Supplemental Appendix For

Food, State Power, and Rebellion: The Case of Maize

Contents

This appendix proceeds in five parts. The first part report different descriptive statistics and figures not reported in the main text. In the second part, a detailed discussion of the data used to create the *Maize area*$_t$ is provided. The third section reports and discusses a large number of GLM sensitivity analyses, illustrating the findings' robustness to a large number of alternative confounders, in addition to GAMs with alternative parameterizations. In the fourth section, ROC curves of the full GLMs and GAMs are reported and discussed, Finally, the fifth section provides additional case-based discussion, including background to the Mau Mau rebellion as well as descriptive maps and correlations illustrating similar correlations in other, contemporary world regions..

Additional Information

Descriptive statistics

Table A1: Maize As Total Caloric Intake For Selected Countries*

Country	Maize (kcal/capita/day)†	Total Intake (kcal/capita/day)‡	Maize % Total
Malawi	1,163.667	2,236.52	**34.224%**
Zambia	927.667	1,967.49	**32.042%**
Mexico	997.333	2,123.56	**31.957%**
Kenya	695.667	1.798.67	**27.890%**
Guatemala	793.667	2,289.96	**25.738%**
Timor-Leste	598.667	2,180.08	**21.544%**
Togo	569.667	2,158.96	**20.877%**
Mozambique	437.667	1,955.3	**18.290%**
Egypt	567.333	2,629.32	**17.748%**
Moldova	548.667	2,689.94	**16.941%**
Paraguay	519.667	2,836.87	**15.482%**
Venezuela	400.667	2,189.09	**15.471%**
Nepal	372	2,230.92	**14.292%**
Bolivia	291.333	1,866.47	**13.501%**
Uganda	242.667	2,006.2	**10.791%**
Mali	271.667	2,276.33	**10.662%**
Ghana	200.667	2,302.29	**8.017%**
Côte d'Ivoire	182.333	2,104.62	**7.973%**
Haiti	184.333	2,323.93	**7.349%**
Panama	158	2371.33	**6.247%**
Pakistan	110.667	1949.4	**5.372%**
Laos	131.667	2,570.84	**4.872%**
Cambodia	101.667	2,054.87	**4.714%**
Philippines	88.667	1,899.94	**4.459%**
Chad	102	2,461.29	**3.979%**
Viet Nam	84.667	2,115.88	**3.848%**
Thailand	93.333	2,617.22	**3.443%**
Azerbaijan	78	2855.55	**2.659%**
Niger	26.333	1,937.64	**1.341%**
Sudan	23	2,237.72	**1.017%**
Albania	20.667	2,924.91	**0.702%**
Hungary	5	2,449.92	**0.204%**
Bangladesh	3.667	2,119.18	**0.173%**
Lithuania	3.667	2,811.39	**0.130%**
Iraq	3	2,582.48	**0.116%**

* All countries in which the FAO conducted surveys

† Average, 2006-08 (FAO estimates)

‡ Data based on FAO surveys conducted in these countries, 1999-2008

Table A2: Summary Statistics of All Variables

	Minimum	Median	Mean	Max	SD
Cell level variables					
$Rebellion_t$	0	0	0.064	1	0.244
$Maize\ area_t$	0	0.035	0.536	24.625	1.548
$Nighttime\ light_t$	0.014	0.034	0.054	1	0.067
$Population_t$	0	4,058	92,201.26	17,726,052	355,013.6
Travel time	0	507	1,185.9	30,033	1,689.56
$Rebellion_{t-1}$	0	0	0.065	1	0.246
$Rebellion\ (spl.)_t$	0	0	0.091	1	0.287
$Temperature_t$	-31.267	10.274	9.858	35.933	13.898
$Drought_t$	0	0	0.295	2.5	0.743
$Drugs_t$	0	0	0.019	1	0.137
Mountains %	0	0	0.224	1	0.352
Border distance	0	310	565.528	10,997	799.979
Cell area	2.56e-06	2,200.146	2,041.419	3,099.948	871.073
$Gems_t$	0	0	0.005	1	0.067
Country level variables					
$Polity2_t$	-10	6	4.454	10	6.140
$New\ state_t$	0	0	0.0003	1	0.018
$Instability_t$	0	0	0.026	1	0.158
Ethnic frac.	0.01	0.333	0.435	0.925	0.258
Religious frac.	0	0.446	0.442	0.783	0.202
$Oil\ production_t$	0	102,000,000	145,960,198	482,766,901	152,508,369
$Gas\ production_t$	0	311.619	1,340.966	5,013.133	1,685.982

Additional Figures

Figure A1: Average Correlations of *Maize area*$_t$ and Constant Measure of Staple Cropland for the Year 2000 by 0.5○ Grid Cell

