

Diverse neighbors and post-conflict recovery at the village level: Evidence from Iraq after ISIL — Online appendix

Lloyd Lyall

June 2021

Contents

1	Regression diagnostics	2
1.1	Diagnostic plots	2
1.2	Table AV. Further outlier sensitivity analysis	3
2	Causal mechanism investigation	5
2.1	Attacks by former ISIL militants	5
2.2	ILA III data on security incidents and θ	6
2.3	ILA III data on militia control and θ	6
2.4	ILA III data on trust and θ	6
2.5	Table AVI. θ scores and recovery at the district level	6
2.6	Table AVII: Driving cleavage investigation	8
3	Robustness	10
3.1	Alternative constructions of θ	10
3.1.1	Table AVIII: Different θ structures	10
3.1.2	Table AIX. Different θ n	12
3.2	Table AX. Alternative constructions of dependent variable	14
3.3	Tables AXI and AXII. Alternative units of analysis (MAUP)	16
3.4	Table AXIII. Specific groups and regions	18
3.5	Table AXIV. Specific population brackets	20
3.6	Table AXV. Alternative causal strategies for spatial dependence	22
3.7	Table AXVI. Alternative ethnicity coding scheme	24
4	ILA III data and ethno-religious plurality coding	26
4.1	Ethno-religious plurality coding	26
4.2	ILA III data other than ethno-religious data	27
5	Territorial control coding	27
6	Additional notes on luminosity data preparation	31
7	References	31

1 Regression diagnostics

1.1 Diagnostic plots

Figure A8 reports scale-location, normal Q-Q, residual-fitted, and Cook's distance plots for the SARAR model.

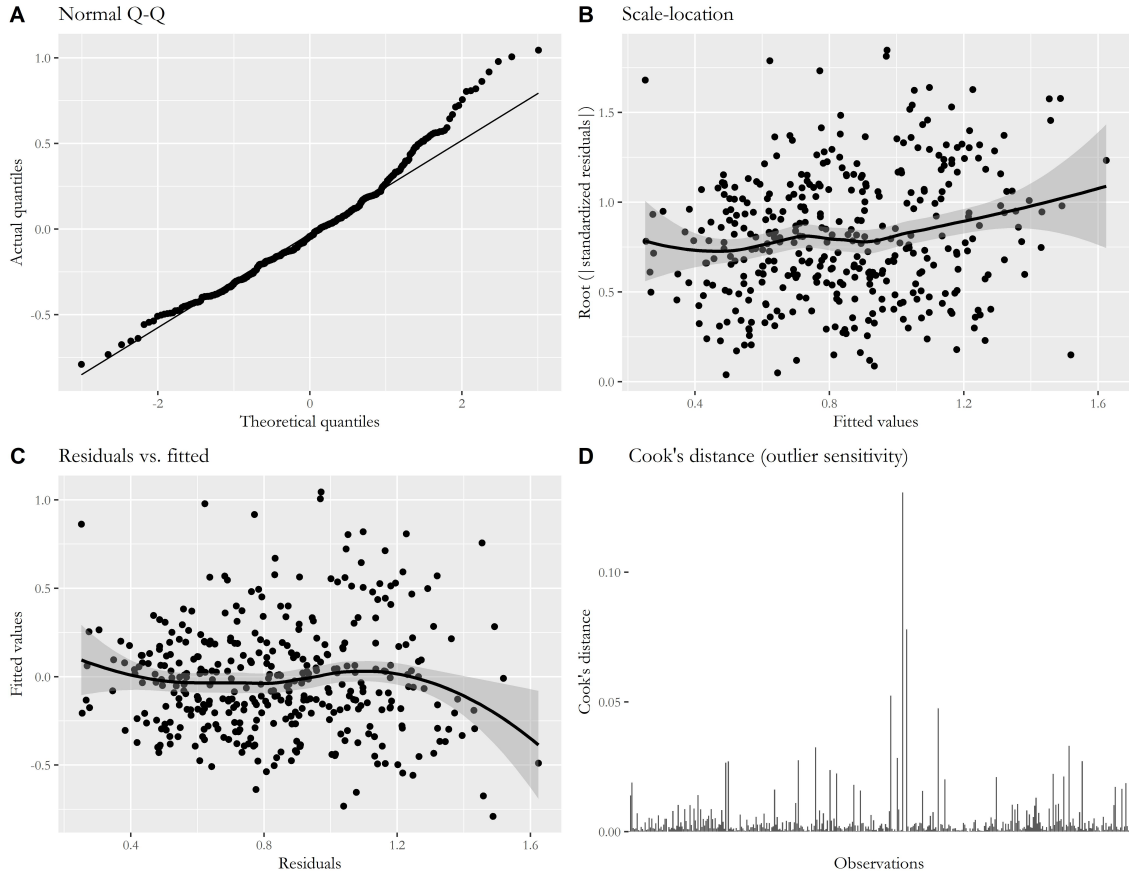


Figure A8. Regression diagnostic plots

The **Cook's distance plot** shows that all points have a Cook's distance (C) less than 0.15, well below the thresholds of 1 (Cook and Weisberg, 1982) and 0.85 (McDonald, 2002) that are commonly suggested as cause for concern. This suggests that the model is not sensitive to outliers. However, it is often advisable to examine the effect of observations with relatively high Cook's distance scores nonetheless; Section 1.2 below therefore performs further outlier sensitivity analysis.

The **scale-location plot** suggests some evidence of heteroskedasticity. The Kelejian and Prucha (2010) heteroskedasticity-robust SARAR model is selected for this reason.

The remaining two plots appear generally healthy: the residuals vs. fitted plot is generally linear, suggesting a linear relationship. The normal-QQ plot suggests the residuals are close to normally distributed.

1.2 Table AV. Further outlier sensitivity analysis

Even though Cook's distances are small for all observations, it is still advisable to investigate points with relatively high Cook's distance scores. The first model in table AV removes the two most influential observations (as can be seen in figure A8 panel D, there are two observations with notably higher Cook's distance scores than the rest). The second model in table AV removes all observations with Cook's distance $> 4/n$, the most aggressive Cook's distance threshold this author is aware of. In each case, after removing the selected observations from the dataset, Model 7 of table IV (herein the 'preferred specification') was re-run on the remaining observations. θ remains highly significant in both case and the effect size is relatively stable.

Table AV. Outlier sensitivity analysis

	2 most influential obs. removed	All obs. with $C > 4/n$ removed
(intercept)	3.563*** (0.440)	3.791*** (0.386)
Neighbor diversity θ	-0.662*** (0.119)	-0.702*** (0.103)
Spatial dependency controls		
λ	0.137 [†] (0.081)	0.128 [†] (0.071)
ρ	0.491*** (0.090)	0.428*** (0.090)
Reconstruction logistics controls		
Occupation length	-0.001 (0.010)	0.009 (0.009)
Date liberated	0.002 (0.010)	-0.010 (0.009)
Dist. to nearest highway	-0.008*** (0.002)	-0.007*** (0.002)
Pre-invasion population/economic controls		
Log(dist. to nearest city)	-0.438*** (0.056)	-0.444*** (0.052)
Log(2012 population)	-0.011 (0.022)	-0.011 (0.019)
Log(2012 population density)	-0.141** (0.047)	-0.106* (0.042)
Log(light per capita before occupation)	-0.133*** (0.019)	-0.106*** (0.016)
Growth before occupation	0.430 (0.338)	0.449 (0.297)
Ethno-religious group dummies		
Arab Shia Muslim	0.039 (0.195)	0.065 (0.158)
Christian	0.016 (0.155)	0.059 (0.172)
Kurd Sunni Muslim	0.174* (0.076)	-0.067 (0.081)
Kurd Shia Muslim	-0.455 (0.281)	-0.434 [†] (0.242)
Kurd Yazidi	-0.037 (0.090)	-0.069 (0.078)
Shabak Shia Muslim	-0.216 [†] (0.117)	-0.177 [†] (0.101)
Shabak Sunni Muslim	-0.149 (0.152)	-0.227 (0.145)
Turkmen Shia Muslim	0.263* (0.130)	0.207 [†] (0.121)
Turkmen Sunni Muslim	-0.204 (0.128)	-0.176 (0.122)
Num. obs.	377.000	348.000
Parameters	20.000	20.000
Log likelihood	-85.691	-19.051
AIC	213.742	80.671

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; [†] $p < 0.1$

2 Causal mechanism investigation

This section supplies some of the empirical results referenced in the discussion section of the text.

2.1 Attacks by former ISIL militants

Figure A9 compares attacks attributable to former ISIL militants from December 2017 to December 2018 with settlement θ scores at the district level. For each district, the median θ score from settlements in that district is used. The attacks plot presents a negative correlation between ISIL attacks and θ scores even with the Rutbah district outlier is removed ($\text{cor} = -0.261$) and with Rutbah, Haweejah and Daquq removed ($\text{cor} = -0.357$).

