Supplementary Information For

State Capacity, Insurgency and Civil War: A Disaggregated Analysis

Contents

This supplementary file is set as follows. We begin by providing a discussion of the data, methods, and summary statistics of the variables used in the paper, as well as the percentage of area affected by war onset or ongoing war for each country that experienced at least one respective event. We then conduct a series of robustness tests corresponding to Table 2 in the main paper. We conclude by providing a detailed illustration of the effect of state capacity as a *predictive* indicator of civil war onset by comparing different forecasting models with and without nighttime light.

Materials and Methods

The geo-located data used for this analysis were obtained from the PRIO-Grid dataset (Tollefsen et al., 2012). PRIO-Grid measures a variety of spatial data at the 0.5x0.5 decimal degree resolution at the equator, which decreases with higher latitudes to account for the Mercator Projection. This dataset thereby allows one to capture the variation of specific geographic and economic phenomena globally (excluding oceans, Antarctica, the Arctic) at the very local level. Crucially, the main variables of interest for the intrastate analyses are measured at the *cell*-, and not country-, level.

The Nighttime Light Indicator

Night light data for the entire terrestrial globe were obtained from NASA and measured at five year intervals between 1992 and 2007. The unit of measurements is the pixel, or a square of 0.008 x 0.008 decimal degrees (approximately 1 km x 1 km around the equator that increases in size as one moves up toward the poles), which varies between cells based on cell area (an additional motivation to include cell area in all our cell-level analyses). We coded our nighttime light indicator by counting all the pixels that had some degree of luminosity within a given cell for the years in question, under the assumption that a higher number of lighten areas corresponds to more areas where the regime *can* and *wants to* be (more) present, all else equal. The number of luminous pixels was then aggregated to the cell, district, or country level (depending on the unit of interest in each particular analysis).

Note that using nighttime light indicator to approximate different aspects of development is not without problems (Huang et al., 2014). The effect of nighttime light might vary between different datasets, and the manner in which nighttime light data are collected is also relevant. We discuss this issue theoretically and illustrate the validity of our specific indicator as an approximation of different subnational measures in the main paper. However, we add additional validation exercises in this Supporting Information file as well. First, to illustrate the validity of this measure, Figure A1 compares different operationalizations of nighttime light in our dataset to the actual distribution of nighttime light in India. As this

figure illustrates, apart from being theoretically defensible, our operationalization closely resembles the actual distribution of nighttime light within different countries and regions, even compared with alternative operationalizations. Importantly, our measure is also *time varying*, which allows us to capture the effects of state capacity over time.

Second, Figure A2 reports correlations between nighttime light and tax revenue by district in Ecuador, as well as how well nighttime light predicts tax revenue in these districts. As also shown in the cases of Brazil, Ghana, and India presented in the main paper, nighttime light provide a reasonable approximation of tax revenue (an established measure of state capacity) data measured at the highly disaggregated level, both when the real and predictive values are concerned. Finally, we also replicate our subnational logit models using an alternative indicator of nighttime light emissions obtained from the PRIO-Grid dataset (Tollefsen et al., 2012) in Table A9 below. In every case, the results suggest that our nighttime light measure is a reasonable approximation of state capacity and presence.

Dependent and Independent Variables

Civil war measures, as well as % mountainous area, urbanization, travel time, distance to nearest border, cell area, population, and gross cell product were obtained from the PRIO-Grid project (Tollefsen et al., 2012), with the latter three variables logged and lagged prior to being included in the models.[1] Civil war was defined inclusively as any cell that experienced at least 25 annual combatant battle deaths resulting from intrastate conflict during a given year, and coded zero otherwise. This definition means that no deaths resulting from collateral damages or civilian casualties are counted under this definition. Our spatial lag variable measures whether civil war onset occurred in any adjacent cell during the same year.

The cell-level variables for population and gross cell product were originally measured by Nordhaus (2006) for the years 1995 and 2000 and then interpolated to the yearly level using a last value carried forward approach. Mountainous area and urbanization measures included the percent of a given cell's mountainous or urban coverage, respectively, as coded
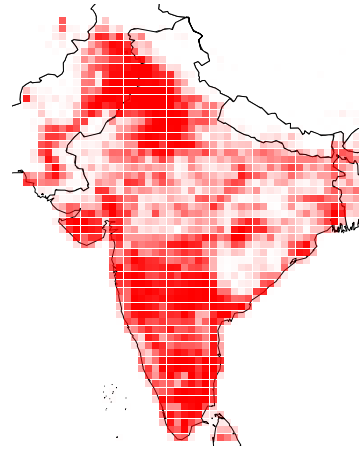
---

[1]Conflict measures were originally obtained from Gleditsch et al. (2002).

4

Figure A1: Night Light Measures for India At the PRIO-Grid Cell Level (2013)
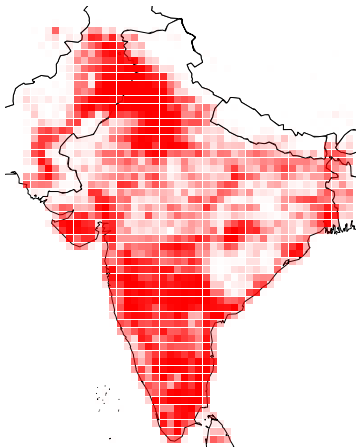
(a) Night light map of India (negative)
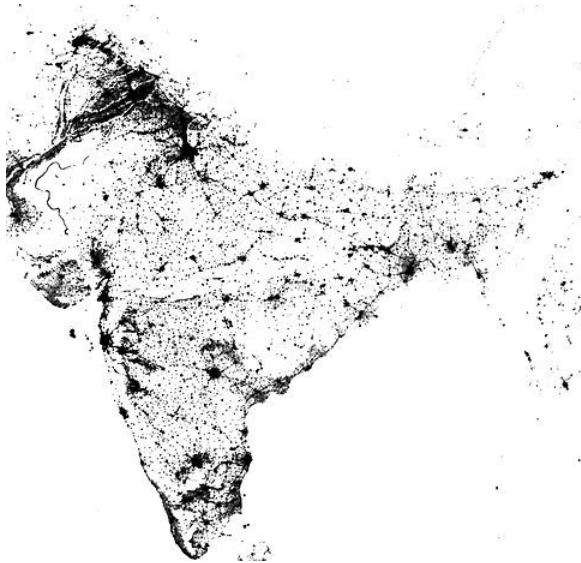
(c) Night light by grid cell (mean)

(b) Night light by grid cell (sum)

(d) Night light by grid cell (mean luminosity)
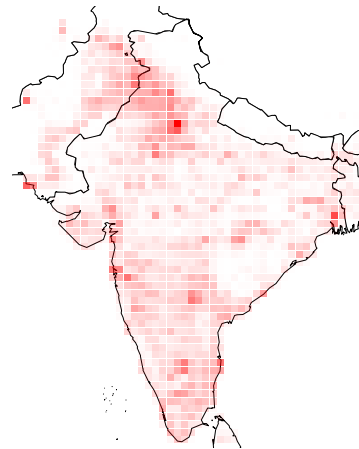
Figure A2: Correlations between Nighttime Light and Tax Revenue by District in Ecuador, 2012

Real nighttime light values   Predicted values based on nighttime light
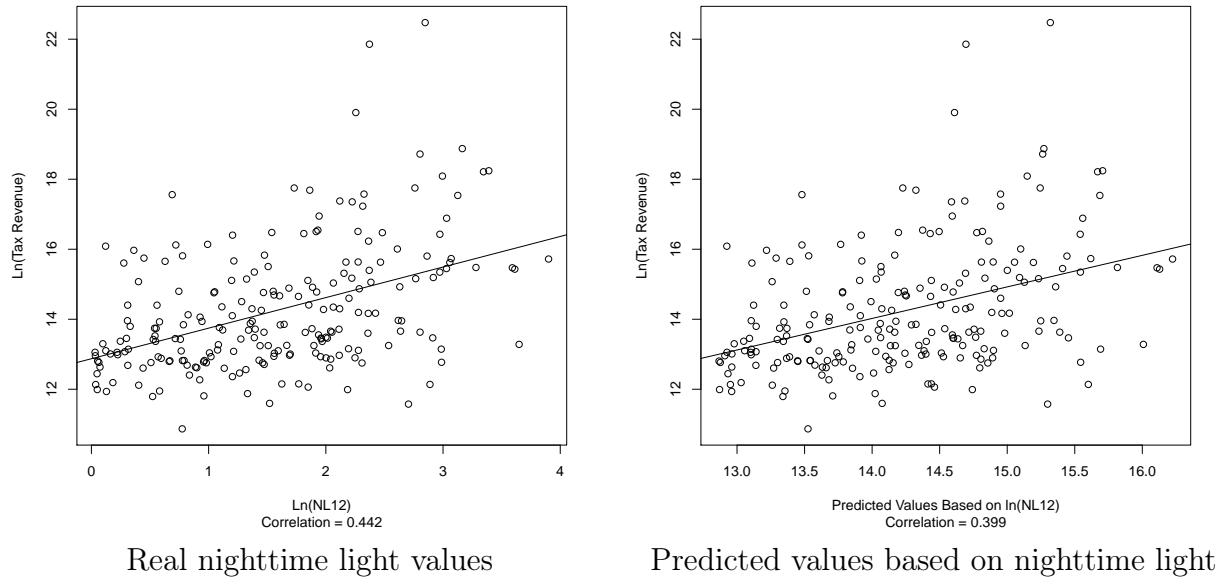
by the Globcover 2009 project (Bontemps, Defourny and Van Bogaert, 2009). Because PRIO-Grid data on mountainous area was available only for the years 2000-2008, we re-merged this covariate into the dataset for the entire period analyzed here (1992-2008) under the assumption that little variation in this variable occurred during this time.