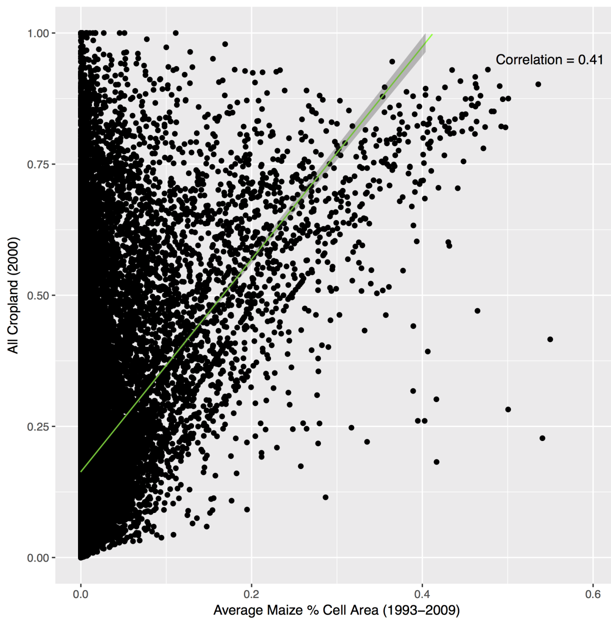


Figure A2: GAM Model Diagnostic – Baseline Specification

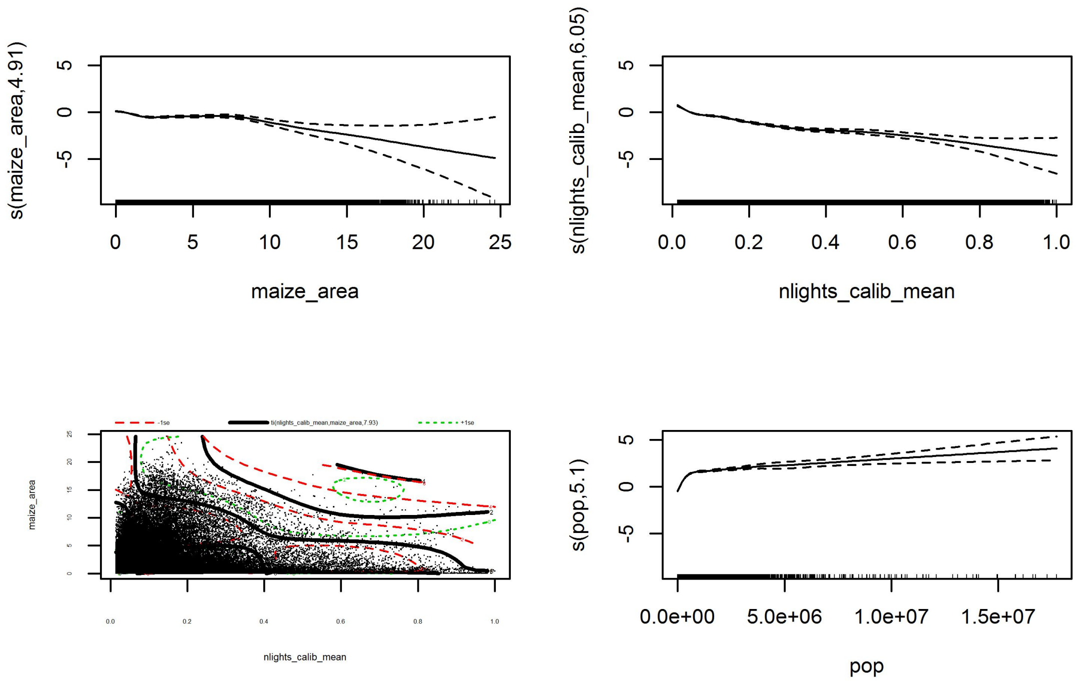
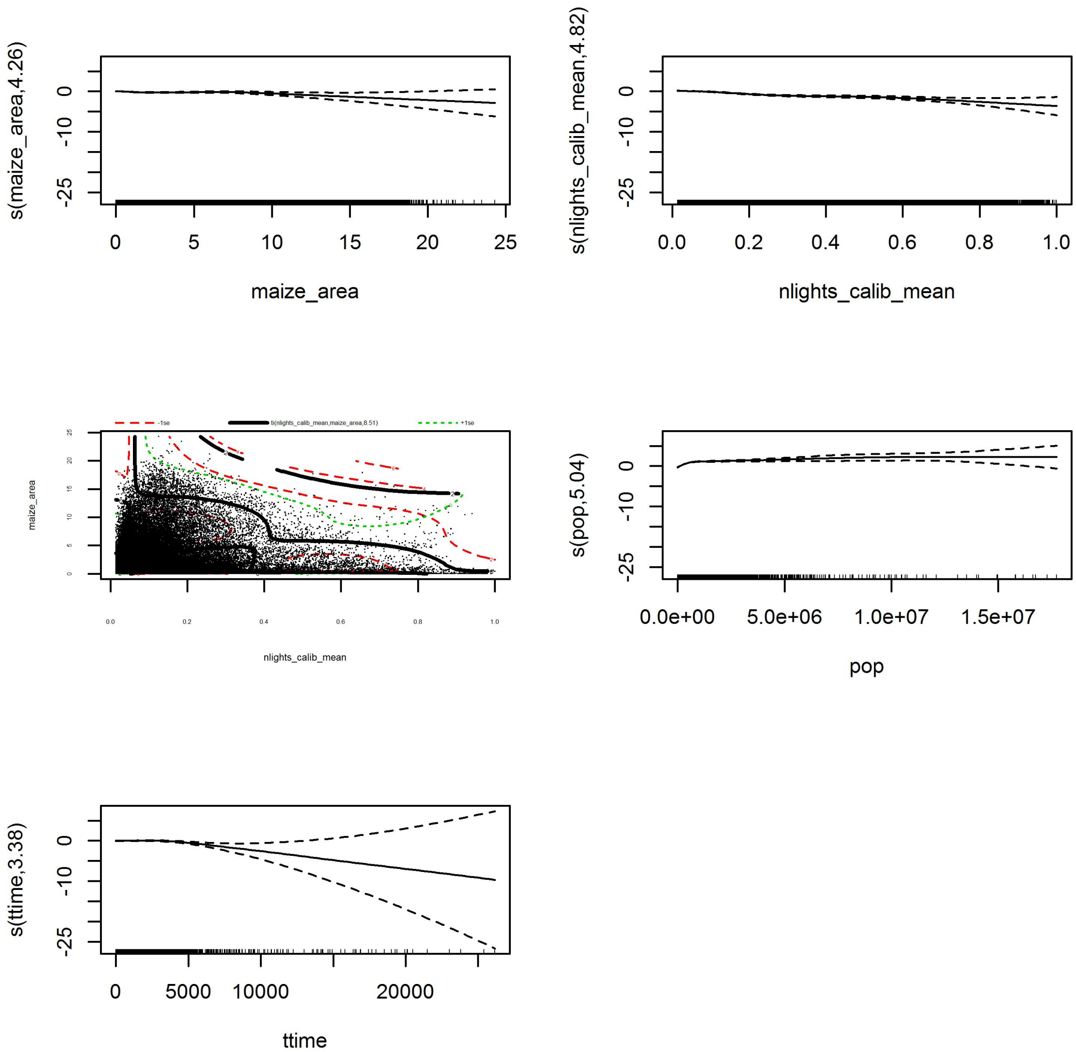


Figure A3: GAM Model Diagnostic – Full Specification



Discussion of the Data Used to Construct *Maize area*$_t$

Data for constructing this continuous *Maize area*$_t$ indicator were obtained from Ray et al. (2012), and measured at the highly localized, ∼0.08∘ grid level, or approximately 9km x 9km at the equator (Ray et al., 2012).[1] First, Ramankutty et al. (2008) created a global cropland map for year 2000. They had two sources of data: (a) Two different global satellite-based land cover data merged together (specifically, BU-MODIS and GLC2000); and (b) National and subnational census data on cropland area. The authors used regression to train the satellite land cover data against the census data, and then map cropland areas at 5 min resolution (0.08 degrees). In a second step, they further adjusted maps (scale up or down all pixels within an administrative unit) to exactly match their census data.

Monfreda, Ramankutty and Foley (2008) then used the cropland map described at (1) as a spatial reference to disaggregate not only area, but also yield data within each administrative unit to the 0.08 degree level for the same year (2000). They used crop yield census data by country, province, or district (depending on availability), although the authors relied on the former, and used the latter (province and district level data) only to ensure that country level data were accurate. As the compliers explain, "We chose to use the subnational data only if the total was between 50% and 200% of the FAOSTAT's national total. Otherwise, we simply used the reported national figures from FAOSTAT" (Monfreda, Ramankutty and Foley, 2008, 9). Most often, sub-national data were used to ensure that the models assigning grown hectares of each crop to each pixel were accurate, but considering the computational challenges and data collection issues with local census data, sub-national data served as a complement rather than a substitute for national level FAO data (see Figure 3 on pg. 10, Monfreda et al. 2008, for a conceptual scheme of this process).

