SUPPLEMENTARY INFORMATION: GEOGRAPHY, OUTCOME, AND CASUALTIES

Summary statistics for the main independent variables

The following table gives an overview of the main independent variables, as well as their distributions, and definitions. Please note that all variables except TBI were taken from a post-1970 subset of the Lyall and Wilson (2009) study.

	Mean	Std. Dev.	Minimum	Maximum	N. Obs	Definition
TBI	0.67	0.10	0.49	0.93	65	Main independent variable
REGIME	-1.54	6.46	-10.00	10.00	65	Polity2 score*
SUPPORT	1.03	0.85	0.00	2.00	65	Codes external support for the rebels*
POWER	-1.82	1.63	-4.76	2.82	65	Cumulative index of state capabilities (Log)*
ENERGY	-0.83	1.69	-6.05	2.33	65	Incumbent energy use divided by population (Log)*
OCCUPY	0.09	0.29	0.00	1.00	65	Dummy coding if country is was occupied*
ELEVATION	899.20	1087.31	-20.75	4902.00	65	Average elevation in the conflict area (Log)*
DISTANCE	471.13	1616.52	1.00	12598.47	65	Km from the incumbent capital to the conflict country*
COLDWAR	0.60	0.49	0.00	1.00	65	Dummy variable for the Cold War*
MECH.	2.95	1.07	1.00	4.00	65	Soldiers per motorized vehicle*
HELI.	0.26	0.44	0.00	1.00	65	Dummy variable for whether incumbent uses helicopters*
TRADE	-3.46	0.89	-5.99	-0.47	65	Exports+Imports as a share of GDP (Log)*
LANG.	7.75	7.17	1.00	27.00	65	Number of languages in the conflict country*

Table 1. All variables marked with "*" have been taken from the Lyall and Wilson (2009) datset. *Mech, Polity, Power,* and *Energy* reflect the situation in the last prewar year.

Ruling out endogeneity in the TBI indicator

Figure 1 shows non-parametric Gaussian kernel estimates for the TBI values based on 1990 and 2010 population figures from the GPW dataset. Changes between 1990 and 2010 in the TBI distributions are marginal, both for conflict and non-conflict countries. KS tests also did not indicate systematic differences for the 1990 and 2010 distributions with $p = 0.94$ for the conflict cases and $p = 0.89$ for the non-conflict cases.

Two possible explanations for this non-effect spring to mind: either internally displaced persons (IDP) do not generally move toward the capital city or the remote periphery, but remain at medium distances from the capital city, or IDPs return to their original settlement areas once the fighting is over. The fact that no systematic variation can be found is important for ruling out out endogeneity in the empirical analysis.

Ruling out x in the explanatory variables

Table 2 shows correlations between the main independent variables. Please note that the variables POWER and HELI. show the strongest correlation at 0.62. No stronger correlation between the explanatory variables can be seen in the data.

Date: November 20, 2013.