The political regime measure Poilty2 was obtained from the Polity IV dataset (Marshall, Jaggers and Gurr, 2013). Military expenditure data were obtained from the Correlates of War (COW) project (Singer, Bremer and Stucky, 1972), and logged. Oil as percent of GDP was obtained from the Oil and Gas dataset complied by Ross (2011) and was lagged one year to handle potential endogeneity issues. We used Fearon and Laitin's (2003) standards to code the variables Polity (binary), anocracy, and new state[2] for the period analyzed here. We also incorporated their constant measures of ethnic and religious fractionalization without any additional changes, whenever these indicators were used (see Fearon and Laitin, 2003, for details on each variable's measurement).

---

[2]New state indicators for three countries that were formed between 1999 and 2008 were added: East Timor, Montanegro, and Kosovo.

The Split Population Weibull Model

In addition to logit models used in the first and second stage of the main paper (as well as the robustness and prediction analyses reported below), we relied on the Weibull model to account for duration related effects. Specifically, we chose to rely on the Weibull model because it is parametric, meaning that a split population framework can be applied to the data. The split-population Weibull (SPW) allows us to adjust the survival model coefficient estimates to account for the excess of zero values in our sample – recall that we have only 113 civil war onset events in the entire sample – by utilizing two stages of analysis. In the first, or risk, stage, 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 (Box-Steffensmeier and Jones, 2004; Svolik, 2008). In the second, or duration, stage, a Weibull model is used to test the effect of each covariate on the expected frequency of each dependent variable, conditional on a case being "non-immune" based on estimates obtained in the splitting stage.

Table A1: Summary Statistics of All Variables Used in Analysis

	Minimum	Median	Mean	Max	SD
Country Level Analyses					
(Country Covars. Only)					
Nighttime light (country)[2]	8.041	12.810	12.900	18.180	2.0119
Civil war onset	0.000	0.000	0.019	1.000	0.138
Prior war	0.000	0.000	0.208	1.000	0.406
Per capita income[1,2,3]	5.279	7.863	7.896	9.934	1.091
Population[1,2,3]	6.109	9.111	9.206	14.030	1.427
Mountainous (%)[2]	0.000	2.332	2.105	4.557	1.436
Noncontiguous	0.000	0.000	0.154	1.000	0.361
Oil exporter	0.000	0.000	0.147	1.000	0.354
New state	0.000	0.000	0.023	1.000	0.148
Instability[1]	0.000	0.000	0.205	1.000	0.404
Polity 2[1]	-10.000	5.000	2.310	10.000	6.882
Ethnic fractionalization	0.001	0.403	0.412	0.925	0.276
Religious fractionalization	0.000	0.375	0.383	0.783	0.216
Anocracy[1]	0.000	0.000	0.275	1.000	0.447
Polity 2 (binary)[1]	0.000	0.000	0.488	1.000	0.500
Cell Level Analyses					
(Cell and Country Covars.)					
Nighttime light[2]	0.000	0.000	0.256	8.189	2.992
Civil war onset	0.000	0.000	0.000	1.000	0.010
Civil war[1]	0.000	0.000	0.060	1.00	0.246
Gross cell product[1,2]	0.00	0.09	0.51	6.97	0.858
Population[1,2]	0.000	8.301	7.804	16.690	3.696
Mountainous area (%)[†]	0.000	0.000	0.224	1.000	0.352
Distance to border[2]	0.000	5.740	5.600	9.300	1.461
Oil production[1, 2]	0.000	18.440	15.880	20.000	6.003
New state[‡]	0.000	0.000	0.000	1.000	0.018
Political instability	0.000	0.000	0.026	1.000	0.158
Polity 2[1]	-10.000	6.000	4.367	10.00	6.194
Ethnic fractionalization	0.000	0.330	0.435	0.925	0.258
Religious fractionalization	0.000	0.446	0.442	0.783	0.202
Cell area[2]	0.000	7.697	7.399	8.039	0.969
Travel time[2]	0.000	6.230	6.340	10.31	1.250
Anocracy[1]	0.000	1.000	0.782	1.000	0.413
Polity 2 (binary)[1]	0.000	0.000	0.485	1.000	0.500
Nighttime light [2, 3]	0	5.320	4.725	8.189	2.565
Nighttime light (total)[2, 3]	0	7.933	7.364	12.740	2.751
Urban area (%)[1]	0.000	0.000	0.207	51.550	1.218
Nighttime light squared[2]	0.000	0.000	15.52	67.06	2.070
Nighttime light squared[2, 3]	0	28.310	28.910	67.060	21.442
Spatial lag onset	0.000	0.000	0.000	0.5	0.005
Military expenditure[1,2]	0.000	16.030	15.660	20.130	2.618
Calibrated light mean (PRIO)	0.014	0.034	0.054	1.000	0.067
Nighttime light (per capita)	0	0	0.272	9.044	0.311
Country size[2]	3.045	15.960	15.210	17.000	1.557

[1] lagged; [2] natural log; [3] by district

[†]: original data on mountains was available only for the years 2000-2008, but re-merged into the dataset for the entire period analyzed here under the assumption that little variation in this covariate occurred.

[‡]: Indicators for East Timor, Montanegro, and Kosovo (formed between 2003 and 2008) were added to this variable.

Table A2: The Geographical Spread of Conflict Onset in Countries That Experienced At Least One Conflict Onset Event

Country	Total No. Cells	Cells with Con. Onset	Affected Region (%)	Country	Total No. Cells	Cells with Con. Onset	Affected Region (%)
Comoros	5.00	1.00	20.00	Thailand	196.00	2.00	1.02
Lebanon	5.00	1.00	20.00	Niger	402.00	4.00	1.00
Burundi	11.00	2.00	18.18	Congo	112.00	1.00	0.89
Rwanda	9.00	1.00	11.11	Cote d'Ivoire	113.00	1.00	0.88
Djibouti	9.00	1.00	11.11	Somalia	241.00	2.00	0.83
Lesotho	12.00	1.00	8.33	Nigeria	312.00	2.00	0.64
Guinea-Bissau	13.00	1.00	7.69	India	1197.00	7.00	0.58
Azerbaijan	44.00	3.00	6.82	Yemen	178.00	1.00	0.56
Israel	15.00	1.00	6.67	CAR	202.00	1.00	0.50
Haiti	17.00	1.00	5.88	Mali	426.00	2.00	0.47
Sri Lanka	37.00	2.00	5.41	Iran	644.00	3.00	0.47
Croatia	41.00	2.00	4.88	UK	217.00	1.00	0.46
Yugoslavia	44.00	2.00	4.55	Angola	435.00	2.00	0.46
Bosnia Herzegovina	22.00	1.00	4.55	Afghanistan	255.00	1.00	0.39
Senegal	75.00	3.00	4.00	Pakistan	336.00	1.00	0.30
Georgia	33.00	1.00	3.03	Turkey	368.00	1.00	0.27
Myanmar	265.00	8.00	3.02	Ethiopia	372.00	1.00	0.27
Liberia	37.00	1.00	2.70	The DRC	763.00	2.00	0.26
Uganda	80.00	2.00	2.50	Chad	391.00	1.00	0.26
Eritrea	52.00	1.00	1.92	Mexico	825.00	2.00	0.24
Tajikistan	60.00	1.00	1.67	Peru	452.00	1.00	0.22
Philippines	236.00	3.00	1.27	Indonesia	1047.00	2.00	0.19
Iraq	169.00	2.00	1.18	Sudan	840.00	1.00	0.12
Guinea	85.00	1.00	1.18	Russia	12465.00	4.00	0.03
Uzbekistan	191.00	2.00	1.05	USA	4863.00	1.00	0.02

Predicted Probability of Conflict for a Change in Nighttime Light (Country)

Finally, we also illustrate the substantive effect of nighttime light as an approximation of state capacity by showing the first change in the probability of civil war onset at the *country* level. These figures are similar to the plots provided in Fearon and Laitin's (2003) study, which similarly show the effect different variables on the probability of civil war onset. To this extent, Figure A3 shows the predicted probability of civil war onset for a given country during a given year based on the estimates provided in Table 1 of the main paper. The x axis plots total nighttime light emissions (in natural log), and the y shows the predicted probability of civil war onset. Clearly, countries with lower levels of nighttime light are also at a substantively higher risk of experiencing civil war during a given year t. Moreover, the lowest levels of nighttime light emissions in these figures correspond to an average risk of approximately 20-40% of experiencing civil war, which is relatively similar to Fearon and Laitin's (2003) results in respect to GDP per capita (30%). This lends support to our reliance on nighttime light as at least as effective indicator of state capacity as GDP per capita.