Next, building on Monfreda, Ramankutty and Foley (2008), Ray et al. (2012) collected a large number of crop area and yield data sets from 1961 to today at sub-national and

[1] For detailed information on the sources and methods used to compile these data, see, Monfreda, Ramankutty and Foley (2008, 4-9), Ramankutty et al. (2008, 6-10), and Ray et al. (2012, Supplementary Information, 11-15).

national levels, and used the cropland map by Ramankutty et al. (2008) described above as a spatial reference to disaggregate this area and yield data within each administrative unit. Note that in Ray et al. (2012), the authors only report the administrative level data, not the spatial disaggregation. The data used in this article are the row area data used in Ray et al. (2012). As mentioned in the main article, these maize area data measure the total harvested area within a 0.08 degree cell and are expressed in hectares.[2] The grid of staple crop areas was created "by disaggregating the yield from the smallest political unit with available data in the agricultural inventory by distributing the inventory data for each administrative unit uniformly to each pixel [i.e., 0.08 $\circ$ grid] within that administrative unit" (Monfreda, Ramankutty and Foley, 2008, 10), and repeating this process annually over the entire period (Ray et al., 2012, Supplementary Information, 11-12). The crop area in each 0.08$\circ$ grid of the final map was set to zero when no reference to a crop existed in the inventory data. Information on these missing points was then interpolated from the latest five years if at higher administrative units crops reports were present (Ray et al., 2012, Supplementary Information, 12). Finally, to ensure that these local maize are data correspond to my 0.5 degree grid cell year unit of analysis, I aggregated the 0.08 pixel level data to the same 0.5 $\circ$ annual grid level. This was done by summing the total grown hectares within a given 0.08 $\circ$ pixel for all pixels within each 0.5 $\circ$ grid cell for each year in the data wherever such information was available.

[2]Areas harvested multiple times in a given year are hence measured more than once in the data.

Robustness Models

GLM Robustness

To evaluate the sensitivity of my findings, Table A3 reports six robustness models that account for alternative confounders, specifications, and modeling choices corresponding to the full standard logit model reported in Table 1 of the main article. I begin by controlling for additional geospatial confounders in Model A1, specifically the percentage of a given cell that is covered by mountains, the distance from a given grid cell to the nearest border, and cell area to account for the distance between each cell to the equator. I then add indicators for whether a given cell was denoted as gem producing during year t, as well as additional, potentially salient country-level confounders: whether a given state experienced instability in year t (coded according to the guidelines used by Fearon and Laitin, 2003), ethno-linguistic and religious fractionalization (obtained from Fearon and Laitin, 2003), and the annual levels of oil and gas production within a given state—obtained from Ross (2004)—to account for the role of other profitable natural resources at the country level

Next, recall that the models reported in the main paper did not include robust standard errors, as these might generate inferential biases in such a large sample (King and Roberts, 2015). Model A3 illustrates the robustness of my statistical findings to this issue by re-estimating the full logit model, clustering standard errors by grid cell. Next, it is possible that the effects within grid cells might be influenced by observed and unobserved country-level factors. To illustrate the robustness of my statistical findings to this concern, Model A4 re-stimates the full logit model with the inclusion of fixed effects by country, which isolate only within-country variations across and within grid cells annually. Yet, it is also possible that such country-specific effects can also be influenced by across-state variations. To this end, Model A5 re-estimates the full logit specification with (normally distributed) random rather than fixed effects by country. Finally, note that my sample likely contains a disproportional number of zeros (only 6.38% of the grid cell years in my 1993–2008 sample had experienced rebellion at some point), which might introduce rare events biases into my analysis. To illustrate the robustness of my results to this concern, the final model in Table A3 re-estimates the full logit specification using rare-events logit, which is better designed to handle such issues (King and Zeng, 2001). Crucially, the statistically significant, positive effect of $\textit{Maize area}_t \times \textit{Nighttime light}_t$ holds across every alternative specification, which serves as an additional confirmation to my research hypothesis.

Table A3: Determinants of Rebellions – Robustness Models

	(A1) Geospatial controls	(A2) Extended controls	(A3) CSEs	(A4) CFEs	(A5) CREs	(A6) Rare events
Maize area$_t$	-0.163^{***} (0.013)	-0.129^{***} (0.013)	-0.110^{***} (0.012)	0.034^{**} (0.015)	0.037^{**} (0.015)	-0.118^{***} (0.012)
Nighttime light$_t$	-2.727^{***} (0.212)	-2.721^{***} (0.224)	-3.581^{***} (0.211)	-2.838^{***} (0.266)	-2.800^{***} (0.267)	-3.304^{***} (0.208)
Maize area$_t$ × *Nighttime light*$_t$	0.379^{***} (0.070)	0.292^{***} (0.070)	0.274^{***} (0.071)	0.151^{**} (0.072)	0.139^{*} (0.072)	0.266^{***} (0.070)
Population$_t$[1]	0.303^{***} (0.009)	0.284^{***} (0.009)	0.355^{***} (0.008)	0.196^{***} (0.012)	0.199^{***} (0.012)	0.368^{***} (0.008)
Travel time[1]	0.125^{***} (0.018)	0.112^{***} (0.018)	0.334^{***} (0.017)	0.030 (0.023)	0.038^{*} (0.023)	0.249^{***} (0.016)
Rebellion$_{t-1}$	1.299^{***} (0.027)	1.260^{***} (0.027)	1.288^{***} (0.027)	1.031^{***} (0.028)	1.040^{***} (0.028)	1.282^{***} (0.027)
Rebellion (spl.)$_t$	4.564^{***} (0.031)	4.538^{***} (0.033)	4.619^{***} (0.030)	2.840^{***} (0.033)	2.835^{***} (0.033)	4.638^{***} (0.030)
Temperature$_t$	0.040^{***} (0.002)	0.022^{***} (0.002)	0.049^{***} (0.002)	0.006^{***} (0.002)	0.007^{***} (0.002)	0.032^{***} (0.001)
Drought$_t$	0.111^{***} (0.013)	0.135^{***} (0.013)	0.111^{***} (0.013)	0.125^{***} (0.014)	0.128^{***} (0.014)	0.099^{***} (0.013)
Drugs$_t$	0.333^{***} (0.035)	0.401^{***} (0.035)	0.341^{***} (0.035)	0.186^{***} (0.042)	0.198^{***} (0.042)	0.369 (0.034
Polity2$_t$	0.011^{***} (0.002)	0.005^{***} (0.002)	0.014^{***} (0.002)	-0.038^{***} (0.005)	-0.038^{***} (0.004)	0.011^{***} (0.002)
New state$_t$	−9.442 (49.028)	−8.759 (52.475)	−9.452 (49.659)	−17.021 (2,373.676)	−4.192 (2.809)	334.8^{***} (49.70)
Mountains %	0.558^{***} (0.035)	0.232^{***} (0.037)	–	–	–	–
Border distance[1]	0.035^{***} (0.008)	0.030^{***} (0.008)	–	–	–	–
Cell area[1]	0.729^{***} (0.082)	0.954^{***} (0.086)	–	–	–	–
Gems$_t$	–	−0.155 (0.095)	–	–	–	–
Instability$_t$	–	1.617^{***} (0.044)	–	–	–	–
Ethnic frac.	–	0.714^{***} (0.041)	–	–	–	–
Religious frac.	–	-1.807^{***} (0.066)	–	–	–	–
Oil production$_t$[1]	–	0.042^{***} (0.002)	–	–	–	–
Gas production$_t$[1]	–	-0.087^{***} (0.006)	–	–	–	–
Constant	-14.707^{***} (0.622)	-15.819^{***} (0.652)	-11.846^{***} (0.201)	-8.708^{***} (0.311)	-11.011^{***} (0.683)	-9.873^{***} (0.166)
Observations	440,909	439,981	463,952	463,952	463,952	463,952
Akaike Inf. Crit.	83,307.960	81,006.620	85,827.610	72,482.460	72,812.770	86,165

* indicates $p < .1$; ** indicates $p < .05$; *** indicates $p < .01$.
Variable coefficients are reported with standard errors clustered by grid cell in parentheses. Fixed effects by country and year are included, although not reported here.