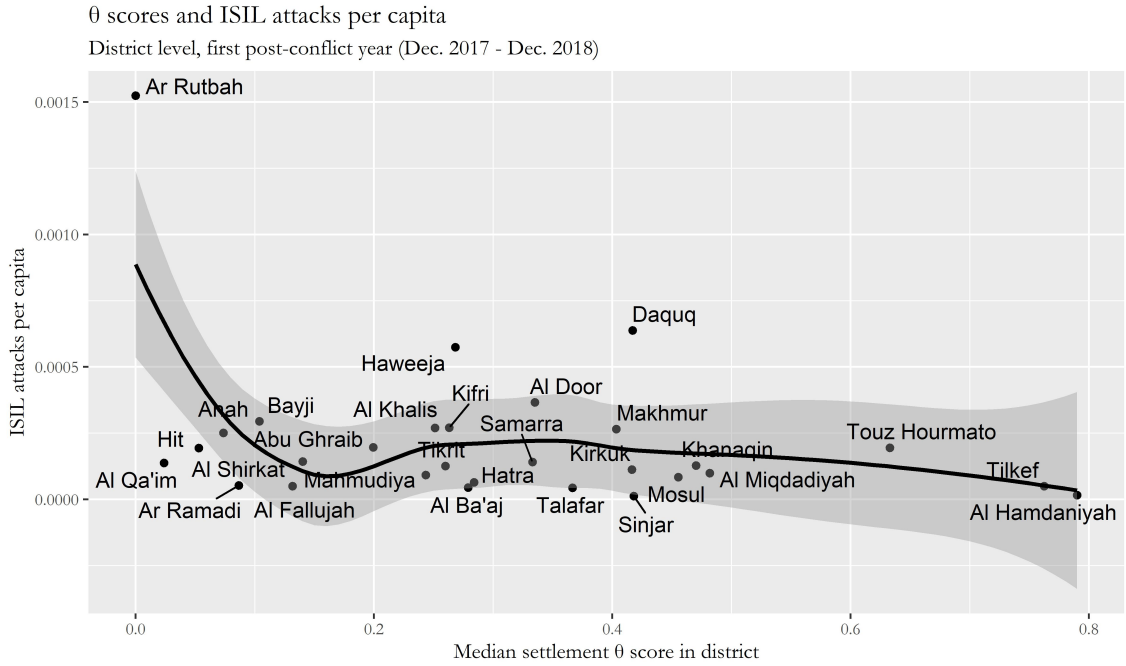


Figure A9. ISIL attacks and θ scores

2.2 ILA III data on security incidents and θ

ILA III question 5.1 asked "In the last three months, did any of the following security incidents occur in this location, (regardless of who the perpetrator was?).

Combining ILA III data on responses to this question with settlement θ scores shows that the θ scores of settlements reporting security incidents of any kind have median = 0.332, mean = 0.376 (N = 129) and the median θ scores of settlements not so reporting have median = 0.366, mean = 0.401 (N = 230). The difference between these groups is not statistically significant (p = 0.317).

2.3 ILA III data on militia control and θ

ILA III question 5.3 asked "who is in charge of controlling this location?"

Combining ILA III responses on this question with settlement θ scores shows that the θ scores of settlements where local militias are present as a controller (the Pershmerga, PMUs, or informal militias) have median = 0.368, mean = 0.422 (N = 224) and the θ scores of settlements where the Iraqi government army or police are present as a controller have median = 0.335, mean = 0.365 (N = 322). The difference between these groups is statistically significant (p = 0.001).

Settlements where militias are in *sole* control have θ scores of median = 0.641, mean = 0.69 (N = 36), and areas where the Iraqi army or police are in sole control have a θ scores of median = 0.275, mean = 0.310 (N = 112). The difference between these groups is also statistically significant (p = 0.001).

Many qualitative sources, such as ICG (2018) and Gaston Derzsi-Horvath (2017) similarly describe the disproportionate presence of sub-state militias in diverse areas.

2.4 ILA III data on trust and θ

ILA III question 6.2 asked "In your subdistrict, are you aware of any of the following... mistrust between different groups".

Settlements reporting some mistrust exists between community members have θ scores of median 0.391, mean 0.440 (N = 70). Settlements reporting no mistrust exists between community members have θ scores of median 0.348, mean 0.394 (N = 265). The difference between these groups is not statistically significant (p=0.08).

2.5 Table AVI. θ scores and recovery at the district level

Table VI presents information on θ scores and recovery at the district level, as well as their corresponding ranks and percentiles. On page 18, this information is compared to district-level rate of return data from p.15 of IOM (2018a).

Table AVI. θ scores and recovery at the district level

	District	Median recovery [DV]	Recovery rank (percentile)	Median theta	Theta rank (percentile)
1	Abu Ghraib	1.790	1 (100%)	0.200	21 (31.03%)
2	Kifri	1.590	2 (96.55%)	0.260	17 (44.83%)
3	Al Khalis	1.430	3 (93.1%)	0.250	19 (37.93%)
4	Kirkuk	1.370	4 (89.66%)	0.420	9 (72.41%)
5	Mahmudiya	1.360	5 (86.21%)	0.240	20 (34.48%)
6	Al Fallujah	1.170	6 (82.76%)	0.130	23 (24.14%)
7	Haweeja	1.160	7 (79.31%)	0.270	16 (48.28%)
8	Al Shirkat	1.130	8 (75.86%)	0.140	22 (27.59%)
9	Bayji	1.060	9 (72.41%)	0.100	24 (20.69%)
10	Al Miqdadiyah	1.020	10 (68.97%)	0.480	4 (89.66%)
11	Makhmur	0.970	11 (65.52%)	0.400	10 (68.97%)
12	Hit	0.960	12 (62.07%)	0.050	27 (10.34%)
13	Khanaqin	0.960	13 (58.62%)	0.470	5 (86.21%)
14	Tikrit	0.830	14 (55.17%)	0.260	18 (41.38%)
15	Ar Ramadi	0.810	15 (51.72%)	0.090	25 (17.24%)
16	Samarra	0.790	16 (48.28%)	0.330	13 (58.62%)
17	Daquq	0.740	17 (44.83%)	0.420	8 (75.86%)
18	Touz Hourmato	0.740	18 (41.38%)	0.630	3 (93.1%)
19	Mosul	0.690	19 (37.93%)	0.460	6 (82.76%)
20	Al Door	0.660	20 (34.48%)	0.330	12 (62.07%)
21	Talafar	0.640	21 (31.03%)	0.370	11 (65.52%)
22	Al Hamdaniyah	0.630	22 (27.59%)	0.790	1 (100%)
23	Anah	0.520	23 (24.14%)	0.070	26 (13.79%)
24	Tilkef	0.470	24 (20.69%)	0.760	2 (96.55%)
25	Al Ba'aj	0.460	25 (17.24%)	0.280	15 (51.72%)
26	Ar Rutbah	0.440	26 (13.79%)	0	29 (3.45%)
27	Hatra	0.420	27 (10.34%)	0.280	14 (55.17%)
28	Sinjar	0.380	28 (6.9%)	0.420	7 (79.31%)
29	Al Qa'im	0.080	29 (3.45%)	0.020	28 (6.9%)

2.6 Table AVII: Driving cleavage investigation

Table AVII seeks to better understand what cleavage is driving the neighbor diversity result. Is this result simply driven by a Sunni-Shia cleavage or a cleavage between groups that fought each other? Or are there penalties to having diverse out-group neighbors even when groups were "on the same side" of the conflict?

Table AVII re-runs the preferred specification on otherwise-identical datasets where θ is calculated with regard to only certain kinds of ethnic, religious, or social cleavages. For example, a θ that considers only the "battle lines" cleavage classifies Sunni Arab and Sunni Turkmen towns as one group, all other towns as another group, and only counts towns of the opposing group as out-group neighbors. Relative to the base θ , this θ will assign higher scores to neighboring towns whose groups were "on opposite sides" of the conflict, and lower scores to neighboring towns of different groups on the same side of the conflict. Similarly, a θ that considers only religious cleavages groups together all groups with the same religion regardless of ethnicity, and thus inflates θ scores for towns with neighbors of a different religion.

The models are named for the type of cleavage the theta construction considers and are described below. The motivation behind Table AVII is that comparisons of θ estimates and model fit statistics between datasets that consider different cleavages in θ construction but are otherwise identical may help illuminate which kinds of social cleavages attract stronger out-group neighbor penalties and result in better-fitting models.

The "**All divisions**" model uses the same θ calculation used in the paper, where all ethno-religious divisions are considered.

The "**battle lines**" model constructs θ where a neighbor is considered "out-group" only if they were on the other 'side' of the ISIL conflict. Sunni Arabs and Sunni Turkmen are one 'side', all other groups are the other 'side'. Of course, the vast majority of Sunni Arabs and Sunni Turkmen did not join ISIL, but ISIL permitted these groups to live in their homes and governed them while it viewed other groups as infidels and attempted to expel and exterminate them. This model considers the purchase of social cleavage.

The "**Sunni/Non-Sunni**" model constructs θ where the only two groups considered are "Sunni" and "Non-Sunni". Accordingly, a neighbor is only considered out-group if on the opposite side of this cleavage (eg. for Non-Sunni settlements, only "Sunni" settlements are out-group neighbors.) "Sunni" includes all ethnic groups that are Sunni regardless of their ethnicity (Sunni Arab, Sunni Kurdish, Sunni Turkmen, and Sunni Shabak)

The "**Religion only**" model constructs θ where the only religious identity is considered. Accordingly, a neighbor is only considered out-group if that neighbor is a different religion (eg. the only divisions considered are Yazidi, Sunni, Shia, and Christian)

Finally, the "**Ethnicity only**" model constructs θ where the only ethnic identity is considered. Accordingly, a neighbor is only considered out-group if that neighbor is a different ethnic group. This means that the only divisions considered are Yazidi, Turkmen, Kurd, Shabak, and Arab. (Note that Yazidis have a unique ethnic and religious identity, so they are considered a separate group in the religion-only and ethnicity-only θ constructions.

