What follows is a discussion of variable coding procedures and robustness tests for *External Territorial Threat, State Capacity, and Civil War*.

1 Notes on Research Design and Coding

In this section, we outline, in considerable detail, how we coded the variables used in both Table I and Table II of the manuscript. The variables for the state capacity models were largely derived from standard models of resource extraction (e.g. Cheibub 1998, Thies 2004, 2010) and civil conflict (Fearon and Laitin 2003, Collier and Hoeffler 2004). However, very few variables come directly from these models, and much of our efforts started from scratch. We detail our methods here.

1.1 The Sample

While our state capacity and civil conflict models were based on multiple studies, we looked carefully at Fearon and Laitin's (2003) conflict models in particular. However, we wanted our sample to be more universal than the sample contained in Fearon and Laitin's seminal study.

Some of the case exclusions in Fearon and Laitin's study were rather unfortunate. For example, Belize and Suriname were excluded by Fearon and Laitin, ostensibly because their respective populations fall below the threshold for which the Polity project begins coding a state's institutional arrangements. However, Belize has faced a credible challenge from Guatemala regarding its sovereignty since Belize's formal independence. Suriname has faced internal challenges to the government's central authority, even resulting in a UCDP-PRIO armed intrastate conflict in 1987. We want to consider these important cases and try to explain them. Further, countries like Qatar and Equatorial Guinea were excluded because their ethnic and religious fractionalization scores were not coded in Fearon's (2003) fractionalization data. Again, the exclusion of these cases is unfortunate, considering the location of both countries and the pressures their respective central governments face in order to retain their authority.

We include in our sample every state in the international system that meets CoW's state system membership requirements (Correlates of War Project 2011) between 1946 and 2007, with only three exceptions: East Timor, Montenegro, and Zanzibar. In the case of East Timor, no ethnic or religious fractionalization estimates were found by Alesina et al. (2003), and none could be reasonably imputed. Montenegro was omitted because of lack of country-years. Zanzibar was omitted for both reasons: lack of fractionalization data and just two country-years in the sample (1963-1964).

All told, we improve Fearon and Laitin's (2003) data by considering 36 new countries over nine additional years.

1.2 Modeling Notes

We refer the reader to our manuscript, which provides, in detail, a description of the coding procedures for the four dependent variables. We also offer the following clarifications in this section of the appendix:

First, we code year 2 and subsequent years as "ongoing" for the conflict variables. For example, the first phase of the Second Laotian War (intrastate war # 756) began in 1963 and ended in 1968. The onset was coded as 1963 for Laos' referent country-year. "Ongoing" values for Laos were coded for 1963, 1964, 1965, 1966, 1967, and 1968. We take the lag of this ongoing variable and omit the observations that equal one. This effectively drops all ongoing conflict observations. One anonymous reviewer asked that we do this because the chances of civil conflict onset are near

zero if there is already a civil conflict ongoing. This is not necessarily true for the sample of CoW intrastate wars, but it is always true in the sample of UCDP-PRIO intrastate armed conflicts.[1] Analyses that do not drop ongoing cases of civil conflict do not differ substantively from those presented in the manuscript.

Second, splines could not be coded for the UCDP-PRIO sample before 1946, given the hard left-censoring of that data set. However, we coded peace years and splines for the sample using CoW's intrastate wars in the years prior to 1946. We take advantage of the longer temporal domain of the CoW project to add this information to the model.

1.3 Independent Variables

We use this section to further describe the coding procedures for our independent variables and provide summary statistics of these variables in Table 1.

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.
CoW Type 4 Civil War Onset	0.012	0.109	0	1
CoW Type 4 Civil War Ongoing	0.038	0.192	0	1
UCDP Intrastate Government Armed Conflict Onset	0.018	0.133	0	1
UCDP Intrastate Government Armed Conflict Ongoing	0.073	0.261	0	1
GDP	8.17	1.12	5.139	11.343
Population	8.651	1.889	2.197	14.097
Mountainous Terrain	2.039	1.408	0	4.557
Non-Contiguous State	0.166	0.372	0	1
Oil	0.126	0.331	0	1
UDS Democracy Score	0.003	0.969	-2.103	2.117
UDS Democracy Score^2	0.940	0.942	0	4.481
Political Instability	0.1	0.3	0	1
Ethnic Fractionalization	0.43	0.264	0	0.930
Religious Fractionalization	0.422	0.238	0	0.86
Territorial Claim	0.403	0.49	0	1
Post-Territorial Claim, Years 1-15	0.247	0.431	0	1
Post-Territorial Claim, Years 16-30	0.156	0.363	0	1
U.S. Economic Aid	11.589	8.042	-19.806	24.212
U.S. Military Aid	7.615	7.712	0	23.337
Inflation	0.004	0.996	-4.369	6.023
Agriculture	-0.082	16.573	-20.797	73.138

1.3.1 GDP

Fearon and Laitin (2003) used real per capita income data drawn from version 5.6 of Penn World Tables with measurements fixed at 1985 constant prices. Originally, our efforts were tailored at

[1]There were eight unique CoW intrastate war onsets over issues of central control for countries that already had an intrastate war ongoing between 1816 and 2007. Three were observed in the period between 1946 and 2007 (Afghanistan 1978 [Saur Revolution, First Afghan Mujahideen Uprising], Iran 1979 [Overthrow of the Shah, Anti-Khomeini Coalition] and Somalia 1991 [First Somalia, Second Somalia]).

simply extending this data forward. However, we could not get data in 1985 constant prices and chose to start over. Here, we detail our coding procedure.

We used version 5.0 of Gleditsch's (2002) expanded trade and real GDP data, which is patterned off Penn World Tables data and, indeed, draws data from it. After downloading this data, we duplicated the deterministic imputation procedure outlined by Fearon and Laitin (2003) in their appendix. Their procedure needed to accomplish multiple things. First, their imputation procedure needed to derive estimates for 1946 to 1949, and project from 1993 to 1999. Our imputation procedure simply needed to fill in small patches of missing data.