Figure A3: Predicted Probability of Civil War Onset - State Level Analysis
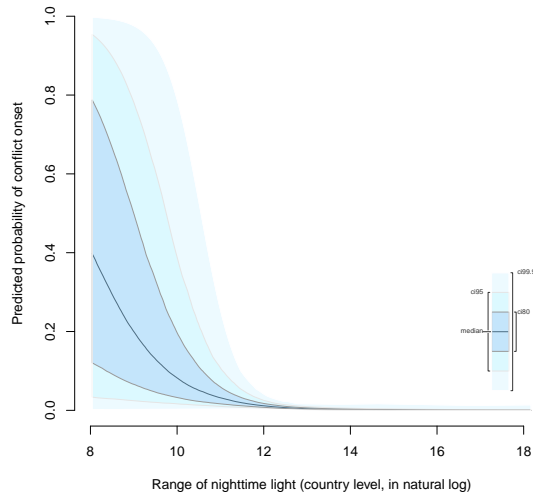
Model 1
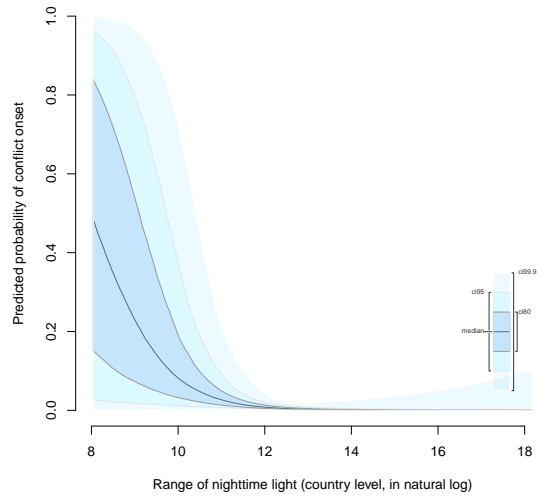
Model 2

Model 3

Robustness Analyses

In this section we demonstrate the robustness of our findings to a large number of alternative model specifications and potential confounders.
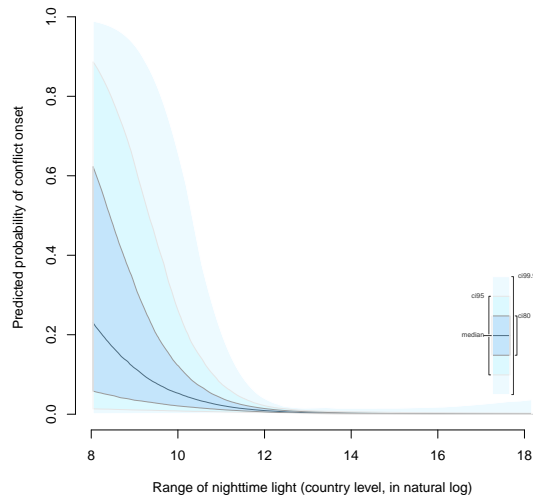
Parsimonious Sensitivity Analysis

We begin this section by running a set of baseline model specifications to show that our results are not driven by the inclusion of some specific controls (e.g., population densities). To this extant, Table A3 replicates the main analyses using only a minimal number of controls for geospatial and temporal dependencies; then with the addition of gross cell product; and finally with the addition of population densities.

Table A3: Civil War Onset – Baseline Specification

	Baseline		With GCP		With Population	
	(1) Civil War	(2) "Ethnic War"	(1) Civil War	(2) "Ethnic War"	(1) Civil War	(2) "Ethnic War"
Nighttime light[2]	0.465* (0.072)	0.477* (0.074)	0.481* (0.077)	0.516* (0.080)	0.426* (0.080)	0.470* (0.082)
Civil war[1]	−0.571† (0.294)	−0.614* (0.305)	−0.408 (0.303)	−0.431 (0.316)	−0.422 (0.302)	−0.446 (0.314)
Population[1,2]	–	–	–	–	0.394* (0.144)	0.353* (0.146)
Gross cell product[1,2]	–	–	−0.116 (0.227)	−0.456† (0.263)	−0.384 (0.258)	−0.703* (0.291)
Mountainous area (%)	2.135* (0.376)	1.761* (0.398)	2.175* (0.398)	1.742* (0.425)	2.051* (0.397)	1.624* (0.424)
Distance to border[2]	−0.378* (0.074)	−0.360* (0.078)	−0.404* (0.077)	−0.370* (0.082)	−0.376* (0.079)	−0.339* (0.084)
Cell area[2]	0.387 (0.295)	0.371 (0.301)	0.291 (0.303)	0.310 (0.319)	−0.348 (0.278)	−0.208 (0.284)
Travel time[2]	−0.701* (0.219)	−0.467* (0.220)	−0.717* (0.237)	−0.555* (0.241)	−0.348 (0.278)	−0.208 (0.284)
Constant	−9.196† (5.120)	−14.887* (5.525)	−9.539† (5.296)	−15.194* (5.818)	−14.345* (5.690)	−19.998* (6.289)
Observations	926,206	629,754	899,821	604,438	899,821	604,438
Akaike Inf. Crit.	1,735.757	1,545.208	1,609.088	1,417.603	1,603.225	1,413.568

Note: †p<0.1; *p<0.05; values in parentheses are robust standard errors clustered by cell-id. Year and country fixed effects were included in each regression, although not reported here.
[1] lagged; [2] natural log

Alternative Grid Cell Level Confounders

Next, we assess the sensitivity of our models to other cell level factors. First, to account for the possibility that the correlations between nighttime light and civil war onset are the result not of state capacity, but of urban development, we replicate the models presented in Table 2 of the main paper with the inclusion of the percent of each cell's urban development levels for a given (lagged) year in Table A4. We are also mindful of the potential effect of conflict "spillovers" from other cells in respect to civil war onset. To account for this possibility, we repeat the models presented in Table 2 with the addition of a spatial lag for civil war onset – coding the number of cells that experienced civil war onset as a fraction of all contingent cells – in Table A5.

Table A4: Civil War Onset – Substate Level Analysis with Urban Development

	(1) Civil War	(2) "Ethnic War"	(3) Civil War
Nighttime light[2]	0.460*	0.488*	0.460*
	(0.081)	(0.083)	(0.081)
Civil war[1]	−0.462	−0.516	−0.465
	(0.310)	(0.323)	(0.309)
Population[1,2]	0.371*	0.346*	0.367*
	(0.143)	(0.148)	(0.143)
Gross cell product[1,2]	−0.597*	−0.734*	−0.598*
	(0.276)	(0.298)	(0.276)
Mountainous area (%)	2.050*	1.639*	2.047*
	(0.400)	(0.426)	(0.400)
Distance to border[2]	−0.368*	−0.336*	−0.368*
	(0.079)	(0.084)	(0.079)
Oil production[1, 2]	−0.080	−0.075	−0.066
	(0.122)	(0.121)	(0.121)
New state	2.726*	2.889*	2.584*
	(1.241)	(1.255)	(1.216)
Political instability	1.154*	1.113*	1.082*
	(0.360)	(0.390)	(0.358)
Polity 2[1]	0.038	0.023	–
	(0.038)	(0.039)	
Anocracy[1]	–	–	−0.224
			(0.359)
Polity 2 (binary)[1]	–	–	0.154
			(0.525)
Cell area[2]	0.002	0.033	0.004
	(0.333)	(0.344)	(0.332)
Travel time[2]	−0.277	−0.196	−0.280
	(0.277)	(0.284)	(0.277)
Urban area (%)[1]	0.096*	0.015	0.096*
	(0.039)	(0.089)	(0.039)
Constant	−12.790*	−18.833*	−12.499*
	(6.225)	(6.755)	(6.197)
Observations	898,421	603,664	898,421
Akaike Inf. Crit.	1,576.360	1,400.499	1,578.999