[1] Natural log

GAM Robustness

Tables A4–A5 re-estimate the GAMs from the main article using P-spline smoothers instead of thin-plate regression smoothers, with the same $k-1$ degrees of freedom. While thin-plate regression smoothers are generally preferred (Wood, 2003), these P-line smoothers combine a B-spline basis, with a discrete penalty on the basis coefficients (Eilers and Marx, 1996). Although this penalty has no exact interpretation in terms of function shape, these P-splines perform almost as well as conventional (thin-plate regression) splines in many standard applications, and can perform better in particular cases where it is advantageous to mix different orders of basis and penalty.

As Tables A4–A5 illustrate, the statistical effect of both the parameteric and nonparameteric parts of the model remain practically unchanged compared with the GAMs reported in Table 1 in the main article. This is further illustrated by Figured A4–A5, which look nearly identical to Figures A2–A3 reported above. Thus, it appears the the results of the GAMs are robust to smoothing parameter choices, which lends additional support to the findings reported in the main paper.

Table A4: GAM Robustness Models

	Baseline	Full
$Maize\ area_t$	See Table A5	
$Nighttime\ light_t$	See Table A5	
$Maize\ area_t \times Nighttime\ light_t$	See Figure A5	
$Population_t{}^1$	See Table A5	
$Travel\ time^1$	See Table A5	
$Rebellion_{t-1}$	5.223***	1.288
	(0.016)	(0.027)
$Rebellion\ (spl.)_t$	–	4.750***
		(0.030)
$Temperature_t$	–	0.030***
		(0.001)
$Drought_t$	–	0.011***
		(0.035)
$Drugs_t$	–	0.426***
		(0.035)
$Polity2_t$	–	0.0001
		(0.002)
$New\ state_t$	–	-9.059
		(49.55)
Constant	-4.989***	-4.901***
	(0.033)	(0.050)
Observations	588,575	463,952
Akaike Inf. Crit.	152,814.8	86,898.2

* indicates $p < .1$; ** indicates $p < .05$; *** indicates $p < .01$.

Variable coefficients are reported with standard errors in parentheses. Fixed effects by year are included, although not reported here

Table A5: Smoothed Parameter Estimates – GAM Robustness Models

	Baseline	**Full**
Maize area$_t$	6.350***	6.297***
	(414.73)	(81.09)
Nighttime light$_t$	7.736***	7.534***
	(2,201.7)	(373.73)
Maize area$_t$ × *Nighttime light*$_t$	6.673***	6.271***
	(43.26)	(24.08)
Population$_t$	7.139***	6.912***
	(5,222.2)	(1,261.77)
Travel time	–	3.759***
		(16.09)

* indicates $p < .1$; ** indicates $p < .05$; *** indicates $p < .01$.
Estimated degrees of freedom are reported with χ^2 values in parentheses.

Figure A4: GAM Robustness Model Diagnostic – Baseline Specification

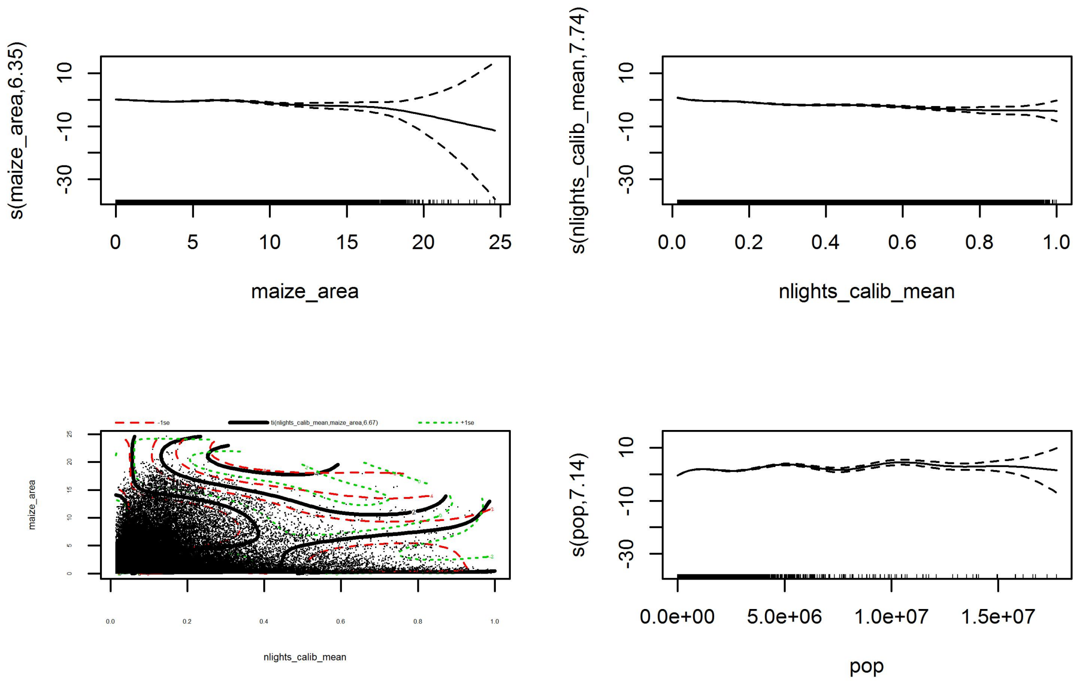
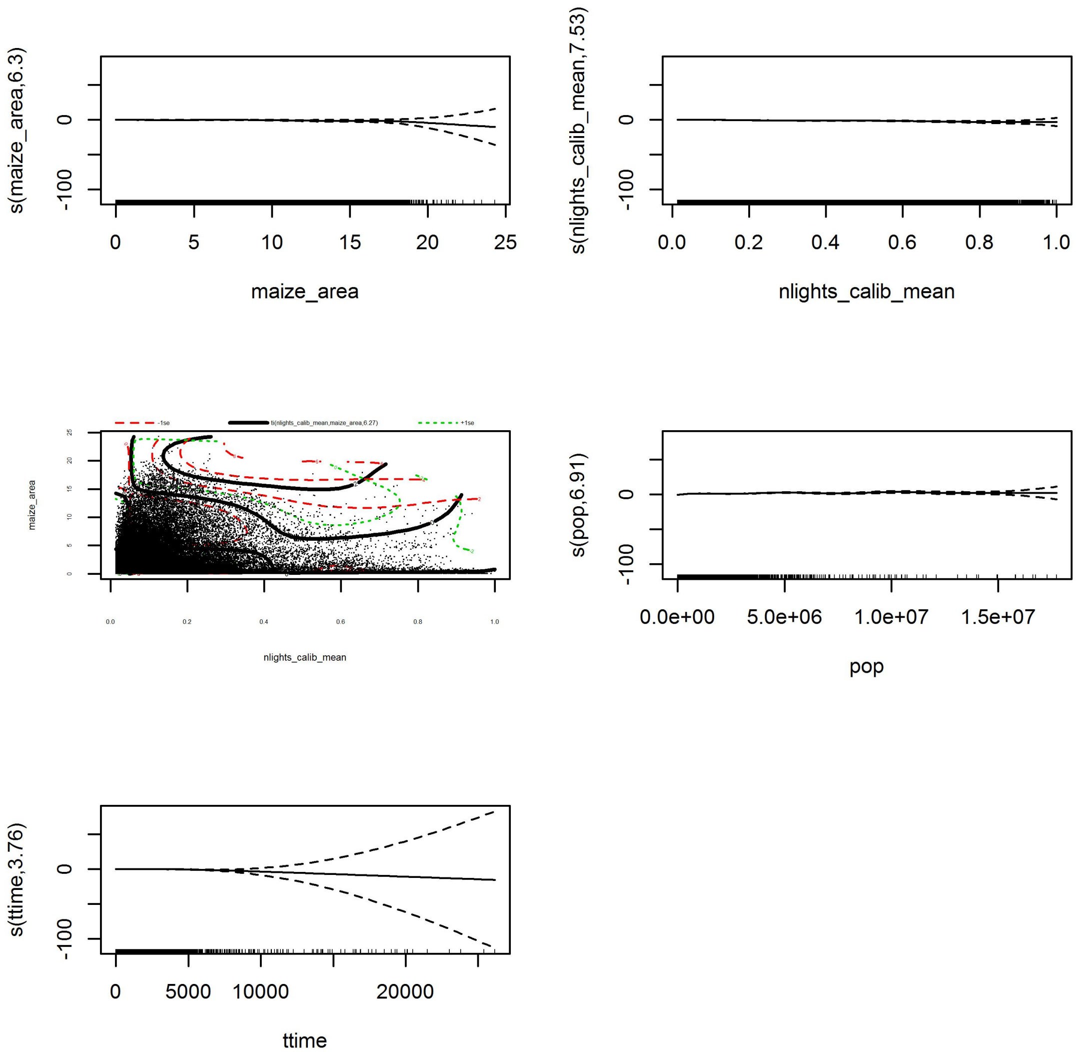