Therefore, like Fearon and Laitin, we regress the natural logarithm of real GDP (from Gleditsch's (2002) data set) on the natural log of the primary energy consumption for a state in the same year (see: Singer, Bremer and Stuckey 1972). Since we were most concerned with prediction, rather than parsimonous models of explanation, we also included country and year fixed effects. The natural log of both the independent and dependent variable correlate moderately ($r = .40$), but the predicted values from the regression model correlate with observed values of the natural logarithm of real GDP very highly ($r = .97$).

After running this model, we used a simple if-else statement to get our GDP variable. If the natural logarithm of the real GDP variable from Gleditsch's data was missing (e.g. country-years before 1950), we used our predicted value. If the natural logarithm of the real GDP variable was observed, we retain it. In Table I and II, this variable is lagged one year.

As a minor note, Penn World Table's (version 7.1) data on China includes the Maddison-Wu revision for macroeconomic indicators otherwise relying on China's official statements of its growth rates (see: Maddison and Wu 2008). This "China version 2" variable is coded as representing China in our sample.

1.3.2 Oil

Fearon and Laitin (2003) consider a state to be an oil state if fuel exports constitute more than a third of total merchandise exports. Our coding of this data largely retains their original observations but must consider the case of states that may meet this threshold after 1999 that did not meet this threshold *in* 1999. Further, we have to consider 36 more countries than those that made it into Fearon and Laitin's sample.

In only one application did we observe a new oil state that was not an oil state in 1999. This was Ecuador, starting in 2000. Looking at Fearon and Laitin's original data, we previously observe Ecuador as an oil exporter through much of the 1990s, though its fuel exports fell below their threshold. However, World Bank data show that Ecuador again met this threshold starting in 2000, and through 2007, but Fearon and Laitin are essentially right when they say, in their appendix, that oil states that go online tend to stay online.

The other issue was coding for 36 new countries that appear in our sample. Some of these were easy. Our introduction of Qatar and Brunei into the sample introduced two new oil states that met this threshold since independence. Others were less intuitive. Carefully combing through the data revealed some brief stages of states being considered "oil states" per this arbitrary criterion. Bahamas can be considered an oil state from independence in 1973 to 1988. Belize can be considered an oil state in 1999. Seychelles can be considered an oil state in 1984 and 1985.

1.3.3 Ethnic Fractionalization, Religious Fractionalization

We discuss in the research design of the manuscript some of the limitations of using Fearon's (2003) ethnic and religious fractionalization data. We opted to use Alesina et al.'s (2003) ethnic and

religious fractionalization data instead because their sample is much more inclusive than Fearon's (2003) data. Our fractionalization measures largely come from Alesina et al. (2003), with the following exceptions.

Alesina et al. (2003) do not code for East Germany, Czechoslovakia, Yemen Arab Republic, Yemen People's Republic, and South Vietnam, and we use Fearon's (2003) estimates instead for these countries. Alesina et al. (2003) do not distinguish between Soviet and post-Soviet Russia, but Fearon (2003) do. Therefore, we use the latter's estimates for Soviet Russia in Table II of the manuscript. Alesina et al. (2003) is used for post-Soviet Russia. Alesina et al. (2003) are missing a religious fractionalization measure for Serbia after the dissolution of Yugoslavia. We use Fearon's estimate for religious fractionalization measures for Serbia after 1990.

Fearon (2003) omit both Sao Tome and Principe and Maldives. For Sao Tome and Principe, Alesina and his colleagues have a religious fractionalization measure but not an ethnic fractionalization measure; they omit the Maldives for both measures. We derive an estimate an ethnic fractionalization value from Ellingsen's (2000) data and use the well-known measure of fractionalization, discussed by Alesina et al. (2003, 158-159).[2] Ellingsen's (2000) data for the ethnic fractionalization of Sao Tome and Principe come from Demographic Yearbook data via the United Nations.

Maldives was a rather simple and uncontroversial imputation. Almost the entirety of the country is Muslim and consists of two ethnic groups: native Maldivians and a tiny, but important, group of ethnic Sinhalese. Our estimates come from the Joshua Project.[3]

1.3.4 New State

Like Fearon and Laitin, we code a state as a "new state" if it is in the first two years of state system membership, using CoW's criterion for entry into the international system.

1.3.5 Mountainous Terrain

We followed Fearon and Laitin's discussion of this variable in their appendix very closely. Their estimates come mostly from the World Bank's "Economics of Civil War, Crime, and Violence" project. We were not able to obtain this original data, but we were able to to replicate Fearon and Laitin's treatment of it.

The geographer A.J. Gerard created estimates for the percentage of mountainous terrain in a country on behalf of this World Bank project. Fearon and Laitin add one to these estimates and take the natural log. For countries that Gerard did not measure, but were included in their sample, Fearon and Laitin consult CIA Factbook data to get the elevation difference between the highest and lowest points in a country. Fearon and Laitin took the natural logarithm of this elevation difference variable and regressed the natural logarithm of Gerard's mountainous terrain percentage on it.

Thereafter, Fearon and Laitin identify a cutoff of 6.5 in the log(elevation difference) variable. Above this threshold, the two variables correlate very highly ($r = .79$). Therefore, for countries not measured by Gerard, Fearon and Laitin code the percent mountainous terrain with an if-else statement. If the data point is missing in Gerard's variable and the natural log of the elevation

[2] Formally, it is computed as 1 minus the sum of squared group shares. It constitutes the probability that two randomly selected individuals from a population belong to different groups:

$$FRACTIONALIZATION_i = 1 - \sum_{i=1}^{n} s_i^2 \quad (1)$$

[3] http://www.joshuaproject.net/countries.php?rog3=MV

difference is above 6.5, Fearon and Laitin estimate the natural log of the percentage of mountainous terrain from the expected value derived from the regression. If the data point is missing on Gerard's variable and the the natural log of the elevation difference is below 6.5, Fearon and Laitin code a state as not being mountainous at all. We mirror this procedure for the 36 new countries that appear in our sample, though we do not make a distinction between the percentage of mountainous terrain that Gerard originally measured and the percentage of mountainous terrain that Fearon and Laitin imputed.

