Chapter 5

Civil Wars Worldwide, 1944-2006

The previous chapter laid the foundation for an empirical understanding of quagmire in civil war. Through its analysis of turning points in the Lebanese Civil War, it illustrated that the path to quagmire aligns with the theoretical expectations developed earlier in the book. Specifically, it provided evidence that the mechanisms of external support as a subsidy and substitution between types of fighting are connected to choices that kept the belligerents locked into civil war. Insofar as the actions of the warring parties in Lebanon and potential foreign backers are concerned, the theory's underpinnings appear valid.

This chapter begins the work of testing the theory empirically, to link the causal factors identified by the theory to the outcome of quagmire. It asks whether foreign interests, the cost of escalation in fighting, and the stakes of conflict work to generate quagmire across civil wars. In what follows, I evaluate whether the theory's predictions account for the presence (and absence) of quagmire in civil wars around the world. To draw systematic inferences, I use statistical analysis to study all civil wars that took place between 1944 and 2006. The list comprises 140 conflicts, in 69 countries and on five continents, ranging from Lithuanian nationalist resistance to Soviet rule from 1944 to 1948 to an insurgency in Chad that

began in 2005. The global scope and long time period of the evidence presented here provide reassurance that the findings are general to the empirical reality of civil war.

The chapter is divided into two parts. In sections 5.1-5.3, I apply the concept of quagmire consistently and systematically across all civil wars in the list. I begin by providing an operational definition that satisfies the theoretical concept's core feature. This establishes a measure of quagmire, so that we can see its distribution in civil wars across the globe and over time. Next, I arrive at the heart of the matter: in Section 5.4, I analyze the incidence of quagmires. Section 5.5 concludes the chapter and reflects on the results.

Where the chapter uses statistical analysis, I provide an overview of the data employed, the intuition behind the methods used, and a non-technical exposition of the results. The chapter's appendices provide a full treatment of the analysis. Appendix A explains the list of civil wars used, contrasts it to other extant lists, and provides details on the other variables used in the statistical analysis. Appendix B contains the full statistical models, results, and robustness checks.

5.1 How to Identify Quagmire

In which civil wars do we observe quagmire? The conceptual definition advanced in Chapter 2 does not provide an easy starting point for such an inquiry. For any given civil war, it calls for a holistic assessment, one that can characterize the situation the belligerents face in interaction with one another. It requires detailed information about each belligerent, across a wide range of areas sufficient to characterize the costs and benefits of continuing conflict. Here, two concerns arise, the first informational, the second analytical.

To establish whether quagmire occurred in even a single war, copious amounts of information are necessary, and at a level of detail that would be difficult to achieve even in the very best case scenario. To collect such information across wars increases the quantity of information required exponentially. To characterize the costs and benefits for each belligerent in the large number of civil wars that occurred during this period would require such a large volume of information, at likely unattainably high levels of granularity. Indeed, it is unclear if such information is accessible or even exists for many civil wars. The relevant records may be classified, exist only in paper form in archives scattered across the world, may have been destroyed in the war, or may not have been recorded in any form in the first place.

Diversity in the nature of the sources available from one war to the next makes comparisons across wars difficult. Evidence deemed sufficient by experts on a particular war may have no direct analogue in another case. Even were detailed information available, the differences in the record of primary documents, contemporaneous reporting, secondary analyses, and memoirs across geographically and temporarily diverse wars make it unlikely that the definition could be applied using types of information that were consistent across wars. The information requirements, then, are insurmountably high.

The nuances of the conceptual definition also raise the possibility that, even when presented with all the evidence, experts on a single war might disagree about whether it fits the definition. And those wars that to some experts seem to fit the concept obviously might to others strain its limits. It may be hard to find standards of evidence to apply to any given war that could convince those who disagree.

The prodigious literature on the American Civil War is an example of the extent to which expert opinion can vary even when record keeping at the time of the war was extensive, the study of the war has continued unabated up to the present, and high-level decision-makers kept diaries and carried on copious correspondences. Shifting from a retrospective to a prospective frame would not provide traction. The accounts of contemporaneous observers are complex and driven by their own narrative concerns, certainly so when it comes to characterizing expectations about a war's duration. There are, for example, domestic political costs to announcing that a war is likely to be long-lasting, or strategic benefits to be gained where the enemy's behavior is concerned by suggesting that victory will be swift.[1] Indeed, there is often little consensus with events underway as to the nature of a conflict – is it a brief violent episode, criminal disorder, potential civil war, or outright civil war?[2] Witness the endless discussion of Iraq's so-called slide into civil war following the U.S. invasion in 2003, a discussion which continued long after

[1]Some contemporaneous assessments of a war's likely course turn out to be preternaturally accurate, but rarely find consensus support at the time. Thus William T. Sherman's passionate reaction to the news of South Carolina's secession from the United States in December 1860:

> "You people of the South don't know what you are doing. This country will be drenched in blood, and God only knows how it will end. It is all folly, madness, a crime against civilization! You people speak so lightly of war; you don't know what you're talking about. War is a terrible thing! You mistake, too, the people of the North. They are a peaceable people but an earnest people, and they will fight, too. They are not going to let this country be destroyed without a mighty effort to save it … Besides, where are your men and appliances of war to contend against them? The North can make a steam engine, locomotive, or railway car; hardly a yard of cloth or pair of shoes can you make. You are rushing into war with one of the most powerful, ingeniously mechanical, and determined people on Earth — right at your doors. You are bound to fail. Only in your spirit and determination are you prepared for war. In all else you are totally unprepared, with a bad cause to start with. At first you will make headway, but as your limited resources begin to fail, shut out from the markets of Europe as you will be, your cause will begin to wane. If your people will but stop and think, they must see in the end that you will surely fail" (Quoted in Foote 1986 [1958]:58-9).

[2]See also Mamdani's (2007) analysis of the political implications of how a conflict is labeled, focusing on Darfur.

the phenomenon transpiring in Iraq had satisfied objective definitional criteria for civil war.[3]

To be clear, the informational and analytical concerns point to a problem not with the concept of quagmire itself but with its empirical application. To compare systematically, we need a way to assess whether a war experiences quagmire which uses comparable information across civil wars. The core of that theoretical concept is that civil wars that experience quagmire continue on past the point at which observers would expect them to conclude; indeed the concept is useful because it allows us to identify obstacles to resolving such wars through victory or negotiations – obstacles that alter the calculus of the belligerents and therefore render their behavior puzzling.

We can build an operational definition from the core feature of the concept by asking, does a war last longer than can be reasonably expected? A civil war in which quagmire occurs should last significantly longer, not than other wars, but than we would otherwise expect that same war to last *had quagmire not occurred*. The point is to determine which portion of the actual length of a civil war is not accounted for by standard factors that determine wars' duration. An operational definition, then, is that civil wars experience quagmire if they last significantly longer than would be expected according to a comprehensive analysis of war duration.

The pragmatic move to an operational definition necessarily sacrifices some of the theoretical definition's conceptual integrity. But as the rest of this chapter shows, the trade-off is well worth it. The operational definition allows us to empirically identify quagmire across any given population of civil wars, and to analyze

[3]See Sambanis (2006).

its determinants rigorously using statistical methods. Of course, this definition will yield a list that is likely to include wars which reasonable observers will contend did not experience quagmire. But the empirical exercise is useful precisely because it is comprehensive and therefore inevitably runs up against problems at the margins. The liminal cases force us to question the operational definition and enrich our understanding, both of what is quagmire and of why it comes to pass.

The operational definition permits the consistent classification of the entries in large civil war lists as having or not having experienced quagmire. Section 5.2 employs data on the length of civil wars in the post-Second World War period. I specify and estimate a reasonable quantitative model with which to assess expected duration. Section 5.3 estimates the predicted length of civil wars in the list based on that model and measures the deviation of the actual from the expected duration. It then uses consistent criteria across all wars to classify this deviation, resulting in a list that separates civil war quagmires from non-quagmires.

5.2 Duration as Building Block

I begin to apply the operational definition of quagmire systematically to all civil wars by specifying and estimating a reasonable statistical model of civil war duration. This section, then, while studying civil war duration, differs from existing research on the subject in a critical respect: the length of war is not an outcome of interest in and of itself, but a tool to allow the systematic identification of quagmire in civil wars. To that end, the regression analysis of civil war duration presented in section 5.2.3 constitutes the basis for assessing wars' expected duration and the gap between it and actual length in section 5.3.

5.2.1 Data

Civil wars represent a fundamental rupture of state sovereignty, pitting challengers against the state. The coding criteria developed by Sambanis (2004) capture this understanding (see Sambanis and Schulhofer-Wohl 2016). The thorough criteria for coding the continuation and end of hostilities make Sambanis' definition of civil war the most suitable one for studying duration empirically.[4] A conflict is counted as a civil war starting in the first year in which it results in at least 500 deaths, provided that there is "effective resistance" on the part of the weaker of the parties to the conflict.[5] The war is then coded as continuing provided that there has been "sustained violence, at least at the minor or intermediate level," defined by a 500-death threshold over the course of any three-year period. In addition to the possibility of the war ending through inactivity, Sambanis also specifies that the end of the war is coded based on either the victory of one of the parties to the conflict, a peace treaty, cease-fire, or truce, or a cessation in fighting. Each type of war termination must be followed by a period of peace of a minimum length specific to it: no period of peace required after a rebel victory, and six months of

[4]So that the results of this chapter can be readily understood when viewed alongside the literature which uses other war lists or definitions of civil war, Appendix B.1.2 provides an overview of the other three lists of civil wars most widely used in research in political science and economics: Fearon and Laitin's (2003) dataset, the UCDP/PRIO Armed Conflict Dataset (Gleditsch et al. 2002; Harbom and Wallensteen 2009), and the Correlates of War Project's Intra-State War Dataset (Sarkees and Wayman 2010)

[5]If the 500-death threshold is not met in a single year, onset is still coded in that year if a 1,000 death threshold is met over the course of the following three years. Sambanis also notes that the 500-death threshold "can be relaxed to a range of 100 to 1,000 because fighting might start later in the year" and that in some cases, if there is a lack of information on deaths during the first year of conflict, onset of civil war can be coded "at the first year of reported large-scale armed conflict, provided that violence continues or escalates in the following years." The weaker of the parties is considered to have engaged in "effective resistance" if it is able to cause at least 100 deaths on the part of the stronger party, a large proportion of which occur during the first year of the war. Evidence that violence is becoming one-sided in the war would imply that the possibility of the civil war ending should be considered, with it having been replaced by another form of political violence.

peace following a government victory, six months following a peace agreement, and two years following a cease-fire, truce, or cessation of hostilities.[6]

I update Sambanis' list, which ended in 1999, through 2006, to arrive at a total of 158 civil wars between 1944 and 2006 (see Appendix B.1.1).[7] I build a dataset of civil war-year observations using this list,[8] but exclude eighteen of the wars in it for the purposes of the analysis that follows. Fourteen of the excluded wars started as coups; these follow a distinct process and have systematically shorter duration than other wars (see, e.g. Houle 2016). The other four are borderline cases due to low numbers of casualties or the nature of the conflict: the Partition of India, 1946-1948; Northern Ireland's Troubles 1971-1998; and the two Palestinian intifadas, 1987-1997, and 2000-. The dataset I use thus covers 140 civil wars, in 69 countries. Table 5.1 presents summary statistics on duration from the data.

5.2.2 Explanations

Studies of the onset of civil wars and peacekeeping have long taken pride of place in the social scientific research on conflict. Nevertheless, a fairly large collection of studies treat duration as the outcome of interest. The difficulty with this literature is that much of it fails to theorize duration as such. Scholars are most often interested in a particular explanatory variable – for example, the effects of international intervention (far and away, the most frequently studied) – and confine

[6]Sambanis defines peace, for the purpose of coding war termination, as the absence of battle-related deaths, or, at a minimum, "fewer than 100 deaths per year" (831).

[7]I apply the most lenient interpretation of Sambanis' criteria to the potential cases of civil war. Seventeen of the 158 wars (10.8%) were ongoing as of December 31, 2006. References to ongoing wars in the text are with respect to that date.

[8]I take the duration of the civil war reported in months in Sambanis' data, and round it to the nearest year, with the duration of wars lasting less than six months rounded up to one year. Appendix B.1.1 discusses the choice of the unit of measurement for civil war duration and contains the list of civil wars.

Table 5.1: Summary Statistics, War Duration (in years)

Category	n	Mean	St. Dev.	Min.	Max.
All wars	140	7.503	7.704	0.083	40.333
By duration:					
under mean	93	2.964	1.960	0.083	7.333
over mean	47	16.484	6.900	7.667	40.333
By region:					
Asia	42	10.046	10.390	0.167	40.333
Europe (including USSR/former USSR)	16	3.313	1.975	0.667	7.333
Latin America	11	10.561	8.699	0.667	28.167
Middle East & N. Africa	21	5.861	5.757	0.083	16.5
Sub-Saharan Africa	50	6.725	5.721	0.083	19.083

their reasoning to developing the logical connection between that variable and duration, with perhaps a nod to controlling for other factors. To leverage these studies – without the benefit of being able distinguish theoretical approaches of interest from the literature, since it does not explicitly employ them – I identify factors that studies identify, either at the theoretical level or via empirical results, as associated with duration.

Based on existing research, six categories of factors potentially associated with duration stand out: a country's geographic features, social characteristics, level of economic development and resources, government capacity, type of conflict, and international environment.[9] During-conflict characteristics of a civil war, be they

[9]This section draws on the logic and empirical results contained in studies of civil war onset and termination in addition to studies of duration itself. Particularly relevant are: Sambanis (2000, 2001, 2002, 2004), Elbadawi and Sambanis (2000, 2002) Balch-Lindsay and Enterline (2000), Regan (2002), Collier and Hoeffler (2004), Fearon and Laitin (2003), Collier, Hoeffler and Sï¿½derbom (2004), Fearon (2004), and Cunningham (2006).

domestic or international, constitute a seventh category, one that receives heavy attention in the literature. This seventh category is problematic, however. The process of ongoing civil war directly affects the factors it encompasses such that they are not independent of war duration. If they were included in the analysis, it would suffer from problems of reverse causality. I therefore exclude this seventh category.

I conduct the statistical analysis of duration based on these six categories. Table 5.2 organizes the variables included in that analysis by category.[10] Below, I explain the connection between each category and duration, as well as the decision to exclude during-conflict factors. Appendix B.1.3 explains the sources used for each variable, and, in the case of a few variables, corrections made to sources' coding.

Geography

A country's physical attributes shape the ways in which a civil war is fought. Belligerents' strategies often depend on the terrain. Mountainous regions prove to be important sanctuaries for rebel groups that may be weak in conventional military terms, providing a natural defensive barrier against the state. After the Algerian government cancelled election results in January 1992 and imposed a state of emergency to prevent an Islamist political party from coming to power, the Islamic Armed Movement and other groups used "mountainous areas" near the capital city of Algiers to organize and develop their military capabilities outside the reach of the government (Martinez 2000:48,69-70). Other geographic features that make an area inaccessible or difficult to traverse – for example, deserts, deep forests,

[10]Unless otherwise noted, variables are measured at the level of the country.

Category	Variable Name	Definition
Geography	*area*	Area in square kilometers
	rugged terrain	Difference between the highest and lowest points in elevation
Society	*ethnic polarization*	RQ index of polarization, minimum value of 0, maximum value of 1, pre-war
Economy	*gdp per capita*	Gross domestic product divided by population, pre-war
Government capacity	*military personnel*	Number of troops, pre-war
	military spending	Defense budget, pre-war
	age of polity	Is the polity more than 20 years old, pre-war?
	executive constraints	Constraints on the chief executive, scale of 1 to 7, from low to high, pre-war
Conflict	*prior civil war*	Did the country experience a civil war in the 20 years prior to the start of the current one?
	secessionist war	Does the conflict concern a dispute over the establishment of a separate polity from some portion of the territory of the country's territory?
International	*borders*	Number of borders shared with other states
	neighboring civil war	Was there a civil war in a neighboring country in the year before the current war started?
	hydrocarbon exporter	Were at least one-third of the country's exports oil or natural gas, pre-war?
	superpower sphere of influence	Does the country border the Soviet Union or any of its successor states; or is it located in Central America?
	Cold War	Did the war start between 1946 and 1991?

Table 5.2: Factors Associated with Duration and Variables Used in the Statistical Analysis

swamps and marshes,[11] or even sheer distance from the centers of state power over which the government must project its power – provide similar advantages. By leveling the power imbalance between a nascent or growing insurgency and the incumbent government, then, difficult geography can lengthen civil war.

I include two measures of geography in the analysis, a country's area and a measure of how rugged is its terrain. The first, *area*, is measured in square kilometers. The second, *rugged terrain*, measures the difference between the highest and lowest points in elevation in the country, in meters.[12] I use Fearon and Laitin's (2003) data for this variable, updating their coding through 2006. Fearon and Laitin code values of this variable as missing for South Vietnam, the Yemen Arab Republic, and the People's Democratic Republic of Yemen. I collected data for these three countries as described in Appendix B.1.3.

Society

The social characteristics of a country also influence warfare, through a variety of channels. At an individual level, willingness to sacrifice for a cause can be a key component of participation and mobilization, but particularly for continuing in the fight given harsh conditions. Social identity proves a strong foundation for this willingness (Sambanis, Schulhofer-Wohl and Shayo 2012). Identity-based conflict can lengthen the time periods over which individuals and organizations expect to produce change. And with individuals' willingness to sacrifice

[11]Hence the notoriety of then-Lieutenant Colonel Francis Marion of the Continental Army, who commanded militia in South Carolina during the American Revolution. His British pursuer, Lieutenant Colonel Banastre Tarleton, labeled this elusive, resourceful insurgent foe the "Swamp-fox" (see Simms 1844:152). Marion's troops and civilian sympathizers enthusiastically took up the moniker.

