac{N_C}{[N_C+N_R(1-\kappa)]^2} < 0$$

And the probability of the insurgents I targeting civilians in rural areas decreases as the cost parameter κ increases.

Insurgent Atrocities in the Capital and Regime Failure

In this section I offer preliminary evidence connecting the effect of atrocities perpetrated by insurgents in capital cities to weakening the targeted regime. This analysis is done in two stages. First, the affect of atrocities in the capital on the hazard of regime transition is estimated using different Cox regression models and examined alongside a variety of controls to account for alternative explanations. Second, the effect of the same indicators on regime durability (i.e. how many years a given regime remains in power) is estimated using negative binomial models. Atrocities perpetrated by insurgents in the capital have a highly significant (to a $p < 0.01$ level) association with regime transition (negative effect) and durability (positive effect), suggesting that – *prima facie* – perpetrating atrocities in the capital is a more effective strategy of hurting the regime in power. Simply put, atrocities perpetrated by insurgents in the capital significantly contribute to regime failure in the 1996-2009 sample. Crucially, the effect of atrocities in the capital is estimated alongside the effect of the total number of atrocities perpetrated by insurgents in the country (i.e. in ll urban, rural, and capital locations), and while the effect of atrocities perpetrated in the capital is highly significant, the effect of atrocities in general is not. This, again, suggests that insurgents would prefer to target the capital because in doing so they get more "bang for their buck," which increases their capabilities to asymmetrically impose higher costs on the (typically stronger) government.

Data and Methods

To estimate the effect of atrocities perpetrated by insurgents in the capital, I rely on additional datasets and measures aggregated at the country level. The dependent variable for the duration analyses was regime failure, *failure* (Geddes, Wright and Frantz, 2014), with the effect of different covariates estimated on the time until transition. The dependent variable for my count models was the regime durability indicator, *durability*, obtained from the Polity IV dataset and defined as "[t]he number of years since the most recent regime change...or the end of transition period defined by the lack of stable political institutions" (Marshall, Jaggers and Gurr, 2013, 16). For both the duration and count models, the main explanatory variable was the (natural log of the) number of atrocities in the capital perpetrated by insurgents in a given country during a given year, *capital atrocities*, building on the same data used in the main paper.

Additional controls were included to account for alternative explanations. To account for the level of democratization, the Polity2 variable, *Polity2* was included in analysis (Marshall, Jaggers and Gurr, 2013). I also included controls for the number of conflict events between or among government and rebel forces, *conflict* (Tollefsen et al., 2012), as well as a measure of all atrocities perpetrated by insurgents within the *entire country, including the capital*, *atrocities*. The use of these variables allows me to account for the possibility that regime transitions and lower regime durability result from a higher incidence of conflict or atrocities perpetrated by insurgents in general, rather than by specifically targeting the capital. The autocorrelation between *capital atrocities* and *atrocities* also increases the probability of a type II error – i.e., failure to reject the null hypothesis in cases where it is wrong – thus serving as an additional robustness measure. To account for the effect of development and population densities, I included controls for gross domestic product by country, *GDP* (WorldBank, 2013), and population, *population* (UNSD, 2013). To account for the importance of natural resources I included an indicator measuring yearly oil production by country,*oil*. To account for the effect of military expenditure in some models, I included a variable coding the nation's military expenditure during a given year (in USD), *mil. exp.*. To account for the effect of rough terrain, I included an indicator denoting the average coverage of a given grid cell that is mountainous for the entire country, *mountain* (Bontemps, Defourny and Van Bogaert, 2009). Last, to account for the potential effect of other type of human rights violations by the government on the propensity of violence in the capital, I included a lagged measure of the CIRI index, *CIRI* (Cingranelli and Richards, 2010).

The effect of these covariates on regime transition was estimated using two types of regression. First,the effect of these indicators on the hazard of transition was estimated using a series of Cox regression models. In these models, a positive coefficient suggests that a given covariates increases the hazard of transition, i.e. reduces the time until transition. Second, because the variable *durability* can only take on non-negative integer values, is bounded at zero, and unbounded above, a set of negative binomial regressions (as discussed in the main paper) was used to estimate the effect of the aforementioned covariates on the number of years during which a given regime is in power.

Results

Table A14 shows the estimates obtained from three Cox survival regressions with different specifications, ranging from a baseline model that only includes the explanatory variable *capital atrocities* and several controls, to a fully specified model. In all regressions, *capital atrocities* is positive and significant (to at least $p < 0.05$ level), suggesting that a higher number of atrocities increases the hazard of regime failure. These results are consistent as one moves from the baseline model to a fully specified model, and taking into account alternative similar explanations such as the effect of atrocities perpetrated by insurgents anywhere in the country. This provide some evidence in support of the argument that targeting the capitals does have an effect on delegitimizing the regime and hastening its untimely removal.