1.3.6 UDS Democracy Score, UDS Democracy Score^2, Instability

We refer the reader to the research design section of our manuscript for discussion of these variables, as well as how they are coded. We lag the UDS Democracy Score and its square term in our analyses. The instability variable is already lagged by virtue of our coding procedure.

1.3.7 Inflation, Agriculture

Inflation and agriculture data come from World Bank estimates, as we mention in the manuscript. We center the agriculture variable on its mean and center the inflation variable at the mean of its natural logarithm. We do this to provide for a meaningful baseline in Table I. Any use of cross-national, time-series inflation data should correct for the effect of positive skew like this. There was not a significant right-hand skew in the agriculture variable to necessitate making a logarithmic transformation.

1.3.8 U.S. Economic Aid, U.S. Military Aid

We refer the reader to the manuscript for our treatment of this variable, though we offer the following clarifications. Greenbook data from the United States Agency for International Development (USAID) are messy and we take care to correct for potential problems. First, USAID Greenbook data list only countries that have ever received aid, which omits observations like Czechoslovakia and East Germany. Clearly, these countries are eligible for aid but did not receive it. Further, USAID does not necessarily sync their data with state system membership, which is why countries like Bahamas can be listed as receiving USAID while still part of the United Kingdom. USAID also sometimes does not specify who exactly was receiving aid, misleading data analysts to assume (for example) that Vietnam was receiving U.S. economic and military aid during the 1960s without ever specifying *which* Vietnam was receiving it. The same problem was observed for China during its civil war. Finally, we followed Baccini and Urpelainen's (2012) treatment of negative values (repayments) that emerge in the economic aid data. We take the natural log of the absolute value of these repayments and multiply it by -1.

1.3.9 UCDP Territorial Conflict

We refer the reader to the manuscript for our treatment of this variable.

1.3.10 Population

We had multiple data sources to use for the population measure, including Correlates of War data (Singer, Bremer and Stuckey 1972) and data contained in Gleditsch's (2002) expanded trade and real GDP data. We chose the Correlates of War data and extracted the *total* population variable for use. We took the natural logarithm of this variable and introduced it into our model as a one-year lag.

1.3.11 Non-Contiguous State

Fearon and Laitin classify a state as "non-contiguous" if there is territory in the state that holds at least 10,000 people and is separated from the territory containing the national capital by either land (e.g. Alaska [U.S.], Kaliningrad [Russia]) or at least 100 kilometers of water (e.g. Greenland [Denmark], Hawaii [U.S.]). We merge this variable from Fearon and Laitin's (2003) data set. For new countries in the sample, we measure carefully whether they satisfy this criteria. Cape Verde counts as non-contigous by this classification (Santo Antao), as do Comoros (Anjouan), Vanuatu (Espiritu Santo), Solomon Islands (Choiseul) and Tonga (Vava'u).

1.3.12 Territorial Claim

As we discuss in the manuscript, we use Huth and Allee's (2002) territorial claims data from 1919-1995 as an indicator of territorial threat in our civil conflict models. The manuscript discusses what the territorial claim data are, as Huth and Allee measure it. Here, we offer further clarification on our extension of these data.

In order to better capture territorial claims that could be threatening to another state, we restrict our focus to claims involving land-contiguous neighbors (see Gibler's (2012) for why this is appropriate). In most applications, this was very clear when looking at Huth and Allee's territorial claim data. In fact, only inter-war Germany and North/South Vietnam required coder judgments of contiguity.

The important thing to address in this appendix is how we dealt with right-censoring of the data. The territorial claims data end in 1995, but our temporal domain extends to 2007. Thus, in order to make the best use of the volume of conflict data we have acquired since Huth and Allee's book went to press, and since the end of MID 3.0's temporal domain (Jones, Bremer and Singer 1996), we chose to look carefully at Huth and Allee's territorial claim data and extend these observations from 1995 to 2007. Our procedure carefully considered their treatment for when claims between states end. All cases ongoing in 1995 were extended to 2007, with these following exceptions. In these cases, we were able to definitively locate when a territorial claim between land-contiguous neighbors ended.

- We code Ecuador and Peru as ending their territorial claim in 1998. The last flare-up between both sides over their disputed border was in 1995, the Cenepa Valley "mini-war". In 1998, both sides reached an agreement on the issue and peace has prevailed between Ecuador and Peru as a result. In short, both sides upheld the previously disputed terms of the 1942 Rio Protocol, though Ecuador's victory in the 1995 clash—its first of any sort against Peru in its history—allowed Ecuador to save face (cf. Simmons 2009).

- We code the U.K. and Ireland as ending their dispute over Northern Ireland with the Good Friday Agreement in 1998.

- Russia has outstanding issues with Estonia at present, which, among other factors (e.g. Romania, China), codes Russia as being involved in a land-contiguous territorial claim through the entirety of our sample. However, several countries involved in land-contiguous territorial claims with Russia settled them after 1995. Ukraine's issue with Russia over the division of the Soviet Black Sea naval fleet was settled on May 28, 1997 (Sherr 1997). Under the terms of this bilateral treaty, Russia and Ukraine agreed to a compromise: both sides split the fleet equally with Russia buying back the more modern ships in the fleet. Ukraine also entered into a lease with Russia for control of the port of Sevastopol.

- Georgia had an outstanding issue of military base rights with Russia, which was coded as beginning in 1993. It ended with an agreement in 2005 (Sokov 2005, Antidze 2007) that closed Russian military bases in Georgia.
- Namibia and Botswana had a dispute over Kasikili/Sedudu Island that was settled by ICJ verdict in 1999 (Johnson 2000).
- Yemen's dispute with Saudia Arabia over the "Empty Quarter" with Saudi Arabia (Rub al Khali) ends with the Treaty of Jeddah in 2000 (Al-Enazy 2002).
- Kazakhstan's border dispute with China was formally resolved in 1998 (Zardykhan 2002).