[12]Following the convention in the literature, I use the log of the elevation difference measure of *rugged terrain* in the analysis (see, e.g., Lyall and Wilson 2009)

and longer time-horizons, political parties and armed organizations may be able to tolerate higher uncertainty over the prospects of winning and higher costs of continuing to fight. In addition, the number and composition of armed organizations, both those opposed to the government and within the government's security forces, often reflect societal divisions. Under military pressure, social ties can help to maintain armed group cohesion (Staniland 2014); such ties may be easier to leverage in highly polarized societies. Fear and animosity across social divides can also stoke armed organizations' ability to recruit new members in the midst of conflict , whether as active fighters, political cadres, or sympathizers who carry out petty tasks, or passive supporters. More than that, the ability to sustain the battle, whether on the part of the government or its opponents, is often linked to each side's ability to feed a narrative of antagonism (Hage 1996).[13]

To capture the broad strokes of these phenomena, I include a measure of the extent to which society can be divided antagonistically along identity lines, the *RQ* index of *ethnic polarization,* following the formula provided by Reynal-Querol (2002).[14] I use Fearon's (2003) cross-sectional data on ethnic groups to provide the number of groups and their respective shares of the population needed to calculate the *RQ* index. In Appendix B.1.3, I explain my use of an ethnic group list and group shares of population that differ from Fearon's in the case of the Republic of Yemen, and the sources of the underlying data I use to construct them. As Fearon does not report any data for Papua New Guinea, South Vietnam, the Yemen Arab Republic, or the People's Democratic Republic of Yemen, I describe

[13] See also Christia (2012) on the use of identity-based narratives to justify alliance formation, with empirics drawn from Afghanistan and Bosnia.

[14] $RQ = 1 - \sum_{i=1}^{N} \left(\frac{1/2-\pi_i}{1/2}\right)^2 \pi_i$, where π_i is the proportion of the population which belongs to group i, and N is the number of groups. See also Montalvo and Reynal-Querol (2005).

the sources I use for the list of groups and groups' shares of population to calculate the *RQ* index for each of these countries. I also describe how I calculate the index for Pakistan prior to the secession of Bangladesh, since Fearon's data for Pakistan are for the period following the secession.

The size of a country's population might also be expected to matter, as larger countries might have more complex social divisions and diversity. However, I do not include a stand-alone measure of population size because population is included via a per capita measure of the size of the economy (see below).

Economy

Levels of economic development are also thought to factor into length of conflict. The opportunity costs to conflict increase the wealthier is a country. As conflict continues, armed organizations should find it more difficult to sustain themselves through recruitment the higher is the level of economic development (Collier, Hoeffler and Söderbom 2004). Pressures for settlement or a quick victory should vary along with economic development in the same way. An additional pathway considers the syzygy between government power and economic development via infrastructural capacity. Higher levels of economic development should be associated with a range of developments which all facilitate the fast prosecution of war: transportation and communications infrastructure, government penetration into society, and manufacturing capacity. All of these should make it easier for governments to defeat an ongoing armed challenge; or at the same time, for an armed challenge to a government to rapidly increase in strength and match and overwhelm a government.[15]

[15]Dunning (2005) illustrates a similar logic for war onset. He attributes Zairian President Mobutu Sese Seko's intentional neglect of infrastructure to the need to make it harder for his gov-

I include gross domestic product in the analysis to capture these arguments, as *gdp per capita*. I also use its square to allow for a curvilinear relationship between the economy and war duration. Wars in countries with moderate levels of economic development, for example, might be shorter than wars in very underdeveloped countries if the gains to infrastructural capacity increase rapidly in GDP. At the same time, wars in developed countries might not be shorter than wars in middle income countries if those gains increase at a diminishing rate. Opportunity costs might operate in a similar way.

Government capacity

The characteristics of a polity significantly affect a government's ability to counter armed rebellion. Institutions can act as constraints on government actions, hampering security forces' latitude to confront the threat using strategy and tactics that they prefer. Speed may also suffer – structures of governmental checks and balances can impede swift planning and swift action. Thus a government's ability to try to quickly defeat an armed threat can be compromised. In an opposite vein, the excesses of autocratic regimes generate grievances that can fuel conflict. The debate on the effects of regime type has carried over from studies of international war (e.g., Reiter and Stam 1998) to internal conflict. Kalyvas (2006:358-62), for example, explains how democracies and autocracies differ in their abilities to obtain information on the populations under their control via denunciation.[16]

Regime type is but one element of government capacity, however. A deeper

ernment to be challenged by rebels. Roads would have allowed peripheral insurgencies swifter and easier access to the capital, and seizing the capital would have been the way to voerthrow the regime given the country's vast size and inherent ungovernability. For a related discussion of the concept of "infrastructural power," see Mann (1984), and also Mann (2008) and Soifer (2008).

[16]See also Zhukov (2007), Lyall (2010), Getmansky (2013), and Byman (2016).

one may be the stability of the state. Is the polity newly constituted? Or does the state have deep roots, having ruled over the same territory and peoples and existed on the international stage for generations upon generations?[17] State strength also plays an important role. A government that can rely on a well-developed security apparatus can respond swiftly and robustly to rebellion, not least by undertaking political initiatives to address the root causes of rebellion. As Pye (1962) explains, "military authorities [in developing countries] often find that they are in control of one of the most effective general purpose organizations in the society and hence they may be called upon, or be compelled by events, to perform the duties of civil authorities."

It is also important to note that aspects of government capacity may be related to civil war duration in unusual ways. Since we consider duration and not the onset of war here, it might be the case that wars last longer in long-established and stable polities. In order for such wars to start in the first place – an event that would be rare in comparative perspective – the opponents of the state would need considerable capacity. And in such a case, the rebellion would be much more difficult to put down, but also not necessarily an easy win for the opposition; and so the war could drag on. In addition, long-standing, stable polities might evolve systematically more complex and bureaucratized forms of government, leading to longer wars prolonged via inter-agency spats, rigid procedures and failure to adapt quickly to the war to overwhelm the insurgents at the beginning. Polities in which government power is more constrained and accountable could face greater constraints in prosecuting the war, due to public opinion, legislative oversight, the domestic politics of opposition, and the like.

[17] Fearon and Laitin (2003) find that a country's status as a new state correlates with civil war onset, for example.

I use five variables that capture the diverse dimensions of government capacity. The first two size up the country's military capacity, measuring the size of its forces via *military personnel* and level of defense spending via *military expenditure*. *Age of polity* distinguishes new states from states that have existed for twenty years, selected as the cut-off because it is roughly one generation. Finally, *executive constraints*, which measures the degree to which the power of the chief executive can be reined in by other governmental actors, captures the extent to which a government functions under the constraints commonly associated with liberal democracies.

Conflict

Characteristics of the conflict have clear bearing on its course. Here I exclude those that might be a product of conflict processes and consider only such characteristics as exist at the outset of the conflict. Chief among these is the legacy of past wars. Is the civil war in question the first to occur in a country, or had the country experienced prior ones in recent history?[18] Previous wars leave behind legacies salient to the conduct of war (see Wood 2008) – political polarization (e.g. Balcells 2012),[19] organizational frameworks and expertise (Daly 2012, 2014), technological expertise, other forms of martial capital including tactical expertise in war-fighting, and stocks of military matériel (Collier and Sambanis 2002; Collier, Hoeffler and Sambanis 2005:7). I capture in *prior war*, a variable that distinguishes countries for which the war in question is the first in twenty years from those

[18]Fearon and Laitin (2003) find that this correlates with civil war onset. Sambanis (2000) notes a link between it and the recurrence of civil war.

[19]More generally, see Lupu and Peisakhin (Forthcoming) on the long-term political effects of repressive violence against individuals.

that had already experienced one during this period.[20] The nature of a civil war – specifically the aims over which it is fought – can also play a key role. Secessionist civil wars, fought to separate a part of the country's territory and establish it as a separate polity, have long been identified as longer lasting than other civil wars, such wars proving less tractable due to behavior on the part of incumbent governments and separatists alike (see, e.g. Toft 2005; Fearon 2004; Walter 2006*a*, 2009). I distinguish these wars from others with the variable *secession*.

International

The international context of a civil war can dramatically influence and alter the conflict's path. Civil war countries that are surrounded by many neighbors may be less stable and more prone to military intervention and myriad other attempts to influence their domestic affairs. Indeed Balch-Lindsay and Enterline (2000:623) reason that as the number of neighbors increases, so too should the "complexity" of the war, and, along with it, the war's length.[21] I therefore include *borders*, which measures the number of adjacent countries. Instability in neighboring countries can spill over into a civil war country and fuel its conflict. War in a neighbor may result in the accumulation of war-making equipment, organizations, and skills in much the same way as would a previous civil war, though the matériel and human capital would be less directly accessible.[22] I distinguish between countries that had a neighbor with an ongoing civil war from those that did not. The variable

[20]The presence of a civil war within the 20 years prior to the war in question is coded strictly for a given polity. Anti-colonial wars, which do are not part of this dataset and are considered civil wars in the colonial empire in question, are not taken into account.

[21]Kathman (2010, 2011) analyzes the motivations for intervention by neighbors and other channels for civil war to spread beyond a single country.

[22]Contributors to Lake and Rothchild (1998) discuss contagion from neighboring countries as a factor in civil war onset (and also diffusion). Balch-Lindsay and Enterline (2000) and Sambanis (2001) test for this in cross-country data.

neighboring civil war is coded 1 if a country had such a neighbor in the pre-war period, 0 otherwise.

The extent to which international powers see interests in a civil war country can also influence the war's course. In the case of civil war countries of critical importance, we might expect powerful states to use all tools at their disposal to end the conflict quickly. I include two measures of international powers' interests. First, is the civil war country a hydrocarbon exporter? The primary role of oil and natural gas in the world economy has made these commodities of crucial strategic importance to international powers. They serve as a linchpin in economic production, a crucial energy resource for human populations. They are also militarily-critical supplies for governments, in maintaining readiness and in sustaining war-making. Second, is the country located in the sphere of influence of one of the two superpowers that have dominated post-war international politics – the U.S. or the Soviet Union, or the Russian Federation following the latter's collapse. For the first, I use the variable *hydrocarbon exporter*, which classifies as such any country for which at least one-third of its exports were fuel. For the second, I code *superpower sphere of influence*, which indicates whether a country borders the U.S.S.R. or its successor states, or is in Central America.

Finally, the international context writ large can exert an influence on a country at war. International rivalries and competition generate motives for intervention, which is associated with longer-running civil wars (Elbadawi and Sambanis 2000; Regan 2002). The nature of that competition and the extent of interference is often a product of the structure of the international system, whether we think of this in terms of the distribution of power across states (Waltz 1979) or the diffusion of ideological polarization (Owen 2010). I therefore treat the Cold War as having

the potential to exercise a systematic influence on civil war dynamics. I define the period as starting in 1946, following the Second World War, and lasting until 1991, with the dissolution of the Soviet Union.[23] The variable *Cold War* indicates whether a civil war started during this period.

During-Conflict Characteristics of Civil Wars

What do the six categories of factors that might influence duration leave out? Scholars have rightly identified characteristics of the war itself, such as battle deaths, the balance of government versus rebel forces, the number of parties to the conflict, foreign interventions, and the use of contraband financing by rebel groups, among others, as likely influences on how long a civil war lasts. These quite plausible relationships are not in dispute. Indeed, as Table 5.3 shows, the prevalent approach is to operationalize during-conflict characteristics and use them in statistical analysis of duration. But it is clear, too, that conflict dynamics affect these factors; in fact, all are measured while the conflict is ongoing, can change as it continues, and are typically affected by the history of the conflict up until that point. As a result, including these during-conflict characteristics of civil wars would violate assumptions built into the statistical model I use and would bias the results (see Lancaster 1990:28-31).

[23] Aid (2010) provides an account of the focus of the U.S. military on countering a Soviet threat even as the Second World War was ending. Although massive changes occurred in the Soviet Union following the fall of the Berlin Wall (and even starting in 1988), I choose the date of the final dissolution of the Soviet Union itself as the end of the Cold War due to the persistence of many Cold War military structures, albeit while experiencing dramatic changes, during those final years. As Odom (1998:275) notes, "The command and control structures of the Warsaw Pact surprisingly survived into 1991."

Period for Variables Used	Studies
Pre-Conflict Only	Elbadawi and Sambanis (2000); Rustad et al. (2008); Montalvo and Reynal-Querol (2010)
Pre-Conflict and During Conflict	Balch-Lindsay and Enterline (2000); Regan (2002); Fearon (2004); Collier, Hoeffler and Söderbom (2004); Akcinaroglu and Radziszewski (2005); Regan and Aydin (2006); Cunningham (2006); Balch-Lindsay, Enterline and Joyce (2008); Brandt et al. (2008); Gent (2008); Lyall and Wilson (2009); Buhaug, Gates and Lujala (2009); Cunningham, Gleditsch and Salehyan (2009); Kirschner (2010); Cunningham (2010); Wucherpfennig et al. (2012); Aydin and Regan (2012); Moore (2012)

Table 5.3: Pre-Conflict and During-Conflict Variables in the Literature

5.2.3 The Length of Civil Wars: Analysis

I estimate a Cox proportional hazards model of civil war duration to use to create a predicted length for each civil war in the dataset.[24] As applied here, this type of regression analysis looks at the probability that civil war continues. We can interpret the estimated coefficients as the effect of each explanatory variable on duration while holding other variables constant. The model includes the explanatory variables discussed above, which represent the six categories of factors associated with duration: geography, society, economy, government capacity, conflict characteristics, and international environment. Table 5.4 lists each variable, the sign of its expected effect on the length of civil war, and the sign of the effect estimated via the regression analysis.[25]

The results are consistent with the theoretical expectations outlined in the pre-

[24]Technical details contained in Appendix B.2.

[25]The signs of the estimated effects listed in 5.4 here are the opposite of those contained in Table B.4, Appendix B.2. The regression coefficients in Table B.4 describe the estimated effect on the "hazard" of civil war ending at any given point in time. They are therefore the opposite of the effect on duration.

Variable	Effect on duration	
	Expected	Estimated
area	+	+ *
rugged terrain	+	+
ethnic polarization	+	+
gdp per capita	+	+ *
gdp per capita, squared	–	– *
military personnel	–	– **
military spending	–	–
age of polity	+	+ **
executive constraints	+	+ ***
prior civil war	+	+ *
secessionist war	+	+
borders	+	+
neighboring civil war	+	+
hydrocarbon exporter	–	–
superpower sphere of influence	–	–
cold war	+	+

* significant at 5%; ** significant at 1%; *** significant at 0.1%

Table 5.4: Duration Analysis Results, Summary

ceding discussion of the six categories of explanations. In Appendix B.2, I estimate a separate model for each category of variable and then include them all in a single model. The discussion here refers to the single model that includes all the categories together. In it, not all the variables have a statistically significant effect on duration. In other words, for the variables that fail to clear this bar, there is a ninety-five percent chance that their impact cannot be distinguished from zero. However, in a multi-variate setting that controls for the impact of other variables that we think are substantively important, it is common for not all the variables to be able to be shown to have an impact. But we must still include them in the model or risk biasing the results. In what follows, I explain the substantive impact of the statistically significant variables: *area*, *gdp per capita*, *age of polity*, *executive constraints*, *prior civil war*, and *military personnel*.

For the continuous variables – area, gdp, and military personnel – I consider the impact of a shift from the mean of the variable to the mean plus one standard deviation. For the dichotomous variables – age of polity and prior civil war, I consider the effect of moving from a value of 0 to a value of 1, i.e. newly established polity to established polity, no prior civil war to prior civil war. For the categorical variable – executive constraints – I consider the effect of a one-category positive shift. Figure 5.1 plots the estimated effects of all six of the above variables on the probability that war continues.

It is important to note that because the regression model used here is non-linear, the effect sizes can vary with the size in the shift considered. Where helpful for the ease of interpretation, I discuss country examples.

Four of the six variables have a positive impact on duration. To start, consider country *area*. Moving from the mean in the data – a country of approximately

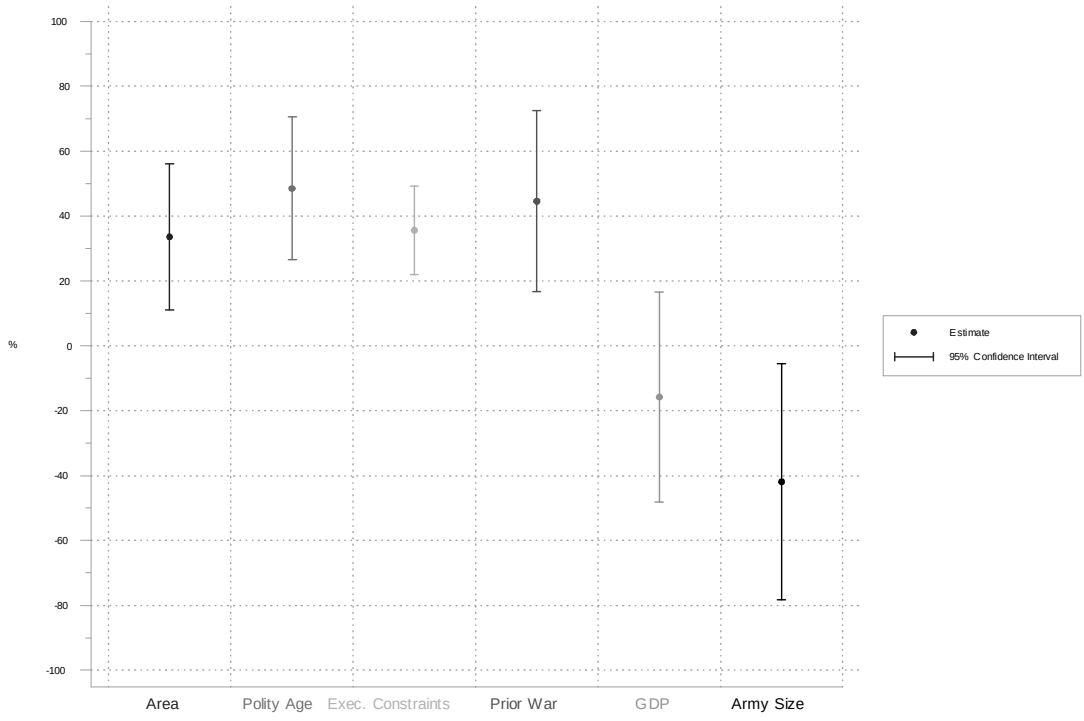

Figure 5.1: Estimated Effect on War Duration

478,487 square kilometers – up one standard deviation, to a country of approximately 2,691,053 square kilometers, increases the probability that civil war continues by 33.6%. The closest shift in the data itself is the difference in area between Papua New Guinea (462,000 square kilometers) and Argentina (2,777,000 square kilometers).

