

Supplementary Information: Geographic Determinants of Indiscriminate Violence

Sebastian Schutte

April 28, 2015

Replication of the figures and results

Full replication data and code are available at http://sebastianschutte.net/?page_id=63. For legal reasons, I won't make the SIGACT data available as it is still classified. In order to verify my codings or relate the covariate information to SIGACT, please download the SIGACT data for Afghanistan from the WikiLeaks website at https://wikileaks.org/wiki/Afghan_War_Diary,_2004-2010. After that, please add a column to the data and add sha256 hashes of the unique report keys into this new column. You can now join the SIGACT events with the covariate information from the replication data based on these hash values and the reportkey_sha256 column in the replication data.

Line-of-sight dataset

I constructed a line-of-sight measurement to account for a tactical particularity of the conflict touched upon above: In areas with limited lines-of-sight due to high densities of natural obstacles, actors might use indirect fire instead of direct fire. Since I coded the first type of attack as selective and the second type as indiscriminate, I needed to control for this factor. Using the digital elevation model (DEM) by Gesch et al. (1999), I calculated the number of surrounding cells that are visible from any location. To keep the computational effort tractable, I resampled the DEM to a cell resolution of 0.05 decimal degrees (about 5km at the equator). This calculation involved the specification of a "horizon" in terms of a maximal distance from the cell under investigation (30 kilometers in this case, as artillery strikes beyond this distance are very unlikely). For all cells within that horizon, Bresenham's (1965) algorithm was used to calculate all cells along a straight line connecting the origin and the target cell. In a second step, elevation levels along this line were used to calculate angles between the cell under investigation and the cells along the line. The number of visible cells was then established by counting the number of cells along this line for which no steeper angle had been calculated for any preceding cell. In this way, cells along this line with no obstructing cell in front of them were established. Since this procedure was repeated for all cells within the horizon, a count of all visible cells was established for each cell in the elevation dataset. The georeferenced data can be downloaded from http://sebastianschutte.net/?page_id=22.

Summary statistics for the main independent variables

The table below shows summary statistics for the insurgencies from the GED dataset.

Variable	Mean	Std. Dev.	Minimum	Maximum
1 Cap/Border dist. (Normalized)	0	0.293	0	1
2 Population	11269	20448.47	0	126452.789
3 Accessibility	118	210.459	0	3490
4 Landcover	9	2.945	0	14
5 Gecon	0	0.076	0	1
6 Elevation	636	560.984	-150.28	2976.56
7 Line-of-sight	309	218.855	1	986
8 Military casualties	9	47.918	0	2363
9 Civilian casualties	1	6.078	0	200

The table below shows summary statistics for the SIGACT dataset.

Variable	Mean	Std. Dev.	Minimum	Maximum
1 Dist Kabul (km)	340.253	185.011	0.124	806.123
2 Population	1511.045	7996.698	3	123730
3 Accessibility	232.581	199.077	0	2715
4 Landcover	9.039	1.522	0	14
5 Gecon	0.28	0.68	0	3
6 Elevation	1358.142	575.471	286	4542
7 Military casualties	0.529	2.904	0	114
8 Civilian casualties	0.323	2.707	0	189

Overview of SIGACT event categories

The table below shows the event categories reported in SIGACT with corresponding numbers of observations. Categories are ranked according to frequency. The column “Category” indicates the (capitalized) name of the category. The “% ISAF” column shows which percentage of the observations were tagged as friendly actions in the data. This overview was an important guideline for the selection of suitable event categories: To ensure external validity of the study and generate large statistical samples, the chosen events needed to be frequent. At the same time, the theoretical concepts of selective and indiscriminate violence needed to apply. These are conflicting demands: Of course, observations where insurgents initiated attacks by indiscriminate means which led to large numbers of civilian casualties would be a better theoretical fit than broad event categories. However, such selection schema also offer numerous coding choices that researchers could select from after looking at different sets of results. Moreover, the exact combination of coding requirements leads to much smaller samples than the focus on high-frequency categories. To strike a balance between theoretical fit, empirical availability, and conceptual rigor, I focused on “direct fire”, the most frequent category in SIGACT, as the control category (“selective violence”). The second most frequent category “IED found/cleared” was omitted for the analysis. This is due to the fact that this event does not speak to the theoretical question at hand. “Indirect fire”, the third most frequent category, was used in the analysis as the treatment category. While “mine strikes”

and “close air support” were also included in the analysis, they do not drive the inferential results as they are comparatively infrequent (321 and 95 observations, respectively). I also considered “escalation of force” and “attack” events for the analysis, but these events do not generally qualify as selective or indiscriminate.