Countries with ongoing (as of 2007), land-contiguous territorial claims, as measured in Huth and Allee's data, include: the United States, Canada, Belize, Guatemala, Venezuela, Guyana, Suriname, Brazil, Bolivia, Chile, Argentina, Uruguay, Spain, Croatia, Slovenia, Russia, Estonia, Latvia, Azerbaijan, Ghana, Togo, Cameroon, Nigeria, Democratic Republic of Congo, Somalia, Ethiopia, South Africa, Lesotho, Morocco, Sudan, Iran, Iraq, Egypt, Syria, Israel, Saudi Arabia, Kuwait, Afghanistan, Tajikistan, China, North Korea, South Korea, India, Bhutan, Pakistan, Bangladesh, Thailand, Laos, Vietnam, Malaysia, and Indonesia.

We considered several potential territorial claims that arose after 1995. However, none met our definitions for inclusion in our sample: (1) between contiguous states, and (2) a qualifying Huth and Allee claim. Further, any errors of omission should bias *against* our results since they would falsely inflate the baseline comparison of no territorial conflict with positive cases of territorial threat.

1.3.13 Post-Territorial Claim Years

We refer the reader to the manuscript for our treatment of this variable.

2 The Imputation Model

Table I of the manuscript uses multiple imputation to address the issue of missingness in the model. Missing data was less of a problem for the dependent variables than they were for two independent variables of interest: the consumer price index (inflation) and the value added of agriculture as a percentage of GDP. Since missingness affected more than 40% of the data, we felt multiple imputation was appropriate.

In particular, we were concerned about missingness on communist countries or countries otherwise aligned with the Soviet Union during the Cold War. As Maddison and Wu (2008) note, relying on China's official statements of growth records can be problematic. Dafoe (2011), and with colleagues (Dafoe, Oneal and Russett 2013), points out how perilous it is to gloss over non-random missingness in analyses using macroeconomic indicators and how improper imputation procedures can further bias results.

Uncertain if these issues were confounding our inferences about the effect of territorial threats on state capacity, we presented analyses using multiple imputation handled by the Amelia II program (Honaker and King 2010). In order to make the analyses tractable, we used a simple imputation model that otherwise conforms to the best practice of using at least all indicators that also appear in the regression model (Meng 1994, Rubin 1976, 1996).

Our primary concern, which motivated our use of multiple imputation, was missing data from communist countries. Therefore, our imputation model includes a variable coding whether and

when a country can be considered communist. This variable is coded as 1 for Cuba (1959-2007), Grenada (1979-1983), East Germany (1954-1990), Poland (1946-1989), Hungary (1949-1990), Czechoslovakia (1948-1990), Albania (1946-1992), Yugoslavia (1946-1992), Bulgaria (1946-1990), Romania (1947-1989), Russia (1946-1990), Yemen People's Republic (1967-1989), China (1949-1989), North Korea (1948-1994), Cambodia (1975-1979), Laos (1975-2007), and (North) Vietnam (1954-2007).

To improve imputation of our proxy for the connective capacity of the state, we draw data from the Relative Political Capacity project (Kugler and Tammen 2012). These variables measure the ratio of extraction relative to what is expected, given a set of economic characteristics for the state in a given year.

The imputation process made sure to impose logical bounds (between 0 and 100) on the government share and agriculture variables. In addition, the data were explicitly specified as time-series, cross-section with a second-order polynomial added to help 𝔸melia's predictive power.

We chose to include the imputed results in Table I of the manuscript because it was appropriate, given the issue of the data. The fact that results are basically identical for the territorial threat variable of interest when casewise-deletion is applied to missing data lends credence to both our argument and the imputation procedure. This is provided in Appendix Table 2.

3 Additional Models and Robustness Tests

We ran additional model specifications for analyses that are presented as Table II in the manuscript. As with the manuscript, all models presented here include standard errors clustered on the country.

One series of tests treats the conflict variables for what they are: rare events. The inclusion of 36 more countries over nine more years in our sample does make our sample more representative than standard insurgency models of conflict, but the inclusion of these countries, in addition to analyses that focus on just armed challenges to the central government's authority, flooded the conflict dependent variables with more 0s than 1s. As Table IV demonstrates, the onset of a CoW intrastate war or a UCDP-PRIO armed conflict over issues of the central government's authority is a very rare event. King and Zeng (2001) noted the problems that rare events introduce in our logistic regression models. We retained normal logistic regression models for the manuscript but estimated the same models using King and Zeng's (2001) rare events logit function in Stata. The results, shown in Appendix Table 3, do not differ from Table II in the manuscript.

Appendix Table 4 introduces ethnic and religious fractionalization measures largely derived from Fearon's (2003) data in lieu of the measures in Table II, which mostly come from Alesina et al.'s (2003) fractionalization data. Countries that Fearon does not code in his data are imputed from Alesina et al. (2003), except for the aforementioned cases of Sao Tome and Principe and Maldives. Appendix Table 4 is very similar to Table II of the manuscript. We observe that countries with a high level of ethnic diversity tend to experience more intrastate armed conflicts over government incompatibility, even if those armed conflicts do not appear to manifest in the higher intensity threshold of war. The effect of the territorial threat variables remain unchanged.

Appendix Table 5 further clarifies the relationship between territorial threat and civil conflict onset. Our argument is that the increase in state capacity that follows a territorial challenge to the state endures even after the claim itself is settled. Thus, states in the immediate wake of territorial peace are *still* unlikely to experience the onset of an armed challenge to the government's authority from within its borders. Central governments whose states have been at territorial peace indefinitely find it more difficult to connect with, and co-opt, the average citizen and sell costly militarization policies toward the public, who become more tolerant of widespread defiance to the government's