The extent to which the polity is established, captured by the variable *polity age*, also has a positive effect on duration. Of the 140 civil wars in the data, 93 (66.4%) occurred in established polities. Moving from a newly established polity to an established polity increases the probability of civil war continuing by 48.5%. Of countries in the data which experienced multiple civil wars, Zaïre even as of its third civil war beginning in 1977 was still a newly established polity. But as of its fourth civil war beginning in 1996, it was an established polity. Cambodia

was a newly established polity when its first civil war began in 1970, but by 1975 and the beginning of its second civil war, it was an established polity. Sudan was a newly established polity as of its first civil war in 1963, but an established polity as of its second civil war in 1983. The Philippines was a newly established polity at the time of its first civil war in 1950, but an established polity when its second civil war began in 1971. Indonesia was a newly established polity for its first three civil wars (beginning in 1950, 1953, and 1956), but by the time of its fourth civil war, starting in 1975, it was an established polity.

Checks and balances on the power of the chief executive also increase the probability that civil war continues. The variable, *executive constraints*, is measured on a scale of 1 to 7. The mean level in the data is nearly 3. Moving from a 3 to a 4 increases the probability that civil war continues by 19.8% (note that for this variable the effects are constant, i.e. any one-category increase in the variable has the same magnitude of effect). To take two examples from the data, Lebanon in 1958 had executive constraints of 3, while by its 1975 civil war it had greater constraints on the executive and scored a 4. Uganda moved in the other direction – as of its 1981 civil war, constraints on the executive stood at 4, while by its 1995 civil war they had fallen to 3.

Whether a country experienced a *prior civil war* also increases civil war duration. 77 of the 140 wars in the data (55%) had experienced a civil war within the twenty years prior to the civil war in question. Moving from no prior civil war to having experienced one already increases the probability that civil war continues by 44.6%.

The size of the economy has mixed effects on war duration. Figure 5.2 shows the more complex, curvilinear estimated effect of *gdp per capita*. For countries with

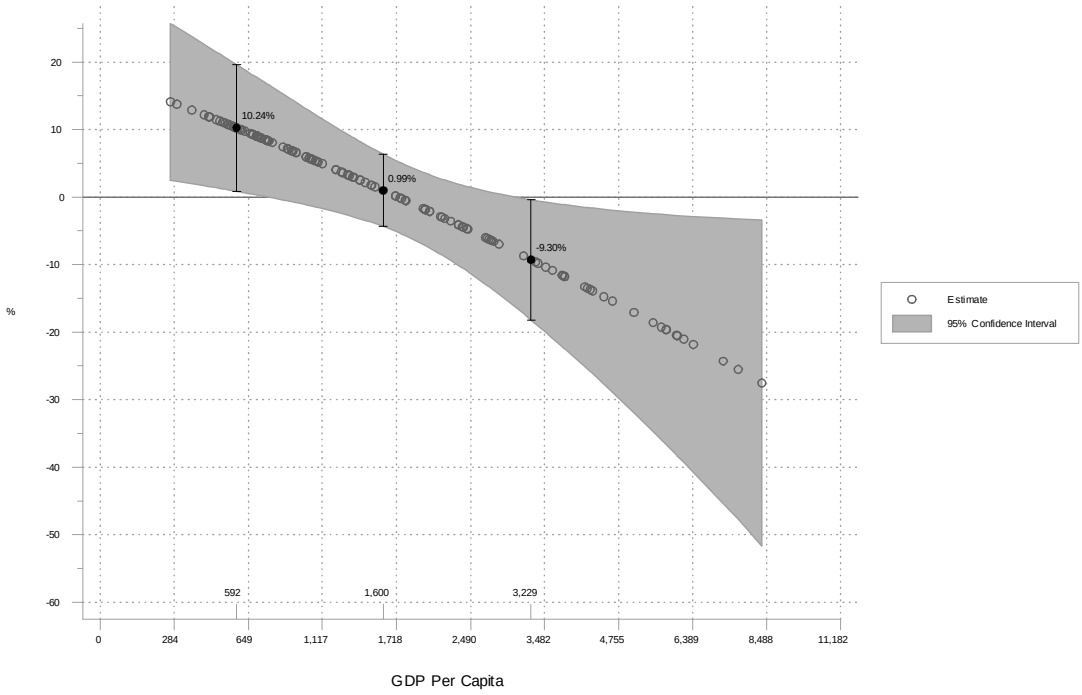

Figure 5.2: Estimated Effect of GDP Per Capita on War Duration

lower levels of per capita income, relatively small increases raise the probability that civil war continues. For example, from 592 dollars per capita, an increase of 100 dollars raises the probability that civil war continues by 10.24%. At 1,600 dollars per capita, close to the mean level in the data (1,595 dollars), this effect falls almost to zero – an increase of 100 dollars raises the probability that civil war continues by 0.99%. This resembles Afghanistan in 1992, which had a GDP per capita of 1,601, moving to Burundi's level of development in 1991 (1,705 dollars per capita). At higher levels of GDP per capita, the effect reverses. For example, at 3,229 dollars per capita (one standard deviation above the mean for the data), an increase of 100 dollars decreases the probability that war will continue by 9.30%. The latter resembles Sri Lanka in 1987, with a GDP per capita of 3,263, moving to Indonesia's level of development in 1990 (3,351 dollars per capita).

Finally, the size of the military decreases duration. The mean level of *military personnel* in the data is 246,992. Moving up one standard deviation, to 897,807, lowers the probability that civil war continues by 41.9%. This is equivalent to moving from Indonesia in 1976 (260,000 military personnel) to Turkey in 1984 (824,000 military personnel). If we again start at the mean but examine a smaller change in the number of military personnel, say, expanding the military by one division (20,000), the effect is to lower the probability that civil war continues by 1.3%.

5.3 Identifying Quagmire: Mind the Gap

The regression analysis described above yields each civil war's predicted length according to standard accounts of duration. I will refer to that predicted length as a war's expected duration. The operationalized definition of quagmire, laid out in section 5.1 above, consists of characterizing the gap between expected duration and a war's actual length. I use a technique developed in applied statistics for medicine to do so, producing a list of civil wars that experienced quagmire.

The gap between a war's expected duration and actual length – known as the regression's residual – is a convenient measure of the duration model's accuracy. But, for any given war, how should we evaluate the size of this gap? It is straightforward to compare it to some statistic calculated based on the gaps of all wars in the dataset, for example their average, or the extent to which they vary from that average. But – without going into statistical theory – this approach fails to take into account that the dataset here is one sample of the population of civil wars. The gap for a war that appears large compared to the average for this dataset

could nevertheless be close to the average across all potential datasets of civil wars. Consequently, we need to be able to compare the gaps that we calculate in this dataset to a theoretical distribution that the statistic of the residual should approximate if it were calculated for the population of civil wars. The residuals calculated from a linear regression model have a theoretical distribution. But the deviance residuals that we would calculate following a duration regression have no theoretical distribution. To establish what constitutes a significant gap between expected duration and actual length, we need to use some other statistic for which we can make this comparison to a known theoretical distribution.

A solution is found in the literature on applied statistics in medicine. Here, Nardi and Schemper (1999) propose a new statistic, the normal deviate residual, which follows a known distribution. This allows us to pick a cut-off point or "critical value" for a test, according to which wars can be judged to have experienced quagmire.

The normal deviate residuals for the duration model confirm that defining quagmire operationally according to the gap between expected and actual duration is quite distinct from looking at war duration. Figure 5.3 graphs war duration against the normal deviate residuals. A negative normal deviate residual implies that a war's actual length exceeds the expected duration. The figure illustrates that the duration of such wars varies widely. In other words, the set of wars that might have experienced quagmire includes wars of many lengths, both long and short.

To complete the operational definition of quagmire, we must select the size of gap that is sufficiently large to indicate that quagmire occurred. The normal deviate residuals, as their name suggests, follow a standard normal distribution;

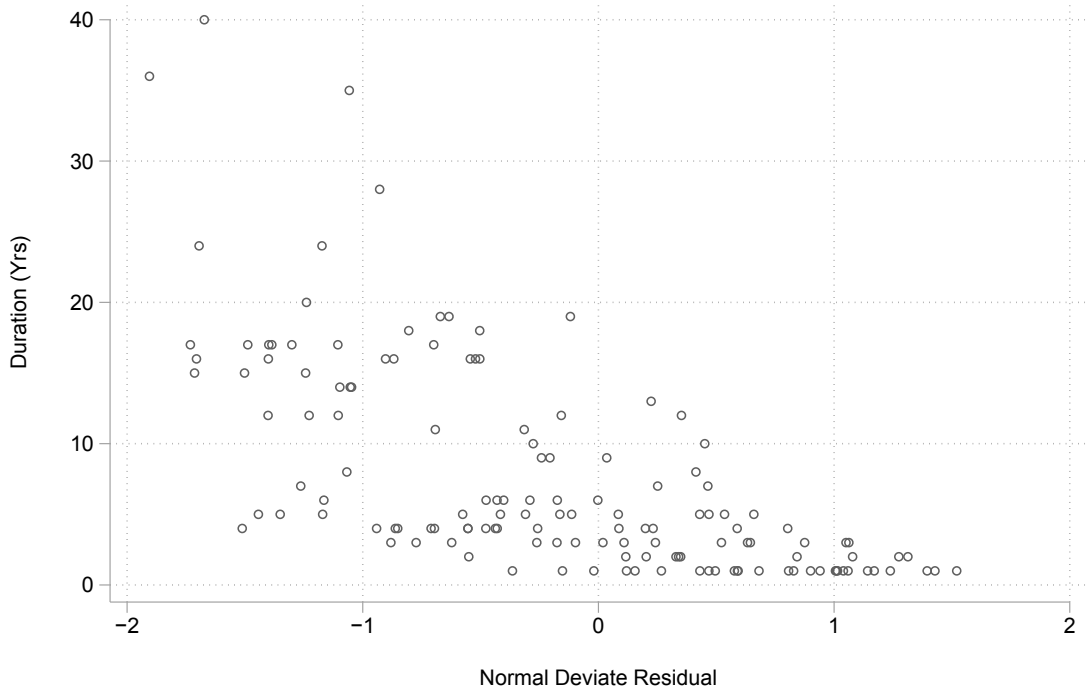

Figure 5.3: War Duration and Normal Deviate Residuals

two-thirds of the data should fall within a standard deviation of the mean, which is zero. A reasonable starting point, then, would be to use one standard deviation from the mean ($+1$, -1) as the cut-off by which to classify quagmire using these residuals. This implements the definition of quagmire in keeping with our intuitive sense that it is a rare phenomenon in civil wars, but one that occurs in a meaningful number of conflicts nevertheless. The operational definition considers quagmire to have occurred in wars the actual length of which exceeds expected duration. This calls for a one-sided test – we want to assess the chances that a given war would have the normal deviate residual in question compared to not having a negative normal deviate residual. In non-technical terms, we want to assess how likely it is that a given war's actual length surpassed the expected duration predicted by the duration model.

I select the 15th percentile of the distribution as the cut-off point. The 15 percentile is slightly more restrictive than one standard deviation from the mean, so using it makes the test a marginally harder one than one using the standard deviation. Using this cut-off, I identify 30 civil wars out of the 140 in the dataset (21.4 percent) as having experienced quagmire.[26] Table 5.5 shows the frequency of quagmire in a region's civil wars. Table 5.6 lists all civil wars that experienced quagmire.

Region	Quagmire	
	instances	% of region's wars
Asia	12	28.6
Europe (incld. USSR/former USSR)	4	25.0
Latin America	2	18.2
Middle East & N. Africa	4	19.0
Sub-Saharan Africa	8	16.0

Table 5.5: Regional Patterns of Quagmire

Substantively, the rationale for selecting the 15th percentile as cut-off is that it minimizes the total chances of errors in categorizing civil wars as having experienced quagmire. It results in a 15 percent chance of failing to reject a civil war as having experienced quagmire when it did not (known in statistics as a type I error). While we could choose a stricter cut-off to reduce this possibility, we must also consider the possibility of erring by failing to reject a civil war as *not* having experienced quagmire when it in fact did do so (a type II error). The 15th percentile cut-off also minimizes the chances of this second type of error, at 12 percent. Note that we have no a priori reason to think that either type of error

[26]For reference, Table 1.1 in Chapter 1 lists all the wars in analysis, highlighting those that experienced quagmire.

is more problematic than the other, so the most reasonable course of action is to reduce the chances of either.[27]

5.4 The Determinants of Quagmire

Having identified quagmire in civil wars, we can now compare this empirical record to the theoretical predictions. In what follows, I review these predictions and explain the data used to assess how the theory fares empirically. The results of the statistical analysis show the theory to be a useful guide for understanding whether quagmire occurs in civil war.

5.4.1 Strategic Interaction and Quagmires: Predictions

Chapter 2 mapped out the ways in which the decision-making of actors external and internal to a civil war is contingent upon one another's choices. I argued that this interplay is the linchpin through which quagmire emerges. The theory abstracted from the messy layers of civil war to analyze three parameters of importance and the relationships among them: foreign interests, the cost to a belligerent of escalating from non-territorial to territorial fighting, and the stakes of the conflict for the belligerents.

Two mechanisms that lead to quagmire emerged: foreign assistance as a subsidy and substitution between territorial and non-territorial fighting. The first mechanism suggests that foreign interests in a civil war country should exceed a threshold level before quagmire can occur; without the requisite importance, outside powers lack the motivation to furnish support to the belligerents. It also

[27] Appendix B.2, Section B.2.2 covers specifics of the effects of changing this cut-off point.

Country	Conflict	Start Year	End Year
Afghanistan	Mujahideen, PDPA	1978	1992
Afghanistan	Taliban vs. Gov't and US/NATO coalition	2001	
Algeria	FIS, AIS, GIA, GSPC	1992	
Angola	UNITA	1975	1991
Azerbaijan	Nagorno-Karabakh	1991	1994
Bangladesh	Chittagong Hills/Shanti Bahini	1974	1997
Burundi	Hutu groups vs. Gov't	1991	
Cambodia	FUNK; Khmer	1970	1975
Cambodia	Khmer Rouge, FUNCINPEC, etc.	1975	1991
Chad	FROLINAT, various ...	1965	1979
Chad	FARF; FROLINAT	1980	1994
Congo-Zaire	RCD, etc. vs. Gov't	1998	
El Salvador	FMLN	1979	1992
Ethiopia	Eritrean war of independence	1974	1991
Greece	EDES/ELAS; EAM	1944	1949
Indonesia	East Timor	1975	1999
Iraq	Kurds; Anfal	1985	1996
Lebanon	Aoun; militias; PLO; Israel	1975	1991
Morocco/Western Sahara	Polisario	1975	1991
Mozambique	RENAMO; FRELIMO	1976	1992
Myanmar/Burma	Communist insurgency	1948	1988
Myanmar/Burma	various ethnic groups; Karen rebellion 2	1960	1995
Peru	Sendero Luminoso, Tupac Amaru	1980	1996
Philippines	MNLF, MILF	1971	2006
Philippines	NPA	1972	1992
Russia	Chechnya 2	1999	
Thailand	Communists (CPT)	1966	1982
Uganda	LRA, West Nile, ADF, etc.	1995	
USSR	Ukraine/UPA	1944	1950
Vietnam	NLF	1960	1975

Table 5.6: Quagmire in Civil Wars, 1944-2006

predicts that beyond that threshold, as foreign interests increase so too does the likelihood of quagmire. The second mechanism suggests the importance of the cost of escalation, not in absolute terms but rather relative to the stakes. This ratio defines the domestic war environment. It predicts tandem dynamics that link the cost of escalation and the stakes to the prospects for quagmire. These dynamics kick in only once the foreign interests threshold is surpassed. In other words, the interaction between the domestic war environment and foreign interests is at the core of the predictions. As the cost of escalation increases, the likelihood of quagmire should rise, but mirroring this prediction, that same likelihood should fall as the stakes increase.

5.4.2 Measuring Foreign Interests and the Domestic War Environment

I aim to determine, using the statistical analysis, whether the theoretical concepts capture consistent patterns in the data. Here, I describe the variables I use to get at foreign interests, the cost of escalation, and the stakes of the conflict. Each is a vast area. I use variables that capture their essential characteristics. And I develop clusters of variables for each concept. This allows us to undertake a holistic analysis. By including a cluster of variables for each concept in the analysis, we can have confidence that consistent patterns in the results are linked to the theoretical concepts themselves. I will be less concerned with the results for any single variable and will instead emphasize whether the theoretical predictions about the concepts are borne out by the analysis.

The theory assumes that foreign states' decision of whether to provide support

to an internal actor turns on the level of interests at play in the civil war country. We therefore need to look at core ideas that get at what are the most weighty of these interests and how to measure different levels of them. To do so, I narrow the empirical by focusing on the interests of major international powers – rather than those of *any* foreign state – because of the preponderance of resources that they can bring to bear in any type of action that might affect a civil war country. The concept of interests is notoriously difficult to define with analytic precision, but clearly centers on a country's ability to secure itself from external threats and to maintain internal prosperity.[28] For governments throughout history, energy resources have been a critical element of the ability to maintain economic growth and project military power, defensive or offensive. Major powers have thus had an interest in securing access to energy resources, the most important of which have been hydrocarbons during the period of the empirical analysis.

The context of the international system also allows us to narrow the focus additionally. Following the end of the Second World War, and through nearly the end of the twentieth century, bipolarity was one of the defining characteristics of the international system (see Waltz 1979). It is also reasonable, therefore, to examine external interests in civil war countries by focusing on the major groupings of powers – the West and the East – as blocs that possessed unparalleled capacities for action on the international stage. Although the shift to a unipolar system after the fall of the Soviet Union meant a restructuring of international competition with respect to great powers (Wohlforth 2009, 1999), international politics at other levels, for example concerning regional issues, has continued to reflect the

[28] An assessment by the U.S. Secretary of Defense, for example, reflected these priorities: "a threat affects U.S. vital interests: If it threatens the survival of the United States or its key allies; If it threatens critical US economic interests; If it poses the danger of a future nuclear threat" (see OSD 1995:5).

broad alliances present during the Cold War and a dynamic of competition.[29] It is also important to note the consequences of the competitive dynamic: a country in which either bloc has some interests holds, by virtue of that fact, some level of interest for the other bloc.