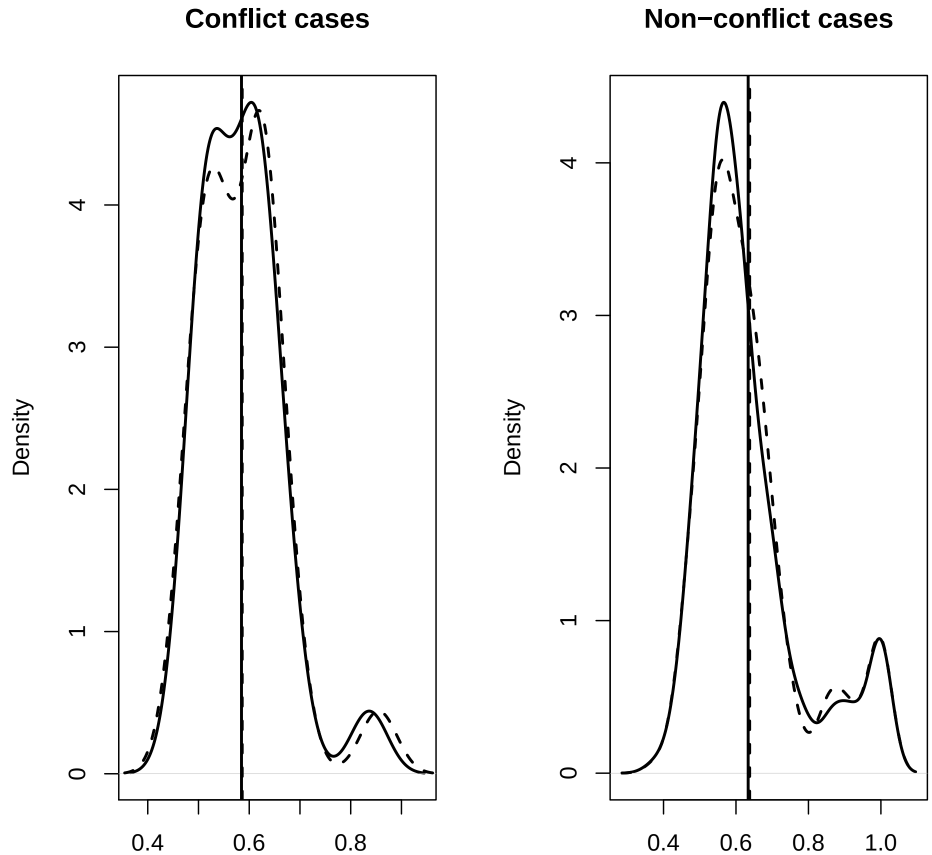


Figure 1. Densities of TBI values for conflict and non-conflict countries for 1990 (solid lines) and 2010 (dotted lines). The vertical lines mark mean values. Note that conflict countries have a marginally smaller TBI on average, but there is very little change over time despite the fact that conflicts took place.

	TBI	REG.	SUPP.	POWER	ENERGY	OCC.	ELEV.	DIST.	COLD.	MECH.	HELI.	LANG.	TRADE
TBI	1.00												
REG.	-0.15	1.00											
SUPP.	0.15	-0.30	1.00										
POWER	-0.10	0.17	-0.10	1.00									
ENERGY	0.01	0.18	-0.05	0.44	1.00								
OCC.	-0.00	0.27	0.12	0.36	0.22	1.00							
ELEV.	0.11	0.00	0.20	0.21	0.09	0.20	1.00						
DIST.	-0.01	0.22	0.00	0.47	0.28	0.48	0.07	1.00					
COLD.	0.04	-0.03	-0.08	0.04	-0.12	0.04	-0.01	-0.12	1.00				
MECH.	0.22	-0.03	-0.10	0.18	0.18	0.27	0.19	0.16	-0.21	1.00			
HELI.	-0.04	0.07	-0.15	0.62	0.29	0.29	0.04	0.29	-0.01	0.32	1.00		
LANG.	-0.12	0.04	0.28	-0.17	-0.26	-0.09	-0.09	-0.04	-0.06	-0.08	-0.08	1.00	
TRADE	-0.01	-0.01	0.04	-0.35	-0.05	-0.02	-0.56	-0.16	0.10	-0.14	-0.10	0.05	1.00

Table 2. Correlation matrix for the main independent variables. Note that correlations do not exceed 0.62 outside the main diagonal.

In order to prevent the variables POWER and HELI. from driving the statistical results, I excluded each variable from the full models and found that this did not change the results substantively (see table 3).

Table 3.