Note: †p<0.1; *p<0.05; values in parentheses are robust standard errors clustered by cell-id. Year and country fixed effects were included in each regression, although not reported here.
[1] lagged; [2] natural log

Table A5: Civil War Onset – Substate Level Analysis with Spatial Lag War Onset

	(1) Civil War	(2) "Ethnic War"	(3) Civil War
Nighttime light[2]	0.480*	0.520*	0.478*
	(0.086)	(0.087)	(0.085)
Civil war[1]	−0.513	−0.555†	−0.513
	(0.318)	(0.330)	(0.318)
Population[1,2]	0.312*	0.271†	0.308*
	(0.155)	(0.157)	(0.155)
Gross cell product[1,2]	−0.400	−0.713*	−0.399
	(0.280)	(0.315)	(0.280)
Mountainous area (%)	2.361*	1.870*	2.355*
	(0.407)	(0.431)	(0.407)
Distance to border[2]	−0.433*	−0.388*	−0.433*
	(0.087)	(0.093)	(0.087)
Oil production[1, 2]	−0.054	−0.054	−0.045
	(0.119)	(0.120)	(0.118)
New state	2.764*	2.957*	2.623*
	(1.237)	(1.261)	(1.207)
Political instability	0.971*	1.031*	0.887*
	(0.399)	(0.423)	(0.395)
Polity 2[1]	0.060	0.047	–
	(0.043)	(0.044)	
Anocracy[1]	–	–	−0.162
			(0.378)
Polity 2 (binary)[1]	–	–	0.282
			(0.569)
Cell area[2]	−0.137	−0.519	−0.142
	(1.009)	(0.826)	(1.008)
Travel time[2]	−0.556†	−0.389	−0.559†
	(0.305)	(0.312)	(0.305)
Spatial lag onset	−83.905	−83.824	−83.838
	(65,243.130)	(69,419.260)	(65,129.960)
Constant Constant	−11.093	−13.619	−10.600
	(10.860)	(9.593)	(10.840)
Observations	823,272	551,697	823,272
Akaike Inf. Crit.	1,477.431	1,307.167	1,481.141

Note: †p<0.1; *p<0.05; values in parentheses are robust standard errors clustered by cell-id. Year and country fixed effects were included in each regression, although not reported here.
[1] lagged; [2] natural log

Alternative Country Level Confounders

We also account for the potential effect of increased military activity and military spending the year before a civil war onset by replicating these models with the addition of (lagged) military expenditure in Table A6. Additionally, to show that our findings are not the driven by the fact that our sample includes a large number of countries that did not experience conflict, we replicate Table 2 on a subsample consisting solely of countries that experienced civil war onset at some point during are temporal period (1992-2008) in Table A7. Finally, we report a set of models that include an indicator measuring the size of each given country to account for the possibility that the effect of nighttime light is the result of geographic vastness (i.e. that there are very few populated regions in the country, and hence that conflict frequents these locations) or the fact that civil war is more likely in smaller countries in Table A8 (because this indicator is not time varying, country fixed effects were not used).

Table A6: Civil War Onset – Substate Level Analysis with Military Expenditure

	(1) Civil War	(2) "Ethnic War"	(3) Civil War
Nighttime light[2]	0.446*	0.491*	0.446*
	(0.080)	(0.083)	(0.080)
Civil war[1]	−0.468	−0.502	−0.474
	(0.311)	(0.325)	(0.311)
Population[1,2]	0.387*	0.343*	0.383*
	(0.144)	(0.146)	(0.144)
Gross cell product[1,2]	−0.389	−0.705*	−0.390
	(0.260)	(0.294)	(0.260)
Mountainous area (%)	2.084*	1.654*	2.083*
	(0.398)	(0.426)	(0.398)
Distance to border[2]	−0.375*	−0.338*	−0.375*
	(0.079)	(0.084)	(0.079)
Oil production[1, 2]	−0.065	−0.065	−0.051
	(0.119)	(0.120)	(0.118)
New state	2.467*	2.642*	2.336†
	(1.228)	(1.248)	(1.206)
Political instability	0.032	0.019	–
	(0.038)	(0.040)	
Polity 2[1]	−0.026	−0.040*	–
	(0.019)	(0.020)	
Anocracy[1]	–	–	−0.187
			(0.359)
Polity 2 (binary)[1]	–	–	0.079
			(0.522)
Cell area[2]	−0.006	0.033	−0.004
	(0.324)	(0.343)	(0.323)
Travel time[2]	−0.337	−0.194	−0.340
	(0.278)	(0.283)	(0.278)
Military expenditure[1,2]	−0.251	−0.182	−0.260
	(0.236)	(0.255)	(0.234)
Constant	−8.542	−15.415†	−8.123
	(7.658)	(8.354)	(7.610)
Observations	891,834	597,123	891,834
Akaike Inf. Crit.	1,568.939	1,390.204	1,571.371

Note: †p<0.1; *p<0.05; values in parentheses are robust standard errors clustered by cell-id. Year and country fixed effects were included in each regression, although not reported here.
[1] lagged; [2] natural log

Table A7: Civil War Onset – Substate Level Analysis, Civil War Countries Only

	(1) Civil War	(2) "Ethnic War"	(3) Civil War
Nighttime light[2]	0.442*	0.488*	0.442*
	(0.080)	(0.083)	(0.080)
Civil war[1]	−0.479	−0.516	−0.482
	(0.310)	(0.323)	(0.309)
Population[1,2]	0.393*	0.350*	0.390*
	(0.144)	(0.147)	(0.144)
Gross cell product[1,2]	−0.410	−0.725*	−0.411
	(0.261)	(0.294)	(0.261)
Mountainous area (%)	2.071*	1.641*	2.068*
	(0.398)	(0.425)	(0.398)
Distance to border[2]	−0.373*	−0.335*	−0.373*
	(0.079)	(0.084)	(0.079)
Oil production[1, 2]	−0.078	−0.075	−0.065
	(0.121)	(0.121)	(0.120)
New state	2.692*	2.888*	2.558*
	(1.236)	(1.255)	(1.214)
Political instability	1.142*	1.113*	1.073*
	(0.359)	(0.390)	(0.358)
Polity 2[1]	0.038	0.023	–
	(0.038)	(0.039)	
Anocracy[1]	–	–	−0.210
			(0.359)
Polity 2 (binary)[1]	–	–	0.153
			(0.525)
Cell area[2]	−0.008	0.031	−0.006
	(0.324)	(0.343)	(0.323)
Travel time[2]	−0.343	−0.199	−0.346
	(0.278)	(0.283)	(0.278)
Constant	−13.296*	−18.883*	−13.011*
	(6.203)	(6.749)	(6.176)
Observations	434,061	237,309	434,061
Akaike Inf. Crit.	1,390.662	1,254.525	1,393.321

Note: [†]p<0.1; *p<0.05; values in parentheses are robust standard errors clustered by cell-id. Year and country fixed effects were included in each regression, although not reported here.
[1] lagged; [2] natural log

Table A8: Civil War Onset – Substate Level Analysis With Country Size

	(1) Civil War	(2) "Ethnic War"	(3) Civil War
Nighttime light[2]	0.183*	0.205*	0.201*
	(0.068)	(0.069)	(0.068)
Civil war[1]	0.639*	0.425	0.581*
	(0.257)	(0.275)	(0.261)
Population[1,2]	0.626*	0.693*	0.632*
	(0.109)	(0.117)	(0.112)
Gross cell product[1,2]	−0.974*	−1.085*	−0.942*
	(0.226)	(0.256)	(0.225)
Mountainous area (%)	0.831*	0.726*	0.866*
	(0.297)	(0.314)	(0.297)
Distance to border[2]	−0.377*	−0.349*	−0.373*
	(0.074)	(0.079)	(0.074)
Oil production[1, 2]	0.023	0.020	0.022
	(0.022)	(0.024)	(0.022)
New state	2.304*	2.460*	2.148*
	(0.791)	(0.810)	(0.799)
Political instability	1.279*	1.054*	1.141*
	(0.314)	(0.337)	(0.318)
Polity 2[1]	−0.022	−0.036†	
	(0.018)	(0.019)	
Anocracy[1]	–	–	−0.562*
			(0.259)
Polity 2 (binary)[1]	–	–	−0.263
			(0.303)
Cell area[2]	0.191	0.078	0.172
	(0.367)	(0.381)	(0.367)
Travel time[2]	−0.337	−0.073	−0.320
	(0.241)	(0.246)	(0.241)
Country size[2]	−0.261*	−0.229*	−0.245*
	(0.093)	(0.117)	(0.093)
Constant	−10.064*	−11.989*	−10.033*
	(3.443)	(3.777)	(3.485)
Observations	898,421	603,664	898,421
Akaike Inf. Crit.	1,539.327	1,371.345	1,534.546