Table 2: The Effect of Territorial Threat on State Capacity

	Model 1 *Military Personnel* (*1947-2007*)	**Model 2** *Government Share* (*1950-2007*)
GDP (lagged)	0.266***	-0.182***
	(0.062)	(0.050)
Population (lagged)	0.897***	-0.076**
	(0.036)	(0.027)
Oil	0.152	0.091
	(0.144)	(0.099)
Ethnic Fractionalization	-0.678*	-0.183
	(0.264)	(0.199)
Religious Fractionalization	-0.620*	0.212
	(0.242)	(0.156)
Mountainous Terrain	0.033	0.019
	(0.037)	(0.027)
UDS Democracy Score (lagged)	-0.286***	0.014
	(0.067)	(0.047)
UDS Democracy Score^2 (lagged)	0.032	0.004
	(0.052)	(0.037)
Instability	0.089	-0.014
	(0.062)	(0.045)
Inflation (lagged)	0.058**	0.001
	(0.021)	(0.022)
Agriculture (lagged)	-0.006	-0.000
	(0.005)	(0.003)
U.S. Economic Aid	-0.010	-0.002
	(0.008)	(0.004)
U.S. Military Aid	0.004	-0.008*
	(0.006)	(0.004)
Ongoing Civil War (lagged)	0.136	0.218†
	(0.143)	(0.128)
UCDP Non-Territorial Conflict (lagged)	0.319†	0.038
	(0.173)	(0.058)
UCDP Territorial Conflict (lagged)	0.327*	0.393***
	(0.153)	(0.112)
Observations	3776	3716

Standard errors in parentheses
† p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table 3: The Effects of External Territorial Threat on Intrastate Conflict Onset (Rare Events Logit)

	Model 1 COW Civil War Onset *1946-2007*	**Model 2** COW Civil War Onset *1946-2007*	**Model 3** UCDP Intrastate Armed Conflict Onset *1946-2007*	**Model 4** UCDP Intrastate Armed Conflict Onset *1946-2007*
GDP (lagged)	-0.303†	-0.314*	-0.174†	-0.172†
	(0.157)	(0.153)	(0.102)	(0.101)
Population (lagged)	0.170*	0.153*	0.078	0.066
	(0.066)	(0.070)	(0.059)	(0.061)
Mountainous Terrain	0.202†	0.213*	0.175**	0.188**
	(0.108)	(0.106)	(0.067)	(0.066)
Non-Contiguous State	-0.202	-0.146	-0.577*	-0.535†
	(0.399)	(0.405)	(0.274)	(0.278)
Oil	0.375	0.413	0.300	0.323
	(0.360)	(0.356)	(0.274)	(0.266)
UDS Democracy Score (lagged)	-1.091**	-1.178**	-0.498*	-0.537**
	(0.386)	(0.403)	(0.206)	(0.206)
UDS Democracy Score^2 (lagged)	-1.220***	-1.251***	-0.840***	-0.840***
	(0.370)	(0.368)	(0.203)	(0.200)
Political Instability	0.821**	0.838**	0.717***	0.723***
	(0.273)	(0.269)	(0.211)	(0.215)
Ethnic Fractionalization	-0.022	0.040	1.045*	1.071*
	(0.641)	(0.662)	(0.455)	(0.443)
Religious Fractionalization	-0.093	-0.118	-0.300	-0.296
	(0.589)	(0.581)	(0.419)	(0.414)
Territorial Claim	-0.997***	-1.470***	-0.134	-0.487*
	(0.292)	(0.351)	(0.191)	(0.230)
Post-Territorial Claim, Years 1-15		-0.650†		-0.540*
		(0.375)		(0.256)
Post-Territorial Claim, Years 16-30		-0.727†		-0.424
		(0.422)		(0.268)
Observations	7,989	7,989	7,721	7,721

Standard errors in parentheses
† p<0.10, * p<0.05, ** p<0.01, *** p<0.001
Peace years, and cubic splines are omitted to save space.

Table 4: The Effects of External Territorial Threat on Intrastate Conflict Onset

	Model 1 COW Civil War Onset *1946-2007*	**Model 2** COW Civil War Onset *1946-2007*	**Model 3** UCDP Intrastate Armed Conflict Onset *1946-2007*	**Model 4** UCDP Intrastate Armed Conflict Onset *1946-2007*
GDP (lagged)	-0.286†	-0.286†	-0.186†	-0.182
	(0.163)	(0.158)	(0.111)	(0.111)
Population (lagged)	0.170**	0.149*	0.078	0.063
	(0.065)	(0.069)	(0.060)	(0.061)
Mountainous Terrain	0.214*	0.234*	0.178**	0.192**
	(0.107)	(0.106)	(0.064)	(0.063)
Non-Contiguous State	-0.234	-0.181	-0.751**	-0.720*
	(0.409)	(0.414)	(0.276)	(0.282)
Oil	0.348	0.378	0.388	0.412
	(0.350)	(0.347)	(0.277)	(0.269)
UDS Democracy Score	-1.155**	-1.239**	-0.548**	-0.584**
	(0.386)	(0.399)	(0.213)	(0.212)
UDS Democracy Score^2	-1.280***	-1.297***	-0.906***	-0.897***
	(0.363)	(0.358)	(0.198)	(0.194)
Political Instability	0.820**	0.841**	0.724***	0.731***
	(0.270)	(0.267)	(0.217)	(0.221)
Ethnic Fractionalization (Fearon)	0.033	0.213	0.546	0.637†
	(0.510)	(0.533)	(0.400)	(0.383)
Religious Fractionalization (Fearon)	0.274	0.246	0.152	0.142
	(0.633)	(0.620)	(0.484)	(0.476)
Territorial Claim	-1.025***	-1.515***	-0.138	-0.514*
	(0.286)	(0.356)	(0.192)	(0.230)
Post-Territorial Claim, Years 1-15		-0.700†		-0.589*
		(0.388)		(0.251)
Post-Territorial Claim, Years 16-30		-0.769†		-0.463†
		(0.421)		(0.265)
Observations	7,988	7,988	7,721	7,721

Standard errors in parentheses
† p<0.10, * p<0.05, ** p<0.01, *** p<0.001
Peace years, and cubic splines are omitted to save space.

authority. We use Appendix Table 5 to illustrate that the baseline category matters in explaining the relationship between interstate territorial threat and intrastate armed conflict.