Table AVII. Investigation of salient cleavages

	All divisions	Battle lines	Sunni/Non-Sunni	Religious only	Ethnic only
(intercept)	3.485*** (0.452)	3.186*** (0.455)	3.056*** (0.450)	3.007*** (0.456)	3.447*** (0.439)
Neighbor diversity θ	-0.679*** (0.122)	-0.687*** (0.171)	-0.670*** (0.176)	-0.510*** (0.152)	-0.723*** (0.118)
Spatial dependency controls					
λ	0.131 (0.084)	0.147 (0.090)	0.123 (0.091)	0.121 (0.093)	0.140 [†] (0.081)
ρ	0.499*** (0.094)	0.536*** (0.087)	0.546*** (0.081)	0.561*** (0.097)	0.479*** (0.087)
Reconstruction logistics controls					
Occupation length	-0.001 (0.010)	-0.003 (0.011)	-0.003 (0.011)	-0.005 (0.011)	-0.003 (0.010)
Date liberated	0.002 (0.011)	0.003 (0.011)	0.004 (0.011)	0.006 (0.011)	0.004 (0.010)
Dist. to nearest highway	-0.008*** (0.002)	-0.007** (0.002)	-0.007** (0.002)	-0.007** (0.002)	-0.008*** (0.002)
Pre-invasion population/economic controls					
Log(dist. to nearest city)	-0.426*** (0.058)	-0.375*** (0.058)	-0.360*** (0.058)	-0.371*** (0.059)	-0.423*** (0.056)
Log(2012 population)	-0.009 (0.022)	-0.009 (0.022)	-0.007 (0.022)	-0.006 (0.023)	-0.014 (0.022)
Log(2012 population density)	-0.143** (0.048)	-0.143** (0.048)	-0.144** (0.048)	-0.145** (0.049)	-0.137** (0.047)
Log(light per capita before occupation)	-0.135*** (0.019)	-0.138*** (0.019)	-0.135*** (0.019)	-0.135*** (0.020)	-0.138*** (0.019)
Growth before occupation	0.391 (0.342)	0.364 (0.346)	0.330 (0.347)	0.452 (0.348)	0.387 (0.340)
Ethno-religious group dummies					
Arab Shia Muslim	0.042 (0.201)	0.023 (0.210)	0.027 (0.211)	0.048 (0.213)	0.050 (0.197)
Christian	0.028 (0.159)	-0.102 (0.160)	-0.006 (0.166)	-0.005 (0.169)	-0.055 (0.154)
Kurd Sunni Muslim	0.178* (0.079)	0.257** (0.083)	0.175* (0.081)	0.209* (0.082)	0.191* (0.078)
Kurd Shia Muslim	-0.462 (0.288)	-0.326 (0.291)	-0.304 (0.292)	-0.385 (0.292)	-0.457 (0.286)
Kurd Yazidi	-0.048 (0.093)	-0.104 (0.094)	-0.072 (0.096)	-0.113 (0.096)	-0.035 (0.091)
Shabak Shia Muslim	-0.184 (0.119)	-0.274* (0.121)	-0.197 (0.127)	-0.223 [†] (0.127)	-0.188 (0.117)
Shabak Sunni Muslim	-0.134 (0.155)	-0.242 (0.157)	-0.108 (0.166)	-0.171 (0.163)	-0.136 (0.153)
Turkmen Shia Muslim	0.100 (0.121)	0.209 (0.128)	0.238 [†] (0.131)	0.073 (0.123)	0.108 (0.120)
Turkmen Sunni Muslim	-0.008 (0.120)	-0.096 (0.120)	-0.149 (0.120)	-0.111 (0.121)	0.017 (0.120)
Num. obs.	379.000	379.000	379.000	379.000	379.000
Parameters	20.000	20.000	20.000	20.000	20.000
Log likelihood	-95.209	-110.399	-117.246	-109.930	-93.758
AIC	232.764	263.145	276.838	262.206	229.862

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; [†] $p < 0.1$.

3 Robustness

This section reports the results of robustness checks referenced in the paper.

3.1 Alternative constructions of θ

Tables AVIII and AIX re-estimate the preferred specification on datasets that use alternative constructions of the neighbor diversity score θ , to investigate whether the results are sensitive to the construction of θ . The results show the neighbor diversity score θ is robust to different θ construction strategies.

3.1.1 Table AVIII: Different θ structures

Table AVIII investigates several different constructions of θ . The models are named for their θ construction, and each θ construction is described below.

”Inverse log with ERF” is the θ score used in the paper.

$$\text{Inverse log w/ ERF} : \theta_k = \left[\sum_{n=1}^5 \frac{1}{\log(\text{distance}_{k,n} + 1)} \right] \times \left[2 - \sum_{n=1}^5 p_n^2 \right]$$

”Inverse with ERF” is θ score used in the paper, except that inverse distances are not logged.

$$\text{Inverse w/ ERF} : \theta_k = \left[\sum_{n=1}^5 \frac{1}{\text{distance}_{k,n}} \right] \times \left[2 - \sum_{n=1}^5 p_n^2 \right]$$

”Inverse log” is a θ score built using only the left-hand expression of the theta score used in the paper (it considers only the distance to out-group neighbors $n_1 \dots n_5$, but not the heterogeneity among $n_1 \dots n_5$.)

$$\text{Inverse log} : \theta_k = \sum_{n=1}^5 \frac{1}{\log(\text{distance}_{k,n} + 1)}$$

”Inverse” is a θ score akin to ”Inverse log” but does not log distances.

$$\text{Inverse} : \theta_k = \sum_{n=1}^5 \frac{1}{\text{distance}_{k,n}}$$

Table AVIII. Alternative constructions of theta

	Inverse log w/ ERF	Inverse w/ ERF	Inverse log	Inverse
(intercept)	3.485*** (0.452)	3.213*** (0.434)	3.278*** (0.445)	3.025*** (0.437)
Neighbor diversity θ	-0.679*** (0.122)	-0.753*** (0.140)	-0.723*** (0.154)	-0.682*** (0.159)
Spatial dependency controls				
λ	0.131 (0.084)	0.153 [†] (0.084)	0.161 [†] (0.085)	0.153 [†] (0.087)
ρ	0.499*** (0.094)	0.488*** (0.086)	0.492*** (0.083)	0.502*** (0.087)
Reconstruction logistics controls				
Occupation length	-0.001 (0.010)	-0.003 (0.010)	-0.001 (0.010)	-0.003 (0.011)
Date liberated	0.002 (0.011)	0.003 (0.011)	0.002 (0.011)	0.004 (0.011)
Dist. to nearest highway	-0.008*** (0.002)	-0.008*** (0.002)	-0.007** (0.002)	-0.008*** (0.002)
Log(dist. to nearest city)	-0.426*** (0.058)	-0.390*** (0.056)	-0.384*** (0.056)	-0.362*** (0.057)
Pre-invasion population/economic controls				
Log(2012 population)	-0.009 (0.022)	-0.007 (0.022)	-0.007 (0.022)	-0.007 (0.023)
Log(2012 population density)	-0.143** (0.048)	-0.145** (0.048)	-0.146** (0.048)	-0.144** (0.048)
Log(light per capita before occupation)	-0.135*** (0.019)	-0.135*** (0.019)	-0.139*** (0.019)	-0.133*** (0.019)
Growth before occupation	0.391 (0.342)	0.416 (0.343)	0.396 (0.346)	0.427 (0.347)
Ethno-religious group dummies				
Arab Shia Muslim	0.042 (0.201)	0.030 (0.201)	0.031 (0.203)	0.030 (0.205)
Christian	0.028 (0.159)	0.121 (0.165)	0.051 (0.165)	0.043 (0.167)
Kurd Sunni Muslim	0.178* (0.079)	0.226** (0.079)	0.261** (0.082)	0.250** (0.082)
Kurd Shia Muslim	-0.462 (0.288)	-0.403 (0.288)	-0.338 (0.291)	-0.356 (0.292)
Kurd Yazidi	-0.048 (0.093)	-0.065 (0.092)	-0.077 (0.093)	-0.100 (0.093)
Shabak Shia Muslim	-0.184 (0.119)	-0.122 (0.124)	-0.187 (0.123)	-0.201 (0.124)
Shabak Sunni Muslim	-0.134 (0.155)	-0.046 (0.160)	-0.116 (0.160)	-0.119 (0.162)
Turkmen Shia Muslim	0.100 (0.121)	0.180 (0.123)	0.225 [†] (0.128)	0.215 [†] (0.128)
Turkmen Sunni Muslim	-0.008 (0.120)	0.040 (0.123)	0.003 (0.123)	0.009 (0.124)
Num. obs.	379.000	379.000	379.000	379.000
Parameters	20.000	20.000	20.000	20.000
Log likelihood	-95.209	-98.933	-104.763	-106.623
AIC	232.764	240.213	251.871	255.593

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; [†] $p < 0.1$.

3.1.2 Table AIX. Different θ n

Table AIX uses the "Inverse w/ ERF" θ structure used in the paper, but varies the number of settlements n which θ considers as the set of out-group neighbors. The models are again named for their θ construction. In model $n=3$ θ considers only the three nearest out-group neighbors to target settlement k :

$$Model\ n = 3 : \theta_k = \left[\sum_{n=1}^3 \frac{1}{\log(distance_{k,n} + 1)} \right] \times \left[2 - \sum_{n=1}^5 p_n^2 \right]$$

Similarly, model $n=4$ considers the nearest 4 out-group neighbors, $n=5$ the nearest five neighbors, and so on.