Note: †p<0.1; *p<0.05; values in parentheses are robust standard errors clustered by cell-id. Year fixed effects were included in each regression, although not reported here.
[1] lagged; [2] natural log

Alternative IV Conceptualizations

In this subsection we test the sensitivity of our results to alternative conceptualizations of the independent variable. To account for the possibility that the relationship between state capacity and violence is curvilinear, as some previous studies have suggested (Koren, Forthcoming; Ulfelder, 2012), we replicate our main logit models with the addition of a quadratic term (which is insignificant in all models). We also account for the possibility that our findings are driven by the nighttime light indicator we chose to use by employing an alternative, annual variable measuring the average levels of nighttime light in a given grid cell (calibrated to a maximum value of one) obtained from the PRIO-Grid dataset (Tollefsen et al., 2012). We chose to limit our original indicator to vary over a five-year period so as to account for potential endogeneities, because in some years a given cell might have especially high levels of nighttime light emissions that are *temporary*—e.g., due massive construction projects—that do not persist over time. We show that our findings are robust to this decision by employing PRIO-Grid's alternative measure in Table A9. The next set of robustness models accounts for situations where relatively high nighttime light emission levels might result from factors such as high population densities or high urban concentrations, even in countries and regions with low public good provision levels and low tax collection. To do so, we normalize our nighttime light indicator by population size, which better accounts for the possibility that high nighttime light emission are caused by reasons other than high state capacity levels. The results of this analysis are provided in Table A11.

Table A9: Civil War Onset – Alternative Nighttime Light Indicator

	(1) Civil War	(2) "Ethnic War"	(3) Civil War
Calibrated light mean (PRIO)[2]	6.628* (1.629)	6.828* (1.781)	6.642* (1.628)
Civil war[1]	−0.463 (0.307)	−0.514 (0.320)	−0.469 (0.306)
Population[1,2]	0.628* (0.141)	0.596* (0.145)	0.625* (0.141)
Gross cell product[1,2]	−1.017* (0.301)	−1.176* (0.323)	−1.019* (0.302)
Mountainous area (%)	1.932* (0.396)	1.534* (0.423)	1.930* (0.396)
Distance to border[2]	−0.364* (0.080)	−0.338* (0.086)	−0.364* (0.080)
Oil production[1, 2]	−0.077 (0.122)	−0.071 (0.121)	−0.066 (0.122)
New state	2.590* (1.253)	2.756* (1.271)	2.471* (1.224)
Political instability	1.167* (0.360)	1.108* (0.391)	1.099* (0.358)
Polity 2[1]	0.039 (0.038)	0.022 (0.039)	–
Anocracy[1]	–	–	−0.226 (0.356)
Polity 2 (binary)[1]	–	–	0.224 (0.523)
Cell area[2]	−0.160 (0.284)	−0.162 (0.289)	−0.157 (0.284)
Travel time[2]	−0.618* (0.267)	−0.541* (0.274)	−0.620* (0.267)
Constant	−9.724[†] (5.889)	−15.418* (6.491)	−9.472 (5.870)
Observations	879,106	596,062	879,106
Akaike Inf. Crit.	1,596.836	1,423.606	1,599.425

Note: [†]p<0.1; *p<0.05; values in parentheses are robust standard errors clustered by cell-id. Year and country fixed effects were included in each regression, although not reported here.
[1] lagged; [2] natural log

Table A10: Civil War Onset – Substate Level Analysis with Curvilinear Nighttime Light

	(1) Civil War	(2) "Ethnic War"	(3) Civil War
Nighttime light[2]	0.498*	0.545*	0.495*
	(0.174)	(0.180)	(0.174)
Nighttime light squared[2]	−0.008	−0.009	−0.008
	(0.022)	(0.023)	(0.022)
Civil war[1]	−0.482	−0.519	−0.485
	(0.310)	(0.323)	(0.309)
Gross cell product[1,2]	−0.375	−0.682*	−0.377
	(0.277)	(0.316)	(0.277)
Population[1,2]	0.391*	0.345*	0.387*
	(0.144)	(0.147)	(0.144)
Mountainous area (%)	2.045*	1.615*	2.044*
	(0.403)	(0.430)	(0.403)
Distance to border[2]	−0.375*	−0.339*	−0.375*
	(0.079)	(0.085)	(0.079)
Oil production[1, 2]	−0.080	−0.077	−0.067
	(0.121)	(0.121)	(0.121)
New state	2.674*	2.867*	2.545*
	(1.231)	(1.250)	(1.211)
Political instability	1.141*	1.114*	1.074*
	(0.359)	(0.390)	(0.358)
Polity 2[1]	0.038	0.023	–
	(0.038)	(0.039)	
Anocracy[1]	–	–	−0.206
			(0.360)
Polity 2 (binary)[1]	–	–	0.160
			(0.525)
Cell area[2]	0.020	−0.017	0.019
	(0.328)	(0.355)	(0.334)
Travel time[2]	−0.353	−0.209	−0.355
	(0.279)	(0.285)	(0.279)
Constant	−12.189*	−18.865*	−12.998*
	(6.137)	(6.746)	(6.172)
Observations	898,421	603,664	898,421
Akaike Inf. Crit.	1,580.527	1,400.392	1,583.200

Note: [†]p<0.1; *p<0.05; values in parentheses are robust standard errors clustered by cell-id. Year and country fixed effects were included in each regression, although not reported here.
[1] lagged; [2] natural log

Table A11: Civil War Onset – Substate Level Analysis, Nighttime Light Normalized Per Capita

	(1) Civil War	(2) "Ethnic War"	(3) Civil War
Nighttime light (per capita)[2]	2.688*	2.527*	2.686*
	(0.364)	(0.384)	(0.364)
Civil war[1]	−0.518	−0.570	−0.521
	(0.308)	(0.321)	(0.307)
Population[1,2]	0.611*	0.579*	0.608*
	(0.142)	(0.148)	(0.142)
Gross cell product[1,2]	−0.353	−0.607*	−0.353
	(0.259)	(0.295)	(0.260)
Mountainous area (%)	1.969*	1.560*	1.966*
	(0.395)	(0.421)	(0.395)
Distance to border[2]	−0.399*	−0.378*	−0.399*
	(0.078)	(0.083)	(0.078)
Oil production[1, 2]	−0.070	−0.063	−0.059
	(0.119)	(0.118)	(0.119)
New state	2.540*	2.699*	2.421*
	(1.228)	(1.247)	(1.206)
Political instability	1.129*	1.092*	1.063*
	(0.359)	(0.389)	(0.357)
Polity 2[1]	0.038	0.023	–
	(0.038)	(0.039)	
Anocracy[1]	–	–	−0.207
			(0.358)
Polity 2 (binary)[1]	–	–	0.188
			(0.526)
Cell area[2]	0.097	0.124	0.098
	(0.339)	(0.354)	(0.338)
Travel time[2]	−0.568*	−0.502	−0.570*
	(0.274)	(0.281)	(0.274)
Constant	−13.272*	−17.520*	−12.995*
	(6.233)	(6.828)	(6.210)
Observations	840,443	546,576	840,443
Akaike Inf. Crit.	1,588.313	1,412.377	1,590.979

Note: [†]p<0.1; [*]p<0.05; values in parentheses are robust standard errors clustered by cell-id. Year and country fixed effects were included in each regression, although not reported here.
[1] lagged; [2] natural log

Province Level Analysis

Next, we examine whether our findings remain consistent if we rely on a different sub-national aggregation – the province level. First, one might argue that more than focusing on the geographically disaggregated distribution of nighttime light, the emphasis should be placed on the effect of administrative units in governing the provision of electricity. To show that the findings are robust to our decision to focus on the grid cell rather than the administrative unit-level, we conduct a robustness analysis in which all measures were aggregated at the province/district level (using the natural log of both average nighttime light per district and the sum of light per district). Note that while the effect of nighttime light is robust to this change in the unit of analysis, the effect of mountains now becomes insignificant. We then repeat the curvilinear nighttime light analysis discussed above using the same province-year sample in Table A13.