Figure A5: GAM Robustness Model Diagnostic – Full Specification



Discussion of ROCs

This section reports receiver-operator characteristic curves (ROCs), which plot the ratio of how many times the model correctly predicted to how many times the model incorrectly predicted a given event based on some probability threshold.[3] From this perspective, a perfect model will have an area-under-curve (AUC) that equals one, while a random "coin-

[3]Again, bifurcated at 0.5 (Ward, Greenhill and Bakke, 2010).

flip" models will have an AUC of 0.5. Again, to avoid overfitting out-of-sample data are used (Ward, Greenhill and Bakke, 2010). The AUCs for both model (GLM and GAM) are identical, as are their 95% confidence intervals, at 0.990 (0.989 $\Leftrightarrow$ 0.991) suggesting that, at least in this case, GLM and GAM are as effective at forecasting rebellion activity. The upside is that both models have very high AUC values, which makes them both an effective tool for forecasting local rebellion activity at year t.

While the PR curves used in the main paper are preferred for highly class-imbalanced data as the ones used here, these ROCs allow us to statistically test if the addition of the interaction term $Maize\ area_t \times Nighttime\ light_t$ adds to our ability to forecast rebellions. Indeed, as Figures A6 and A6 illustrate, adding $Maize\ area_t$ and $Maize\ area_t \times Nighttime\ light_t$ provides a highly statistically significant, even if substantively modest (which is, again, arguably unsurprising, considering the very large size of the sample), improvement in predictive power: the Z values based on a DeLong, DeLong and Clarke-Pearson (1988) test for nested models were 9.22 for the GLM and for the GAM 9.32.. This provides additional evidence to suggest that indeed, the role of staples and their interactive effect with state power on rebellion incidence is empirically valid.

Figure A6: Comparison of Out-of-Sample ROC Curves (No *Maize area*$_t$, *Maize area*$_t$ $\times$ *Nighttime light*$_t$)