The results show that the the neighbor diversity score finding is robust to different values for n , the number of neighbors θ considers.

Table AIX. Alternative choices for n in the theta calculation

	n=3	n=4	n=5	n=6	n=7	n=8
(intercept)	3.404*** (0.459)	3.553*** (0.458)	3.485*** (0.452)	3.508*** (0.456)	3.453*** (0.455)	3.374*** (0.460)
Neighbor diversity θ	-0.596*** (0.125)	-0.732*** (0.132)	-0.679*** (0.122)	-0.722*** (0.132)	-0.668*** (0.126)	-0.616*** (0.130)
Spatial dependency controls						
λ	0.122 (0.087)	0.116 (0.085)	0.131 (0.084)	0.132 (0.085)	0.131 (0.086)	0.131 (0.088)
ρ	0.512*** (0.095)	0.506*** (0.094)	0.499*** (0.094)	0.511*** (0.098)	0.520*** (0.099)	0.533*** (0.100)
Reconstruction logistics controls						
Occupation length	-0.001 (0.011)	-0.001 (0.010)	-0.001 (0.010)	-0.000 (0.010)	-0.001 (0.011)	-0.001 (0.011)
Date liberated	0.002 (0.011)	0.002 (0.011)	0.002 (0.011)	0.001 (0.011)	0.002 (0.011)	0.002 (0.011)
Dist. to nearest highway	-0.009*** (0.002)	-0.008*** (0.002)	-0.008*** (0.002)	-0.008*** (0.002)	-0.008*** (0.002)	-0.008*** (0.002)
Log(dist. to nearest city)	-0.413*** (0.058)	-0.435*** (0.059)	-0.426*** (0.058)	-0.423*** (0.058)	-0.416*** (0.058)	-0.404*** (0.059)
Pre-invasion population/economic controls						
Log(2012 population)	-0.006 (0.022)	-0.009 (0.022)	-0.009 (0.022)	-0.010 (0.022)	-0.011 (0.022)	-0.013 (0.022)
Log(2012 population density)	-0.150** (0.048)	-0.144** (0.048)	-0.143** (0.048)	-0.140** (0.048)	-0.140** (0.048)	-0.138** (0.048)
Log(light per capita before occupation)	-0.136*** (0.019)	-0.135*** (0.019)	-0.135*** (0.019)	-0.135*** (0.019)	-0.137*** (0.019)	-0.136*** (0.019)
Growth before occupation	0.433 (0.345)	0.407 (0.342)	0.391 (0.342)	0.391 (0.342)	0.369 (0.342)	0.412 (0.344)
Ethno-religious group dummies						
Arab Shia Muslim	0.045 (0.205)	0.045 (0.202)	0.042 (0.201)	0.054 (0.203)	0.056 (0.204)	0.052 (0.207)
Christian	-0.059 (0.159)	-0.001 (0.158)	0.028 (0.159)	0.028 (0.160)	0.042 (0.162)	0.028 (0.164)
Kurd Sunni Muslim	0.173* (0.080)	0.165* (0.079)	0.178* (0.079)	0.176* (0.079)	0.195* (0.079)	0.191* (0.080)
Kurd Shia Muslim	-0.464 (0.290)	-0.468 (0.287)	-0.462 (0.288)	-0.349 (0.287)	-0.379 (0.287)	-0.331 (0.289)
Kurd Yazidi	-0.073 (0.094)	-0.049 (0.093)	-0.048 (0.093)	-0.052 (0.093)	-0.070 (0.093)	-0.078 (0.094)
Shabak Shia Muslim	-0.279* (0.118)	-0.201 [†] (0.118)	-0.184 (0.119)	-0.160 (0.121)	-0.140 (0.123)	-0.150 (0.126)
Shabak Sunni Muslim	-0.176 (0.156)	-0.120 (0.156)	-0.134 (0.155)	-0.150 (0.155)	-0.122 (0.157)	-0.154 (0.158)
Turkmen Shia Muslim	0.072 (0.122)	0.061 (0.121)	0.100 (0.121)	0.091 (0.121)	0.106 (0.121)	0.095 (0.122)
Turkmen Sunni Muslim	-0.098 (0.120)	-0.074 (0.119)	-0.008 (0.120)	-0.001 (0.120)	-0.006 (0.121)	-0.032 (0.121)
Num. obs.	379.000	379.000	379.000	379.000	379.000	379.000
Parameters	20.000	20.000	20.000	20.000	20.000	20.000
Log likelihood	-100.284	-95.774	-95.209	-95.164	-96.229	-100.295
AIC	242.915	233.893	232.764	232.674	234.805	242.937

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; [†] $p < 0.1$.

3.2 Table AX. Alternative constructions of dependent variable

Table AX tests the robustness of the out-group neighbors result to different conceptions of the dependent variable, the fraction of pre-invasion light recovered. It re-estimates the preferred specification on datasets that use the alternative constructions of the dependent variable described below. Table AX shows the neighbor diversity result is robust to alternate dependent variables that measure the fraction of pre-invasion light recovered for each settlement relative to that settlement's liberation date, although the effect size is smaller. The neighbor diversity result is also robust to alternate fixed endpoint dates for all settlements.

The first four models re-run the preferred specification on datasets where the dependent variable is calculated with reference to the time passed *relative to each settlement's liberation date*. The "3 months" specification calculates the log fraction of pre-invasion light recovered by each settlement three months after *that settlement's* liberation. Formally, for settlement k , the dependent variable DV takes on:

$$DV_k = \log \left[\frac{\text{Light level}_k \text{ at date of } k's \text{ liberation} + 3 \text{ months}}{\text{Light level}_k \text{ in November 2013}} + 1 \right]$$

The "6 months" model does the same for 6 months after each settlement's liberation. "12 months" and "24 months" do the same for 12 and 24 months. Importantly, note that the "24 months" model contains a smaller number of observations, because not all settlements had been liberated for 24 months when data was collected. Only settlements liberated for at least 24 months are included in this specification.

The final two models re-run the preferred specification on datasets where the dependent variable is calculated with reference to the same date for all settlements (as in the paper), but vary this date. "March 2018" calculates the log fraction of pre-invasion light recovered by March 2018 (3 months after the last Iraqi settlement was liberated). Formally:

$$DV_k = \log \left[\frac{\text{Light level}_k \text{ in March 2018}}{\text{Light level}_k \text{ in November 2013}} + 1 \right]$$

"June 2018" does the same for June 2018 (6 months after the last Iraqi settlement was liberated).

Table AX. Alternative constructions of the dependent variable

	3 months	6 months	12 months	24 months	March 2018	June 2018
(intercept)	1.034** (0.323)	1.329*** (0.400)	1.622*** (0.403)	0.985 [†] (0.577)	2.994*** (0.453)	3.103*** (0.465)
Neighbor diversity θ	-0.214** (0.082)	-0.301** (0.106)	-0.367*** (0.108)	-0.555*** (0.161)	-0.641*** (0.123)	-0.595*** (0.126)
Spatial dependency controls						
λ	0.463*** (0.061)	0.466*** (0.080)	0.267** (0.088)	0.116 (0.143)	0.361*** (0.080)	0.324*** (0.083)
ρ	0.120 (0.099)	0.333*** (0.092)	0.417*** (0.094)	0.593*** (0.132)	0.465*** (0.071)	0.493*** (0.068)
Reconstruction logistics controls						
Occupation length	0.005 (0.007)	0.006 (0.009)	0.003 (0.009)	-0.016 (0.013)	-0.003 (0.010)	-0.004 (0.011)
Date liberated	0.001 (0.007)	0.005 (0.009)	0.011 (0.009)	0.042** (0.014)	0.004 (0.011)	0.005 (0.011)
Dist. to nearest highway	0.001 (0.002)	-0.003 (0.002)	-0.006** (0.002)	-0.003 (0.003)	-0.008*** (0.002)	-0.008*** (0.002)
Log(Dist. to nearest big city)	-0.120** (0.038)	-0.219*** (0.051)	-0.311*** (0.052)	-0.274*** (0.077)	-0.367*** (0.058)	-0.392*** (0.060)
Pre-invasion population/economic controls						
Log(2012 population)	-0.007 (0.020)	-0.001 (0.022)	-0.008 (0.021)	0.005 (0.029)	0.012 (0.023)	0.013 (0.023)
Log(2012 population density)	-0.082* (0.041)	-0.091 [†] (0.046)	-0.060 (0.045)	-0.128* (0.063)	-0.170*** (0.049)	-0.168*** (0.050)
Log(light per capita before occupation)	-0.079*** (0.015)	-0.088*** (0.018)	-0.084*** (0.018)	-0.141*** (0.026)	-0.142*** (0.020)	-0.146*** (0.020)
Growth before occupation	-0.110 (0.295)	0.135 (0.336)	-0.018 (0.322)	-0.247 (0.495)	0.465 (0.355)	0.370 (0.357)
Ethno-religious group dummies						
Arab Shia Muslim	-0.009 (0.125)	-0.059 (0.169)	0.004 (0.176)	-0.092 (0.192)	-0.084 (0.202)	-0.033 (0.208)
Christian	-0.050 (0.126)	-0.088 (0.149)	-0.028 (0.147)	-0.156 (0.181)	-0.028 (0.163)	-0.046 (0.165)
Kurd Sunni Muslim	0.180** (0.058)	0.208** (0.072)	0.232** (0.071)	0.191* (0.081)	0.000 (0.080)	0.076 (0.082)
Kurd Shia Muslim	0.091 (0.256)	0.056 (0.284)	0.002 (0.272)	-0.196 (0.276)	-0.268 (0.298)	-0.355 (0.300)
Kurd Yazidi	-0.101 (0.069)	-0.100 (0.084)	-0.108 (0.084)	-0.225 [†] (0.121)	-0.046 (0.094)	-0.051 (0.096)
Shabak Shia Muslim	-0.063 (0.091)	0.063 (0.109)	-0.056 (0.108)	-0.234 (0.235)	0.183 (0.121)	0.138 (0.123)
Shabak Sunni Muslim	0.005 (0.125)	0.047 (0.146)	-0.019 (0.143)	-0.282 (0.236)	0.067 (0.159)	0.175 (0.161)
Turkmen Shia Muslim	0.123 (0.107)	0.176 (0.120)	0.392*** (0.114)	0.302 (0.272)	0.277* (0.125)	0.209 [†] (0.126)
Turkmen Sunni Muslim	0.006 (0.104)	0.025 (0.118)	-0.078 (0.113)	-0.028 (0.128)	-0.227 [†] (0.124)	-0.075 (0.125)
Num. obs.	379.000	379.000	379.000	213.000	379.000	379.000
Parameters	20.000	20.000	20.000	20.000	20.000	20.000
Log likelihood	-18.838	-73.519	-66.420	-52.582	-114.100	-124.495
AIC	80.022	189.384	175.187	149.540	270.547	291.337