	Dependent variable:			
	outcome		casualties	
	ordered logistic		*negative binomial*	
	(1)	(2)	(3)	(4)
REGIME	0.115**	0.090*	−0.030	−0.031
	(0.049)	(0.047)	(0.027)	(0.027)
SUPPORT	0.755**	0.674*	1.104***	1.100***
	(0.382)	(0.369)	(0.220)	(0.221)
ENERGY	−0.359*	−0.174	−0.018	−0.046
	(0.204)	(0.199)	(0.099)	(0.109)
OCCUPATION	1.695	2.112*	−1.757***	−1.793***
	(1.281)	(1.231)	(0.647)	(0.664)
ELEVATION	−0.00002	−0.0001	−0.00005	−0.0001
	(0.0004)	(0.0004)	(0.0002)	(0.0002)
DISTANCE	0.0004	0.001*	0.0003	0.0002
	(0.0003)	(0.0004)	(0.0003)	(0.0004)
COLDWAR	−1.332**	−1.243**	1.684***	1.590***
	(0.555)	(0.565)	(0.327)	(0.326)
MECH.	−0.212	0.028	0.469***	0.539***
	(0.246)	(0.249)	(0.160)	(0.153)
LANG.	0.068	0.067	−0.024	−0.015
	(0.043)	(0.043)	(0.023)	(0.023)
log(TRADE)	−0.454	−0.430	−0.451**	−0.395*
	(0.404)	(0.397)	(0.206)	(0.206)
log(POPULATION)	−0.521***	−0.080	−0.125	−0.185
	(0.152)	(0.144)	(0.157)	(0.200)
TBI	−1.280***	−1.561***	83.246***	81.455***
	(0.136)	(0.025)	(21.205)	(21.690)
TBI^2			−61.875***	−59.930***
			(15.730)	(16.083)
HELI.	1.361**		0.463	
	(0.677)		(0.428)	
POWER		−0.278		0.178
		(0.186)		(0.184)
Constant			−21.636***	−20.285***
			(7.271)	(7.285)
Note:				*p<0.1; **p<0.05; ***p<0.01

Removing the SUPPORT variable

As an additional robustness test, I removed the *SUPPORT* variable, as foreign support to rebel organizations may be provided conditionally on military success in ongoing conflicts. Removing this variable weakens the effects of the main independent variable for the full and optimized outcome models, but it does not affect the results for the severity models (see the table on the next page).