Rank	Category	% ISAF	N Obs.
1	DIRECT FIRE	5	16286
2	IED FOUND/CLEARED	0	8369
3	INDIRECT FIRE	5	7229
4	IED EXPLOSION	0	7022
5	OTHER	9	4684
6	MEDEVAC	94	3293
7	UNEXPLODED ORDNANCE	0	2770
8	CACHE FOUND/CLEARED	100	2739
9	ESCALATION OF FORCE	100	2267
10	ATTACK	25	2267
11	SAFIRE	0	1695
12	MEETING	0	1404
13	MEETING - DEVELOPMENT	0	988
14	IED SUSPECTED	0	893
15	ACCIDENT	6	834
16	TURN IN	0	813
17	MEETING - SECURITY	0	753
18	DETAINED	0	682
19	IED FALSE	0	550
20	AMBUSH	2	537
21	DETAINEE TRANSFER	99	517
22	INTERDICTION	0	488
23	MEDEVAC (LOCAL NATIONAL)	100	428
24	OTHER (HOSTILE ACTION)	4	417
25	FRAGO	0	404
26	PLANNED EVENT	0	403
27	QA/QC PROJECT	0	400
28	TRANSFER	0	399
29	MINE FOUND/CLEARED	0	396
30	SURVEILLANCE	23	369
31	PATROL	88	364
32	IED AMBUSH	0	350
33	MINE STRIKE	0	321
34	ANP TRAINING	0	282
35	PREMATURE DETONATION	0	237
36	DEMONSTRATION	0	236
37	PSYOP	100	189
38	DETAIN	100	185
39	IED HOAX	0	185
40	MEDEVAC PATIENT TRANSFER	100	160

Continued on next page

Table 1 – continued from previous page

Rank	Category	% ISAF	N Obs.
41	MEDCAP	0	160
42	UNKNOWN EXPLOSION	0	155
43	SNIPER OPS	39	154
44	IDF INTERDICTION	100	137
45	OTHER OFFENSIVE	100	132
46	CAS	100	123
47	RELEASED	0	110
48	KIDNAPPING	0	109
49	PROPAGANDA	0	100
50	MURDER	0	99
51	CLOSE AIR SUPPORT	100	95
52	PROJECT START	0	88
53	EQUIPMENT FAILURE	0	81
54	PROJECT CLOSEOUT	0	81
55	CORDON/SEARCH	100	80
56	ARTY	100	77
57	RPG	0	76
58	GREEN-GREEN	100	72
59	DELIBERATE ATTACK	100	69
60	MEDEVAC (OTHER)	100	64
61	DETAINEE RELEASE	98	60
62	ERW/TURN-IN	100	58
63	NATURAL DISASTER	0	55
64	CONVOY	86	53
65	ARREST	100	50
66	EVIDENCE TURN-IN/RECEIVED	100	50
67	SNOW AND ICE REMOVAL	0	49
68	ASSASSINATION	0	48
69	RAID	86	44
70	COUNTER MORTAR FIRE	100	41
71	ARSON	0	41
72	THEFT	0	40
73	CHECKPOINT RUN	0	37
74	RECON	42	33
75	SMALL UNIT ACTIONS	100	32
76	OTHER DEFENSIVE	100	30
77	CARJACKING	0	30
78	SECTARIAN VIOLENCE	0	30
79	VOGE	0	29
80	RECONNAISSANCE	0	28
81	CRIMINAL ACTIVITY	0	27
82	ERW RECOVERED	100	24
83	POLICE ACTIONS	100	24
84	SMUGGLING	0	22