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; [†] $p < 0.1$.

3.3 Tables AXI and AXII. Alternative units of analysis (MAUP)

This section evaluates the robustness of the preferred specification to the modifiable areal unit problem (MAUP). MAUP problems arise if regression results are not consistent when the unit of analysis is defined at different levels of aggregation.

In each specification in tables AXI and AXII, a different population and density threshold was used to define settlements, the dataset was then re-constructed using that unit of analysis, and the preferred specification was then run on that dataset. Of course, different units of analysis result in a different number of settlements in the dataset. All other inclusion and exclusion criteria other than settlement population and density thresholds remained the same in each case. Tables AXI and AXII show that the diverse neighbors result is robust to different units of analysis, and its effect size relatively consistent across them.

Table AXI reports summarized regression results from 13 possible units of analysis. **Table AXII** reports full regression results for 5 representative models selected from table AXI. The models are named for the unit of analysis in the dataset they are run on; d is the minimum population density threshold per square kilometer and p is the minimum population density threshold.

Table AXI. Summarized model results for different units of analysis

	Unit of analysis	Theta estimate	No. obs.	Model AIC
1	d=100, p=200	-0.583 (0.168)***	417	337.879
2	d=100, p=500	-0.583 (0.168)***	417	337.879
3	d=100, p=1000	-0.636 (0.178)***	254	168.304
4	d=200, p=200	-0.679 (0.122)***	379	232.764
5	d=200, p=500	-0.679 (0.122)***	379	232.764
6	d=200, p=1000	-0.62 (0.139)***	240	113.04
7	d=500, p=1000	-0.597 (0.159)***	214	115.11
8	d=750, p=200	-0.636 (0.141)***	245	152.221
9	d=750, p=500	-0.636 (0.141)***	245	152.221
10	d=750, p=1000	-0.676 (0.151)***	189	98.156
11	d=1000, p=200	-0.553 (0.163)***	180	98.931
12	d=1000, p=500	-0.553 (0.163)***	180	98.931
13	d=1000, p=1000	-0.553 (0.163)***	180	98.931

d = minimum pop. density threshold (per square kilometer), p = minimum pop. threshold
A technical glitch in sphet package makes it not possible to run (d=500, p=200) and (d=500, p=500)

Table AXII. Alternative units of analysis (MAUP)

	d=100, p=200	d=200, p=500	d=500, p=1000	d=750, p=200	d=1000, p=1000
(intercept)	4.716*** (0.612)	3.485*** (0.452)	5.061*** (0.741)	4.131*** (0.726)	3.680*** (0.860)
Neighbor diversity θ	-0.583*** (0.168)	-0.679*** (0.122)	-0.597*** (0.159)	-0.636*** (0.141)	-0.553*** (0.163)
Spatial dependency controls					
λ	0.180* (0.086)	0.131 (0.084)	0.196* (0.098)	0.169* (0.068)	0.151* (0.069)
ρ	0.462*** (0.087)	0.499*** (0.094)	0.469† (0.258)	0.324** (0.111)	0.215 (0.156)
Reconstruction logistics controls					
Occupation length	-0.008 (0.010)	-0.001 (0.010)	-0.018 (0.011)	-0.008 (0.011)	-0.009 (0.012)
Date liberated	0.006 (0.010)	0.002 (0.011)	0.015 (0.012)	0.004 (0.012)	0.006 (0.012)
Dist. to nearest highway	-0.004 (0.003)	-0.008*** (0.002)	-0.002 (0.004)	-0.002 (0.004)	0.003 (0.005)
Log(dist. to nearest city)	-0.384*** (0.057)	-0.426*** (0.058)	-0.422*** (0.065)	-0.343*** (0.060)	-0.381*** (0.065)
Pre-invasion population/economic controls					
Log(2012 population)	-0.017 (0.021)	-0.009 (0.022)	0.009 (0.028)	-0.039 (0.031)	-0.048 (0.035)
Log(2012 population density)	-0.088* (0.040)	-0.143** (0.048)	-0.168* (0.066)	-0.040 (0.075)	0.060 (0.093)
Log(light per capita before occupation)	-0.114*** (0.019)	-0.135*** (0.019)	-0.150*** (0.025)	-0.143*** (0.024)	-0.139*** (0.028)
Growth before occupation	-0.037 (0.353)	0.391 (0.342)	0.524 (0.430)	0.568 (0.421)	0.955* (0.456)
Ethno-religious group dummies					
Arab Shia Muslim	0.142 (0.242)	0.042 (0.201)	0.066 (0.168)	-0.097 (0.177)	0.026 (0.209)
Christian	-0.127 (0.222)	0.028 (0.159)	-0.145 (0.214)	-0.088 (0.201)	-0.087 (0.187)
Kurd Sunni Muslim	0.063 (0.087)	0.178* (0.079)	0.072 (0.097)	0.022 (0.092)	-0.059 (0.102)
Kurd Shia Muslim	-0.579† (0.323)	-0.462 (0.288)			
Kurd Yazidi	-0.181* (0.090)	-0.048 (0.093)	-0.095 (0.101)	-0.118 (0.100)	-0.223* (0.098)
Shabak Shia Muslim	-0.273† (0.147)	-0.184 (0.119)	-0.298 (0.195)	-0.206 (0.162)	-0.309† (0.168)
Shabak Sunni Muslim	-0.162 (0.175)	-0.134 (0.155)	-0.344 (0.301)	-0.285 (0.198)	-0.384† (0.214)
Turkmen Shia Muslim	-0.109 (0.123)	0.100 (0.121)	0.125 (0.173)	0.075 (0.148)	0.035 (0.152)
Turkmen Sunni Muslim	0.302† (0.168)	-0.008 (0.120)	0.174 (0.156)	-0.054 (0.182)	-0.128 (0.174)
Num. obs.	417.000	379.000	214.000	245.000	180.000
Parameters	20.000	20.000	19.000	19.000	19.000
Log likelihood	-147.879	-95.209	-36.596	-55.422	-28.090
AIC	337.879	232.764	115.110	152.221	98.931

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.1$. d = minimum pop. density threshold (per km²), p = minimum pop. threshold

3.4 Table AXIII. Specific groups and regions

Table AXIII re-estimates the preferred specification on only settlements of different groups and regions. The first model ('Disputed region') estimates the preferred specification only for settlements falling in the region 'disputed' between the Kurdish regional government and the Iraqi government, based on boundaries from Smith and Shadarevian (2017). The second model ('Uncontested region') repeats this exercise for only settlements outside this region. The third model ('All') reports the results for all settlements for reference. The fourth model ('Sunni Arab only') repeats the exercise for only those settlements with a Sunni Arab majority; the final model ('Non-Sunni Arab only') repeats the exercise only for settlements with a non-Sunni Arab majority. In each case, θ scores are still calculated using the full dataset.

Table AXIII shows the neighbor diversity result is robust in all cases except under the 'non-Sunni only' aggregation, where it is negative but not significant. This non-Sunni Arab only estimate still moves in the expected direction, and its insignificance is probably due to the relatively small number of non-Sunni Arab settlements (N=90). Table AXII investigates the purchase of the neighbor diversity result among non-Sunni Arab settlements using a design that is better suited to the small non-Sunni Arab sample size.

A technical error in the `sphet` package (Piras 2010) makes it not possible to run the Kelejian and Prucha (2010) model on the non-Sunni Arab only aggregation, so I use the analogous SARAR model but without robustness to heteroskedasticity available under the R `spdep` package (Bivand et. al. 2005) for all models in this table. The results from the heteroskedasticity-robust and non-heteroskedasticity-robust models are very similar.