	Dependent variable:								
	Outcome					Casualties			
	ordered logistic					*negative binomial*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
REGIME	0.063		0.063	0.067	0.060	-0.107^{***}		-0.116^{***}	-0.120^{***}
	(0.043)		(0.043)	(0.042)	(0.041)	(0.028)		(0.028)	(0.026)
POWER	-0.575^{***}		-0.575^{***}	-0.401^{*}	-0.691^{***}	0.280		0.288	
	(0.223)		(0.223)	(0.211)	(0.253)	(0.219)		(0.221)	
ENERGY	−0.181		−0.180	−0.203		−0.060		0.014	
	(0.209)		(0.209)	(0.208)		(0.120)		(0.121)	
OCCUPY	2.857^{**}		2.847^{**}	3.138^{***}	2.591^{**}	−0.742		−0.574	−0.441
	(1.284)		(1.283)	(1.210)	(1.128)	(0.700)		(0.715)	(0.620)
ELEV.	0.0001		0.0001	0.0002		0.0002		0.0001	
	(0.0004)		(0.0004)	(0.0004)		(0.0002)		(0.0002)	
COLDWAR	-1.336^{**}		-1.333^{**}	-1.373^{**}	-1.151^{**}	1.517^{***}		1.450^{***}	1.217^{***}
	(0.563)		(0.562)	(0.559)	(0.549)	(0.366)		(0.368)	(0.357)
MECH.	−0.300		−0.298	−0.329		0.454^{**}		0.360^{**}	0.583^{***}
	(0.274)		(0.274)	(0.274)		(0.177)		(0.176)	(0.161)
HELI.	1.665^{**}		1.663^{**}	1.632^{**}	1.369^{*}	0.857		0.230	
	(0.814)		(0.813)	(0.816)	(0.746)	(0.524)		(0.518)	
LANG.	0.091^{**}		0.091^{**}	0.099^{**}	0.096^{**}	−0.014		0.018	
	(0.040)		(0.040)	(0.040)	(0.040)	(0.025)		(0.025)	
ENERGY	−0.509		−0.508	−0.533	-0.559^{*}	−0.189		−0.211	−0.147
	(0.403)		(0.403)	(0.396)	(0.330)	(0.234)		(0.232)	(0.176)
log(POPULATION)	−0.116		−0.118	−0.230		-0.527^{**}		−0.331	
	(0.147)		(0.147)	(0.148)		(0.224)		(0.229)	
DISTANCE	0.001	0.001^{*}	0.001		0.001^{*}	−0.0004		−0.0003	
	(0.0004)	(0.0003)	(0.0004)		(0.0004)	(0.0004)		(0.0004)	
TBI		-1.134^{***}	-0.131^{***}	0.124	−0.701		91.928^{***}	55.157^{**}	55.762^{**}
		(0.136)	(0.027)	(0.079)	(0.851)		(26.789)	(24.320)	(23.382)
TBI^2							-65.922^{***}	-38.888^{**}	-38.988^{**}
							(19.747)	(17.968)	(17.252)
Constant						12.051^{***}	-21.023^{**}	−8.681	−12.724
						(2.575)	(8.978)	(8.117)	(7.843)
(CUTPOINT 1-2)	-2.8267^{***}	-4.2307^{***}	-5.9505^{***}	$-2.1855^{\cdot}$					
	(0.1915)	(0.1296)	(0.1883)	(1.1121)					
(CUTPOINT 2-3)	-0.8117^{**}	−1.0207	-2.7635^{***}	1.0055					
	(0.2861)	(0.5640)	(0.5525)	(1.3596)					
Observations	65	65	65	65	65	57	57	57	57
θ						0.675^{***} (0.107)	0.478^{***} (0.074)	0.693^{***} (0.111)	0.678^{***} (0.108)
Akaike Inf. Crit.						1, 273.503	1, 281.787	1, 275.439	1, 263.141

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 4.

	Dependent variable:						
	Outcome *ordered logistic*				Casualties *negative binomial*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
REGIME		0.128**	0.131**	0.119**		−0.012	−0.059**
		(0.052)	(0.051)	(0.049)		(0.028)	(0.027)
SUPPORT		0.971**	0.944**	0.972**		0.953***	0.853***
		(0.390)	(0.388)	(0.380)		(0.221)	(0.207)
POWER		−0.503**	−0.347	−0.671**		0.244	
		(0.231)	(0.222)	(0.266)		(0.196)	
ENERGY		−0.220	−0.250			−0.127	
		(0.219)	(0.222)			(0.107)	
OCCUPY		1.964	2.387*	2.000		−1.702**	−0.967
		(1.394)	(1.334)	(1.249)		(0.674)	(0.589)
ELEVATION		0.0001	0.0001			0.0001	
		(0.0005)	(0.0005)			(0.0002)	
DISTANCE	0.001**	0.0004		0.001		−0.00001	
	(0.0003)	(0.0004)		(0.0004)		(0.0004)	
COLDWAR		−1.036*	−1.073*	−1.022*		1.701***	1.605***
		(0.595)	(0.591)	(0.577)		(0.349)	(0.336)
MECH.		−0.022	−0.066			0.498***	0.587***
		(0.288)	(0.290)			(0.174)	(0.160)
HELI.		1.818**	1.801**	1.753**		0.376	
		(0.866)	(0.877)	(0.820)		(0.459)	
LANG.		0.078*	0.086**	0.081*		−0.031	
		(0.044)	(0.043)	(0.042)		(0.024)	
TRADE.		−0.579	−0.614	−0.607*		−0.371*	−0.323**
		(0.419)	(0.413)	(0.336)		(0.210)	(0.162)
log(POPULATION)		−0.163	−0.286*			−0.474**	
		(0.167)	(0.168)			(0.225)	
URBAN TBI	−2.903***	−8.103***	−8.398***	−7.820***	139.952***	50.237*	27.342
	(0.139)	(0.047)	(0.099)	(0.813)	(31.716)	(28.421)	(26.495)
URBAN TBI^2					−93.582***	−33.972*	−16.922
					(21.957)	(19.678)	(18.247)
(CUTPOINT 1-2)	−2.8267***	−4.2307***	−5.9505***	−2.1855·			
	(0.1915)	(0.1296)	(0.1883)	(1.1121)			
(CUTPOINT 2-3)	−0.8117**	−1.0207	−2.7635***	1.0055			
	(0.2861)	(0.5640)	(0.5525)	(1.3596)			
(Constant)					−41.325***	−8.497	−5.710
					(11.393)	(9.775)	(9.457)
Observations	63	63	63	63	55	55	55
θ					0.517*** (0.082)	0.899*** (0.150)	0.826*** (0.136)
Akaike Inf. Crit.					1,235.327	1,219.239	1,211.414
Note:							*p<0.1; **p<0.05; ***p<0.01