Table A12: Civil War Onset – Substate Level Analysis, Administrative District

	Average nighttime light			Total nighttime light		
	(1) Civil War	(2) "Ethnic War"	(3) Civil War	(1) Civil War	(2) "Ethnic War"	(3) Civil War
Nighttime light[2,3]	0.577* (0.103)	0.587* (0.106)	0.578* (0.103)	0.244* (0.066)	0.250* (0.068)	0.244* (0.066)
Civil war[1]	−0.573† (0.345)	−0.637† (0.361)	−0.581† (0.344)	−0.566† (0.338)	−0.603† (0.354)	−0.574† (0.338)
Population[1,2]	0.164 (0.187)	0.063 (0.189)	0.158 (0.187)	0.413* (0.190)	0.313 (0.195)	0.409* (0.190)
Gross cell product[1,2]	−0.237 (0.305)	−0.422 (0.336)	−0.239 (0.305)	−0.150 (0.294)	−0.310 (0.326)	−0.150 (0.295)
Mountainous area (%)	1.118* (0.509)	1.124* (0.530)	1.114* (0.508)	0.987* (0.503)	1.078* (0.532)	0.986* (0.503)
Distance to border[2]	−0.099 (0.102)	−0.139 (0.108)	−0.100 (0.102)	−0.117 (0.101)	−0.162 (0.105)	−0.117 (0.101)
Oil production[1,2]	−0.094 (0.123)	−0.088 (0.122)	−0.076 (0.121)	−0.075 (0.119)	−0.071 (0.119)	−0.064 (0.119)
New state	2.913* (1.328)	3.120* (1.349)	2.783* (1.305)	2.625* (1.309)	2.805* (1.326)	2.531† (1.293)
Political instability	1.203* (0.364)	1.113* (0.393)	1.130* (0.363)	1.169* (0.363)	1.094* (0.392)	1.110* (0.361)
Polity 2[1]	0.038 (0.038)	0.021 (0.039)	–	0.037 (0.038)	0.021 (0.040)	–
Anocracy[1]	–	–	−0.209 (0.369)	–	–	−0.174 (0.365)
Polity 2 (binary)[1]	–	–	0.077 (0.530)	–	–	0.151 (0.529)
Cell area[2]	0.415 (0.338)	0.526 (0.367)	0.420 (0.338)	0.077 (0.312)	0.154 (0.338)	0.080 (0.312)
Travel time[2]	0.410 (0.386)	0.222 (0.397)	0.404 (0.386)	0.185 (0.388)	−0.016 (0.403)	0.180 (0.388)
Constant	−7.235 (6.379)	−7.396 (6.674)	−7.022 (6.356)	−5.231 (6.403)	−4.523 (6.650)	−5.022 (6.390)
Observations	37,630	29,653	37,630	37,630	29,653	37,630
Akaike Inf. Crit.	1,226.477	1,094.091	1,229.157	1,248.984	1,115.692	1,251.684

Note: †p<0.1; *p<0.05; values in parentheses are robust standard errors clustered by district. Year and country fixed effects were included in each regression, although not reported here.
[1] lagged; [2] natural log; [3] by district

Table A13: Civil War Onset – Substate Level Analysis with Curvilinear Nighttime Light (District Level)

	(1) Civil War	(2) "Ethnic War"	(3) Civil War
Nighttime light[2,3]	0.590*	0.614*	0.589*
	(0.223)	(0.229)	(0.223)
Nighttime light squared[2,3]	−0.002	−0.004	−0.002
	(0.028)	(0.030)	(0.028)
Gross cell product[1,2]	−0.227	−0.399	−0.231
	(0.338)	(0.375)	(0.338)
Civil war[1]	−0.575†	−0.641†	−0.583†
	(0.346)	(0.362)	(0.345)
Population[1,2]	0.163	0.061	0.157
	(0.187)	(0.190)	(0.187)
Mountainous area (%)	1.113*	1.114*	1.110*
	(0.514)	(0.534)	(0.514)
Distance to border[2]	−0.101	−0.142	−0.101
	(0.104)	(0.110)	(0.104)
Oil production[1,2]	−0.094	−0.089	−0.076
	(0.123)	(0.122)	(0.122)
New state	2.908*	3.108*	2.780*
	(1.329)	(1.349)	(1.305)
Political instability	1.203*	1.114*	1.130*
	(0.364)	(0.393)	(0.363)
Polity 2[1]	0.038	0.021	–
	(0.038)	(0.040)	
Anocracy[1]			−0.208
			(0.369)
Polity 2 (binary)[1]			0.078
			(0.530)
Cell area[2]	0.414	0.523	0.419
	(0.338)	(0.367)	(0.338)
Travel time[2]	0.407	0.217	0.402
	(0.388)	(0.399)	(0.388)
Constant	−7.201	−7.313	−6.996
	(6.400)	(6.701)	(6.376)
Observations	37,630	29,653	37,630
Akaike Inf. Crit.	1,228.473	1,096.073	1,231.154

Note: †p<0.1; *p<0.05; values in parentheses are robust standard errors clustered by cell-id. Year and country fixed effects were included in each regression, although not reported here.
[1] lagged; [2] natural log [3] by district

Random Effects Models

In this subsection we repeat the analyses presented in Table 2 and including random effects at the cell and, separately, country levels in Tables A14 and A15, respectively. These random effects assume a hierarchical structure in which some cell and country specific factors – such as spatial dependencies, population spillovers, etc. – might still independently influence the incidence of civil war onset (Gelman and Hill, 2007). The use of these random effects therefore takes into consideration the possibility that civil war is inherently more likely to arise in some cells. Note that these models include the time-invariant indicators *ethnic fractionalization* and *religious fractionalization* because no country fixed effects were included in these models. Hence, these indicators, which are constant over time, are not automatically dropped from analysis.

Table A14: Civil War Onset – Substate Level Analysis with Grid-Cell Random Effects

	(1) Civil War	(2) "Ethnic War"	(3) Civil War
Nighttime light[2]	0.135* (0.080 ⇔ 0.192)	0.234* (0.191 ⇔ 0.271)	0.233* (0.136 ⇔ 0.301)
Civil war[1]	0.597* (0.167 ⇔ 0.974)	0.198 (-0.224 ⇔ 0.593)	0.322 (-0.026 ⇔ 0.603)
Population[1,2]	0.367* (0.329 ⇔ 0.411)	0.557* (0.520 ⇔ 0.588)	0.324* (0.301 ⇔ 0.349)
Gross cell product[1,2]	-0.491* (-0.569 ⇔ -0.401)	-0.939* (-1.059 ⇔ -0.793)	-0.776* (-0.992 ⇔ -0.457)
Mountainous area (%)	0.722* (0.522 ⇔ 0.917)	0.878* (0.559 ⇔ 1.352)	1.133* (0.895 ⇔ 1.378)
Distance to border[2]	-0.452* (-0.493 ⇔ -0.404)	-0.260* (-0.345 ⇔ -0.182)	-0.259* (-0.316 ⇔ -0.216)
Oil production[1, 2]	-0.036* (-0.051 ⇔ -0.019)	-0.044* (-0.061 ⇔ -0.029)	-0.036* (-0.058 ⇔ -0.017)
New state	2.477* (0.873 ⇔ 4.086)	2.366* (0.704 ⇔ 3.833)	2.847* (1.285 ⇔ 4.067)
Political instability	1.073* (0.414 ⇔ 1.618)	0.889* (0.499 ⇔ 1.307)	1.538* (1.121 ⇔ 1.892)
Ethnic fractionalization	1.671* (1.377 ⇔ 1.980)	0.704* (0.166 ⇔ 1.161)	1.548* (1.158 ⇔ 1.989)
Religious fractionalization	-0.728* (-1.123 ⇔ -0.347)	-0.301* (-0.602 ⇔ -0.015)	-0.513* (-1.020 ⇔ -0.097)
Polity 2[1]	-0.025* (-0.042 ⇔ -0.005)	-0.047* (-0.061 ⇔ -0.034)	–
Anocracy[1]	–	–	-0.554* (-0.736 ⇔ -0.282)
Polity 2 (binary)[1]	–	–	-0.031 (-0.210 ⇔ 0.124)
Cell area[2]	-0.408* (-0.545 ⇔ -0.270)	-0.086 (-0.209 ⇔ 0.076)	-0.015 (-0.142 ⇔ 0.092)
Travel time[2]	-0.271* (-0.366 ⇔ -0.209)	-0.028 (-0.143 ⇔ 0.106)	-0.452* (-0.616 ⇔ -0.323)
Constant	-8.638* (-10.016 ⇔ -7.243)	13.955* (-15.461 ⇔ -12.757)	-12.099* (-12.933 ⇔ -11.243)
Observations	892,913	603,664	892,913
Deviance Inf. Crit.	1,497.061	1,357.978	1,243.579