Table A15 similarly shows the estimates obtained from three separate negative binomial regressions with similar specifications to those used in the Cox models, ranging from the baseline to full specification. Here, *capital atrocities* is negative and significant (to a $p < 0.01$ level), suggesting that more atrocities perpetrated by insurgents in the capital translate into lower levels of regime durability. These results are consistent across all specifications, as one move's from the baseline to the full model specification, regardless of the addition of different relevant controls. These findings, again, provide evidence to suggest that killing civilians in the capital, specifically, reduces regime durability, which suggests that insurgents could inflict higher costs on the regime by targeting the capital compared with other locations. In substantive terms, a change in *capital atrocities* from the 25th to the 75th percentile in the

Full model presented in Table A15 translates to an average change of approximately -11% in regime durability, with a 95% confidence interval of -5.868% ⇔ -15.732%.[10] Similarly, a change in *capital atrocities* from the 5th to 95th percentile in the same model translates to an average change of approximately -14% in regime durability, with a 95% confidence interval of -7.080% ⇔ -23.460%.[11]

[10]Or -1.5 years with a 95% confidence interval of -0.905 years ⇔ -2.104 years.

[11]Or -2.298 years with a 95% confidence interval of -1.055 years ⇔ -4.169 years.

References

Bagozzi, Benjamin E. 2015. "Forecasting Civil Conflict with Zero-Inflated Count Models." *Civil Wars* 17(1):1–24.

Bontemps, Sophie, Pierre Defourny and Eric Van Bogaert. 2009. "Globcover 2009. Products Description and Validation Report." European Space Agency.

Cingranelli and Richards. 2010. "Human Rights Dataset." `http://www.humanrightsdata.org/`.

DeLong, Elisabeth R., David M. DeLong and Daniel L. Clarke-Pearson. 1988. "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach." *Biometrics* 44:837–845.

Fjelde, Hanne and Lisa Hultman. 2014. "Weakening the Enemy: A Disaggregated Study of Violence against Civilians in Africa." *Journal of Conflict Resolution* 58(7):1230–1257.

Freedom House. 2017. "Freedom of the Press Report." `https://freedomhouse.org/report-types/freedom-press`.

Geddes, Barbara, Joseph Wright and Erica Frantz. 2014. "Autocratic Breakdown and Regime Transitions." *Perspectives on Politics* 12(2): forthcoming.

King, Gary and Margaret E Roberts. 2014. "How Robust Standard Errors Expose Methodological Problems They Do Not Fix, and What to Do About It." *Political Analysis* .

Koren, Ore. Forthcoming. "Means to an End: Pro-Government Militias as a Predictive Indicator of Strategic Mass Killing." *Conflict Management and Peace Science* .

Koren, Ore and Anoop Sarbahi. Forthcoming. "State Capacity, Insurgency and Civil War: A Disaggregated Analysis." *International Studies Quarterly* .

Marshall, M. G., T. R. Jaggers and K. Gurr. 2013. "Polity iv project: Political regime characteristics and transitions, 1800-2012." Technical Report.

Martinez-Cesena, Eduardo Alejandro, Pierluigi Mancarella, Mamadou Ndiaye and Markus Schläpfer. 2015. "Using Mobile Phone Data for Electricity Infrastructure Planning." `http://arxiv.org/ftp/arxiv/papers/1504/1504.03899.pdf`.

PITF. 2009. "Political Instability Task Force Worldwide Atrocities Event Data Collection Codebook Version 1.0B2." `http://eventdata.parusanalytics.com/data.dir/atrocities.html`.

Raleigh, Clionadh and Haavard Hegre. 2009. "Population size, concentration, and civil war. A geographically disaggregated analysis." *Political Geography* 28:224–238.

Ross, Michael L. 2011. "Oil and Gas Data, 1932-2011." `http://hdl.handle.net/1902.1/20369UNF:5:dc22RlDasveOTAJvwIjBTA==V2`.

Shubik, Martin and Robert James Weber. 1981. "System Defense Games: Colonel Blotto, Command and Control." *Naval Research Logistics Quarterly* 28(2):281–287.

Sundberg, Ralph and Erik Melander. 2013. "Introducing the UCDP Georeferenced Event Dataset." *Journal of Peace Research* 50(4):523–532.

Tollefsen, Andreas Forø, Håvard Strand and Halvard Buhaug. 2012. "PRIO-GRID: A Unified Spatial Data Structure." *Journal of Peace Research* 49(2):363–374.

UNSD. 2013. "National Accounts." `http://unstats.un.org/unsd/snaama/dnlList.asp`.

Ward, Michael D., Brian D. Greenhill and Kristin M. Bakke. 2010. "The Perils of Policy by P-Value: Predicting Civil Conflicts." *Journal of Peace Research* 47(4):363–375.

Weidmann, Nils B. 2016. "A Closer Look at Reporting Bias in Conflict Event Data." *American Journal of Political Science* 60(1):206–218.

World-Bank. 2013. "World Development Indicators." `http://data.worldbank.org/data-catalog/world-development-indicators`.