Urban Distance

The table above shows regression results for outcome and casualties based on a different operationalization of the TBI. Instead of distances to the capital city, distances to the nearest major city were calculated based on Nelson (2008). Please note that the results are essentially identical to the ones reported in the paper based on distances to the capital city. Due to computational limitations, I omitted the insurgency cases for Russia and the Soviet Union, which resulted in 63 instead of 65 observations for the outcome analysis an 55 instead of 57 for the casualty analysis.

Binary dependent variable models

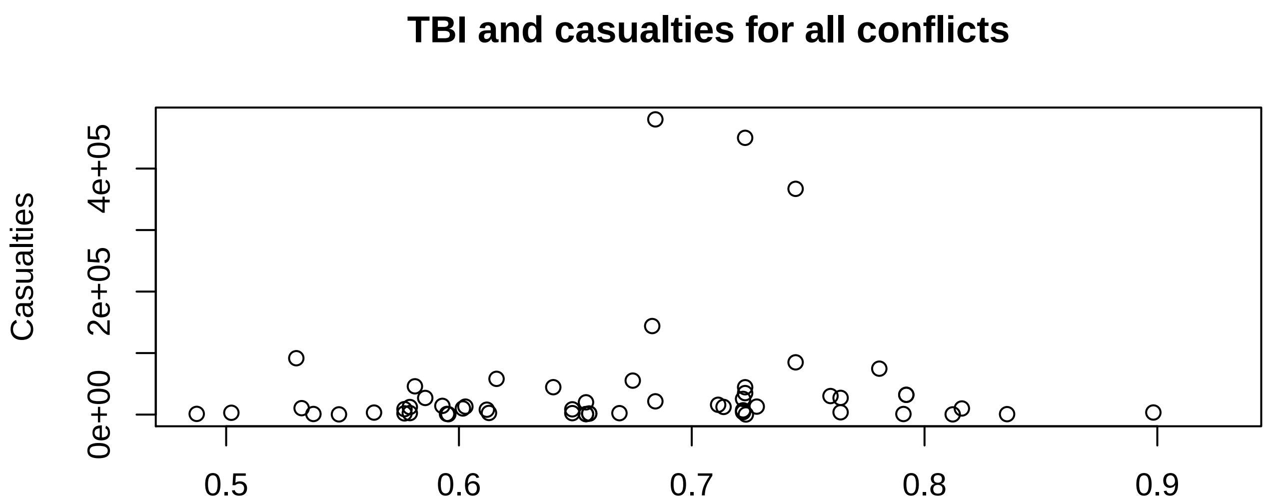
As an additional robustness check, I ran binary dependent variable models to predict incumbent success and defeat in insurgencies. Please note that the TBI variable is not significant, but the directions of the estimates correspond to the Ordinal Logit model in the paper. The ordinal information provided in the original study is therefore necessary for statistically significant results (see table below).