Note: *p<0.05; values in parentheses are 95% credible intervals. Year fixed effects were included in each regression, although not reported here. Models were estimated using 210,000 MCMC iterations, of which the first 200,000 were discarded.
[1] lagged; [2] natural log

Table A15: Civil War Onset – Substate Level Analysis with Country Random Effects

	(1) Civil War	(2) "Ethnic War"	(3) Civil War
Nighttime light[2]	0.258* (0.173 ⇔ 0.344)	0.365* (0.318 ⇔ 0.408)	0.306* (0.241 ⇔ 0.352)
Civil war[1]	−0.032 (-0.597 ⇔ 0.706)	−0.072 (-0.679 ⇔ 0.558)	−0.049 (-0.225 ⇔ 0.401)
Population[1,2]	0.336* (0.268 ⇔ 0.402)	0.370* (0.332 ⇔ 0.423)	0.293* (0.183 ⇔ 0.376)
Gross cell product[1,2]	−0.340* (-0.562 ⇔ -0.108)	−0.455* (-0.582 ⇔ -0.330)	−0.221 (-0.567 ⇔ 0.056)
Mountainous area (%)	1.514* (1.292 ⇔ 1.713)	1.089* (0.739 ⇔ 1.521)	1.103* (0.848 ⇔ 1.364)
Distance to border[2]	−0.409* (-0.482 ⇔ -0.317)	−0.379* (-0.435 ⇔ -0.322)	−-0.391* (-0.449 ⇔ -0.298)
Oil production[1, 2]	−0.004 (0.037 ⇔ 0.033)	−0.040 (-0.078 ⇔ 0.012)	−-0.057* (-0.091 ⇔ -0.021)
New state	2.283* (0.384 ⇔ 4.134)	2.598* (0.603 ⇔ 4.668)	1.867 (-0.283 ⇔ 3.876)
Political instability	1.326* (1.005 ⇔ 1.776)	1.394* (1.024 ⇔ 1.722)	1.57* (1.038 ⇔ 1.886)
Ethnic fractionalization	2.456* (0.921 ⇔ 3.943)	1.870* (0.389 ⇔ 3.581)	2.185* (0.903 ⇔ 3.422)
Religious fractionalization	1.909* (0.224 ⇔ 3.718)	0.362 (-1.493 ⇔ 2.319)	0.206 (-1.409 ⇔ 1.665)
Polity 2[1]	−0.010 (-0.045 ⇔ 0.026)	−-0.032* (-0.072 ⇔ 0.007)	–
Anocracy[1]	–	–	−0.759* (-1.141 ⇔ -0.489)
Polity 2 (binary)[1]	–	–	−0.315 (-0.837 ⇔ 0.580)
Cell area[2]	0.235* (0.106 ⇔ 0.370)	−0.413* (-0.562 ⇔ -0.203)	−0.216* (-0.271 ⇔ -0.142)
Travel time[2]	−0.489* (-0.592 ⇔ -0.377)	-0.031 (-0.174 ⇔ 0.115)	-0.099 (-0.283 ⇔ 0.075)
Constant	−15.589* (-17.090 ⇔ -13.984)	−16.204* (-0.175 ⇔ 0.115)	−10.919* (-13.478 ⇔ -7.744)
Observations	892,913	603,664	892,913
Deviance Inf. Crit.	1,353.766	1,242.538	1,670.877

Note: *p<0.05; values in parentheses are 95% credible intervals. Year fixed effects were included in each regression, although not reported here. Models were estimated using 210,000 MCMC iterations, of which the first 200,000 were discarded.
[1] lagged; [2] natural log

Finally, to further illustrate that nighttime light could be used as a *disaggregated* indicator of state capacity in respect to conflict, it is worth contrasting the effect of our nighttime light indicator against another measure of state capacity at the highly localized level. To this extent, we estimate four logit models using a subsample of grid cells located solely in the Indian state of Andhra Pradesh. We estimate both baseline and full specifications, and in each specification we report one model that includes the number of health centers by sub district for 2012, and again, this time using our nighttime light indicator instead. Using the number of health centers as an indicator of state capacity and the broader empirical framework correspond to our third validation exercise reported in the main paper (hence, the number of health centers is constant for 2012). Because, luckily, the onset of civil war is an exceptionally rare event, we include instead an indicator measuring whether civil war – i.e. a conflict with at least 25 combatant casualties – afflicted the region or not during a given year. Because they are constant across all cells, state level indicators were not included in analysis. Additionally, due to the relatively small sample size and the limitations on the number of degrees of freedom it entails, we replaced the use of year fixed effects with a control variable for a linear time trend. The cell area indicator was also omitted because it shows almost no variation between different grid cells due to the geographic size of Andhra Pradesh.

Table A16 shows the estimates of two models, Baseline and Full, accounting for the relationship between nighttime light and civil war alongside the aforementioned alternative measure of state capacity and a variety of cell-level confounders included in the previous models. The effect of health centers, although statistically insignificant, is nevertheless positive in Models 1 and 3. The effect of nighttime light is again positive and statistically significant. This again suggests that nighttime light is an effective indicator of state capacity and its localized relationship with civil war, as both state capacity indicators used here show a positive association with civil war at the grid cell level.

Table A16: Civil War – Substate Level Analysis, Andhra Pradesh

| | Baseline | | Full | |
	Model 1	Model 2	Model 3	Model 4
N. health centers[2]	0.103	–	0.099	–
	(0.183)		(0.191)	
Nighttime light[2]	–	0.624*	–	0.752*
		(0.183)		(0.213)
Civil war[1]	−16.619	−16.811	−16.626	−16.851
	(687.966)	(678.543)	(687.146)	(675.290)
Population[1,2]	0.058	−0.619*	0.081	−0.732*
	(0.148)	(0.250)	(0.153)	(0.281)
Gross cell product[1,2]	−0.736*	−0.778*	−0.927*	−0.805*
	(0.321)	(0.325)	(0.366)	(0.376)
Mountainous area (%)	–	–	0.403	1.007
			(0.806)	(0.829)
Distance to border[2]	–	–	0.538	−0.481
			(0.836)	(0.913)
Travel time[2]	–	–	−0.408	−0.361
			(0.405)	(0.407)
Time trend	0.470*	0.423*	0.479*	0.413*
	(0.049)	(0.050)	(0.050)	(0.052)
Constant	−921.608	−827.093	−949.944	−794.629
	(694.880)	(685.821)	(694.721)	(684.091)
Observations		1,424		
Akaike Inf. Crit.	485.682	474.324	490.370	478.032

Note: [†]p<0.1; *p<0.05; values in parentheses are robust standard errors clustered by cell-id.
[1] lagged; [2] natural log

Forecasting Localized Conflict With Nighttime Light

Beyond statistical significance, a valid benchmark model of conflict must also be able to reasonably *forecast* civil war onset (Brandt, Freeman and Schrodt, 2011; Ward, Greenhill and Bakke, 2010; Koren, Forthcoming). We thus additionally estimate the effect of nighttime light as a *predictive* indicator of civil war onset. We begin by calculating the predicted probability of civil war onset for each cell-year in our sample based on the coefficient estimates obtained from each model. This is done by subtracting an observation's predicted probability from one. We then repeat this exercise using a similar model that does not include nighttime light, but otherwise remains unchanged. Both models are estimated on a sample for the years 1992-2006 (presented in Tables A17). We then use the estimates obtained from each model to forecast civil war onset on out-of-sample data for the years 2007-2008.

The predictive strength of each model with and without the inclusion of nighttime light emissions was measured by calculating the area under the receiver operator characteristic (ROC) curve (Ward, Greenhill and Bakke, 2010) for the 2007-2008 sample. Additionally, the ROCs for Models 1,2, and 3 (with nighttime light) are presented in Figure A4, with the area under the curve (AUC) reported for each model and 95% confidence intervals in parenthesis. We then evaluate whether the AUC for each model that included the nighttime light indicator is larger from the AUC a model that does not include this indicator, and test whether this difference is statistically significant using a nonparametric significance test developed by DeLong, DeLong and Clarke-Pearson (1988). The difference in the AUC, model fit, and Delong et al. test scores for each model are reported in Table A4. The addition of nighttime light produces a significant improvement in the predictive power for each of the three models. Crucially, the strength of each model to forecast civil war onset is shown using out-of-sample data, i.e. data that were not used to estimate these models in the first place. AIC scores also confirm this conclusion by favoring in every case models that include nighttime light. This suggests that by approximating state capacity, nighttime light levels are an effective *predictive* indicator of civil war onset.