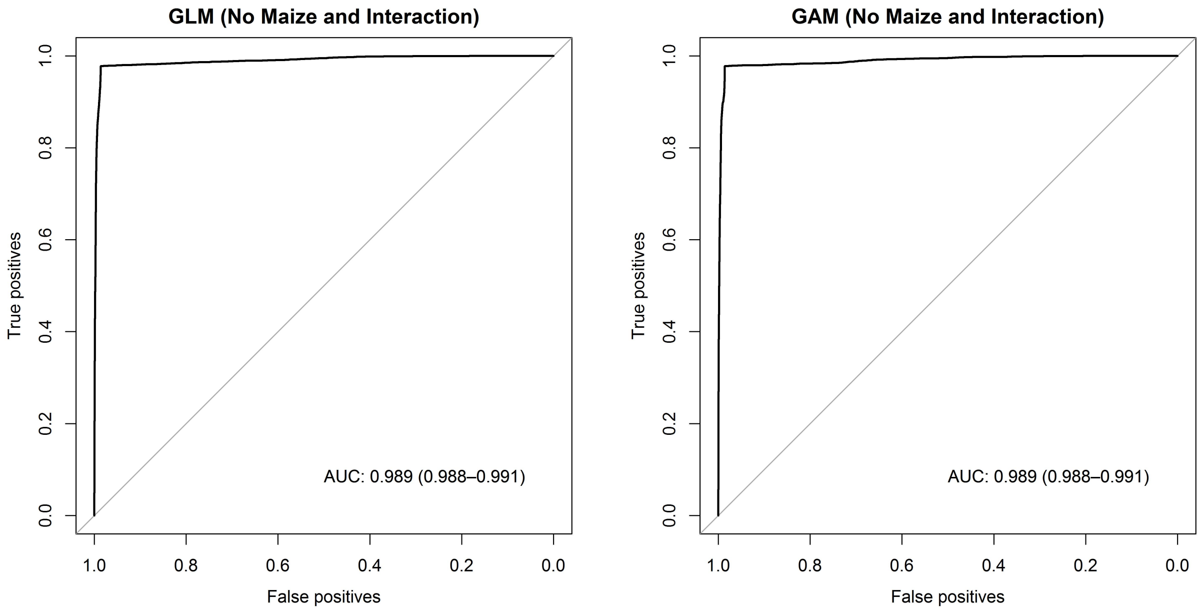
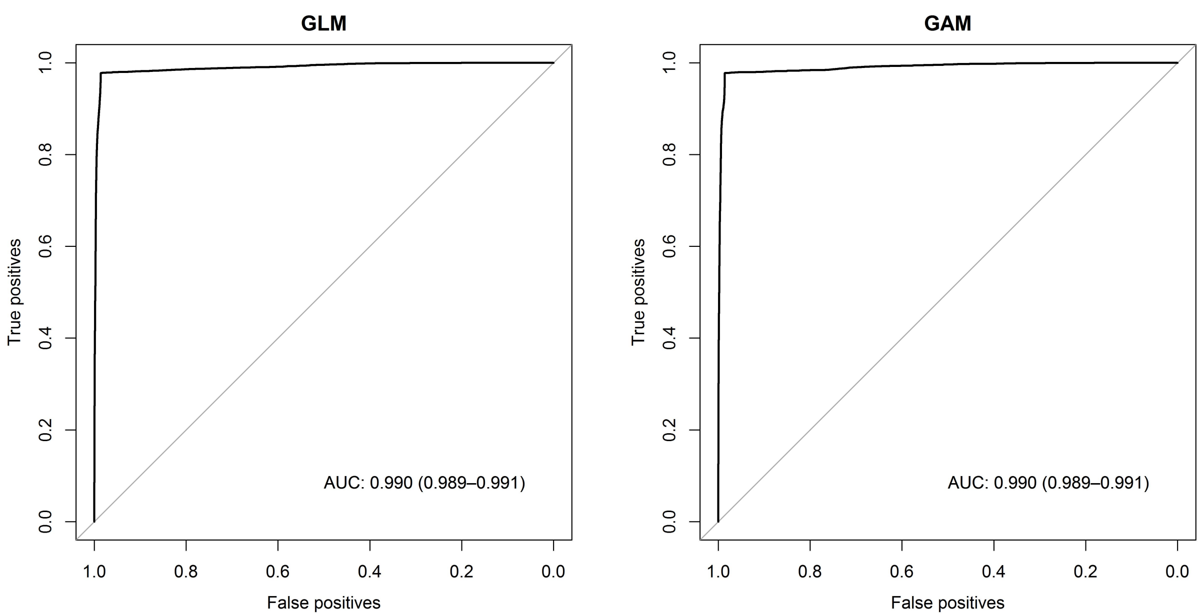


Figure A7: Comparison of Out-of-Sample ROC Curves (Full Models)



Additional Case Study Information

Background of the Mau Mau Rebellion

The Mau Mau rebellion that began in 1952 was one of Britain's most violent decolonization wars (Bennett, 2013). The Colony and Protectorate of Kenya, originally the East Africa

Protectorate, had been a British colony since 1895, although private companies operated in the region since the 1840s. The conflict all but ended in 1956, after the rebel leader Dedan Kimathi was captured and executed, although limited skirmishes and small-scale raids continued to occur until the end of the decade. Casualty estimates for the rebellion, including both combatants and civilians, range from 5,000 to 20,000 deaths (Bennett, 2013, 18–19). It was a brutal rebellion involving severe human rights violations by both British and Mau Mau troops (Branch, 2007). British counterinsurgency operations were extensive and included the fortifications of villages and police posts, large-scale raids and patrols, widespread imprisonment of suspects in transitional camps, and even civilian killings and witch-hunting of Mau Mau oath-takers (Luongo, 2006).

Grievances, especially those related to agriculture and the distribution of land, were a major motivation for the rebellion. During the war, many Kikuyu, Embu and Meru—the major ethnic groups in Kenya—were pushed to join Mau Mau because they were excluded from the means of achieving self-sustenance within the colonial political economy by settler farmers, colonial administrators, and other "land-hungry" patrons (Branch, 2007). Even prior to the rebellion, these aggrieved individuals rejected the leadership of chiefs within their respective groups, men who in practice served as local administrators under the colonial regime, by refusing to obey and sometimes even directly attacking them. The politics of protest also became radicalized, especially after those who served in the British armed forces during the Second World War returned to Kenya. Having served overseas, these veterans rejected the belief, encouraged by the British, that a European is better than an African (Branch, 2007).

The Mau Mau rebellion started as a low-intensity conflict, with sporadic attacks against local chiefs and police stations (Bennett, 2013). With security forces stretched thinly across the region and the loyalist auxiliary Home Guard poorly organized, however, the majority of Kenyans could not be protected from Mau Mau and related violence. As a result, the low-intensity emergency quickly morphed into a full-scale civil war, with the Mau Mau escalating both the number and scale of its attacks, and adding British military forces and European farmers to its list of targets. The Mau Mau initiated loyalty oaths, which many civilians—even if they did not support the rebellion's aims—took as to avoid being labeled as loyalists (Branch, 2007). The Mau Mau, however, did not always resort to terror to obtain support. Its appeal to widespread grievances, especially the desire to expand access to land and food resources, made Mau Mau a popular cause so long that there was no viable alternative. Violence against civilians was used more to gain the begrudging endorsement of waverers and guarantee compliance, or to influence European farmers (Branch, 2007).

The British response was initially unsuccessful. Massive waves of arrests failed to halt the violence, and patrols and raids achieved no tangible results. This led many to criticize the colonial government as having no real strategy to handle the increasingly violent rebellion (Bennett, 2013). Slowly, British military retaliation gained momentum. More local police guards were hired, and arrests became more effective at weakening the Mau Mau. The British also started a system of "screening camps," designed to weed out Mau Mau oath takers and other supporters. At the same time, Mau Mau attacks became more violent. On March 26, 1953, the Mau Mau attacked both Lari village and the Naivsha police station. In Lari, the Mau Mau massacred 120 civilians, while the raid on Naivsha resulted with the release of many prisoners and the loss of weapons and munitions, which heavily embarrassed the government (Bennett, 2013, 17-18). The British responded by heavy arming of the Home Guard and the deployment of the Buffs and Devons brigades, which signaled to beginning of an organized counterinsurgency campaign.

Descriptive Figures from other World Regions

Figures A8–A13 report the distribution of average maize, nighttime light, and rebellion levels (averaged by grid cell) as well as the correlations between each of the former two and the latter (by grid-cell year) for three countries—Burundi, Afghanistan, and Colombia—representing three distinct world regions, Africa, Asia, and Latin America. These figures serve to complement the maps reported in the Mau Mau case analysis using different sources of data (which were also used in analysis). As these maps and plots illustrate, noticeable overlap existed between maize levels and nighttime light levels on the one hand, and the frequency of rebellion activity on the other. This serves as additional, case-specific evidence for the argument developed in the main article.