I define the West as all countries that joined the North Atlantic Treaty Organization (NATO) alliance before the end of the Cold War. I define the East as the combination of all countries that were at one time members of the Warsaw Pact, Yugoslavia, the People's Republic of China, Cuba, and any of their successor states if applicable.

The West has had a strong interest in promoting global trade, including developing export markets for its goods. This has included the promotion of liberal political institutions (though not necessarily democracy) as a foundation for building capitalist economies.[30] As Assistant Secretary of the Navy for the United States, Franklin D. Roosevelt argued in 1913, "our national defense must extend all over the western hemisphere, must go out a thousand miles into the sea, must embrace the Philippines and over the seas wherever our commerce may be."[31]

The East, in contrast, traditionally defined its interests more narrowly as the maintenance of its military power and the security of its territory (see Feste 1992). This stems from the importance of Eurasia to the growth and maintenance of strategic power in international politics (see Mackinder 1904; Kissinger 1954; Brzezinski 1991; Gray 2004; Diamond 1997). The Soviet Union, later Russia, and to an extent China, therefore constitute targets for other states by virtue of their loca-

[29]In fact, as Brooks and Wohlforth (2008) point out, regional competition in the post-Cold War era is often mistaken as evidence for balancing against the U.S.

[30]See Odom and Dujarric (2004).

[31]Quoted in Borneman (2012:76).

tions.[32] Although security is also an important component of the West's interests, the United States' dominant role has meant a different emphasis on how security needs have translated into strategy, given the U.S.' protected geographic position, buffered by two oceans.[33]

I noted above the particular importance of commercial interests for the West, especially in the development of export markets. A civil war country's imports from the West can therefore be used as an indicator of Western interests in that country. Keeping in mind international competition, it is also likely to be a decent indicator of Eastern bloc interests to the extent that those are driven by the West's priorities. I measure a civil war country's imports from NATO members using the Correlates of War project's bilateral trade data (Barbieri, Keshk and Pollins 2008). I construct the variable *NATO imports* as the log of the average of the civil war country's imports from NATO members in a given year.[34] Due to endogeneity concerns, I use a time-invariant pre-war version of this variable.

Next, to capture the security component of external interests, I include a measure of the geographic proximity of the civil war country to the great powers of the post-Second World War era – the permanent five members of the United Nations Security Council. The variable *great power proximity* measures the minimum distance between the civil war country and these five countries.[35] As explained above, proximity is most likely to capture the interests of the East, so from among

[32] On the defense of Eurasia as central to the Russian foreign policy, see Clover (2016).

[33] Nuclear weapons also play an important role, with the U.S. nuclear umbrella guaranteeing Western European security (see Debs and Monteiro 2016).

[34] The import variables are calculated according to the formula:

$\ln\left(imports_{itJ}\right) = \ln\left(1 + \frac{1}{N_{Jt}} \sum_{j \in J_t} imports_{ijt}\right)$, where J_t is the set of countries in bloc J with observed data in year t, and imports are measured in millions of current USD.

[35] Computed from Weidmann, Kuse and Gleditsch (2010).

those five states, the USSR/Russian Federation and China. I include all five countries, however, since this may be a indicator of some level of security interests for the U.S., United Kingdom, and France, too. And again, recall that the international context makes this a decent indicator of the level of interests that a state may have in the civil war country in that competition creates interests where other states have interests.

Finally, international powers during the period of the analysis have had a strong interest in maintaining access to energy resources, specifically oil and natural gas. Whether a civil war country is a significant of these two resources is therefore likely to powerfully elevate its strategic importance as evaluated by external powers. I therefore include the *hydrocarbon exporter* variable described above in the duration analysis.

The theory considers the influence of two dimensions of the domestic war environment on belligerents' decision-making. The first, the cost of escalation in the level of fighting, captures the nature of war-fighting in the country in question. In particular, how much more costly is fighting designed to take and hold territory than fighting that lacks territorial aims? A country's physical and social geography form the key determinants of the cost of escalation. Many types of terrain that are difficult to pass through give an advantage to defenders, with elevated positions and urban areas being of particular consequence. Additionally, if an armed actor were to successfully dislodge her opponent and take a piece of territory, her ability to hold it would depend on the nature of any population that fell under her control as a consequence. If ethnicity divides the warring groups, holding territory populated by co-ethnics would be far less costly than trying to maintain control over territory populated by the other ethnic group, whose members could

supply the enemy with intelligence, might even become combatants and engage in acts of sabotage or guerilla attacks, and could help the enemy infiltrate the area as a precursor to a new battle for control over it. Social geography also plays an important role even in the case of an ethnically homogenous population. More densely populated areas would provide the enemy with greater cover by making it more difficult to distinguish civilians from enemy combatants. This increases the likelihood of enemy infiltration from outside, as well as that of civilian resistance from within. Opposing combatants and their supporters would be able to operate more easily and impose a higher cost on the armed actor that had taken control of the area than were the area it seized sparsely populated. For similar reasons, ethnic polarization, and its attendant residential ethnic segregation, also creates greater challenges for an armed actor seeking to take territory and maintain control over it. Heavily congested, structurally dense urban areas would also give the enemy cover by interfering with the armed actor's ability to surveil the population under its control.

I employ four variables to measure of the cost of escalation. Two come from the duration analysis described above – *rugged terrain* and *ethnic polarization*. I also include a measure of *population density*[36] and a measure of the population share of urban-dwellers, *urbanization*.[37] Per the discussion above, we should expect the cost of escalation to rise as each of these four variables increases in size, and therefore the likelihood of quagmire to increase.

The stakes of conflict constitute the second dimension of the domestic war en-

[36]I construct *population density* from Maddison's (2010) population data and and Banks' (2008) area data. Maddison does not provide complete temporal coverage for several countries in the dataset, so I interpolate and extrapolate population data where necessary to fill in missing values for 14 observations in two countries. See Appendix B.1.3 for further details.

[37]To construct *urbanization*, I use Banks' (2008) data on the size of the urban population.

vironment, reflecting the value of what each actor stands to gain if it were to win the war. Rather than representing spoils per se, the stakes can be thought of as the boost that winning would give to an actor's ability to fashion policies to its liking following the end of the war. If the two actors represent a central government and a secessionist region, the relevant policies might concern the autonomy of regions. In the extreme case, if the secessionist region were to unambiguously win the war, it would in effect be putting in place its preferred level of autonomy as independence from the central government. Alternatively, if the two actors represent the government and a challenger to it, any number of policies could be relevant, but the core issue would be whether the government remains in power or is replaced by the challenger. The side that wins a civil war will gain (or preserve) the ability to use government institutions and authority to consolidate its power, helping it to deter future challenges. Such future implications of victory that are therefore part of the stakes of war create a powerful incentive for actors to fight (see Powell 2012). Beyond the interaction with the current challenger, the incumbent can also obtain reputational benefits from victory. Should it win, it can cultivate a reputation for strength that will allow it to implement its preferred policies even in the face of what would otherwise have become opposition to it from new quarters. Walter (2006*a*,*b*) finds evidence for this logic in her analysis of ethnic group demands for self-determination and of government responses to separatist groups (for the theoretical basis of the argument, see Kreps and Wilson 1982).

These two illustrative scenarios highlight that the expected durability of the outcome of the conflict as a key aspect of the stakes. Not only might durability stem from the creation of a reputation by the winner, as mentioned above, but it might also be linked to structural factors largely outside of the reach of the actors

themselves. If, for example, the winner took or retained power in a country that was inherently unstable, the present victory might mean little for its ability to maintain its position in the future.

I include four variables already described in the duration analysis to get at the stakes via the durability of victory. First, *polity age*. How firmly rooted is the country itself, both within the population and in the international system? Second, *executive constraints*. The more democratic is a country, with executive constraints being a key element in this, the more it should be able to resolve political disputes through institutions without the resort to violence. Doyle and Sambanis (2006) provide a full discussion of this mechanism in explaining the higher durability of participatory peace following civil war – that is, peace with a modicum of democratic institutions – relative to a minimalist sovereign peace defined as the absence of war. Third, *prior war* within the twenty years. Has the country experienced civil war in recent times, understood socially, hence the use of a generation as the time period? Fourth, is the war over *secession*?

The link between the first two of these four variables and the stakes is straightforward. As a polity becomes more entrenched and as the constraints on executive power in government increase, the less likely future violent conflict becomes, and therefore the higher are the stakes of the current war. So as these two variables increase in value, the likelihood of quagmire should decrease.

The third and fourth variables are more complex. The experience of civil war in the recent past indicates, on the one hand, the possibility of lower stakes, since countries that have already experienced civil war are much more likely to see future civil wars. On the other hand, decision-makers and indeed ordinary citizens may well expect the current war to be more decisive if it comes on the heels of

a previous war than if it had occurred in isolation; less chance of a future war implies higher stakes in the current one.

If a war revolves around secession, this might imply, perhaps counterintuitively, that the stakes are lower than they would be if it were it a campaign to replace the central government. Should an incumbent government lose a secessionist war, it would still retain power in the rest of the country. And a defeat of secessionists could very well be accompanied by central government concessions towards the region in the form of autonomy. Indeed, for the elites involved, the consequences of a failed attempt at secession are not always grave. For example, in 2000, Russia elevated a former secessionist leader in Chechnya to head the regional government there, in effect making a former adversary its proxy ruler (Souleimanov 2015).[38] Such positive prospects for leaders aside, secession could also increase the stakes of conflict, since a successful result would likely be more durable than the successful result of a center-seeking war. Our expectations are therefore mixed when it comes to these two variables. *Prior war* and *secession* might each indicate lower stakes, and as a result lead us to predict a higher likelihood of quagmire; or might describe a situation with higher stakes than otherwise, with the resulting prediction of a lower likelihood of quagmire.

Table 5.7 lists the explanatory variables used in the analysis, organized by the theoretical concepts that they are designed to capture. It also lists the theory's predictions, both for the concepts and the individual variables used in the analysis. Appendix B.1 discusses coding details and the sources used for each variable. Nearly all variables are measured at the country level and during the pre-war period, the exception being whether the civil war in question was fought over

[38]Though it was perlious to hold that position. Kadyrov subsequently became president of the region in 2003 via an unfree election. He was assassinated some seven months later (Meyers 2004).

secession.

Concept (Prediction)	Variable	Prediction
Foreign interests (+)		
	NATO imports	+
	great power proximity	+
	hydrocarbon exporter	+
Domestic War Environment		
Cost of Escalation (+)		
	population density	+
	urbanization	+
	ethnic polarization	+
	terrain	+
Stakes (–)		
	polity age	–
	executive constraints	–
	prior war	+ / –
	secession	+ / –

Note: Predictions apply only after foreign interests threshold reached. See Section 5.4.1 above.

Table 5.7: Expected Effects on the Likelihood of Quagmire, Summary

The attentive reader will notice that not all of the variables within the domestic war environment category map exclusively onto the concepts of the cost of escalation or the stakes. That some can be interpreted as applying to both concepts certainly appears to be a limitation of the analysis. But it is not as problematic as it might seem. First, while it would be ideal to find such variables, in practice doing so at the level of the country or war is often not possible. In other words, this may be a natural difficulty when it comes to cross-country statistical analysis. Rather than forego the analysis because of it, we can still see what we can learn from it even with such a limitation. Second, and more importantly, the theory's tandem predictions for the cost of escalation and the stakes often imply that a vari-

able can be interpreted as representing either one without contradiction. As long as increasing values of a variable represent both increasing cost of escalation and decreasing stakes, then the prediction is that the likelihood of quagmire increases. Similarly, decreasing cost of escalation and increasing stakes share the prediction of the likelihood of quagmire decreasing.

Any variable hypothesized to increase the cost of escalation but which also might increase the risk of a new war is tied to the stakes of conflict via the durability of victory mechanism. Two variables included in the analysis – *rugged terrain* and *ethnic polarization* – might fill both roles. This alternative interpretation does not alter the predictions about their effect on the likelihood of quagmire. If each one decreases the stakes by making a future war more probable, we would expect quagmire to be more likely. This is the same expectation we already have about the effect of each variable as a measure of an increasing cost of escalation.

5.4.3 The Incidence of Quagmire: Analysis

I estimate a logit model of quagmire. The dependent variable is dichotomous and indicates whether the civil war in question experienced quagmire. I include as explanatory variables those described above and listed in Table 5.7. I also include interactions between each of the foreign interests variables and the domestic war environment variables to capture the hypothesized threshold effect of foreign interests. Appendix B.2 presents the statistical analysis in detail.

Figures 5.4 and 5.5 graph the estimated marginal effects of the explanatory variables – that is, the change in the probability of quagmire in civil war produced by an infinitessimal shift in the explanatory variable in question – for dichotomous

variables, though, the shift from 0 to 1 – while holding all others constant.[39] Figures 5.6-5.9 graph the estimated marginal effects of the domestic war environment variables across the values of each of the foreign interests variables. The results of the analysis provide clear support for theory of quagmire's empirical expectations.

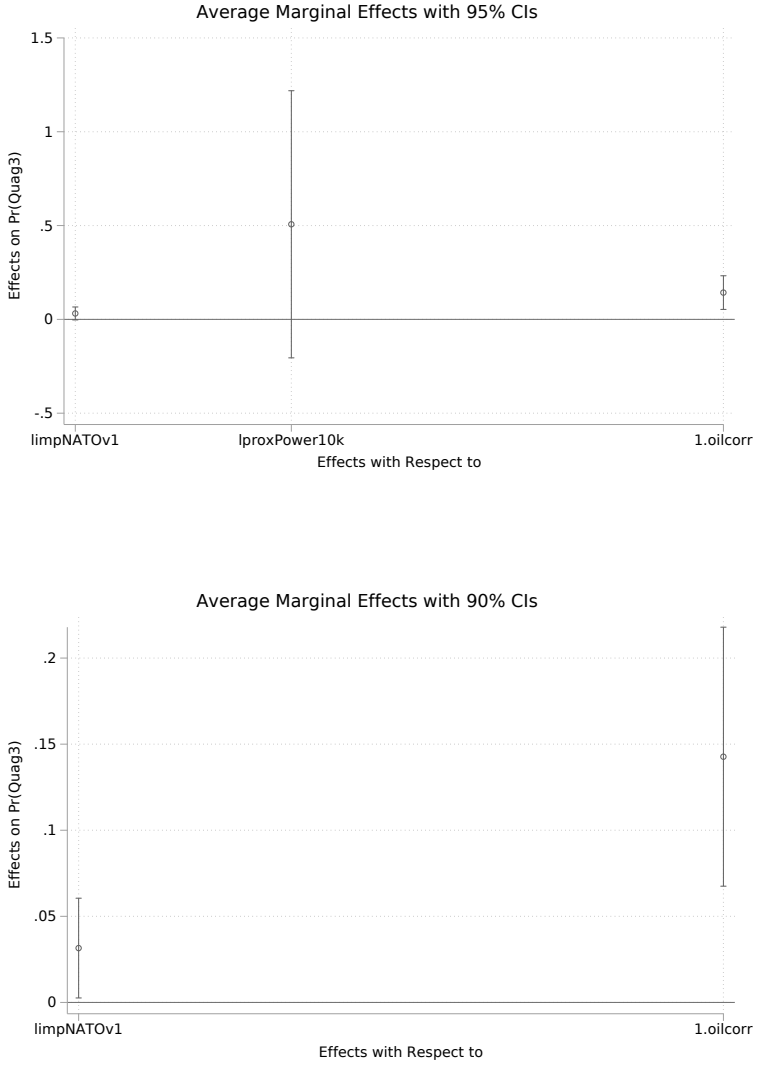

Figure 5.4: Marginal Effects on the Probability of Quagmire, Interests Variables, 90% and 95% Confidence Intervals

[39]Each variable's average marginal effect is plotted. This statistic is defined as the average of the marginal effects calculated for "each observation at its observed values." See Long and Freese (2014:242-6) for further discussion of this and other types of marginal effects.

Of first note is that, as predicted, the probability of quagmire increases as foreign interests increase. Figure 5.4 shows the positive and statistically significant estimated marginal effects of *NATO imports* and *hydrocarbon exporter*.[40] The former increases the probability of quagmire by 3.2%,[41] while going from being a non-hydrocarbon exporter to being one increases the probability of quagmire by 14.3%.

Figure 5.5 shows that the domestic war environment variables do not, on their own, have statistically significant effects on the probability of quagmire. However, taking into account the interactions with the interests variables, the full results do support the theoretical expectations of, first, a threshold effect of foreign interests; and, second, of an increasing probability of quagmire as the cost of escalation increases and as the stakes decrease. We can observe this by examining the marginal effects of the domestic war environment variables across the levels of each foreign interests variable.

To begin, we examine the interaction between foreign interests as *NATO imports* and the domestic war environment. Three of the four variables that capture the cost of escalation show the predicted marginal effect once the level of *NATO imports* crosses a threshold level: *rugged terrain*,[42] *ethnic polarization*, and *population density* (Figure 5.6).[43] The two variables that capture the stakes of conflict and for which we have unambiguous expectations each show the predicted marginal

[40]While the estimated marginal effect of *great power proximity* is positive, it is not statistically significant.

[41]Statistically significant at the 10% level.

[42]Statistically significant at the 10% level.

[43]The marginal effects for the interactions between *NATO imports* on the one hand, and *rugged terrain*, *ethnic polarization*, and *population density*, on the other hand, are also consistent with the theory's predicted negative relationship between the stakes and quagmire. These cost of escalation variables also representing decreasing stakes due to a positive relationship with the prospect of future war.

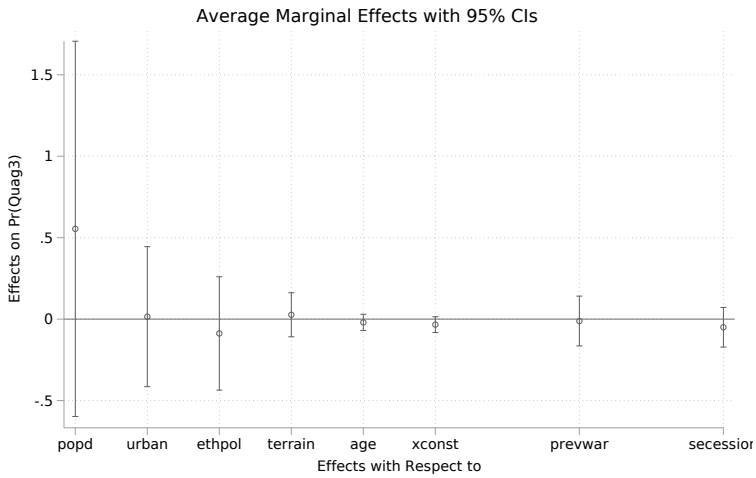

Figure 5.5: Marginal Effects on the Probability of Quagmire, Domestic War Environment Variables, 95% Confidence Interval

effect: *age of polity* and *executive constraints* (Figure 5.7). A third stakes variable, *secession*, exhibits a negative relationship with the probability of quagmire,[44] which is consistent with our expectations if we interpret that variable as increasing the stakes. Neither of the other two war environment variables – *urbanization* and *prior war* – show a relationship contrary to expectations.