Table AXIII. Different regions and groups

	Disputed region	Uncontested region	All	Sunni Arab only	Non-Sunni Arab only
(intercept)	2.426*** (0.690)	4.687*** (0.562)	3.559*** (0.431)	4.306*** (0.483)	1.415 (0.879)
Neighbor diversity θ	-0.543*** (0.162)	-0.878*** (0.202)	-0.685*** (0.117)	-0.811*** (0.130)	-0.180 (0.244)
Spatial dependency controls					
λ	0.041 (0.119)	0.094 (0.082)	0.105 (0.069)	0.098 (0.076)	-0.006 (0.165)
ρ	0.491*** (0.110)	0.387*** (0.083)	0.444*** (0.057)	0.433*** (0.071)	0.379* (0.164)
Reconstruction logistics controls					
Occupation length	-0.031 [†] (0.017)	0.025* (0.012)	-0.001 (0.010)	0.025* (0.012)	-0.044** (0.016)
Date liberated	0.033 [†] (0.017)	-0.021 [†] (0.012)	0.002 (0.010)	-0.025* (0.012)	0.052** (0.017)
Dist. to nearest highway	-0.004 (0.004)	-0.011*** (0.003)	-0.008*** (0.002)	-0.011*** (0.002)	0.009 [†] (0.005)
Log(dist. to nearest city)	-0.440*** (0.083)	-0.472*** (0.080)	-0.436*** (0.054)	-0.402*** (0.060)	-0.362** (0.112)
Pre-invasion population/economic controls					
Log(2012 population)	-0.011 (0.034)	-0.009 (0.030)	-0.009 (0.022)	-0.022 (0.024)	0.028 (0.049)
Log(2012 population density)	-0.115 [†] (0.067)	-0.176** (0.067)	-0.143** (0.048)	-0.132** (0.051)	-0.152 (0.103)
Log(light per capita before occupation)	-0.155*** (0.030)	-0.120*** (0.024)	-0.135*** (0.019)	-0.118*** (0.020)	-0.179*** (0.042)
Growth before occupation	0.298 (0.454)	0.551 (0.526)	0.401 (0.342)	0.777* (0.379)	-1.164 [†] (0.638)
Ethno-religious group dummies					
Christian	-0.056 (0.163)		0.025 (0.158)		-0.592* (0.294)
Kurd Sunni Muslim	0.187* (0.089)	0.320 [†] (0.189)	0.173* (0.077)		-0.107 (0.258)
Kurd Shia Muslim	-0.493 [†] (0.290)		-0.468 (0.290)		-0.806* (0.363)
Kurd Yazidi	-0.033 (0.108)		-0.051 (0.090)		-0.648* (0.288)
Shabak Shia Muslim	-0.228 [†] (0.123)		-0.188 (0.117)		-0.694* (0.273)
Shabak Sunni Muslim	-0.177 (0.158)		-0.137 (0.154)		-0.702* (0.282)
Turkmen Shia Muslim	0.091 (0.123)		0.098 (0.122)		-0.304 (0.293)
Turkmen Sunni Muslim	-0.014 (0.124)		-0.010 (0.121)		-0.204 (0.287)
Arab Shia Muslim		0.064 (0.180)	0.051 (0.193)		
Num. obs.	202	177	379	290	89
Parameters	21	15	22	13	21
Log Likelihood	-36.882	-29.871	-73.488	-48.529	-3.625
AIC (Linear model)	132.143	106.230	241.659	158.490	48.394
AIC (Spatial model)	115.763	89.742	190.976	123.058	49.251
LR test: statistic	20.380	20.488	54.682	39.432	3.143
LR test: p-value	0.000	0.000	0.000	0.000	0.208

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; [†] $p < 0.1$

3.5 Table AXIV. Specific population brackets

Table AXIV re-estimates the preferred specification on only settlements within specific population brackets. The first model estimates the preferred specification only for settlements with a population between 500 and 1,000 people; the second repeats this exercise for only settlements between 1,000 and 3,000 people; the third for only settlements greater than 3,000 people. The

results show that the neighbor diversity effect is robust across different settlement sizes, although its significance and effect size varies. The effect appears to be largest and most significant among small villages, and weakest among larger towns.

Table AXIV. Different population brackets

	500 - 1000	1000 - 3000	3000 - 1e+06	all larger than min
(intercept)	6.438*** (1.194)	2.868*** (0.831)	4.176*** (0.702)	3.485*** (0.452)
Neighbor diversity θ	-0.847*** (0.189)	-0.767*** (0.193)	-0.598*** (0.167)	-0.679*** (0.122)
Spatial dependency controls				
λ	0.274* (0.118)	0.142 (0.192)	0.103 (0.157)	0.131 (0.084)
ρ	0.463* (0.197)	0.696** (0.255)	0.228 (0.383)	0.499*** (0.094)
Reconstruction logistics controls				
Occupation length	0.026 (0.017)	-0.015 (0.014)	0.005 (0.011)	-0.001 (0.010)
Date liberated	-0.023 (0.017)	0.019 (0.014)	-0.009 (0.012)	0.002 (0.011)
Dist. to nearest highway	-0.008* (0.004)	-0.010** (0.003)	-0.006* (0.003)	-0.008*** (0.002)
Log(dist. to nearest city)	-0.308** (0.099)	-0.490*** (0.084)	-0.498*** (0.085)	-0.426*** (0.058)
Pre-invasion population/economic controls				
Log(2012 population)	-0.450** (0.145)	0.113 (0.084)	-0.022 (0.040)	-0.009 (0.022)
Log(2012 population density)	-0.098 (0.090)	-0.224** (0.076)	-0.127 (0.085)	-0.143** (0.048)
Log(light per capita before occupation)	-0.082* (0.035)	-0.162*** (0.027)	-0.162*** (0.037)	-0.135*** (0.019)
Growth before occupation	0.012 (0.634)	1.406** (0.535)	0.668 (0.635)	0.391 (0.342)
Ethno-religious group dummies				
Arab Shia Muslim	0.091 (0.193)	0.222 (0.206)		0.042 (0.201)
Christian	0.665* (0.335)	-0.226 (0.217)	0.084 (0.204)	0.028 (0.159)
Kurd Sunni Muslim	0.110 (0.122)	0.120 (0.111)	-0.011 (0.174)	0.178* (0.079)
Kurd Yazidi	-0.169 (0.319)	-0.096 (0.139)	-0.087 (0.099)	-0.048 (0.093)
Shabak Shia Muslim	-0.281 (0.179)	-0.071 (0.208)	-0.308 (0.209)	-0.184 (0.119)
Shabak Sunni Muslim	0.356 (0.252)	-0.281 (0.225)		-0.134 (0.155)
Turkmen Shia Muslim	-0.048 (0.241)	0.101 (0.189)	0.203 (0.180)	0.100 (0.121)
Turkmen Sunni Muslim	-0.338 (0.258)	0.258 (0.161)	-0.444* (0.192)	-0.008 (0.120)
Kurd Shia Muslim		-0.420 (0.283)		-0.462 (0.288)
Num. obs.	140.000	141.000	99.000	379.000
Parameters	19.000	20.000	17.000	20.000
Log likelihood	-42.337	-26.487	-0.837	-95.209
AIC	129.007	99.974	43.230	232.764

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.1$

3.6 Table AXV. Alternative causal strategies for spatial dependence

Table AXV examines the robustness of the preferred specification to a variety of alternate strategies for controlling for spatial dependence. It runs the preferred specification on models that use simple OLS ('OLS'), clustered standard errors ('Clustered'), district-level fixed effects ('FE'), a simultaneous spatial autoregressive model ('Simultaneous SAR'; used in article), a spatial error model ('Spatial error'), and a spatial lag model ('Spatial lag'). The results show that the neighbor diversity effect is robust to different strategies to control for spatial dependence, and its effect size is relatively consistent across them.