Figure A8: Average Maize Densities (left), Nighttime Light Levels (center), and Rebellion Frequency (right) – Burundi

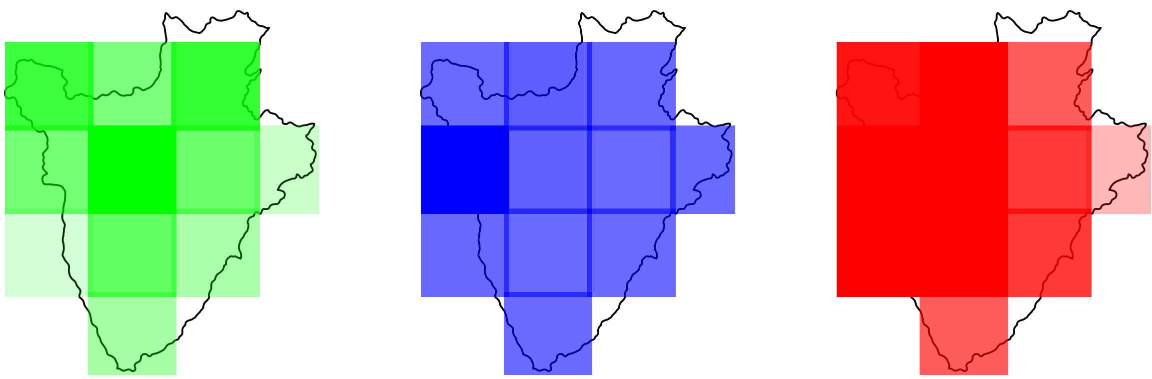
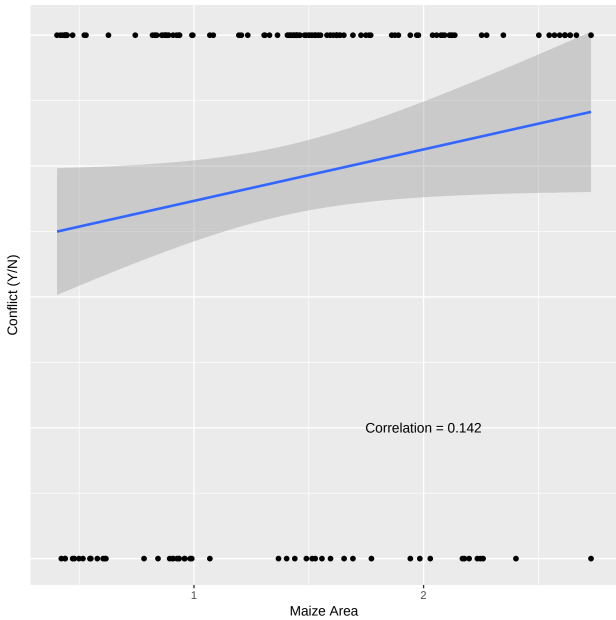
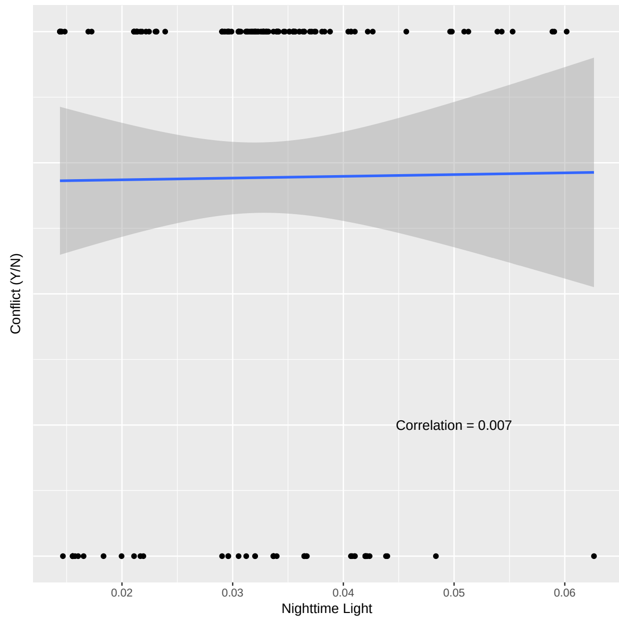


Figure A9: Correlations between *Maize area*$_t$, *Nighttime light*$_t$ and *Rebellion*$_t$ by 0.5○ Grid Cell – Burundi



Maize and Rebellion



Nighttime Light and Rebellion

Figure A10: Average Maize Densities (left), Nighttime Light Levels (center), and Rebellion Frequency (right) – Afghanistan

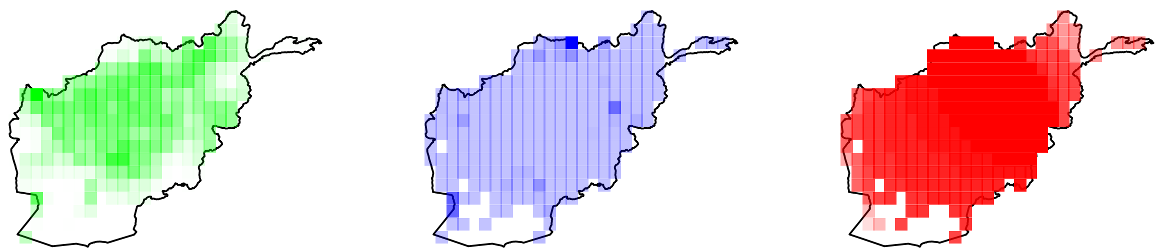
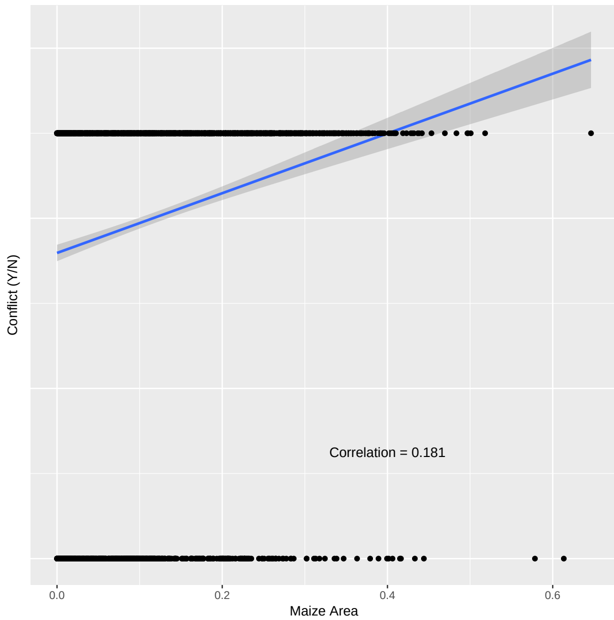
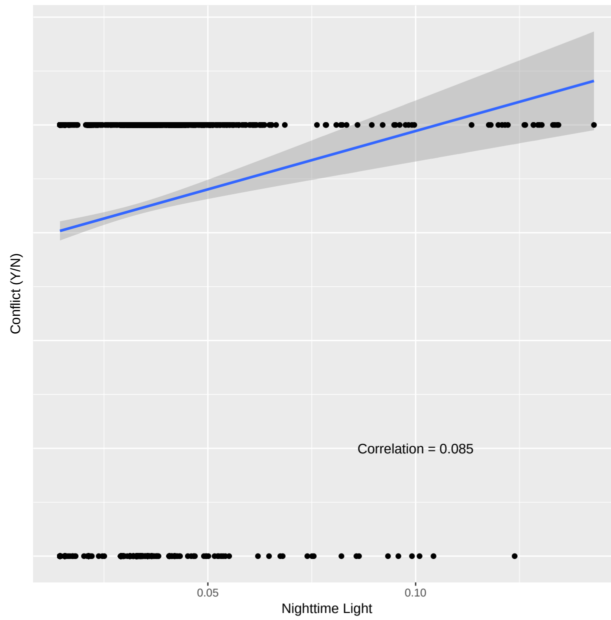