Next, we look at the interaction between foreign interests as *great power proximity* (Figure 5.8). Viewed across the values of great power proximity, *rugged terrain* has the anticipated positive marginal effect.[45] Consistent with the results for *NATO imports*, we again find that *secession* has a negative marginal effect,[46] which accords with our expectations if *secession* increases the stakes of civil war. The only result that might run counter to the theory's predictions concerns *urbanization*. As depicted in Appendix B.2, at the very highest levels of *great power proximity*, it exhibits a negative marginal effect. However, for all other levels it shows no sta-

[44]It is negative and statistically significant at 10% level for a small portion of the range of *NATO imports*.

[45]Statistically significant at the 10% level.

[46]Statistically significant at the 10% level.

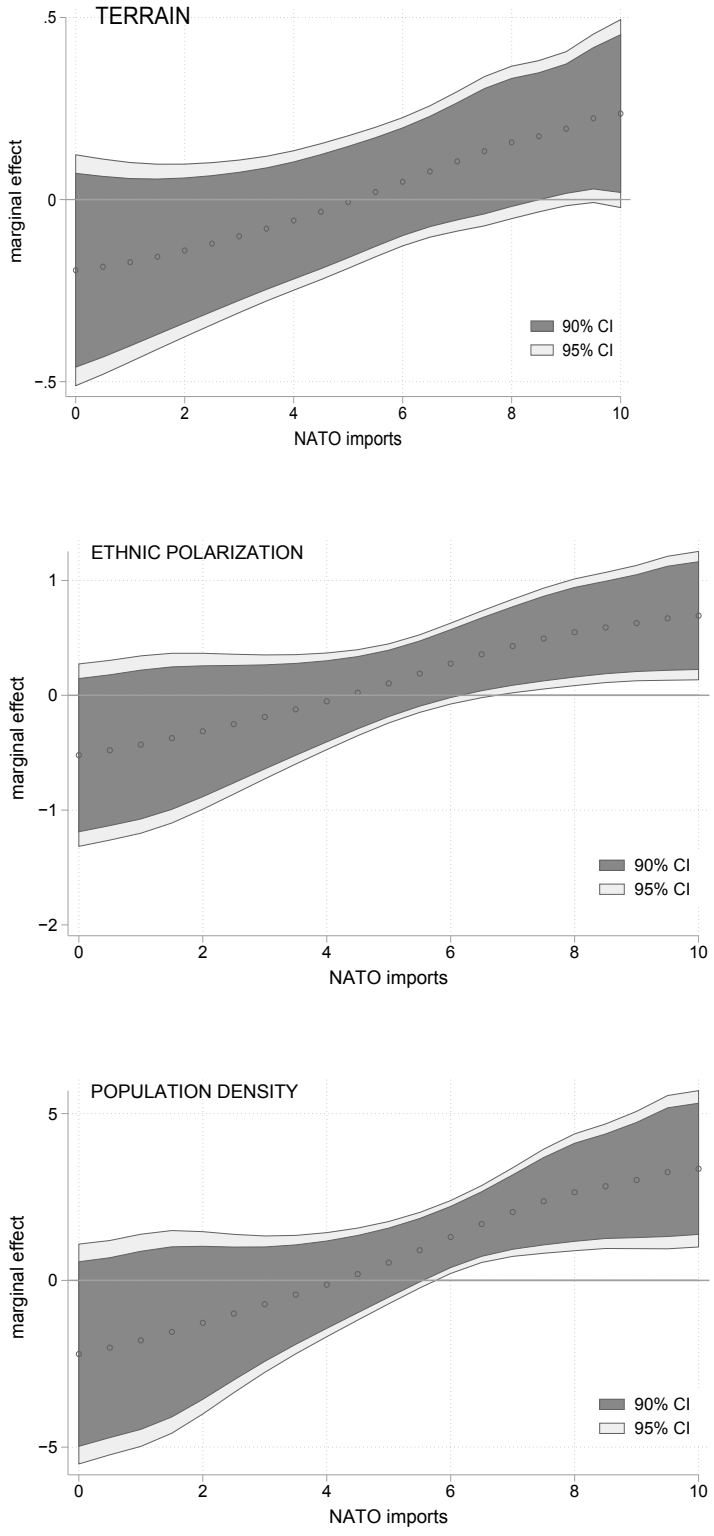

Figure 5.6: Marginal Effects on the Probability of Quagmire, Cost of Escalation Variables Interacted with Foreign Interests as NATO Imports

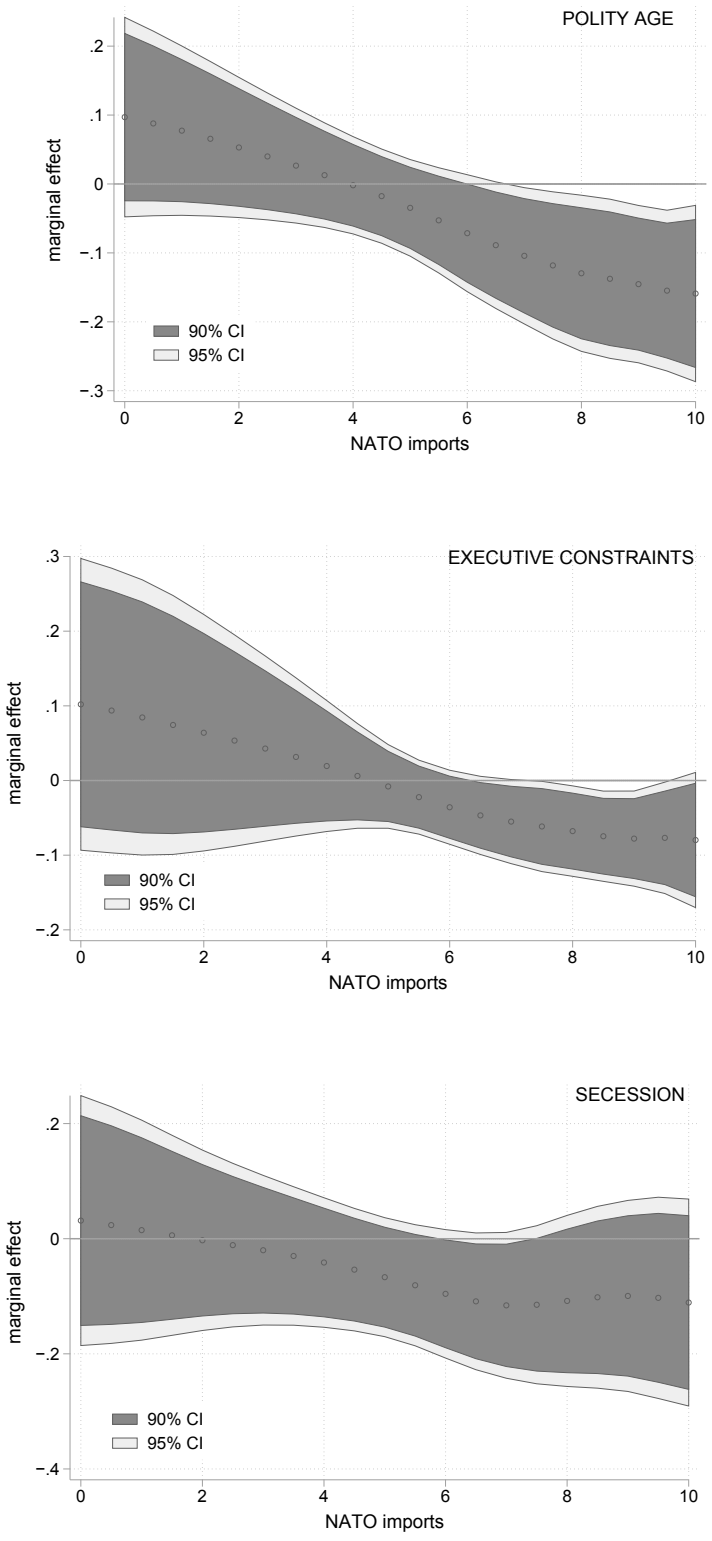

Figure 5.7: Marginal Effects on the Probability of Quagmire, Stakes Variables Interacted with Foreign Interests as NATO Imports

tistically significant relationship with the probability of quagmire.

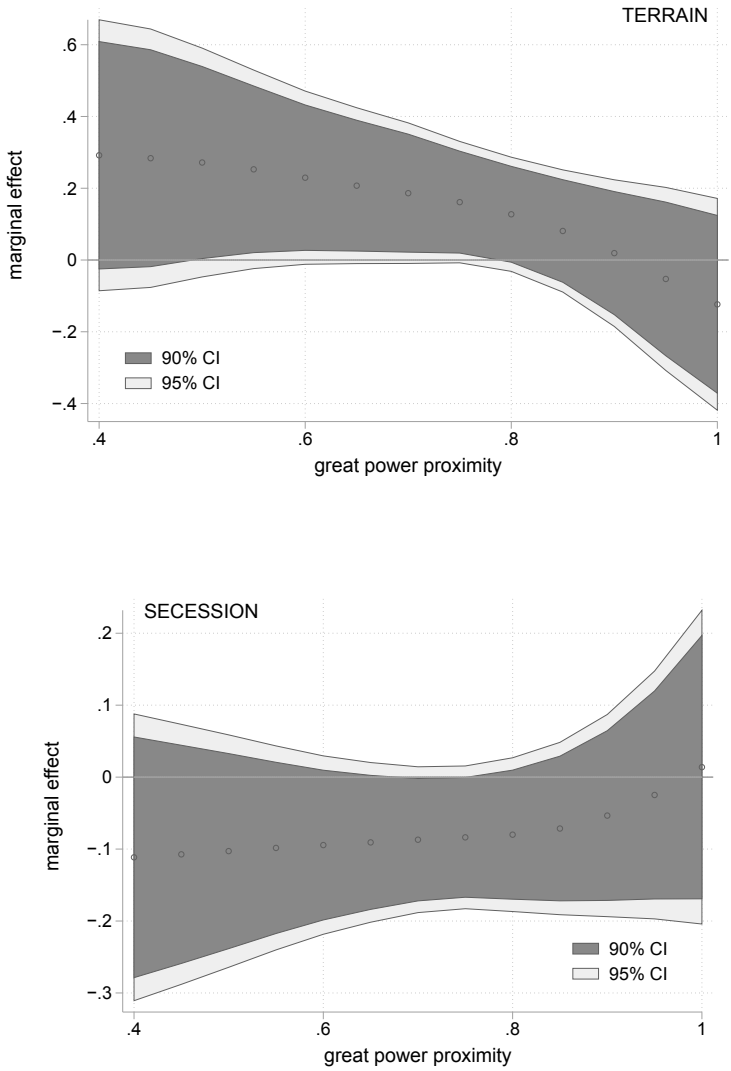

Figure 5.8: Marginal Effects on the Probability of Quagmire, Interactions with Foreign Interests as Great Power Proximity

Finally, we view the marginal effects compared over *hydrocarbon exporter* status (Figure 5.9). The results here show, for one each of the cost of escalation and stakes variables, the expected threshold effect of interests and the expected direction of the relationship between the domestic war environment variables and the likelihood of quagmire. If exporter status is positive, *rugged terrain* has a positive

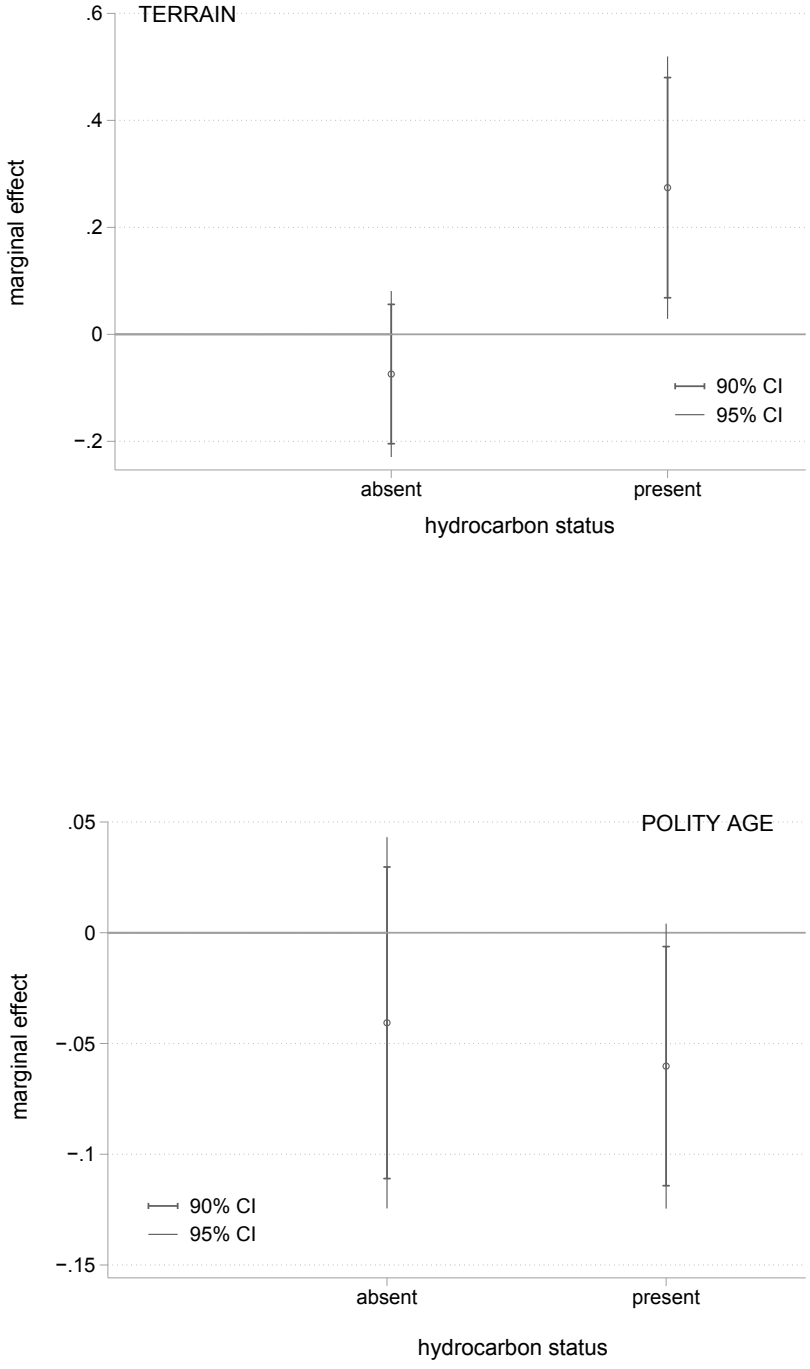

Figure 5.9: Marginal Effects on the Probability of Quagmire, Interactions with Foreign Interests as Hydrocarbon Exporter Status

and statistically significant effect and *polity age* has a negative and statistically significant marginal effect[47]; neither variable has a statistically significant effect for non-exporter countries. Though not pictured here, the results also include the expected threshold effect for *prior war*. Specifically, it has a statistically significant marginal effect if exporter status is positive and no statistically significant effect otherwise. The direction of the effect is negative, but recall that the expectation for prior war was ambiguous. The salient aspect of the results here is not the direction of the marginal effect but finding the threshold effect itself, i.e. that there is a statistically significant marginal effect only in interaction with *hydrocarbon exporter* status.

Two additional results merit discussion. Though not pictured here, these are included in Appendix B.2. With non-exporter status, *executive constraints* and *secession* each have a negative and statistically significant marginal effect, which changes to positive and is still statistically significant with exporter status. The initial negative results fit the predicted relationship between increasing stakes and quagmire, though we do not see a threshold effect of foreign interests, at least when operationalized as *hydrocarbon exporter* status. And the positive results when that exporter status changes could be due to the effect of that variable individually outweighing any effect of *executive constraints* or *secession*, and do not necessarily indicate relationships counter to the theory's predictions. The result for urbanization is consistent with this possibility – with non-exporter status, urbanization has a negative and statistically significant marginal effect[48], but no statistically significant effect with exporter status. The effects of hydrocarbon exporter status may therefore wash out the possible independent effect of urbanization.

[47] The latter at the 10% level.

[48] At the 10% level.

* * *

Across multiple variables that proxy for the domestic war environment, and using three different measures of foreign interests, we see results consistent with the theory of quagmire. A threshold effect of foreign interests is evident, as is positive effect of the cost of escalation and the corresponding negative effect of the stakes of conflict. In addition, for *secession*, one of the variables for which we had ambiguous predictions, the results are consistent across all three measures of foreign interest. The predicted effects are present for an array of variables designed to capture the same concept.

5.5 Discussion

The statistical investigation of quagmire in civil wars is challenging. This chapter makes two contributions that confront that challenge. It tackles the question of defining quagmire consistently across civil wars, conflicts which may vary considerably when it comes to the conduct of warfare, not to mention their geographic, historical, political, economic, and social circumstances. With the cross-country framework that is suited to examining the theory's general applicability to understanding quagmire come considerable limits on analytical precision. Even so, statistical analysis demonstrates that patterns consistent with the theory are in evidence and so provides confidence in the empirical case for the theory.

By identifying quagmire in civil wars in a systematic fashion, I also hope to lay the groundwork for future research. The list, whatever its flaws, can serve as a starting point for the analysis of the causes of quagmire and its effects. The exercise of creating it also demonstrates that intuitive political concepts – often loosely

and unsystematically flung about in public discourse and policy circles– can nevertheless be studied quantitatively. Fortuitously for the study of quagmire, an existing technique could be applied from biomedical statistics. Other similarly intuitive but unsystematically defined concepts, then, may well be fruitfully studied quantitatively, if we search broadly for the necessary tools.