Table A17: Civil War Onset – Substate Level Analysis For Forecasting (1992-2006)

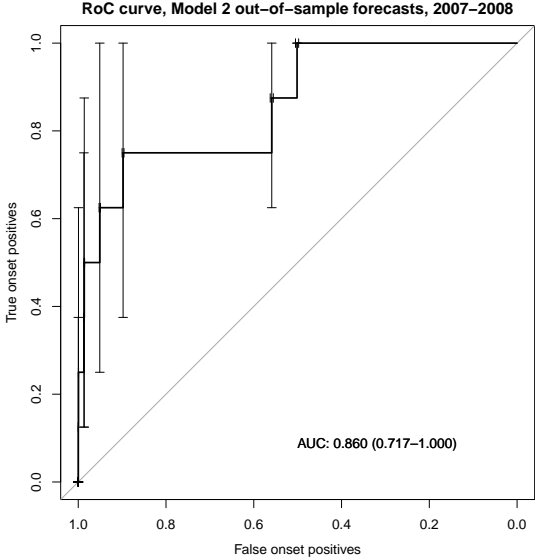|	(1) Civil War		(2) "Ethnic War"		(3) Civil War	
Nighttime light[2]	0.342*	–	0.387*	–	0.344*	–
	(0.083)		(0.086)		(0.083)	
Civil war[1]	−0.480	−0.530	−0.431	−0.492	−0.514	−0.564†
	(0.327)	(0.325)	(0.338)	(0.336)	(0.323)	(0.321)
Gross cell product[1,2]	−0.332	−0.328	−0.637*	−0.568†	−0.343	−0.338
	(0.269)	(0.263)	(0.305)	(0.299)	(0.270)	(0.265)
Population[1,2]	0.503*	0.715*	0.478*	0.706*	0.512*	0.726*
	(0.157)	(0.154)	(0.160)	(0.159)	(0.156)	(0.154)
Mountainous area (%)	1.892*	1.716*	1.546*	1.392*	1.929*	1.748*
	(0.421)	(0.414)	(0.446)	(0.437)	(0.419)	(0.412)
Distance to border[2]	−0.398*	−0.423*	−0.366*	−0.404*	−0.393*	−0.418*
	(0.083)	(0.084)	(0.088)	(0.088)	(0.083)	(0.084)
Oil production[1, 2]	−0.017	−0.011	−0.016	−0.008	0.033	0.039
	(0.117)	(0.116)	(0.117)	(0.115)	(0.110)	(0.108)
New state	2.382*	2.178†	2.580*	2.359†	2.112†	1.934†
	(1.207)	(1.190)	(1.225)	(1.208)	(1.171)	(1.157)
Political instability	1.164*	1.159*	1.167*	1.152*	1.111*	1.112*
	(0.365)	(0.364)	(0.397)	(0.396)	(0.370)	(0.368)
Polity 2[1]	0.013	0.014	−0.001	0.00004	–	–
	(0.041)	(0.041)	(0.043)	(0.043)		
Anocracy[1]	–	–	–	–	−0.778†	−0.772†
					(0.406)	(0.404)
Polity 2 (binary)[1]	–	–	–	–	0.318	0.381
					(0.666)	(0.667)
Cell area[2]	−0.095	−0.178	−0.082	−0.195	−0.093	−0.174
	(0.311)	(0.284)	(0.330)	(0.293)	(0.312)	(0.286)
Travel time[2]	−0.357	−0.713*	−0.185	−0.602*	−0.342	−0.698*
	(0.297)	(0.293)	(0.303)	(0.301)	(0.296)	(0.292)
Constant	−10.328*	−9.765*	−13.034*	−12.128*	−10.826*	−10.247*
	(3.076)	(2.922)	(3.475)	(3.282)	(3.154)	(3.002)
Observations	783,143	783,143	530,028	530,028	787,767	787,767
Akaike Inf. Crit.	1,420.277	1,436.185	1,278.117	1,297.796	1,429.597	1,445.787

Note: †p<0.1; *p<0.05; values in parentheses are robust standard errors clustered by cell-id. Year and country fixed effects were included in each regression, although not reported here.
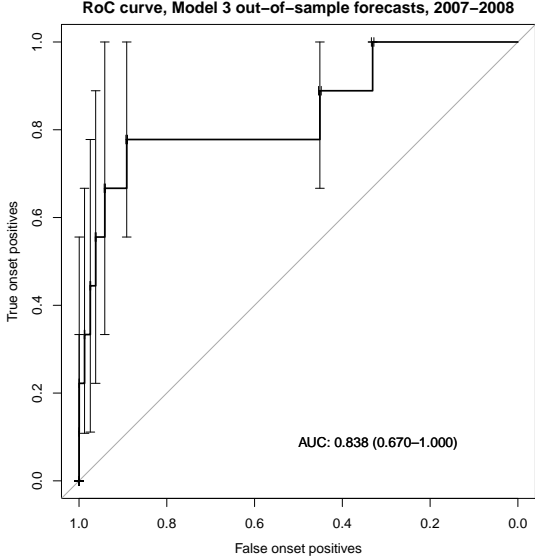[1] lagged; [2] natural log

Figure A4: ROC Curve For Each Model, Substate Level Analysis (2007-2008)
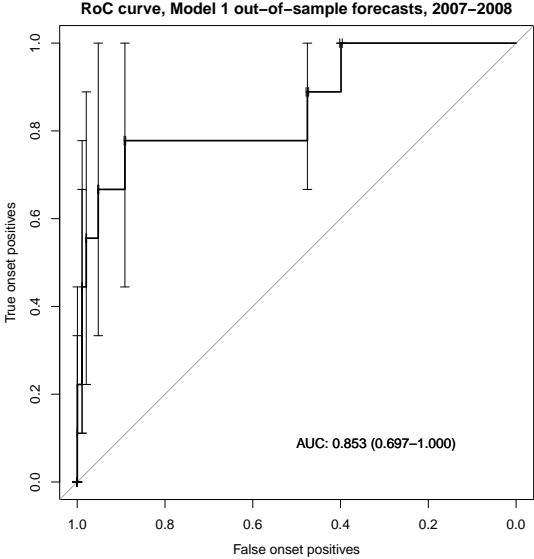
ROC: Model 1 (2007-2008)

ROC: Model 2 (2007-2008)

ROC: Model 3 (2007-2008)

Table A18: Difference in Area Under Curve For Each Model

| | (1) Civil War | | | (2) "Ethnic War" | | | (3) Civil War | |
	NL	No NL	‖	NL	No NL	‖	NL	No NL
AUC	85.3%	82.7%	‖	86.0%	82.0%	‖	83.8%	80.9%
AIC	1,420.277	1,436.185	‖	1,278.117	1,297.796	‖	1,429.597	1,445.787
DeLong et al. test	2.202*			2.351*			2.597*	

Note: [†]p<0.1; [*]p<0.05.

References

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

Box-Steffensmeier, Janet M. and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists.* Cambridge: Cambridge University Press.

Brandt, Patrick T., John R. Freeman and Philip A. Schrodt. 2011. "Real Time, Time Series Forecasting of Inter- and Intra-state Political Conflict." *Conflict Management and Peace Science* 28(1):41–64.

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.

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

Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel Hierarchical Model.* Cambridge: Cambridge University Press.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg and Håvard Strand. 2002. "Armed Conflict 1946-2001: A New Dataset." *Journal of Peace Research* 39(5):615–637.

Huang, Qingxu, Xi Yang, Bin Gao, Yang Yang and Yuanyuan Zhao. 2014. "Application of DMSP/OLS nighttime light images: A meta-analysis and a systematic literature review." *Remote Sensing* 6(8):6844–686.

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

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

Nordhaus, William D. 2006. "Geography and macroeconomics: New data and new findings." *Proceedings of the National Academy of Sciences* 103(10):3150–3517.

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

Singer, J. David, Stuart Bremer and John Stucky. 1972. Capability Distribution, Uncertainty, and Major Power War, 1820-1965. In *Peace, War, and Numbers*, ed. Bruce Russett. Beverly Hills: Sage.

Svolik, Milan W. 2008. "Authoritarian Reversals and Democratic Consolidation." *American Political Science Review* 102(2):153–168.

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

Ulfelder, Jay. 2012. "Forecasting Onsets of Mass Killing." Paper prepared for presentation at the annual Northeast Political Methodology Meeting.

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.