In this table, the baseline category is states that have been at territorial peace for either 1-20 years or, similar to Table II in the manuscript, 1-30 years. In Model 1, we observe that states with ongoing territorial claims are unlikely to have an intrastate war. The coefficient for states that have been at territorial peace for more than twenty years is positive, but not significant. Model 2 uses the 30 year baseline and finds that not only are states with territorial claims less likely to have intrastate wars (when compared to the baseline of states that have not had them for 1-30 years), but states without territorial threats for more than 30 years are *more* likely to have intrastate wars when compared to this baseline category. This illustrates well Model 2 in Table II in the manuscript.

Models 3 and 4 of Appendix Table 5 uses these same baseline adjustments on the sample of UCDP-PRIO intrastate armed conflicts. We observe findings that clarify the effect of territorial threat on intrastate armed conflict onset. Model 3 shows that states with ongoing territorial claims are no more or less likely to experience the onset of an intrastate armed conflict at a lower level intensity *when compared to the baseline category of those states within twenty years of settling a territorial claim.* However, states more than twenty years removed from a salient claim to the state's territory are *more* likely to experience the onset of an armed conflict when compared to that same baseline category. Model 4 reveals a similar story using the 30-year baseline category. instead of the 20-year baseline category.

The story remains the same. States that have salient territorial threats see an increase in state capacity that decreases the likelihood of an internal challenge to the government's authority at *both* higher *and* lower levels of intensity. We use Appendix Table 5 to clarify how the baseline category matters in our statements about the relationship among territorial threat, state capacity, and intrastate conflict.

Appendix Table 6 tests whether the results of the civil war models from Table II in the manuscript differ when we also consider different types of civil war onset. We did this at the request of an anonymous reviewer who questioned that our analyses on government civil conflict onsets is a departure from what is best practice in understanding civil conflict onset. While we have good reason to believe that the theory we offer about external threat and state capacity is applicable to an understanding of armed challenges for control of the government, rather than secessionist conflicts, we examined whether our argument holds when we consider just the "Type 5" civil war onsets, or both the "Type 4" and "Type 5" wars. The "Type 5" onsets more closely resemble the old "territorial" civil wars from older versions of the Correlates of War civil war data set.

As Appendix Table 6 shows, our expectations about the relationship among external threat, state capacity, and civil war onset are largely affirmed in analyses of just the "Type 5" civil wars, as well as both types of civil wars. Model 1 and Model 2 offer support consistent with our argument even when we just look at civil wars over local issues. As we reiterated in our analyses of Appendix Table 5, the baseline category matters in our assessment of the effect of territorial threat on civil war onset. When the baseline is better specified to be countries who have been at territorial peace for over thirty years, the negative relationship between territorial threat and civil war onset becomes more apparent. When both types of civil wars are considered, we actually find stronger evidence for our argument than what we present in Table II in the manuscript.

Appendix Table 7 estimates Table II from the manuscript but only for the sample of states in sub-Saharan Africa. We did this at the request of an anonymous reviewer, who (rightly) noted that our models may fall into the same trap that ensared much of Tilly's war-making/state-marking arguments. Most notably, Tilly's arguments about state development in Europe have not traveled

Table 5: The Effects of External Territorial Threat on Intrastate Conflict Onset

	Model 1 COW Civil War Onset *1946-2007*	**Model 2** COW Civil War Onset *1946-2007*	**Model 3** UCDP Intrastate Armed Conflict Onset *1946-2007*	**Model 4** UCDP Intrastate Armed Conflict Onset *1946-2007*
GDP (lagged)	-0.308*	-0.319*	-0.171†	-0.174†
	(0.156)	(0.151)	(0.102)	(0.101)
Population (lagged)	0.160*	0.153*	0.073	0.067
	(0.069)	(0.069)	(0.060)	(0.061)
Mountainous Terrain	0.212*	0.221*	0.186**	0.190**
	(0.106)	(0.106)	(0.066)	(0.067)
Non-Contiguous State	-0.264	-0.202	-0.622*	-0.569*
	(0.401)	(0.404)	(0.279)	(0.278)
Oil	0.351	0.376	0.305	0.308
	(0.360)	(0.355)	(0.269)	(0.268)
UDS Democracy Score	-1.171**	-1.241**	-0.542**	-0.560**
	(0.389)	(0.404)	(0.206)	(0.206)
UDS Democracy Score^2	-1.273***	-1.321***	-0.858***	-0.876***
	(0.365)	(0.364)	(0.199)	(0.203)
Political Instability	0.818**	0.829**	0.714***	0.716***
	(0.272)	(0.269)	(0.214)	(0.215)
Ethnic Fractionalization	0.060	0.080	1.081*	1.087*
	(0.658)	(0.650)	(0.448)	(0.445)
Religious Fractionalization	-0.073	-0.099	-0.282	-0.306
	(0.581)	(0.580)	(0.413)	(0.414)
Territorial Claim	-0.820*	-0.791*	0.080	0.013
	(0.350)	(0.320)	(0.232)	(0.210)
Post-Territorial Claim, Years 20+	0.378		0.429†	
	(0.331)		(0.236)	
Post-Territorial Claim, Years 31+		0.705*		0.498*
		(0.332)		(0.219)
Observations	7,988	7,988	7,721	7,721

Standard errors in parentheses
† p<0.10, * p<0.05, ** p<0.01, *** p<0.001
Peace years, and cubic splines are omitted to save space.