The statistical analysis of the determinants of quagmire constituted a difficult test of the theory of quagmire. The dependent variable was the unexplained variation from a duration analysis. Many of the explanatory variables had already been used in the duration analysis – unlikely candidates to explain variation that they had already failed to explain. And yet we still obtained a clear pattern of results: across three quite different measures of foreign interests a range of variables to capture the domestic war environment, we saw a clear threshold effect of foreign interests, and the predicted relationships between foreign interests, the cost of escalation, and the stakes, and the outcome of quagmire.

The findings are not intended as the final word on quagmire in civil war. But they provide confidence that the concept can be evaluated generally and that the theory of quagmire this book develops provides explanatory traction. The mechanisms highlighted by the analysis of Lebanon in the previous chapter should be considered seriously well beyond the borders of that small country and the temporal confines of the last decade and a half of the Cold War.

Chapter 6 moves the empirical exploration forward by maintaining the comparative element of this chapter's analysis but going beyond its necessarily broad treatment of the concepts of foreign interests and the domestic war environment. In it, I analyze the experience of quagmire in Chad's civil war between 1960 and 1979 and contrast it with Yemen's 1994 civil war, which might have shared that

fate but in the end side-stepped it. This comparison provides a study of the paths to quagmire that is deeper than a cross-country statistical analysis permits, yet complements the analysis and results just presented here by maintaining a dialogue with them.

Appendix B

Cross-Country Data and Statistical Analysis

B.1 Data

B.1.1 Civil War List

While there is no standard unit of measurement for civil war duration in the literature, the choice is a significant one. Many quantitative analyses specify the start and end dates of civil wars down to the month, if not the day, and measure duration in these units.[1] However, as Montalvo and Reynal-Querol (2010) note, measuring duration in months or days, which some researchers do on the grounds that they are therefore measuring the dependent variable with greater accuracy (see Gates and Strand 2004; Cunningham 2010), is in fact likely to introduce far greater measurement error than if duration is approximated in years, due to the "extreme unreliability" of the estimates of deaths that researchers use in coding the beginning and end of wars. Table B.1 provides the list of civil wars

[1] Regan (2002), Collier, Hoeffler and Söderbom (2004), DeRouen and Sobek (2004), Cunningham (2006), and Gent (2008) measure duration in months, while Balch-Lindsay and Enterline (2000), Balch-Lindsay, Enterline and Joyce (2008), Buhaug, Gates and Lujala (2009), and Cunningham (2010) measure it in days, among others. Other studies using a yearly metric for duration include Elbadawi and Sambanis (2000) and Fearon (2004).

used in the empirical analyses in this chapter, which is an update of Sambanis' (2004) war list from 1999 through to the end of 2006. Durations for ongoing wars are therefore censored at the end of 2006. The tables lists the length up to that cut-off; the statistical analysis treats censored observations appropriately.

Table B.1: Civil Wars, 1944-2006

War No.	Country	Dates (Year Month)				Conflict Name	Duration	
		Start		End			Months	Yrs Coded
1	Afghanistan	1978	4	1992	2	Mujahideen, PDPA	167	14
2	Afghanistan	1992	2	1996	9	Taliban v. Burhanuddin Rabbani	56	5
3	Afghanistan	1996	9	2001	10	United Front v. Taliban	62	5
4	Afghanistan	2001	10			Taliban vs. Gov't, US/NATO coalition	63	5
5	Algeria	1962	7	1963	1	post-independence strife	7	1
6	Algeria	1992	1			FIS, AIS, GIA, GSPC	180	15
7	Angola	1975	11	1991	5	UNITA	187	16
8	Angola	1992	9	1994	11	UNITA	27	2
9	Angola	1994		1997	12	Cabinda; FLEC	48	4
10	Angola	1997	1	2002		UNITA	72	6
11	Argentina	1955	6	1955	9	Peron v. military	4	1
12	Argentina	1975		1977		Montoneros, ERP, Dirty War	36	3
13	Azerbaijan	1991	4	1994	10	Nagorno-Karabakh	43	4
14	Bangladesh	1974	1	1997	12	Chittagong Hills/Shanti Bahini	288	24
15	Bolivia	1952	4	1952	4	MNR rebellion in La Paz	1	1
16	Bosnia	1992	3	1995	11	Rep. Srpska/Croats	45	4
17	Burundi	1965	10	1969	4	Hutu uprising	43	4
18	Burundi	1972	4	1972	12	Hutu uprising	9	1
19	Burundi	1988	8	1988	8	Org. massacres on both sides	1	1
20	Burundi	1991	11			Hutu groups v. govt	182	15
21	Cambodia	1970	3	1975	1	FUNK; Khmer	59	5
22	Cambodia	1975	5	1991	10	Khmer Rouge, FUNCINPEC, etc	198	17
23	Central African Rep.	1996	4	1997	1	Factional fighting	10	1
24	Chad	1965	10	1979	6	FROLINAT, various...	164	14
25	Chad	1980	3	1994	6	FARF; FROLINAT	172	14
26	Chad	1994	6	1997	8	FARF; FROLINAT	39	3
27	Chad	2005	12			MDJT, SCUD, RLD vs. Gov't	13	1
28	China	1946	3	1949	1	PLA	35	3
29	China	1947	2	1947	3	Taiwanese v. Nationalist soldiers	2	1
30	China	1950	2	1951	4	re-annexation	15	1
31	China	1956	3	1959	3	Tibetan uprising	37	3
32	China	1967	1	1968	9	Red Guards	21	2
33	Colombia	1948	4	1966		La Violencia	225	19
34	Colombia	1978	11			FARC, ELN, drug cartels, etc	338	28
35	Congo – Brazzaville	1993	5	1997	10	Lissouba v. Sassou-Nguesso	54	5
36	Congo – Brazzaville	1998	8	1999	11	Cobras v. Ninjas	16	1
37	Congo-Zaire	1960	7	1965	11	Katanga, Kasai, Kwilu, Eastern	65	5
38	Congo-Zaire	1967	7	1967	11	Kisangani mutiny	5	1
39	Congo-Zaire	1977	11	1978	5	FLNC; Shabba 1 & 2	7	1
40	Congo-Zaire	1996	10	1997	5	AFDL (Kabila)	8	1
41	Congo-Zaire	1998	8			RCD, etc v. govt	101	8
42	Costa Rica	1948	3	1948	4	NLA	2	1
43	Croatia	1992	2	1995	12	Krajina, Medak, Western Slavonia	47	4
44	Cuba	1958	6	1959	1	Castro revolution	8	1
45	Cyprus	1963	11	1967	10	GC-TC civil war	48	4

Continued on next page

Table B.1 – continued from previous page

War No.	Country	Dates (Year Month)				Conflict Name	Duration	
		Start		End			Months	Yrs Coded
46	Cyprus	1974	7	1974	8	TCs; GCs; Turkish invasion	2	1
47	Djibouti	1991	11	1994	12	FRUD	38	3
48	Dominican Republic	1965	4	1965	9	Mil. coup	6	1
49	Egypt	1994	2	1997	9	Gamaat Islamiya; Islamic Jihad	44	4
50	El Salvador	1979	10	1992	1	FMLN	148	12
51	Ethiopia	1974	1	1991	5	Eritrean war of independence	209	17
52	Ethiopia	1976	1	1988	4	Ogaden; Somalis	148	12
53	Ethiopia	1978	3	1991	5	Ideological; Tigrean	159	13
54	Georgia	1991	1	1992	7	South Ossetia	19	2
55	Georgia	1992	7	1994	1	Abkhazia (& Gamsakhurdia)	19	2
56	Greece	1944	12	1949	10	EDES/ELAS; EAM	58	5
57	Guatemala	1966	10	1972	12	Communists;	75	6
58	Guatemala	1978	3	1994	4	Communists; Indigenous	194	16
59	Guinea-Bissau	1998	6	1999	5	Vieira v. Mane mutiny	12	1
60	Haiti	1991	10	1995	12	Cedras v. Aristide	51	4
61	India	1946	7	1948	6	Partition and ethnic rioting	24	2
62	India	1984	2	1993	12	Sikhs	119	10
63	India	1989				Naxalites (CPI-M; PWG; MCC)	216	18
64	India	1989	12			Kashmir	205	17
65	India	1990	11			Assam; Northeast States	194	16
66	Indonesia	1950	7	1950	12	Rep. S. Moluccas	6	1
67	Indonesia	1953	9	1953	11	Darul Islam	3	1
68	Indonesia	1956	12	1960	12	Darul Islam, PRRI, Permesta	49	4
69	Indonesia	1975	12	1999	10	East Timor	287	24
70	Indonesia	1976		1978		OPM (West Papua)	36	3
71	Indonesia	1990		1991		Aceh	24	2
72	Indonesia	1999		2005	8	Aceh	80	7
73	Iran	1978	9	1979	12	Khomeini	16	1
74	Iran	1979	3	1984	7	KDPI (Kurds)	65	5
75	Iraq	1959	6	1959	10	Shammar	5	1
76	Iraq	1961	9	1970	3	KDP, PUK (Kurds)	103	9
77	Iraq	1974	3	1975	4	KDP, PUK (Kurds)	14	1
78	Iraq	1985	1	1996	10	Kurds; Anfal	142	12
79	Iraq	1991	3	1993	12	Shiite uprising	34	3
80	Iraq	2003	3			US/Coalition occupation, Iraqi civil war	46	4
81	Israel	1987	1	1997	1	Intifada; Palestinian conflict	121	10
82	Israel	2000	9			Intifada; Palestinian conflict	76	6
83	Ivory Coast	2002	9	2005	4	Coup, Gov't vs. Forces Nouvelles	32	3
84	Jordan	1970	2	1971	7	Fedeyeen/Syria v. govt	18	2
85	Kenya	1963		1967		Shifta war (Somalis)	60	5
86	Kenya	1991	10	1993	9	Rift valley ethnic violence	24	2
87	Korea	1948	10	1949	4	Yosu Rebellion	7	1
88	Laos	1960	10	1973	2	Pathet Lao	149	12
89	Lebanon	1958	5	1958	9	Nasserites v. Chamoun	5	1
90	Lebanon	1975	4	1991	9	Aoun; militias; PLO; Israel	198	17
91	Liberia	1989	12	1990	11	Doe v. rebels	12	1
92	Liberia	1992	1	1997	7	NPLF; ULIMO; NPF; LPC; LDF	67	6
93	Liberia	1999	5	2003	8	anti-Taylor resistance	52	4

Continued on next page

Table B.1 – continued from previous page

War No.	Country	Dates (Year Month)				Conflict Name	Duration	
		Start		End			Months	Yrs Coded
94	Mali	1990	6	1995		Tuaregs; Maurs	67	6
95	Moldova	1991	12	1992	7	Transdniestria	8	1
96	Morocco	1975	12	1991		W. Sahara/Polisario	193	16
97	Mozambique	1976		1992	7	RENAMO; FRELIMO	199	17
98	Myanmar/Burma	1948	9	1951	7	Karen rebellion 1	35	3
99	Myanmar/Burma	1948	9	1988		Communist insurgency	484	40
100	Myanmar/Burma	1960		1995	3	Various ethnic groups; Karen rebellion 2	423	35
101	Namibia	1973	1	1989	5	SWAPO; SWANU; SWATF	197	16
102	Nepal	1996	2			CPN-M/UPF (Maoists)	131	11
103	Nicaragua	1978	10	1979	7	FSLN	10	1
104	Nicaragua	1981	2	1990	4	Contras & Miskitos	111	9
105	Nigeria	1967	1	1970	1	Biafra	37	3
106	Nigeria	1980	12	1985	4	Muslims; Maitatsine rebellion	53	4
107	Oman	1971	1	1975	3	Dhofar rebellion	51	4
108	Pakistan	1971	3	1971	12	Bangladesh secession	10	1
109	Pakistan	1973	1	1977	7	Baluchistan	55	5
110	Pakistan	1994	11	1999	10	MQM, Sindhis v. Mohajirs	60	5
111	Papua New Guinea	1988	11	1998	4	BRA (Bougainville)	114	10
112	Paraguay	1947	3	1947	8	Febreristas, Libs, Comms	6	1
113	Peru	1980	7	1996	12	Sendero Luminoso, Tupac Amaru	198	17
114	Philippines	1950	9	1952	7	Huks	23	2
115	Philippines	1971		2006		MNLF, MILF	432	36
116	Philippines	1972	9	1992	9	NPA	241	20
117	Russia	1994	12	1996	8	Chechnya 1	21	2
118	Russia	1999	9			Chechnya 2	88	7
119	Rwanda	1963	11	1964	2	Tutsi uprising	4	1
120	Rwanda	1990	9	1993	8	Hutu vs. Tutsi groups	36	3
121	Rwanda	1994	4	1994	7	RPF; genocide	4	1
122	Senegal	1989	1	1999	12	MFDC (Casamance)	132	11
123	Sierra Leone	1991	3	1996	11	RUF, AFRC, etc.	69	6
124	Sierra Leone	1997	5	2001	5	post-Koroma coup violence	49	4
125	Somalia	1988	5	1991	1	SSDF, SNM (Isaaqs)	33	3
126	Somalia	1991	5			post-Barre war	188	16
127	South Africa	1976	6	1994	4	ANC, PAC, Azapo	215	18
128	Sri Lanka	1971	4	1971	5	JVP	2	1
129	Sri Lanka	1983	7	2002	2	LTTE, etc.	224	19
130	Sri Lanka	1987	9	1989	12	JVP II	28	2
131	Sri Lanka	2003				LTTE, etc.	48	4
132	Sudan	1963	10	1972	3	Anya Nya	102	9
133	Sudan	1983	7	2002	7	SPLM, SPLA, NDA, Anyanya II	229	19
134	Sudan	2003				Darfur, SLA, JEM, etc.	48	4
135	Syria	1979	6	1982	2	Muslim Brotherhood	33	3
136	Tajikistan	1992	4	1997	6	Popular Democratic Army; UTO	63	5
137	Thailand	1966	1	1982	12	Communists (CPT)	204	17
138	Thailand	2004	1			Pachani; Pulo, BRN, RKK, GMIP	36	3
139	Turkey	1984	8	2000	2	PKK (Kurds)	187	16
140	Uganda	1966	5	1966	6	Baganda rebellion	2	1
141	Uganda	1978	10	1979	4	Tanzanian war	7	1

Continued on next page

Table B.1 – continued from previous page

War No.	Country	Dates (Year Month)				Conflict Name	Duration	
		Start		*End*			*Months*	*Yrs Coded*
142	Uganda	1981	1	1987	1	NRA/Museveni, etc	73	6
143	Uganda	1990	1	1992	7	Kony (pre-LRA)	31	3
144	Uganda	1995	1			LRA, West Nile, ADF, etc.	144	12
145	United Kingdom	1971	1	1998	4	Northern Ireland	328	27
146	USSR	1944		1947		Latvia/LTSPA, etc.	36	3
147	USSR	1944		1948		Lithuania/BDPS	48	4
148	USSR	1944		1948		Estonia/Forest Brthers	48	4
149	USSR	1944		1950		Ukraine/UPA	72	6
150	Vietnam	1960	1	1975	4	NLF	184	15
151	Yemen	1994	5	1994	7	South Yemen	3	1
152	Yemen AR	1948	2	1948	3	Yahaya rebellion	2	1
153	Yemen AR	1962	11	1970	6	Royalists	92	8
154	Yemen PR	1986	1	1986	1	Faction of Socialist Party	1	1
155	Yugoslavia	1991	5	1991	12	Croatia/Krajina	8	1
156	Yugoslavia	1998	3	1999	3	Kosovo	13	1
157	Zimbabwe	1972	12	1979	12	ZANU, ZAPU	85	7
158	Zimbabwe	1983	1	1987	12	Ndebele guerillas	60	5

B.1.2 Comparing Civil War Lists

The criteria upon which the list of civil wars used in this chapter is based (Sambanis 2004), are the most specific, theoretically-based guidance on coding the beginning and end of civil wars among any of the definitions contained in the civil war lists most widely used in the literature in political science and economics. I illustrate the contrast here by reviewing the definitions used by Fearon and Laitin (2003), the Uppsala/PRIO Armed Conflict database, and the Correlates of War project (COW) and indicating the practical consequences of the differences in the definitions used.

Fearon and Laitin (2003), who define civil wars as conflicts which result in a total of at least 1,000 deaths, operationalize continuing war as the maintenance of a yearly average of 100 deaths across the war years. Other than ending through inactivity, civil wars as coded by Fearon and Laitin also terminate due to a victory by either side, or "[a] wholesale demobilization, truce, or peace agreement" with a two-year period of peace afterwards (76). Authors using the Uppsala/PRIO Armed Conflict database have separated civil wars from lower levels of armed conflict by some form of a 1,000-death threshold, either for the conflict as a whole as in Fearon and Laitin (2003) or for its most violent year, with war continuing as long as no two-year period occurs during which there are fewer than 25 deaths in each of the two years (see Cunningham 2006; Buhaug, Gates and Lujala 2009).[2] Finally, the Correlates of War project (COW) has coded conflicts resulting in a total of at least 1,000 battle deaths as civil wars, but without specifying how the start

[2]Buhaug, Gates and Lujala (2009:556): "Consequently, we merged units that have identical ID codes, incompatibility, location, and main actors, and are separated by less than twenty-four months of inactivity (these lapses are treated as part of the ongoing conflict)"; Nunn and Qian (2011) use the 1,000-death threshold for the conflict as a whole in their category of "big war," although they code the incidence of conflict using UCDP, not start and end dates to it.

and end dates of a war are determined (see Sarkees 2000:129).[3]

One need look no further than the country that appears first alphabetically in civil war lists to see the results of these differences in coding rules. Sambanis (2004) lists three wars in Afghanistan, the first from 1978 to 1992, the second from 1992 to 1996, and the third from 1996 to 2001. While listing the same 1978 to 1992 war, Fearon (2004), updating Fearon and Laitin's (2003) war list, codes only one war starting in 1992 and ongoing through 1999. The Uppsala/PRIO data as used by Buhaug, Gates and Lujala (2009) consolidates this further, with a single war coded starting in 1978 and ongoing through 2003. Finally, version 3.0 of the Correlates of War project's intra-state war list (Sarkees 2000) includes a war from 1978 to 1992, but none following it, although the data are coded through 1997.