Continued on next page

Table 1 – continued from previous page

Rank	Category	% ISAF	N Obs.
85	TESTS OF SECURITY	0	22
86	NEGLIGENT DISCHARGE	0	19
87	BLUE-BLUE	100	18
88	GREEN-BLUE	100	16
89	UAV	100	16
90	SUPPORTING CF	0	15
91	CASEVAC	100	14
92	DOWNED AIRCRAFT	0	13
93	ENEMY ACTION	0	13
94	VETCAP	0	13
95	TRIBAL FEUD	0	12
96	REFUGEES	0	12
97	BORDER OPS	100	11
98	VEHICLE INTERDICTION	100	11
99	LOOTING	0	11
100	VANDALISM	0	11
101	IED THREAT	0	10
102	BLUE-GREEN	100	10
103	HARD LANDING	0	9
104	INSURGENT VEHICLE	0	9
105	REPETITIVE ACTIVITIES	0	8
106	AIR MOVEMENT	0	8
107	COUNTER INSURGENCY	100	8
108		0	7
109	SEARCH AND ATTACK	100	7
110	RESUPPLY	0	7
111	COUNTER MORTAR PATROL	100	7
112	TRIBAL	100	7
113	COUNTER NARCOTIC	100	6
114	BLUE-WHITE	100	6
115	GREEN-WHITE	100	6
116	SABOTAGE	0	6
117	DRUG OPERATION	0	6
118	ANA-ON-ANP	100	6
119	DEFECTING	0	5
120	IDF COUNTER FIRE	100	5
121	CCA	40	5
122	SERMON	0	5
123	EXTORTION	0	5
124	SURRENDERING	0	4
125	SUPPORTING AIF	0	4
126	FOOD DISTRIBUTION	0	4
127	PSYOP (WRITTEN)	100	4
128	MOVEMENT TO CONTACT	100	4

Continued on next page

Table 1 – continued from previous page

Rank	Category	% ISAF	N Obs.
129	FINANCE	0	3
130	AIR ASSAULT	0	3
131	POLICE INTERNAL	100	3
132	SHOW OF FORCE	100	2
133	TCP	100	2
134	BREACHING	100	2
135	NONE SELECTED	50	2
136	BLUE-ON-WHITE	100	2
137	PSYOP (TV/RADIO)	100	2
138	INTERNAL SECURITY FORCES	100	2
139	DRUG VEHICLE	0	2
140	AMF-ON-ANA	100	2
141	THREAT	0	1
142	ELICITATION	0	1
143	SECURITY BREACH	0	1
144	REPORTED LOCATION	0	1
145	NARCOTICS	0	1
146	AMNESTY	0	1
147	COUNTER TERRORISM	100	1
148	RECRUITMENT (WILLING)	0	1
149	GRAFFITI	0	1
150	POISONING	0	1
151	MUGGING	0	1
152	BLACK LIST	0	1
153	NBC	0	1

Table 2: Regression results from the SIGACT analysis. The estimated models predict indiscriminate violence as a function of distance to the capital and show a positive effect on incumbent indiscriminate violence and a negative effect on insurgent indiscriminate violence. In this case, only direct and indirect fire were used to code the event categories.

	<i>Dependent variable:</i>			
	Insurgent indiscriminate		Incumbent indiscriminate	
	(1)	(2)	(3)	(4)
Dist. Pak. (km)	-0.003*** (0.0002)	-0.003*** (0.0002)	-0.007*** (0.001)	-0.006*** (0.001)
Dist. Kabul (km)	-0.003*** (0.0002)	-0.003*** (0.0002)	0.002* (0.001)	0.002** (0.001)
Elevation	-0.0003*** (0.00004)	-0.0003*** (0.00004)	-0.0004 (0.0002)	-0.0004** (0.0002)
Population	0.00001*** (0.00000)	0.00001*** (0.00000)	-0.0001 (0.0001)	
Line-of-sight	-0.0003** (0.0001)		-0.001 (0.001)	
GECON	-0.476*** (0.031)	-0.476*** (0.031)	0.051 (0.193)	
Urban dist.	0.001*** (0.0001)	0.001*** (0.0001)	0.0002 (0.0004)	
Landcover	0.197*** (0.012)	0.192*** (0.012)	-0.228*** (0.056)	-0.235*** (0.056)
Prev. violence	-0.017*** (0.001)	-0.018*** (0.001)	0.013** (0.006)	0.013** (0.006)
Constant	-0.714*** (0.144)	-0.653*** (0.140)	1.670** (0.835)	1.701*** (0.606)
Observations	22,192	22,192	1,193	1,193
Log Likelihood	-12,324.430	-12,326.610	-705.659	-707.187
Akaike Inf. Crit.	24,668.870	24,671.220	1,431.318	1,426.374

Note:

*p<0.1; **p<0.05; ***p<0.01

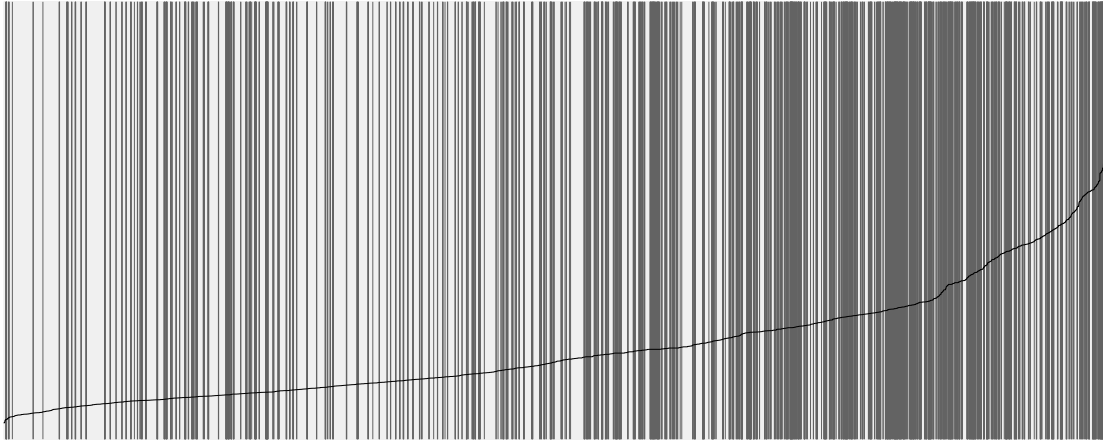
Table 3: Robustness test with the casualty counts of the SIGACT data. Note the quadratic relationship between distance from capital that corresponds to the findings from the GED analysis. The small coefficients are due to distances being coded in kilometers instead of being normalized.

	<i>Dependent variable:</i>		
	Civilian casualties (SIGACT)		
	(1)	(2)	(3)
Dist. Kabul (km)	-0.002** (0.001)	-0.003** (0.001)	-0.005*** (0.001)
Dist. Kabul ²	0.00000* (0.00000)	0.00000** (0.00000)	0.00001*** (0.00000)
Friendly cas.	0.367*** (0.043)	0.369*** (0.043)	0.331*** (0.045)
Enemy cas.	0.025** (0.012)	0.025** (0.012)	0.055*** (0.012)
Population	0.00004*** (0.00000)	0.00004*** (0.00000)	
Landcover	0.322*** (0.026)	0.322*** (0.026)	
GECON	-0.129* (0.066)	-0.128* (0.066)	
Pashtun	-0.057 (0.093)		
Hazara	-0.879*** (0.331)		
Constant	-3.964*** (0.315)	-4.002*** (0.300)	-0.575*** (0.122)
Observations	28,919	28,919	28,920
Log Likelihood	-11,855.140	-11,858.240	-11,995.720
θ	0.037*** (0.001)	0.037*** (0.001)	0.033*** (0.001)
Akaike Inf. Crit.	23,730.290	23,732.480	24,001.450

Note:

*p<0.1; **p<0.05; ***p<0.01

Insurgent violence (random subset of 10,000 events)



Incumbent violence (3,633 events)

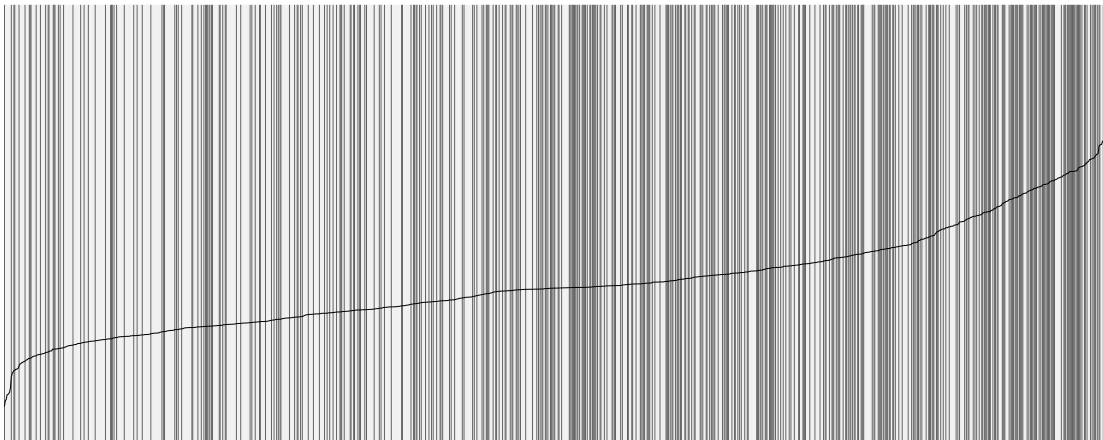


Figure S1: The separation plots above visualize in-sample predictive performance of the binary dependent variable models for the SIGACT study. Ideally All vertical lines (instances of indiscriminate violence) would be placed on the right, i.e. where the predicted probabilities of indiscriminate violence are highest (horizontal line). Note that the in the case of insurgent violence, the corresponding model performs better than in the case of incumbent violence.