Table A13: A Comparison of Models 4 and 4B

	Model 4	**Model 4B**
Count Stage		
Capital	0.721***	–
	(0.135)	
Civil conflict	0.626***	0.634***
	(0.117)	(0.117)
Atrocities (ln)	0.417***	0.415***
	(0.040)	(0.039)
Urban	−0.008	0.007
	(0.011)	(0.011)
Oil[1]	0.006	−0.003
	(0.010)	(0.010)
GCP[1]	0.049	0.040
	(0.087)	(0.087)
Border distance[1]	−0.074	−0.065
	(0.053)	(0.053)
Polity2	−0.001	−0.003
	(0.012)	(0.012)
Population[1]	0.081	0.120**
	(0.053)	(0.053)
Travel time[1]	0.052	0.042
	(0.085)	(0.085)
Cell area[1]	0.097	0.074
	(0.124)	(0.124)
Constant	−4.362***	−4.435***
	(1.065)	(1.066)
Inflation Stage		
Civil conflict	−1.517***	−1.495***
	(0.131)	(0.130)
Atrocities	−4.382***	−4.342***
	(0.455)	(0.442)
Urban	−0.115***	−0.107***
	(0.017)	(0.017)
Oil[1]	0.008	0.004
	(0.011)	(0.011)
GCP[1]	0.494***	0.460***
	(0.103)	(0.103)
Border distance[1]	0.202***	0.212***
	(0.055)	(0.055)
Polity2	−0.009	−0.009
	(0.012)	(0.012)
Population[1]	−0.637***	−0.604***
	(0.060)	(0.059)
Travel time[1]	−0.085	−0.092
	(0.100)	(0.099)
Cell area[1]	0.211	0.203
	(0.145)	(0.144)
Constant	8.954***	8.724***
	(1.237)	(1.231)
N	803,795	
AIC	14,011.51	14,038.15

Note: *p<0.1; **p<0.05; ***p<0.01; year fixed effects included in each regression, though not reported here. All independent variables are lagged by one year.

[1] Natural log

Table A14: Cox Proportional Hazard Model Estimates of Regime Failure, 1996-2009

	(Baseline)	(Medium)	(Full)
Capital atrocities	1.453***	1.394***	1.440**
	(0.470)	(0.504)	(0.585)
Atrocities	−0.175	−0.144	−0.158
	(0.114)	(0.111)	(0.128)
Conflict	0.002	0.001	0.00002
	(0.001)	(0.002)	(0.003)
Polity2	0.044	0.069	−0.006
	(0.044)	(0.048)	(0.064)
GDP[1]	−0.956***	−0.796**	−0.760**
	(0.224)	(0.338)	(0.382)
Population[1]	−0.051	0.162	0.353
	(0.151)	(0.226)	(0.241)
Oil[1]	–	0.062	0.068
		(0.041)	(0.046)
Mil. exp.[1]	–	−0.448***	−0.434***
		(0.154)	(0.162)
Mountains	–	–	0.305
			(0.834)
CIRI	–	–	0.154*
			(0.092)
N	775	714	700
AIC	327.75	309.65	283.07
Log Likelihood	−157.872	−146.826	−131.532
Wald Test	35.460*** (df = 6)	37.290*** (df = 8)	29.340*** (df = 10)
LR Test	36.046*** (df = 6)	41.296*** (df = 8)	36.021*** (df = 10)
Score (Logrank) Test	35.641*** (df = 6)	35.409*** (df = 8)	31.749*** (df = 10)

Note: *p<0.1; **p<0.05; ***p<0.01.
[1] Natural log

Table A15: Negative Binomial Model Estimates of Regime Durability, 1996-2009

	(Baseline)	(Medium)	(Full)
Capital atrocities	-0.318^{***}	-0.332^{***}	-0.282^{***}
	(0.082)	(0.082)	(0.082)
Atrocities	0.010^{*}	0.008	0.009
	(0.006)	(0.006)	(0.006)
Conflict	0.0005^{*}	0.001^{**}	0.001^{***}
	(0.0002)	(0.0002)	(0.0003)
Polity2	-0.020^{***}	-0.028^{***}	-0.053^{***}
	(0.004)	(0.004)	(0.007)
GDP[1]	0.632^{***}	0.772^{***}	0.721^{***}
	(0.017)	(0.035)	(0.040)
Population[1]	0.086^{***}	0.201^{***}	0.186^{***}
	(0.014)	(0.030)	(0.032)
Oil[1]	–	-0.027^{***}	-0.025^{***}
		(0.004)	(0.004)
Mil. exp[1]	–	-0.040^{*}	-0.025
		(0.021)	(0.024)
Mountains_t	–	–	0.091
			(0.091)
CIRI_t	–	–	0.056^{***}
			(0.011)
Constant	-3.925^{***}	-6.247^{***}	-6.225^{***}
	(0.284)	(0.509)	(0.555)
N	2,036	1,837	1,765
AIC	15,891.92	14,300.35	13,721.06

Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors clustered by country, and year fixed effects included in each regression, though not reported here. All independent variables are lagged by one year.

[1] Natural log