Table B.2 compares the coverage of the four main datasets of civil wars in greater detail, using the only period common to all four, 1946-1997. For the purpose of comparison, I have gone back to the original datasets and reported the number of wars and countries at war in each dataset for the period, counting coups but excluding anticolonial wars, which are included in Fearon (2004) and Cunningham (2006) but not in Sambanis' (2004) data or in Buhaug, Gates and Lujala (2009).[4] Differences in coding have stark implications not only for the measurement of the length of wars, but also at a more basic level, for which wars are

[3]Sarkees (2000:126) claims that earlier versions of the COW data required annual battle deaths of at least 1,000 to code a civil war and that a new war was to be coded should fighting resume following a cease-fire that lasted at least six months. Both criteria would provide some guidance for determining whether a war continued during a particular period of time. However, the codebooks for the COW data state only that "at least 1,000 battle deaths [must have] resulted during the civil war" and additionally do not mention coding a new war following cease-fires of at least six months (Singer and Small 1994, 1984). Sambanis (2004:16–19) discusses the COW definition of civil war, ambiguities in the project's coding rules, and the analytical problems that the definition and these ambiguities create.

[4]Even when the criteria for inclusion in the count of wars and countries used in Table B.2 match those referred to in other dataset publications, the figures will not match due to the time restriction.

included, and even for which countries ever make it into the analysis. Although Fearon and Laitin, the lenient definition of civil war in Buhaug, Gates and Lujala, Cunningham, and COW ostensibly all use the same definition of civil war as a conflict resulting in a total of 1,000 deaths over its course, the differences in the datasets with respect to how to code continuing conflict result in important discrepancies between the datasets: there are 110 wars in 67 countries in Fearon, 119 wars in 67 countries using the lenient definition in Buhaug, Gates and Lujala, 111 wars in 61 countries in Cunningham, and 104 wars in 55 countries in the COW data. As can be seen in the fourth column of the table, even when the numbers of countries are close, this does not imply that the coverage is the same; the same applies to wars.

The discrepancy between the figures derived from Buhaug, Gates, and Lujala's lenient definition and Cunningham's definition is of particular concern since the two studies use the same initial dataset and apply the same coding rules to it. Note also that although the strict coding definition of civil war used by Buhaug, Gates, and Lujala yields coverage of the same number of countries as the definition of civil war used by Cunningham, the definitions are substantively different and the lists of countries covered do not match (see Table B.2). The numbers of wars and countries reported in Montalvo and Reynal-Querol (2010) provide another example of discrepancies resulting from the use of the UCDP/PRIO dataset. Montalvo and Reynal-Querol report 117 civil wars in 74 countries between 1960 and 1999, using a definition of civil war as a conflict that causes at least 25 deaths per year, with a two-year period of inactivity resulting in coding a new war start (133). This is the same definition employed by Buhaug, Gates and Lujala (2009:556) to include civil wars and lower level armed conflicts. But Buhaug, Gates and Lujala's

data show 212 such conflicts in 91 countries between 1960 and 1999. If coups are excluded, these numbers drop to 179 conflicts in 78 countries. It is only possible to get close to the same numbers reported by Montalvo and Reynal-Querol if not only coups are excluded, but also "parallel conflicts" (this variable is not discussed by Buhaug, Gates and Lujala in their published article but is available in their dataset). But even after "parallel conflicts" are dropped, discrepancies remain; the data then show 118 conflicts in 76 countries.

Dataset	Wars	Countries	Country Coverage vs. Sambanis: Omissions (Additions)
Update of Sambanis (2004)	143	75	
Fearon (2004)	110	67	Congo-Brazzaville, Kenya, Namibia, Egypt, Syria, Israel, Oman, Thailand
UCDP/PRIO†			
BGL (2009), strict††	101	61	Haiti, UK, Croatia, Cyprus, Moldova, Mali, Senegal, Central African Rep., Kenya, Djibouti, Namibia, Egypt, Jordan, Bangladesh, Papua New Guinea (Chile, France)
BGL (2009), lenient	119	67	Haiti, Croatia, Cyprus, Moldova, Mali, Central African Rep., Kenya, Djibouti, Namibia, Jordan, Papua New Guinea (Chile, France, Malaysia)
Cunningham (2006)	111	61	Haiti, Dominican Rep., Croatia, Cyprus, Moldova, Mali, Central African Rep., Kenya, Djibouti, Zimbabwe, Namibia, Egypt, Jordan, Oman, Papua New Guinea (France)
COW, Version 3.0	104	55	Haiti, UK, Croatia, Cyprus, Moldova, Mali, Senegal, Central African Rep., Kenya, Djibouti, S. Africa, Namibia, Morocco, Algeria, Egypt, Syria, Israel, Oman, S. Korea, Bangladesh, Nepal (Chile, Romania)

Note: All anticolonial civil wars contained in Fearon and Cunningham are excluded to make them comparable in scope to the other war lists, as is the Hyderbad conflict in Cunningham's data. Wars starting from 1944 on are included in the figures for Sambanis and COW version 3 and from 1945 on in the figures for Fearon, as these wars/countries are still included from 1946 on in BGL and Cunningham.

† Country totals given here for datasets based on UCDP/PRIO are one higher than would be obtained from tabulating the data to correct for these datasets treating the Yemen Arab Republic and the Republic of Yemen as a single country.

†† Strict: any conflict that caused at least 1,000 deaths in its most deadly year is a civil war. The lenient definition sets the threshold at 1,000 deaths over the entire conflict.

Table B.2: Comparison of Civil War Lists for the Period 1946-1997

B.1.3 Explanatory Variables

Sources

Table B.3, below, lists sources for the explanatory variables used in the statistical analyses of duration and the incidence of quagmire. and their sources.

Finding Missing Values and Other Corrections to Rugged Terrain and Ethnic Polarization Data

For a small number of countries, values of *rugged terrain* and of the group population shares used to construct the *ethnic polarization* index are missing in the original data source. Below, I discuss the additional sources I used to replace missing data with accurate values. In the discussion of the missing group population shares, I also indicate how I corrected the list of groups used in two cases.

Fearon and Laitin (2003) code their elevation difference variable – which I use as the basis for *rugged terrain* – as missing for the Republic of Vietnam, the Yemen Arab Republic, and the People's Democratic Republic of Yemen. I used two sources to find accurate information with which to replace the missing data. For the Republic of Vietnam, I examined topographical maps contained in the *Indochina Atlas* (CIA 1970). For the Yemen Arab Republic and the People's Democratic Republic of Yemen, the best available topographic maps come from a series produced by the Soviet military at a 1:200,000 scale, the *Topograficheskai a karta mira masshtaba* (GUGK 1970). These maps are in Russian only. Although I do not read Russian, with the help of a basic table on the Cyrillic alphabet, I was able to read

[5]Time-invariant coding of NATO member states, excluding states that joined NATO after 1991. Thus covers Belgium, Canada, Denmark, France, Iceland, Italy, Luxembourg, the Netherlands, Norway, Portugal, the United Kingdom, the United States, Greece, Turkey, German Federal Republic (1955-1991) and Germany (1991-2006), and Spain.

Table B.3: Explanatory Variables

Variable	Definition	Source
Area	Area in square kilometers	Banks (2008)
Elevation difference	Natural log of the difference in elevation between the highest and lowest points in the country (100 m)	Fearon and Laitin (2003), with corrections made using GUGK (1970) and missing data filled in using CIA (1970).
Population density	Density of population in thousands of people per square kilometer	Calculated using population data from Maddison (2010) and area data from Banks (2008)
Urbanization	Ratio of urban to rural population	
Ethnic polarization	Ethnic polarization calculated according to the *RQ* index	Calculated based on data in Fearon (2003), with corrections made and missing data filled in using Schrock et al. (1966), Maddison (2010), Corstange (2011), CIA (1985), and Lewis (2009).
GDP per capita	Gross domestic product per capita, in 1990 Geary-Khamis dollars	Maddison (2010)
Military personnel	Number of troops, pre-war	Singer, Bremer and Stuckey (1972)
Military spending	Defense budget, pre-war	Singer, Bremer and Stuckey (1972)
Age of polity	Is the polity 20 years old? Pre-war	Correlates of War Project (2017)
Executive constraints	Constraints on the chief executive, scale of 1 to 7, from low to high, pre war	Polity IV Project (2016)
Prior civil war	Did a civil war had taken place in the country in the 20 years before current war started? Dichotomous	Civil war list
Secessionist war	Does the conflict concern a dispute over the establishment of a separate polity from some portion of the country's territory? Dichotomous	Updated from Sambanis (2004)
Borders	Number of borders shared with other states	Stinnett et al. (2002)
Neighboring civil war	Was there a civil war in a neighboring country in the year before the current war started? Dichotomous	Civil war list
Hydrocarbon exporter	Were at least one-third of the country's exports oil or natural gas? Pre-war, dichotomous	Updated from Sambanis (2004)
Superpower sphere of influence	Does the country border the Soviet Union or any of its successor states; or is it located in Central America	n/a
NATO imports	Natural log of the average pre-war volume of the civil war country's imports from NATO members (in millions of current USD), 5-year lag[5]	Calculated using Barbieri, Keshk and Pollins (2008)
Great power proximity	Minimum distance between country and the five permanent members of the UN Security Council	Weidmann, Kuse and Gleditsch (2010)
Cold War	Was the year of observation between 1946 and 1991?	n/a

place names in the set of maps for each Yemen and calculate values for *elevation difference* accordingly.

Fearon's (2003) data on group shares of countries' populations does not cover all countries needed for the *ethnic polarization* data series. Although many countries which no longer existed at the time the dataset was constructed are covered, it still excludes the Republic of Vietnam, the Yemen Arab Republic, and the People's Democratic Republic of Yemen. In addition, although in most cases of the dissolution of a country the dataset contains values for the original country and its successors, only one set of entries appears for Pakistan. This is problematic for this chapter's analysis, since Pakistan's ethnic composition prior to the war of 1971 in which Bangladesh seceded was different from that following the secession; the source data report only the post-secession group list. The dataset also contains an observation for Papua New Guinea, but no groups are listed and therefore group population share is coded as missing. Finally, based on the research I conducted to find group lists and population shares for the pre-unification Yemens, I concluded that the group list for the unified Republic of Yemen did not adequately reflect ethnic divisions in the country. As a result, I also updated its group list and population shares used to calculate *ethnic polarization*.

To obtain a list of the appropriate groups and group population shares with which to construct the *RQ* index of polarization for the five countries listed above, I turned to a variety of sources.

For the Republic of Vietnam, I used Schrock et al. (1966), an exhaustive report published by the Department of the Army on *Minority Groups in the Republic of Vietnam*.

For the Republic of Yemen, Fearon reported the only group as Arabs (96% of

the population). In fact, an important sectarian division exists in Yemen between Zaydi and Shafi'i Muslims. Since sectarian divisions are used as the basis for group lists in other countries (for example, Iraq and Lebanon), it is consistent to develop the group list for Yemen using this division as a starting point. Indeed, it seems that the sectarian division has typically not been used as a basis for defining groups in Yemen due to the relative absence of conflict between the groups. However, it should be clear that the absence of *mobilization* along a cleavage does not indicate the absence of distinct groups within the population. I used a rough estimate of the Zaydi and Shafi'i population shares for the Republic of Yemen and their geographic location from Corstange (2011) in combination with information on the population shares of Jews and Afro-Arabs in the YAR and on the populations of the YAR and PDRY (CIA 1982; CIA 1985; CIA 1987; Maddison 2010) to calculate the group population shares for the Republic of Yemen and the YAR. Using these, I calculated the *RQ* index of *ethnic polarization* for both countries. Note that based on my revised group list, the *RQ* index for the Republic of Yemen (0.91) is radically different than it would have been had I used the group list from Fearon's data (0.15). I then used Corstange (2011, 2016) to determine that the Arab population of the PDRY was almost exclusively Shafi'i, and calculated the *RQ* index accordingly.

For Pakistan, I used Fearon's group lists and population shares for Bangladesh and Pakistan and population data from Maddison (2010) to construct group shares for a unified Pakistan prior to the secession of Bangladesh. I then calculated the *RQ* index based on these group shares.

Finally, for Papua New Guinea, I calculated the *RQ* index based on a list of groups and group shares of population derived from Lewis (2009).

B.2 Statistical Analyses and Robustness Checks

B.2.1 Duration as Building Block

I estimate a Cox proportional hazards model of civil war duration:

$$\lambda(t; \mathbf{X}_i) = \lambda_0(t) e^{(\mathbf{X}_i' \boldsymbol{\beta})} \tag{B.1}$$

where $\lambda(t)$ represents the hazard function at time t, $\lambda_0(t)$ the baseline hazard function, and $\mathbf{X}_i$ the vector of the explanatory variables for observation i. Here, the hazard is of a civil war ending.

The Cox model, which is semi-parametric, makes no assumptions about the shape of the baseline hazard. Instead, the baseline hazard is treated as a nuisance parameter and drops out of the model, although it can be estimated separately. The advantage of this approach when compared to a fully parametric one is that it focuses the analysis on the relationship between the explanatory variables and the outcome, without making potentially untenable assumptions about the baseline hazard rate. This is particularly important in the analysis of civil wars, as the literature provides contradictory guidance about duration dependency.[6] A parametric model, in contrast, offers potential gains in the efficiency of the estimator, although these tend to be quite small, and the ability to make out of sample predictions. However, in the case of an inappropriate distributional assumption, any gains in efficiency will be lost and, more importantly, parameter estimates can be inconsistent (see Meyer 1990; Narendranathan and Stewart 1993).[7] Since I use the

[6] One segment of the literature highlights war-weariness, indicating that civil wars become more and more likely to end the longer they continue. However, arguments about war economies and conflict traps suggest that the opposite may also be true – the longer civil wars last, the more likely they are to create a situation in which peace becomes difficult to create.

[7] Misspecification of the baseline hazard rate in a parametric model is likely to especially affect

duration model to generate a list of quagmires, and that analysis rests on the relationships between civil war duration and the explanatory variables of interest, the disadvantages of using a parametric model likely outweigh any benefits when compared to the Cox model (see Box-Steffensmeier and Jones 2004:47,85-7).[8]

I address possible unobserved heterogeneity in the data by using a variance-corrected model (Lin and Wei 1989), clustering on the country in which the civil war took place. This takes into account possible correlation among failure times for observations from the same country.[9] An additional correction stems from the Cox model's treatment of time as continuous, with the implication that no two observations in the data fail at the same time. Even if civil wars truly end at different times, the use of any unit of measure of war duration is likely to create ties, particularly when the unit of observation is the war-year, as in the data here. I use Efron's (1977) method for handling ties,[10] and show the robustness of the results to two alternate methods for handling ties in Section B.2.4, one of which allows the possibility of truly tied failure times. A final, primary concern in estimating the model is whether the proportional hazards assumption holds. I test the assumption for the specifications estimated and report the detailed results of these tests in Section B.2.4.

the consistency of the estimated coefficients for time-varying explanatory variables.

[8]Although the models presented in this section maintain the assumption of proportional hazards, the robustness checks in Section B.2.4 indicate non-proportionality in some covariates. I demonstrate that the results are robust to incorporating non-proportionality in the model. However, the inclusion of time-varying covariates indicates against using a parametric model given that the shape of the baseline hazard is not clearly established by theory. Section B.2.4 does consider the appropriateness of using a parametric model. The results obtained by estimating a Weibull model, the distributional assumption of which may be largely consistent with the data, confirm the results of the Cox model presented here.

[9]A different method for addressing unobserved heterogeneity is the estimation of a random effects model. I show in Section B.2.4 that estimation of such a model fails to reject the null hypothesis of no unobserved war-specific heterogeneity.

[10]Efron's method is an approximation of the exact partial likelihood method, which takes into account all possible orderings of tied failures.

Note that the signs of the estimated coefficients here are the opposite from the signs of the estimated effects presented in the main text. This is due simply to a change made so that interpretation of the estimated effects would be straightforward in the main text. The Cox model estimates the effect on the hazard of civil war ending. Estimated coefficients are therefore positive if they increase the probability of war ending, and negative if they decrease that probability. In the main text I reported the sign of the estimated effect on duration; in the main test, positive coefficients indicate a decreased probability of war ending and negative coefficients an increased probability of war ending.

B.2.2 Identifying Quagmire

Figure B.1 shows a histogram of the normal deviate residuals calculated from the model in Section B.2.1 above, with the normal distribution plotted as a curve against it. The normal deviate residuals from the duration model are a reasonable fit for the distribution.