Table AXV. Alternative strategies to control for spatial dependence

	OLS	Clustered	FE	General SARAR	Spatial error	Spatial lag
(intercept)	3.695*** (0.390)	3.695*** (0.451)	2.446*** (0.626)	3.559*** (0.431)	3.718*** (0.424)	3.287*** (0.380)
Neighbor diversity θ	-0.720*** (0.097)	-0.720*** (0.125)	-0.666*** (0.153)	-0.685*** (0.117)	-0.693*** (0.120)	-0.675*** (0.093)
Spatial dependency controls						
λ				0.105 (0.069)		0.194*** (0.039)
ρ				0.444*** (0.057)	0.476*** (0.037)	
Reconstruction logistics controls						
Occupation length	-0.001 (0.008)	-0.001 (0.013)	-0.018 [†] (0.010)	-0.001 (0.010)	-0.002 (0.010)	-0.000 (0.008)
Date liberated	0.001 (0.008)	0.001 (0.014)	0.021* (0.010)	0.002 (0.010)	0.003 (0.010)	0.001 (0.008)
Dist. to nearest highway	-0.007*** (0.002)	-0.007* (0.003)	-0.005 [†] (0.002)	-0.008*** (0.002)	-0.008*** (0.002)	-0.006** (0.002)
Log(dist. to nearest city)	-0.484*** (0.045)	-0.484*** (0.059)	-0.143 (0.122)	-0.436*** (0.054)	-0.463*** (0.052)	-0.416*** (0.045)
Pre-invasion population/economic controls						
Log(2012 population)	-0.019 (0.025)	-0.019 (0.020)	0.004 (0.023)	-0.009 (0.022)	-0.012 (0.022)	-0.004 (0.024)
Log(2012 population density)	-0.102* (0.051)	-0.102* (0.045)	-0.137** (0.050)	-0.143** (0.048)	-0.139** (0.048)	-0.126** (0.049)
Log(light per capita before occupation)	-0.132*** (0.019)	-0.132*** (0.024)	-0.120*** (0.020)	-0.135*** (0.019)	-0.136*** (0.019)	-0.129*** (0.018)
Growth before occupation	0.538 (0.368)	0.538 (0.465)	0.028 (0.359)	0.401 (0.342)	0.447 (0.342)	0.342 (0.351)
Ethno-religious group dummies						
Arab Shia Muslim	0.152 (0.128)	0.152 (0.107)	0.027 (0.125)	0.051 (0.193)	0.079 (0.197)	-0.002 (0.126)
Christian	0.004 (0.156)	0.004 (0.058)	0.021 (0.155)	0.025 (0.158)	0.009 (0.159)	0.061 (0.148)
Kurd Sunni Muslim	0.057 (0.071)	0.057 (0.085)	0.154* (0.076)	0.173* (0.077)	0.167* (0.078)	0.092 (0.067)
Kurd Shia Muslim	-0.464 (0.327)	-0.464*** (0.065)	-0.739* (0.304)	-0.468 (0.290)	-0.486 [†] (0.289)	-0.418 (0.310)
Kurd Yazidi	-0.118 (0.083)	-0.118 (0.074)	-0.035 (0.125)	-0.051 (0.090)	-0.058 (0.092)	-0.089 (0.079)
Shabak Shia Muslim	-0.249* (0.110)	-0.249*** (0.057)	-0.328** (0.126)	-0.188 (0.117)	-0.205 [†] (0.118)	-0.184 [†] (0.105)
Shabak Sunni Muslim	-0.135 (0.155)	-0.135 [†] (0.077)	-0.125 (0.153)	-0.137 (0.154)	-0.153 (0.155)	-0.092 (0.147)
Turkmen Shia Muslim	0.134 (0.137)	0.134 (0.177)	0.206 (0.128)	0.098 (0.122)	0.084 (0.122)	0.176 (0.130)
Turkmen Sunni Muslim	-0.067 (0.130)	-0.067 (0.245)	0.021 (0.135)	-0.010 (0.121)	-0.016 (0.121)	-0.036 (0.123)
R ²	0.422		0.560			
Adj. R ²	0.393		0.499			
Num. obs.	379	379	379	379	379	379
R ² (full model)		0.422				
R ² (proj model)		0.422				
Adj. R ² (full model)		0.393				
Adj. R ² (proj model)		0.393				
Parameters				22	21	21
Log Likelihood				-73.488	-74.839	-91.757
AIC (Linear model)				241.659	241.659	241.659
AIC (Spatial model)				190.976	191.678	225.514
LR test: statistic				54.682	51.981	18.145
LR test: p-value				0.000	0.000	0.000

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; [†] $p < 0.1$. District fixed effects are not shown.

3.7 Table AXVI. Alternative ethnicity coding scheme

Table AXVI re-estimates the preferred specification under an alternative ethno-religious coding scheme. For the "Smith and Shadarevian" model, the Smith and Shadarevian (2017) data is used exclusively to code ethno-religious composition. The "IOM ILA III" model uses the ILA III data and is shown for comparison.

Table AXVI. Ethno-religious coding from Smith and Shadarevian (2017)

	Smith and Shadarevian	IOM ILA III
(intercept)	2.952*** (0.435)	3.485*** (0.452)
θ (based on IOM ILA III)		-0.679*** (0.122)
θ (based on Smith and Shadarevian)	-0.637*** (0.137)	
Spatial dependency controls		
λ	0.173* (0.084)	0.131 (0.084)
ρ	0.495*** (0.088)	0.499*** (0.094)
Reconstruction logistics controls		
Occupation length	-0.013 (0.010)	-0.001 (0.010)
Date liberated	0.015 (0.011)	0.002 (0.011)
Dist. to nearest highway	-0.009*** (0.002)	-0.008*** (0.002)
Log(dist. to nearest city)	-0.408*** (0.058)	-0.426*** (0.058)
Pre-invasion population/economic controls		
Log(2012 population)	-0.016 (0.023)	-0.009 (0.022)
Log(2012 population density)	-0.124* (0.048)	-0.143** (0.048)
Log(light per capita before occupation)	-0.132*** (0.019)	-0.135*** (0.019)
Growth before occupation	0.416 (0.346)	0.391 (0.342)
Ethno-religious group dummies		
Arab Shia Muslim	0.008 (0.203)	0.042 (0.201)
Christian	-0.107 (0.157)	0.028 (0.159)
Kurd Sunni Muslim	0.202* (0.080)	0.178* (0.079)
Kurd Shia Muslim	-0.305 (0.292)	-0.462 (0.288)
Kurd Yazidi	-0.151' (0.092)	-0.048 (0.093)
Shabak Shia Muslim	-0.261* (0.118)	-0.184 (0.119)
Shabak Sunni Muslim	-0.226 (0.155)	-0.134 (0.155)
Turkmen Shia Muslim	0.078 (0.122)	0.100 (0.121)
Turkmen Sunni Muslim	-0.134 (0.120)	-0.008 (0.120)
Num. obs.	379.000	379.000
Parameters	20.000	20.000
Log likelihood	-102.830	-95.209
AIC	248.006	232.764

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ' $p < 0.1$

4 ILA III data and ethno-religious plurality coding

The IOM (2017) ILA III dataset is the primary source of plurality ethno-religious group codings at the settlement level. It is also used to generate statistics on post-conflict administration and social outcomes, which are not used in the empirical analysis but are referenced during discussion. The coordinates of IOM ILA III survey points do not always map directly onto the settlement boundaries drawn by the coding rule on a 1:1 basis: sometimes IOM surveyed several locations within a settlement; other times, no survey points are located within a settlement boundary. This section explains how IOM ILA III data was matched to settlements.

4.1 Ethno-religious plurality coding

Settlement-level data on ethno-religious composition is also based primarily on the ILA III dataset. Given the importance of ethno-religious composition variable to the conclusion of the paper and the central role it plays in the empirical analysis, special effort was made to ensure that ethno-religious codings a) are as accurate and unbiased as possible and b) were generated for all settlements.

1. In a very small number of cases, ILA III surveyors selected the 'only if the largest group cannot be determined, select a second one' option when recording ethno-religious composition. In these cases, the group used is selected randomly from the two groups listed.
2. If more than one ILA III survey point is located within the boundary of a settlement, the plurality group for each survey point is calculated, and the group that is the plurality in the largest number of survey points is used. If two or more groups are tied with an equal number of qualifying survey points where they are pluralities, the group used is selected randomly from among the tying groups.
3. In some cases, no survey points are located within the settlement boundary drawn by the coding rule. This occurs in two scenarios: 1) when the coordinates of the ILA survey point lie just outside the bounds of the settlement drawn by the coding rule (and the survey point is clearly intended to apply to the settlement in question), and 2) when the settlement has not been surveyed. In these cases:
 - If the distance between the settlement centroid and the closest ILA III survey point is less than 10 kilometers, the nearest ILA III survey point is used to code ethno-religious composition for the settlement in question.
 - If the distance between the settlement centroid and the closest ILA III survey point is more than 10 kilometers, ethno-religious mapping from Smith and Shadarevian (2017) is used to code the ethno-religious composition instead. Smith and Shadarevian constructed a mapping of ethno-religious groups in northern Iraq with near-settlement-level granularity based on a combination of open-source intelligence and fieldwork in summer 2016. I prefer to use the ILA III data wherever possible on the assumption that direct survey data is more accurate, but table AXVI in this appendix shows that the results of this paper are robust to ethno-religious codings developed entirely from the Smith and Shadarevian mapping.
 - In two settlements where the ILA survey point(s) would have otherwise been used, IOM surveyors selected "unknown" for the ethno-religious group. In these cases codings from the Smith and Shadarevian mapping are used.
 - Under this approach, 53 settlements (14%) use the Smith and Shadarevian-based coding instead of the ILA III-based coding. 138 settlements (36%) are coded using ILA III survey points with coordinates inside their settlement boundaries as drawn by the coding rule. The remaining 188 settlements do not have an ILA III survey point inside their settlement boundary and are coded using the nearest ILA III survey point. Among settlements for which the nearest ILA III survey point was used, summary statistics for the distance between the ILA III survey point used and the settlement centroid are as follows: median = 2.78km, mean = 3.47km, 75th percentile = 5.41km. These small distances confirm that in many cases the coordinates used are likely intended to apply to

the settlement in question, and the ILA survey point coordinates simply landed outside the boundary drawn by the coding rule. Even where this is not the case, the statistics show that settlements are almost always coded using survey information from a very nearby location.

4.2 ILA III data other than ethno-religious data

For ILA III data other than ethno-religious data (for example, statistics cited in the discussion), the same process was used except that a) if multiple ILA III survey points were inside the settlement boundaries, the one with the largest population was used, and b) settlements in Ba'aj district, where the nearest ILA III survey point is more than 10km away, were simply not considered in the analysis.

5 Territorial control coding

ISIL occupation start dates were hand coded using news reports listing the beginning of ISIL occupation for each settlement. Where news reports identifying the occupation start date for a settlement were not available, settlements were assigned the occupation start date of the closest settlement for which a news report was available. Since ISIL took over large swaths of Iraqi territory quickly - often overrunning entire districts in a matter of days - this approach likely captures the occupation start month correctly for the vast majority of settlements. A table listing the news sources used is included below.