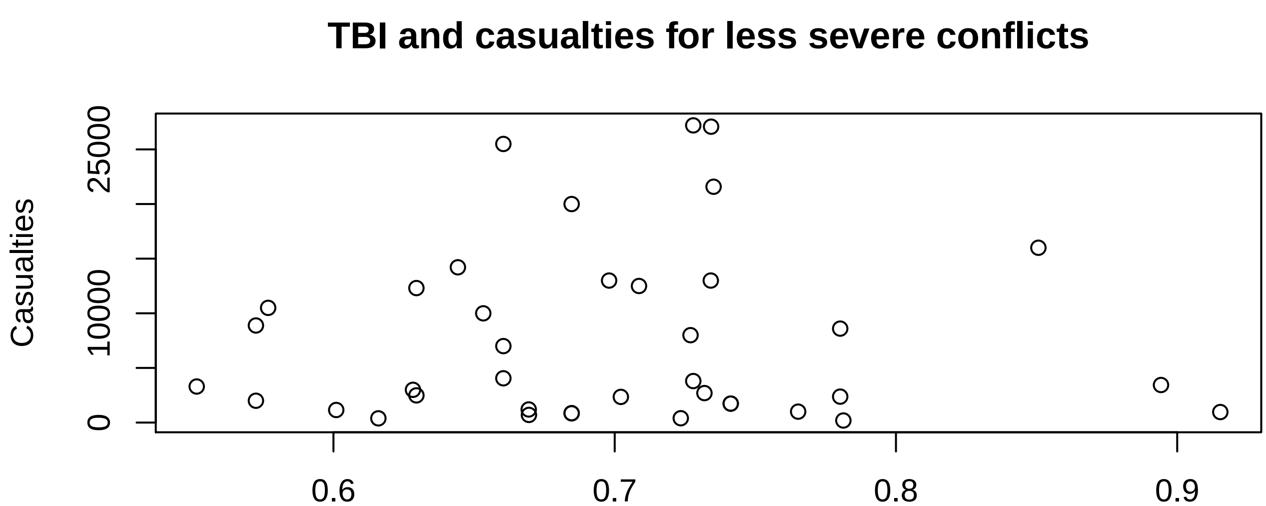
Table 5.

	Dependent variable:			
	win		defeat	
	(1)	(2)	(3)	(4)
REGIME	-0.271^{***}	-0.263^{**}	0.125	0.123
	(0.101)	(0.104)	(0.096)	(0.096)
SUPPORT	-1.926^{**}	-2.143^{***}	1.135^{*}	1.146^{*}
	(0.754)	(0.822)	(0.613)	(0.642)
POWER	1.300^{**}	1.665^{**}	-0.993^{**}	−0.843
	(0.545)	(0.715)	(0.489)	(0.579)
ENERGY	0.440	0.418	−0.557	-0.694^{*}
	(0.298)	(0.302)	(0.340)	(0.388)
OCCUPY	−17.692	−18.514	4.176^{*}	4.010^{*}
	(2,371.910)	(2,287.469)	(2.144)	(2.254)
DISTANCE	-0.006^{*}	-0.006^{*}	0.001	0.0004
	(0.003)	(0.003)	(0.001)	(0.001)
COLDWAR	0.883	1.100	-1.604^{*}	−1.421
	(0.858)	(0.949)	(0.880)	(0.917)
HELI.	-5.997^{***}	-6.319^{***}	0.619	1.115
	(2.063)	(2.214)	(1.497)	(1.599)
LANG.	−0.026	−0.001	0.070	0.075
	(0.064)	(0.072)	(0.057)	(0.058)
TRADE	−0.491	−0.557	-1.277^{**}	-1.477^{**}
	(0.513)	(0.539)	(0.529)	(0.610)
log(POPULATION)		−0.494		−0.534
		(0.565)		(0.608)
TBI		1.677		−5.034
		(4.236)		(5.768)
Constant	2.895	6.826	-9.959^{***}	−2.610
	(2.627)	(7.014)	(3.545)	(7.972)
Observations	65	65	65	65
Log likelihood	−22.775	−22.272	−20.185	−19.615
Akaike Inf. Crit.	67.550	70.543	62.370	65.230
Note:				*p<0.1; **p<0.05; ***p<0.01