Table 6: The Effects of External Territorial Threat on Intrastate Conflict Onset

	Model 1 COW Type 5 Civil War Onset *1946-2007*	**Model 2** COW Type 5 Civil War Onset *1946-2007*	**Model 3** COW Type 4 & 5 Civil War Onset *1946-2007*	**Model 4** COW Type 4 & 5 Civil War Onset *1946-2007*
GDP (lagged)	-0.002	-0.029	-0.269*	-0.272*
	(0.145)	(0.139)	(0.111)	(0.108)
Population (lagged)	0.401***	0.350***	0.270***	0.248***
	(0.096)	(0.092)	(0.064)	(0.064)
Mountainous Terrain	0.398***	0.382**	0.220*	0.224**
	(0.119)	(0.117)	(0.086)	(0.082)
Non-Contiguous State	0.130	0.198	0.335	0.387
	(0.363)	(0.323)	(0.290)	(0.298)
Oil	1.038*	1.162**	0.709*	0.743*
	(0.452)	(0.424)	(0.332)	(0.332)
UDS Democracy Score	-0.012	-0.057	-0.602**	-0.661**
	(0.227)	(0.214)	(0.198)	(0.205)
UDS Democracy Score^2	-0.092	-0.109	-0.598**	-0.608**
	(0.259)	(0.253)	(0.221)	(0.219)
Political Instability	0.483	0.494	0.660*	0.673*
	(0.458)	(0.461)	(0.278)	(0.279)
Ethnic Fractionalization (Alesina et al.)	0.677	0.799	0.269	0.331
	(0.678)	(0.656)	(0.479)	(0.487)
Religious Fractionalization (Alesina et al.)	-0.380	-0.113	0.034	0.059
	(0.560)	(0.508)	(0.437)	(0.426)
Territorial Claim	0.148	-0.753†	-0.526*	-1.008***
	(0.405)	(0.426)	(0.251)	(0.304)
Post-Territorial Claim, Years 1-15		-2.018***		-0.869**
		(0.600)		(0.312)
Post-Territorial Claim, Years 16-30		-0.405		-0.547†
		(0.526)		(0.318)
Observations	8088	8088	7748	7748

Standard errors in parentheses

† p<0.10, * p<0.05, ** p<0.01, *** p<0.001

well to other parts of the globe, including sub-Saharan Africa. These state borders were largely imposed by former colonial masters, and it is unclear if our arguments would necessarily apply to states whose borders are largely artificial.

Nevertheless, as Appendix Table 7 shows, the results are quite similar to the full sample presented in the manuscript. We observe in Model 1 and Model 2 that states in sub-Saharan Africa under conditions of territorial threat are less likely to experience the onset of a civil war over control of the government. Model 2 shows that the expected relationship between territorial threat and state capacity is observed in the immediate wake of settling territorial claims, even in Africa. Meanwhile, Models 3 and 4 suggest that there is no substantive relationship between territorial threat and lower intensity militarized challenges to the government's authority for this region. Thus, states with territorial claims in sub-Saharan Africa are no more or less likely to experience the onset of lower intensity insurrections against the government, but, as we expect, these challenges are unlikely to escalate to civil wars as our theory predicts (and as we demonstrate with the full samples analyzed in the manuscript).

Appendix Table 8 introduces the (lagged) U.S. aid data from Table I in the manuscript to test if external support for states matters more than external threat in explaining intrastate conflict propensities. The models show that states receiving U.S. military aid tend to observe more armed challenges to the government at both high and low levels of intensity. No consistent effect was observed for economic aid. The effects of the threat variables remain unchanged from Table II in the manuscript.

Table 7: The Effects of External Territorial Threat on Intrastate Conflict Onset in Sub-Saharan Africa

	Model 1 COW Civil War Onset *1946-2007*	**Model 2** COW Civil War Onset *1946-2007*	**Model 3** UCDP Intrastate Armed Conflict Onset *1946-2007*	**Model 4** UCDP Intrastate Armed Conflict Onset *1946-2007*
GDP (lagged)	-0.210	-0.144	-0.293†	-0.233
	(0.281)	(0.272)	(0.161)	(0.173)
Population (lagged)	0.302	0.292	-0.034	-0.024
	(0.213)	(0.193)	(0.146)	(0.149)
Mountainous Terrain	0.135	0.189	0.221†	0.228†
	(0.183)	(0.172)	(0.117)	(0.120)
Non-Contiguous State	-0.299	0.061	-0.803	-0.626
	(0.626)	(0.477)	(1.175)	(1.139)
Oil	0.020	-0.431	-0.176	-0.228
	(0.614)	(0.547)	(0.846)	(0.786)
UDS Democracy Score	-2.354*	-2.411†	-0.685	-0.701
	(1.180)	(1.307)	(0.432)	(0.427)
UDS Democracy Score^2	-2.100*	-1.831*	-0.699†	-0.649†
	(0.893)	(0.868)	(0.363)	(0.361)
Political Instability	0.570	0.694	0.723*	0.732*
	(0.438)	(0.463)	(0.288)	(0.299)
Ethnic Fractionalization	0.220	0.916	1.714	1.643
	(1.295)	(1.202)	(1.120)	(1.111)
Religious Fractionalization	-0.173	0.010	-0.631	-0.644
	(1.385)	(1.356)	(0.600)	(0.601)
Territorial Claim	-1.210†	-2.190**	0.092	-0.274
	(0.618)	(0.736)	(0.315)	(0.375)
Post-Territorial Claim, Years 1-15		-1.544*		-0.445
		(0.655)		(0.376)
Post-Territorial Claim, Years 16-30		-1.170†		-0.699*
		(0.658)		(0.355)
Observations	1,833	1,833	1,735	1,735

Standard errors in parentheses
† p<0.10, * p<0.05, ** p<0.01, *** p<0.001
Peace years, and cubic splines are omitted to save space.