ISIL's retreat was more gradual and less uniform than its advance. An increase in public attention surrounding the conflict also meant that much more information was generated surrounding the end dates of ISIL occupation. The first-choice method of coding end dates remained to code settlements' occupation end dates directly using news reports about the liberation of each settlement. The news sources used are listed in the table below. Where such reports were not available, I used conflict mapping from the LiveUAMap project (2019) and the Institute for the Study of War (2019) to code occupation end dates. Both LiveUAMap and the Institute for the Study of War began producing regular control-of-territory mappings partway through the conflict based on local intelligence, social media, military updates, and news reports in several languages.

Table AXVII: News sources for territorial control coding

Governorate	Settlement Name	ISIL Control Event	Date	Source	Note
Al-Anbar	Al-Nikhib	Start	6/24/2014	https://www.longwarjournal.org/archives/2014/06/isis-advances-on-oil-fields.in.php	
	Fallujah	Start	4/1/2014	https://www.telegraph.co.uk/news/worldnews/middleeast/iraq/10550563/Fallujah-falls-under-Al-Qaeda-control-in-blow-for-Iraq-security.html	
		End	6/1/2016		
	Haditha	Start	12/1/2014	https://www.telegraph.co.uk/news/2016/05/03/iraqi-forces-break-through-longest-isl-siege-to-beleaguered-tow/	contested but never fully captured
		End	5/1/2016	https://www.telegraph.co.uk/news/2016/05/03/iraqi-forces-break-through-longest-isl-siege-to-beleaguered-tow/	contested but never fully captured
	Hit	Start	9/1/2014	https://www.aljazeera.com/news/middleeast/2014/10/isl-sunni-militias-battle-iraqi-town-201410281343980744.html	
		End	4/1/2016		
	Hussebia	Start	5/1/2015	https://thehill.com/policy/international/242920-isis-captures-another-town-in-west-iraq?amp	
		End	5/1/2018	https://www.marinecorpstimes.com/news/your-military/2015/06/06/iraqi-troops-militias-repel-is-attacks-in-anbar-province/	
	Karma	Start	3/1/2014	https://www.longwarjournal.org/archives/2014/06/isis-advances-on-oil-fields.in.php	
Al Qaim	Kubaysah	Start	9/14/2014	https://www.longwarjournal.org/archives/2014/06/isis-advances-on-oil-fields.in.php	
	Ramadi	Start	11/1/2014	http://www.hurriyetaidailynews.com/jihadists-launch-onslaught-on-iraqs-ramadi-74645	
		End	1/1/2016		
	Rawa	Start	6/1/2014	https://www.nytimes.com/2014/06/23/world/middleeast/iraq.html	
		End	11/1/2017	https://www.cbsnews.com/news/last-iraqi-town-held-by-isis-recaptured-officials-say/	
	Rutba	Start	6/1/2014	https://www.bbc.com/news/world-middle-east-27960142	
		End	6/1/2016		
	Saqlawia	Start	6/1/2014	https://www.reuters.com/article/uk-iraq-security/u-s-considers-air-strikes-on-iraq-holds-talks-with-iran-idUKKBNOO0LF20140616?rpc=932	
		End	6/1/2016	https://edition.cnn.com/2016/06/04/middleeast/iraq-syria-military-advance-on-isis/	
	Al Qaim	Start	8/1/2014	http://www.niqash.org/en/articles/security/5319	
Arbil		End	10/1/2017		
	Makhmur	Start	8/1/2014	http://www.rudaw.net/english/kurdistan/20112016	
At-Ta'mim		End	8/2/2014	http://www.rudaw.net/english/kurdistan/20112016	
	Al Zab	Start	6/1/2014	http://www.defenddemocracy.org/media-hit/bill-roggio-isis-seizes-more-towns-in-northern-and-central-iraq/	
		End	9/1/2017	https://www.iraqnews.com/iraq-war/iraqi-forces-end-phase-1-hawija-offensive/	

News sources for territorial control coding (*continued*)

Governorate	Settlement Name	ISIL Control Event	Date	Source	Note
	Mahouz / Abbasi	Start	6/11/2014	https://www.longwarjournal.org/archives/2014/06/isis.advances.on.oil.fields.in.php	
	Rashad	Start	6/11/2014	https://www.longwarjournal.org/archives/2014/06/isis.advances.on.oil.fields.in.php	
	Riyadh	Start	6/11/2014	https://www.longwarjournal.org/archives/2014/06/isis.advances.on.oil.fields.in.php	
	Unnamed (44.37987, 33.45453)	Start	5/23/2019	https://www.longwarjournal.org/archives/2014/06/isis.advances.on.oil.fields.in.php	
Babil	Jurf al-Sakhar	Start	3/1/2014	https://www.longwarjournal.org/archives/2014/06/isis.advances.on.oil.fields.in.php	
Diyala	Adhaim	Start	6/14/2014	https://www.longwarjournal.org/archives/2014/06/isis.advances.on.oil.fields.in.php	
	Mansuriyah 1 oil well	Start	6/26/2014	https://www.longwarjournal.org/archives/2014/06/isis.advances.on.oil.fields.in.php	
	Miqdadiyah	Start	6/16/2014	https://www.longwarjournal.org/archives/2014/06/isis.advances.on.oil.fields.in.php	
		End	1/22/2015	https://finance.yahoo.com/news/isis-suffers-heaviest-defeat-iraq-101500786.html	
	Qarah Tapah	End	9/17/2016	https://www.iraqnews.com/iraq-war/joint-forces-liberate-tabba-village-nineveh/	
Hawija	Hawija	Start	6/1/2014	https://www.bbc.com/news/world-middle-east-41346427	
		End	10/1/2017		
Kirkuk		Start	6/1/2014	https://www.bbc.com/news/world-middle-east-41346427	
		End	10/1/2017		
Ninawa	Al Kuwayr	Start	7/3/2014	https://www.longwarjournal.org/archives/2014/06/isis.advances.on.oil.fields.in.php	
		End	7/10/2014	https://www.longwarjournal.org/archives/2014/06/isis.advances.on.oil.fields.in.php	
	Al Qayyarah	Start	6/1/2014	http://www.enro.who.int/irq/iraq-news/who-scales-up-activities-to-support-newly-liberated-qayyarah-iraq.html	
		End	8/1/2016	http://www.enro.who.int/irq/iraq-news/who-scales-up-activities-to-support-newly-liberated-qayyarah-iraq.html	
	Bartella	Start	8/1/2014	http://www.latimes.com/world/middleeast/la-ig-iraq-christians-20140919-story.html	
		End	10/1/2016	https://www.france24.com/en/20161023-video-iraqi-troops-liberate-christian-town-bartella	
	Bashiqa	Start	6/1/2014	http://www.rudaw.net/english/kurdistan/07112016	
		End	11/1/2016	http://www.rudaw.net/english/kurdistan/07112016	
	Bazwaia	End	10/20/2012	https://www.haaretz.com/middle-east-news/iraqi-forces-take-isis-held-tv-station-as-mosul-battle-mounts-1.5455596	
	Mosul	Start	6/1/2014	https://www.bloomberg.com/news/articles/2017-07-09/mosul-liberated-as-islamic-state-faces-total-defeat-in-iraq	
		End	7/1/2017		
	Qaraqosh (Bakhdida)	Start	8/1/2014	http://www.lefigaro.fr/flash-actu/2014/08/07/97001-20140807FILWWW00062-irak-des-jihadistes-prennent-la-plus-grande-ville-chretienne-du-pays.php	

6 Additional notes on luminosity data preparation

This section mentions two additional points about luminosity data preparation.

- Monthly VIIRS images for each June and July 2018 are unavailable. Following Li et. al. (2018), data for these months is interpolated using neighboring months. Using the exclusion criteria from Li et. al. (2018): Pixels with light values > 500 NanoWatt/(cm²sr) are labelled flares; these areas and 10km buffers were masked.
- When applying loess smooths to settlement lighting panels, I set the loess smoothing parameter $= 0.5$. Settlements occupied for fewer than 8 months do not have enough data to estimate a loess smooth with $= 0.5$, so in these cases the raw data is used. The results are not sensitive to other reasonable choices of λ .

7 References

Bivand, Roger, et al (2005) The spdep package. *Comprehensive R Archive Network, Version 2005*: 05-83.

Cook, R. Dennis, and Sanford Weisberg (1982) Residuals and influence in regression. New York: Chapman and Hall, 1982.

Gaston, Erica Andras Derzsi-Horvath (2017) Iraq after ISIL: Sub-State Actors, Local Forces, and the Micro-Politics of Control. *Global Public Policy Institute* (https://www.gppi.net/media/Gaston_Derzsi-Horvath_Iraq_After_ISIL.pdf).

International Organization for Migration, Iraq Mission (2018a) Integrated Location Assessment 3 (<http://iraqdtm.iom.int/ILA3.aspx>).

McDonald, Barry (2002) A teaching note on Cook's distance-a guideline. (<https://mro.massey.ac.nz/handle/10179/4352>)

International Crisis Group (2018) Winning the Post-ISIS Battle for Iraq in Sinjar. *International Crisis Group* Middle East Report 183. (<https://www.crisisgroup.org/middle-east-north-africa/gulf-and-arabian-peninsula/iraq/183-winning-post-isis-battle-iraq-sinjar>).

Li, Xi; Shanshan Liu, Michael Jendryke, Deren Li Chuanqing Wu (2018). Night-time light dynamics during the Iraqi civil war. *Remote Sensing* 10(6): 858.

Piras, Gianfranco (2010) sphet: Spatial models with heteroskedastic innovations in R. *Journal of Statistical Software* 35(1): 1-21.