Casualties and TBI

The bivariate relationship between TBI and casualties corresponds very well to the inverse U-shaped functional form discussed in the paper, but the sample is dominated by three cases with more than 40,000 casualties: the civil war in Cambodia, the anti-Soviet insurgency in Afghanistan, and the People's Mujahideen insurgency in Iran. The inverse U-shaped form is still visible in the data after excluding these cases:





Residual analysis

The table below shows residuals for outcomes and casualties for models 4 and 8, respectively. The data was obtained from Lyall and Wilson (2009). Only those cases are shown for which casualty counts could be established.

Table 6 – Continued on next page

No.	Incumbent	Insurgent	Start	End	Dev. casualties	Dev. outcome
1	Soviet Union	Afghanistan	1980	1989	-283481	0
2	Cambodia	Khmer Rouge	1978	1992	-218836	0
3	Iran	MEK	1979	2001	-166009	0
4	Lebanon	Various	1975	1990	-83286	0
5	Mozambique	RENAMO	1976	1992	-66069	0
6	Cambodia	FUNK	1970	1975	-60710	-2
7	Congo	Cobras, Ninjas	1997	1999	-56896	0
8	El Salvador	FMLN	1979	1992	-30296	-1
9	Uganda	NRA	1981	1987	-22872	-1
10	Nicaragua	FSLN	1978	1979	-15338	-1
11	Tajikistan	UTO	1992	1997	-14847	1
12	Peru	Sendero Luminoso	1980	1999	-8892	0
13	Liberia	NPFL	1989	1997	-7462	0
14	Turkey	Kurds	1983	1999	-6438	-1
15	Sri Lanka	LTTE	1987	1989	-4536	0
16	Russia	Chechens	1994	1996	-3285	0
17	Burundi	FDD	1993	2005	-2311	1
18	Sierra Leone	RUF	1991	1999	-1463	1
19	Azerbaijan	Armenia	1992	1994	-848	0
20	Sri Lanka	LTTE	1983	1987	-580	-1
21	Liberia	LURD/MODEL	1999	2003	-491	-1
22	Pakistan	Baluchi	1973	1977	-180	0
23	India	Sikhs	1984	1994	-40	0
24	Pakistan	Mohajirs	1993	1999	-17	0
25	Papua New	BRA	1988	1998	498	1
26	Moldova	Dniester	1992	1992	1406	-1
27	Ivory Coast	PMIC	2002	2005	2087	0
28	Israel	Palestinian	1987	1993	2588	0
29	Guinea Bissau	Mil factions	1998	1999	2747	-1
30	Burundi	Hutu rebels	1972	1972	3027	1
31	Syria	Muslim Brotherhood	1980	1982	3173	0
32	Ethiopia	Eritrea	1974	1991	4540	0
33	Iraq	Kurds	1991	1991	5431	0
34	Rwanda	ALiR	1994	2000	6755	0
35	Sri Lanka	JVP	1971	1971	7817	1
36	Nicaragua	Contras	1981	1988	7868	0
37	Iraq	Kurds	1980	1988	7958	0
38	Argentina	ERP	1973	1977	8551	1
39	Georgia	Abkhazia	1992	1994	8827	0
40	CAR	Factions	1994	1997	12161	0
41	Chad	Rebels	1994	1998	13402	1
42	Bangladesh	Shanti Bahini	1976	1997	14213	-1