Figure A11: Correlations between *Maize area*$_t$, *Nighttime light*$_t$ and *Rebellion*$_t$ by 0.5○ Grid Cell – Afghanistan



Maize and Rebellion



Nighttime Light and Rebellion

Figure A12: Average Maize Densities (left), Nighttime Light Levels (center), and Rebellion Frequency (right) – Colombia

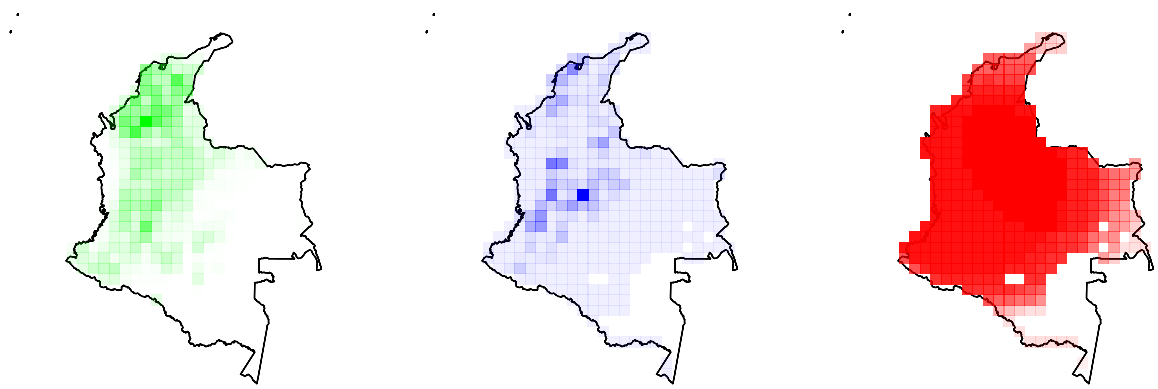
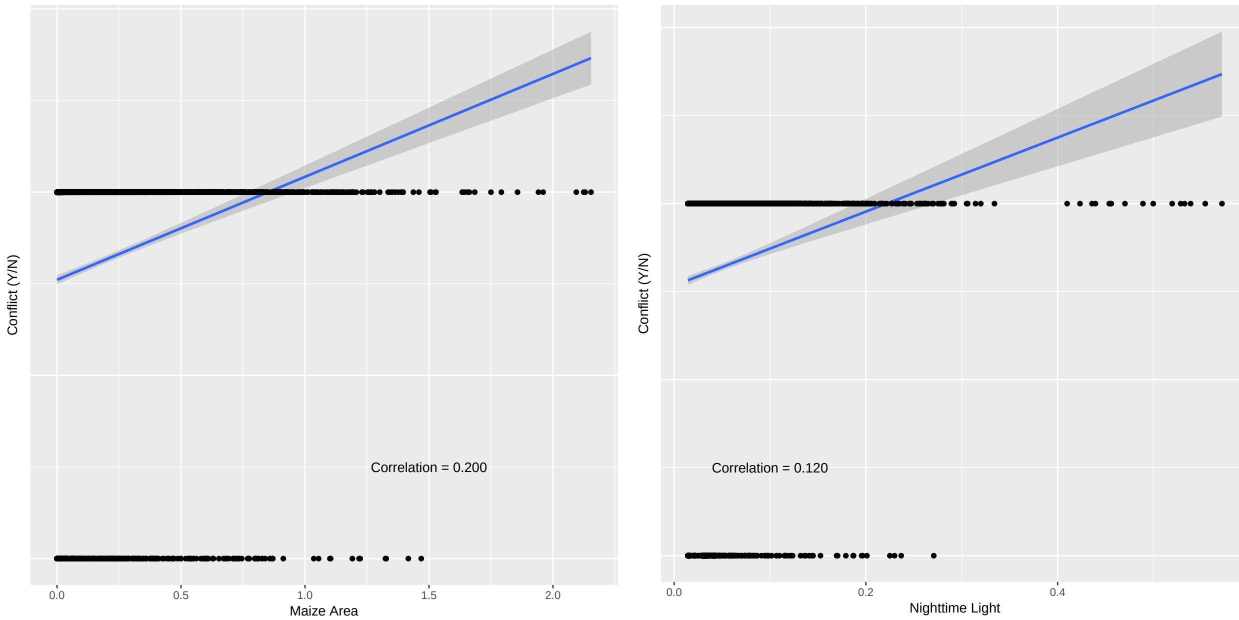


Figure A13: Correlations between *Maize area*$_t$, *Nighttime light*$_t$ and *Rebellion*$_t$ by 0.5○ Grid Cell – Colombia



Maize and Rebellion | Nighttime Light and Rebellion

References

Bennett, Huw. 2013. *Fighting the Mau Mau: The British Army and Counter-Insurgency in the Kenya Emergency.* Cambridge University Press.

Branch, Daniel. 2007. "The enemy within: loyalists and the war against Mau Mau in Kenya." *The Journal of African History* 48(02):291–315.

DeLong, Elizabeth 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* pp. 837–845.

Eilers, Paul HC and Brian D Marx. 1996. "Flexible smoothing with B-splines and penalties." *Statistical science* pp. 89–102.

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

King, Gary and Langche Zeng. 2001. "Logistic regression in rare events data." *Political analysis* 9(2):137–163.

King, Gary and Margaret E Roberts. 2015. "How Robust Standard Errors Expose Methodological Problems They Do Not Fix, and What to Do About It." *Political Analysis* 23:159–179.

Luongo, Katherine. 2006. "If you can't beat them, join them: Government cleansings of witches and Mau Mau in 1950s Kenya." *History in Africa* 33:451–471.

Monfreda, Chad, Navin Ramankutty and Jonathan A Foley. 2008. "Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000." *Global biogeochemical cycles* 22(1).

Ramankutty, N, AT Evan, C Monfreda and JA Foley. 2008. "Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000." *Global Biogeochem. Cycles* 22:Article N. GB1003.

Ray, Deepak K., Navin Ramankutty, Nathaniel D. Mueller, Paul C. West and Jonathan A. Foley. 2012. "Recent patterns of crop yield growth, stagnation, and collapse." *Nature Communications* 3:Article N. 1293.

Ross, Michael L. 2004. "What do we know about natural resources and civil war?" *Journal of peace research* 41(3):337–356.

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.

Wood, Simon N. 2003. "Thin plate regression splines." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65(1):95–114.