Table 8: The Effects of External Territorial Threat on Intrastate Conflict Onset

	Model 1 COW Civil War Onset *1946-2007*	**Model 2** COW Civil War Onset *1946-2007*	**Model 3** UCDP Intrastate Armed Conflict Onset *1946-2007*	**Model 4** UCDP Intrastate Armed Conflict Onset *1946-2007*
GDP (lagged)	-0.327*	-0.337*	-0.179†	-0.175
	(0.160)	(0.159)	(0.107)	(0.107)
Population (lagged)	0.143*	0.136†	0.039	0.033
	(0.069)	(0.070)	(0.063)	(0.064)
Mountainous Terrain	0.177†	0.189†	0.164*	0.175**
	(0.103)	(0.102)	(0.065)	(0.065)
Non-Contiguous State	-0.379	-0.333	-0.662*	-0.628*
	(0.392)	(0.400)	(0.283)	(0.287)
Oil	0.373	0.396	0.335	0.346
	(0.349)	(0.344)	(0.266)	(0.259)
UDS Democracy Score	-1.251**	-1.342**	-0.593**	-0.624**
	(0.402)	(0.413)	(0.210)	(0.209)
UDS Democracy Score^2	-1.243***	-1.274***	-0.813***	-0.812***
	(0.376)	(0.376)	(0.209)	(0.209)
Political Instability	0.770**	0.797**	0.698**	0.704**
	(0.280)	(0.275)	(0.214)	(0.217)
Ethnic Fractionalization	-0.136	-0.074	1.011*	1.039*
	(0.644)	(0.659)	(0.449)	(0.435)
Religious Fractionalization	0.113	0.081	-0.098	-0.100
	(0.595)	(0.586)	(0.435)	(0.429)
U.S. Economic Aid (lagged)	-0.003	-0.004	0.004	0.004
	(0.030)	(0.031)	(0.021)	(0.021)
U.S. Military Aid (lagged)	0.045*	0.046*	0.044**	0.042**
	(0.020)	(0.020)	(0.014)	(0.014)
Territorial Claim	-1.008***	-1.471***	-0.142	-0.443†
	(0.289)	(0.347)	(0.196)	(0.234)
Post-Territorial Claim, Years 1-15		-0.601†		-0.451†
		(0.364)		(0.255)
Post-Territorial Claim, Years 16-30		-0.794†		-0.405
		(0.422)		(0.270)
Observations	7,928	7,928	7,661	7,661

Standard errors in parentheses
† p<0.10, * p<0.05, ** p<0.01, *** p<0.001
Peace years, and cubic splines are omitted to save space.

References

Al-Enazy, Askar Halwan. 2002. "'The International Boundary Treaty' (Treaty of Jeddah) Concluded between the Kingdom of Saudi Arabia and the Yemeni Republic on June 12, 2000." *American Journal of International Law* 96(1):161–173.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth* 8(2):155–194.

Antidze, Margarita. 2007. "Russia Closes Last Military Base in Georgia." *Reuters* November 13.

Baccini, Leonardo and Johannes Urpelainen. 2012. "Strategic Side Payments: Preferential Trading Agreements, Economic Reform, and Foreign Aid." *Journal of Politics* 74(4):932–949.

Cheibub, José Antonio. 1998. "Political Regimes and the Extractive Capacity of Governments: Taxation in Democracies and Dictatorships." *World Politics* 50(3):349–376.

Collier, Paul and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56(4):563–595.

Correlates of War Project. 2011. "State System Membership List, v. 2011." Online, http://correlatesofwar.org.

Dafoe, Allan. 2011. "Statistical Critiques of the Democratic Peace: Caveat Emptor." *American Journal of Political Science* 55(2):247–262.

Dafoe, Allan, John R. Oneal and Bruce Russett. 2013. "The Democratic Peace: Weighing the Evidence and Cautious Inference." *International Studies Quarterly* 57(1):201–214.

Ellingsen, Tanja. 2000. "Colorful Community or Ethnic Witches Brew? Multiethnicity and Domestic Conflict During and After the Cold War." *Journal of Conflict Resolution* 44(2):228–249.

Fearon, James D. 2003. "Ethnic and Cultural Diversity by Country." *Journal of Economic Growth* 8(2):195–222.

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

Gibler, Douglas M. 2012. *The Territorial Peace: Borders, State Development, and International Conflict*. New York, NY: Cambridge University Press.

Gleditsch, Kristian Skrede. 2002. "Expanded Trade and GDP Data." *Journal of Conflict Resolution* 46(5):712–724.

Honaker, James and Gary King. 2010. "What to Do about Missing Values in Time-Series Cross-Section Data." *American Journal of Political Science* 54(2):561–581.

Huth, Paul K. and Todd L. Allee. 2002. *The Democratic Peace and Territorial Conflict in the Twentieth Century*. New York, NY: Cambridge University Press.

Johnson, Constance. 2000. "Case Concerning Kasikili/Sedudu Island (Botswana/Namibia)." *The International Journal of Marine and Coastal Law* 15(4):581–599.

Jones, Daniel M., Stuart A. Bremer and J. David Singer. 1996. "Militarized Interstate Disputes, 1816-1992: Rationale, Coding Rules, and Empirical Patterns." *Conflict Management and Peace Science* 15(2):163–213.

King, Gary and Langche Zeng. 2001. "Logistic Regression in Rare Events Data." *Political Analysis* 9(1):137–163.

Kugler, Jacek and Ronald L. Tammen, eds. 2012. *The Performance of Nations*. Lanham, MD: Rowman and Littlefield.

Maddison, Angus and Harry X. Wu. 2008. "Measuring China's Economic Performance." *World Economics* 9(2):13–44.

Meng, Xiao-Li. 1994. "Multiple-imputation Inferences with Uncongenial Sources of Input." *Statistical Science* 9(4):538–558.

Rubin, Donald B. 1976. "Inference and Missing Data." *Biometrika* 63(3):581–592.

Rubin, Donald B. 1996. "Multiple Imputation after 18+ Years." *Journal of the American Statistical Association* 19(434):473–489.

Sherr, James. 1997. "Russia-Ukraine Rapprochement?: The Black Sea Accords." *Survival* 39(3):33–50.

Simmons, Beth A. 2009. *Territorial Disputes and Their Resolution: The Case of Ecuador and Peru*. United States Institute of Peace.

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

Sokov, Nikolai. 2005. "The Withdrawal of Russian Military Bases from Georgia: Not Solving Anything." *PONARS Policy Memo* 363.

Thies, Cameron G. 2004. "State Building, Interstate and Intrastate Rivalry: A Study of Post-Colonial Developing Country Extractive Efforts, 1975-2000." *International Studies Quarterly* 48(1):53–72.

Thies, Cameron G. 2010. "Of Rulers, Rebels, and Revenue: State Capacity, Civil War Onset, and Primary Commodities." *Journal of Peace Research* 47(3):321–332.

Zardykhan, Zharmukhamed. 2002. "Kazakhstan and Central Asia: Regional Perspectives." *Central Asian Survey* 21(2):167–183.