Continued on next page

Table 6 – continued from previous page

No.	Incumbent	Insurgent	Start	End	Dev. casualties	Dev. outcome
43	Mali	Tuaregs	1989	1995	18374	0
44	Yemen	Socialist Party	1986	1987	18719	0
45	Israel	Syria	1982	1982	23237	1
46	Jordan	Fedeyeen	1970	1970	24853	0
47	DRC	FLNC	1977	1978	28491	1
48	Djibouti	FRUD	1991	1994	32299	0
49	Zimbabwe	ZANU	1972	1979	32770	0
50	Iraq	Kurds	1974	1975	37958	0
51	Morocco	Polisario	1975	1989	44045	0
52	Afghanistan	Taliban	1992	1996	76765	0
53	Somalia	SSDF	1981	1991	85715	1
54	Iran	KDPI	1979	1996	93239	1
55	Chad	Libya/Frolinat	1975	1988	113166	0
56	Afghanistan	Afghans	1978	1979	224999	-1
57	Sudan	SPLM, SPLM-faction	1983	2004	243855	0

The table above shows the results of a residual analysis for casualties and outcomes. Analyzing residuals is beneficial, as the results indicate which subset of cases is best captured by the presented models. The results are sorted by residuals for the prediction of casualties. The table also shows deviations between actual and predicted outcomes according to the introduced ordinal scale. The table conveys an intuition for the abilities and limitations of the econometric analysis and I will briefly discuss cases that represent both. At first glance, the overall prediction of ordinal outcomes seems to work rather well: In 34 of the 57 cases, the predicted outcomes correspond to the the actual outcomes. In 27 cases, a deviation of one between actual and predicted outcomes was found. In only one case does the model predict the opposite of the actual outcome: the Cambodian civil war. The casualty model also underpredicts the number of casualties at a margin of almost 61,000 (see row 6) as well as the severity of the Cambodian insurgency after the Vietnamese invasion (row 2). In these cases, the deviation between actual and predicted severity can be explained by the extraordinary historical context of the Vietnam War and the Cambodian genocide separating the two conflicts temporally. The underlying theory assumes a peripheral insurgent movement and the conventional military forces to be the main dyad in the conflict. The political violence that engulfed much of southeast Asia after the start of US combat operation in Vietnam in 1965 deviates from this assumption. In both Cambodian cases in the sample, superpowers were involved in aiding both sides. Although I control for external rebel support in the regression analysis, the corresponding binary indicator cannot account for the magnitude of the support in these cases. Two cases that are almost perfectly predicted are the insurgencies in Pakistan (rows 22 and 24). In both cases, and especially with regard to the multi-staged Balochistan conflict, the theoretical argument closely matches the empirical reality: a peripheral uprising in Balochistan fighting the geographically remote government in Islamabad. The model correctly predicts a ceasefire to result from these cases, and the actual and predicted severities match closely. The case that seems most difficult to predict (in terms of casualties) is Afghanistan (rows 1 and 56). This effect might nevertheless be due to the coding choice by Lyall and Wilson (2009): The 1979 civil war before the Soviet invasion and the 1980-1989 anti-Soviet insurgency are coded as separate conflicts. The severity of the first is overestimated by the statistical model, while the severity of the second is underestimated. If those geographically congruent and temporally adjacent wars were combined in the sample, the corresponding prediction would be more accurate.

Underlying assumptions as points of departure

The presented results, based both on the "peripheral" and the "urban" variant of the TBI, rely on specific assumptions about the conflict process that should be made explicit here. Of course, for population distance to matter in insurgencies, major parts of the population have to be affected by conflict. Thus, the first assumption is that conflict eventually takes place along the center-periphery line or from the most rural places towards the major cities. Moreover, the design assumes that the quality of violence declines as a function of distance from these power centers and that indiscriminate violence leads to reactive support for the adversary. Testing these assumptions would be natural points of departure for follow-up studies.