Regarding the selection of a critical value for the quagmire test using the normal deviate residuals, as discussed above, a reasonable starting point would be to use one standard deviation from the mean ($+1$, -1), but in a one-sided test. The critical value at 15 percent, -1.04, that falls slightly below one standard deviation from the mean. This critical value minimizes the total chances of errors in categorizing civil wars as having experienced quagmire. The quagmire test has a 15 percent chance of a type I error and also minimizes the probability of a type II error, at 12 percent. The total chance of error in categorizing a civil war as having experienced quagmire is therefore 27 percent. Stricter critical values corresponding to even 10 percent of the distribution, or 5 percent or less, increase the overall

Table B.4: Analysis, Civil War Duration,1944-2006

	(1)	(2)	(3)	(4)	(5)	(6)
rough terrain	−0.175					−0.128
	(0.212)					(0.223)
area	−0.0533					−0.237*
	(0.0778)					(0.100)
gdp per capita		−1.045				−2.190*
		(0.952)				(0.944)
gdp per capita, squared		0.608				1.039*
		(0.391)				(0.411)
ethnic polarization		−0.830				−0.725
		(0.583)				(0.534)
military personnel			0.283			0.834**
			(0.294)			(0.312)
military expenditures			0.0530			0.0981
			(0.0543)			(0.0897)
age of polity			−0.672**			−0.663**
			(0.223)			(0.218)
executive constraints			−0.197***			−0.220***
			(0.0559)			(0.0544)
prior civil war				−0.409*		−0.590*
				(0.190)		(0.257)
secession				−0.0370		−0.120
				(0.163)		(0.196)
borders					−0.236	−0.0690
					(0.165)	(0.220)
neighboring civil war					−0.0276	−0.162
					(0.179)	(0.181)
superpower sphere of influence					0.631**	0.312
					(0.240)	(0.249)
hydrocarbon exporter					0.414*	0.246
					(0.207)	(0.253)
Cold War					−0.112	−0.302
					(0.210)	(0.285)
Obs.	140	140	140	140	140	140

Cox proportional hazards model, robust standard errors (clustered on country)
* significant at 5%; ** significant at 1%; *** significant at 0.1%

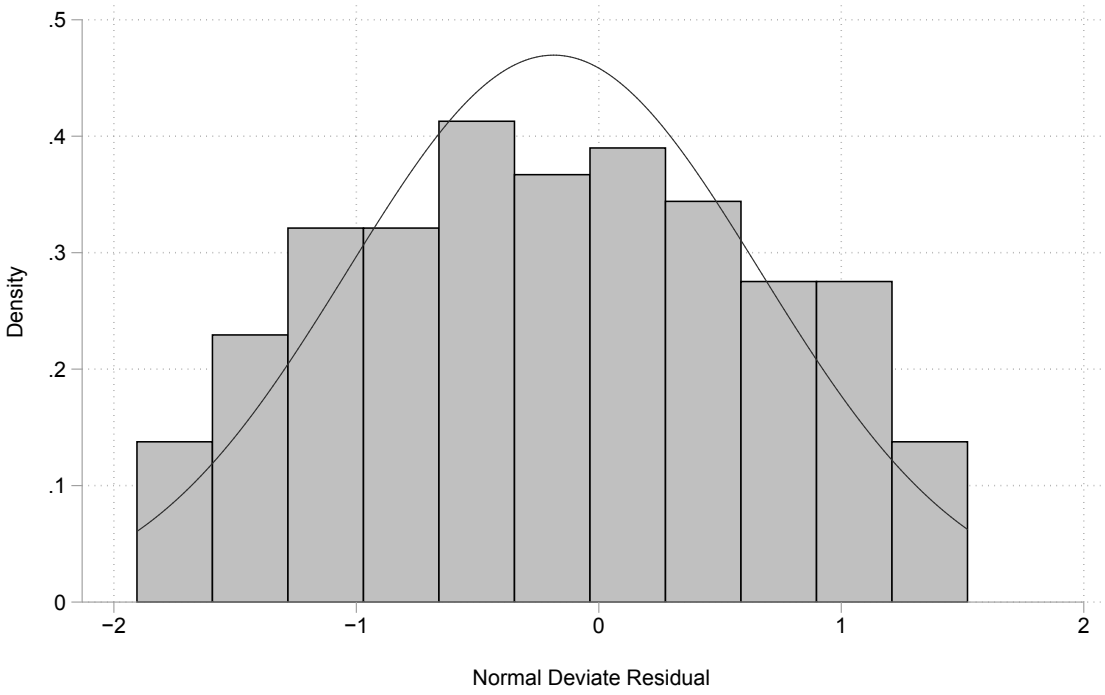

Figure B.1: Histogram, Normal Deviate Residuals

chances of error. For example, selecting a critical value at 5 percent increases the chance of a type II error from 12 to 29 percent; the chance of any error becomes 34 percent.

Table 1.1, Chapter 1, presents highlights the civil wars in which this procedure identifies quagmire as having occurred against the full war list.

B.2.3 The Incidence of Quagmire

I estimate a logit model of the incidence of quagmire in civil war, with robust standard errors clustered on country. All models include each variable independently in addition to the interaction terms.

Table B.5: Analysis, Incidence of Quagmire in Civil War, 1944-2006

		(1)	(2)	(3)
NATO imports				
	× population density	1.213		5.895
		(1.849)		(4.075)
	× urbanization	0.619		3.836+
		(0.796)		(2.065)
	× ethnic polarization	0.698		1.422
		(0.477)		(1.006)
	× rugged terrain	−0.036		0.370
		(0.124)		(0.271)
	× polity age		−0.168∗∗	−0.286+
			(0.054)	(0.155)
	× executive constraints		−0.075	−0.305+
			(0.059)	(0.162)
	× previous war		0.149	−0.187
			(0.204)	(0.324)
	× secession		−0.073	−0.261
			(0.159)	(0.245)
great power proximity				
	× population density	−4.510		1.768
		(21.970)		(32.664)
	× urbanization	−1.986		−31.926
		(17.202)		(24.579)
	× ethnic polarization	−13.509		−9.090
		(10.691)		(14.502)
	× rugged terrain	−8.131∗		−9.267+
		(3.759)		(5.188)
	× polity age		0.907	1.295
			(1.353)	(2.125)
	× executive constraints		1.943	2.064
			(1.430)	(1.920)
	× previous war		1.533	4.805
			(3.541)	(5.616)
	× secession		8.012	6.651
			(5.000)	(5.421)
hydrocarbon exporter				
	× population density	−14.657		30.509
		(11.258)		(37.217)
	× urbanization	3.923		13.191
		(3.895)		(10.431)
	× ethnic polarization	−6.411∗		1.737
		(3.034)		(8.662)
	× rugged terrain	1.356		7.203+
		(1.463)		(3.723)
	× polity age		0.051	−0.888
			(0.464)	(0.870)
	× executive constraints		1.334∗∗	3.781∗∗
			(0.416)	(1.323)
	× previous war		−2.525+	−8.229
			(1.480)	(5.317)
	× secession		4.349∗∗	7.022∗
			(1.513)	(2.773)
Constant		−31.037∗	0.231	−20.583
		(13.474)	(4.289)	(17.332)
Obs.		140	140	140
Pseudo R-squared		0.1133	0.1811	0.3366

B.2.4 Robustness – Duration

The principal concerns about the validity of the results obtained in Section B.2.1 relate to the possibility of unobserved heterogeneity in the data, violations of the Cox model's proportional hazards assumption, sensitivity of the results to the treatment of tied failure times, sensitivity of the results to the treatment of missing data, and sensitivity to changes in the list of wars included in the analysis. I show for each that the results supporting the theory of external assistance as a subsidy are quite robust. Before addressing the concerns, in order, I first estimate a plausibly valid parametric model and demonstrate that any efficiency gains obtained by doing so do not alter the results presented above.

Parametric Estimation

If there is a theoretical basis for assuming a specific functional form for the baseline hazard rate for the data, a parametric estimator would be more efficient than the estimator used in any semi-parametric or non-parametric model (Collett 2003:151; Box-Steffensmeier and Jones 2004:87). As an additional check on the robustness of the results, I therefore estimate a parametric model of civil war duration. To identify a valid parametric distribution, before proceeding with the estimation, I plot the Nelson-Aalen, non-parametric estimate of the cumulative hazard function in Figure B.2.

The cumulative hazard function appears to be monotonically increasing, which would justify the selection of a Weibull model. In addition, the Weibull model's survival function $\left(S(t) = e^{-(\lambda t)^p}\right)$ exhibits a linear relationship with time: $\ln\left(-\ln S(t)\right) = p(\ln\lambda + \ln t)$. This linearity provides another means of confirming the appropriateness of specifying a Weibull distribution for the data (Kalbfleisch

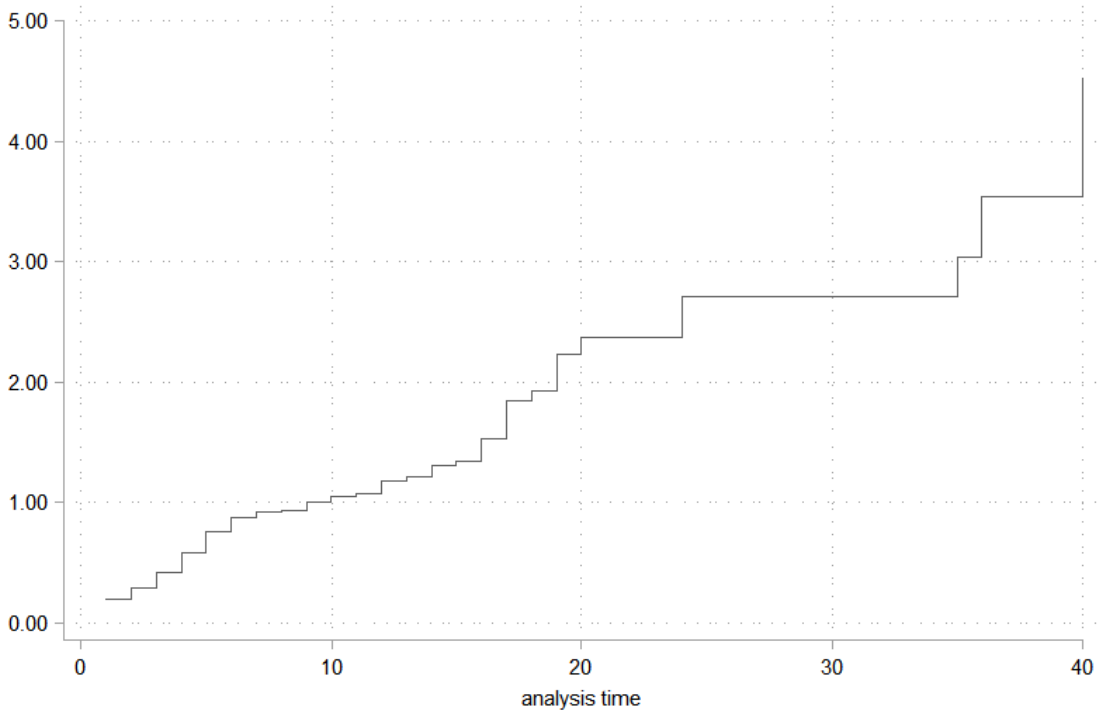

Figure B.2: $\widehat{H}(t)$(Nelson-Aalen)

and Prentice 2002:33). Figure B.3 plots $\ln\left(-\ln(\widehat{S}(t)\right)$ against $\ln t$, where $\widehat{S}(t)$ is the Kaplan-Meier estimate of the survival function.

Based on Figures B.2 and B.3, it therefore seems reasonable to use the Weibull specification. I estimate a Weibull model of civil war duration as:

$$\lambda(t;\mathbf{X}_i) = \lambda p(\lambda t)^{p-1} \tag{B.2}$$

where $\lambda(t)$ represents the hazard function at time t, $\lambda_i = e^{-\mathbf{X}_i'\beta}$, $\mathbf{X}_i$ is the vector containing a constant and the explanatory variables for observation i, and $p > 0$ is the shape parameter. In addition to requiring a monotonic hazard, the Weibull model also assumes proportional hazards. I tested for non-proportionality in the Weibull models reported below following Box-Steffensmeier and Zorn (2001) and

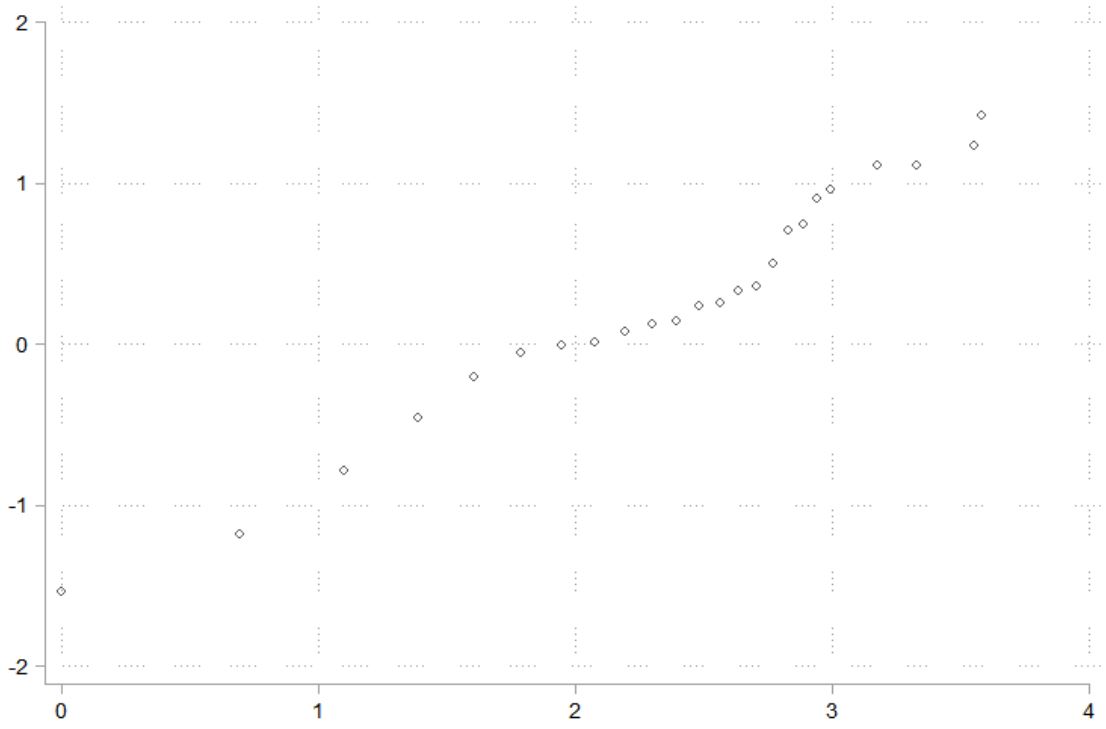

Figure B.3: $\ln\left(-\ln(\widehat{S}(t)\right))$ vs. $\ln(t)$

Carroll (2003). I then re-estimated non-proportional versions, the results of which are congruent with the corresponding non-proportional Cox models reported below. Since the proportional hazards assumption will be tested in detail[11] in the context of the Cox model, below, I do not discuss it further here. Table B.6 reports the results of replicating the Cox proportional hazards estimation of Table B.4, Model 6 using a Weibull model. I include the original Cox model results next to these for ease of comparison.

Unobserved Heterogeneity

The duration analysis (Section B.2.1) addresses the possibility of unobserved heterogeneity in the data by using a variance-corrected model. A second approach commonly used in the literature is to control for unobserved heterogeneity by using a shared frailty model. This modifies the survival function so that it is conditional on a group- or individual-specific effect. In the context of the Cox model, we can then estimate the hazard function as:

$$\lambda(t; \mathbf{X}_{ij}, \omega_i) = \omega_i \lambda_0(t) e^{(\mathbf{X}'_{ij}\boldsymbol{\beta})} \tag{B.3}$$

where as before $\lambda(t)$ represents the hazard function at time t and $\lambda_0(t)$ the baseline hazard function, but now $\mathbf{X}_{ij}$ denotes the vector of the explanatory variables for the jth member of group i, and ω_i is a random effect, or frailty, for group i (Klein 1992:797).[12] As part of the model, we can then estimate θ, the variance of ω_i. If

[11] As Box-Steffensmeier, Reiter and Zorn (2003) note, methods for testing the proportionality assumption are most thoroughly developed for the Cox model. It is therefore preferable to examine the robustness of the results with respect to non-proportionality in the context of the Cox model rather than the Weibull model.

[12] Here, we assume that $\omega_i \sim \Gamma(1, \theta)$ (see Klein 1992). This assumption on ω_i is a standard one (see Greene 2003:797–8). See also the discussion in Van den Berg (2001:3418–22).

Table B.6: Semi-Parametric and Parametric Duration Models Compared

	Model 6 (Cox)	Model 6 (Weibull)
rough terrain	−0.128	−0.101
	(0.223)	(0.243)
area	−0.237∗	−0.222∗
	(0.100)	(0.106)
gdp per capita	−2.190∗	−2.280∗
	(0.944)	(0.975)
gdp per capita, squared	1.039∗	1.123∗∗
	(0.411)	(0.435)
ethnic polarization	−0.725	−0.988
	(0.534)	(0.511)
military personnel	0.834∗∗	0.889∗∗
	(0.312)	(0.313)
military expenditure	0.098	0.071
	(0.090)	(0.090)
age of polity	−0.663∗∗	−0.662∗∗
	(0.218)	(0.238)
executive constraints	−0.220∗∗∗	−0.203∗∗∗
	(0.054)	(0.053)
prior civil war	−0.590∗	−0.464
	(0.257)	(0.250)
secession	−0.120	−0.134
	(0.196)	(0.202)
borders	−0.069	−0.129
	(0.220)	(0.237)
neighboring civil war	−0.162	−0.192
	(0.181)	(0.190)
superpower sphere of influence	0.312	0.311
	(0.249)	(0.255)
hydrocarbon exporter	0.246	0.315
	(0.253)	(0.253)
Cold War	−0.302	−0.414
	(0.285)	(0.290)
Constant		1.479
		(1.062)
ln(p)		0.234∗∗∗
		(0.071)
Obs.	140	140

* significant at 5%; ** significant at 1%; *** significant at 0.1%

$\theta = 0$, group members are independent, and $\omega_i = 1$, indicating no unobserved heterogeneity, and reducing the hazard function to its original form (Equation B.1, above).

A principal concern with the original model estimated without frailty is the possibility of unobserved heterogeneity due to factors specific to each war in the dataset. To test this, I estimate a Cox proportional hazards model with shared frailty of the specification reported in Table B.4, Model 6, with ω_i as a random effect for civil war i. The results indicate that unobserved, war-level heterogeneity is unlikely. Specifically, the likelihood ratio test for the presence of heterogeneity is statistically insignificant; we fail to reject the null hypothesis that $\theta = 0$.[13]

Non-Proportional Hazards

If there are non-proportional hazards for some of the explanatory variables used in the Cox model, this can result in biased and inefficient estimates not only of the coefficients for those variables exhibiting non-proportionality, but for all parameters in the specification (Box-Steffensmeier and Zorn 2001). I perform global and covariate-specific tests for non-proportionality in the main specification estimated in Section B.2.1 (Model 6) using scaled Schoenfeld residuals.[14] Estimation of the non-proportional re-specifications of these models, discussed in detail below, demonstrates the robustness of the results set out in Section B.2.1. In some models, properly accounting for non-proportionality provides even stronger evidence in support of the theory.

I reproduce the full test statistics for Model 6 in Table B.7.[15] In the tests, a

[13] $\overline{\chi}^2(1) = 0.00001$, p-value $= 0.499$.

[14] See Keele (2010) for a thorough discussion of proper implementation of the non-proportional hazards test, given the possibility of false positives due to model misspecification.

[15] Although the global test for Model 6 fails to reject the null hypotheses of proportional hazards,

statistically significant correlation between the scaled residuals and a function of time indicates non-proportionality. Briefly, the global test for the model fails to reject the null hypothesis of proportional hazards.[16] However, as Box-Steffensmeier, Reiter and Zorn (2003:45) note in their applied study of non-proportional hazards, "there are no clear guidelines about the use of global versus local tests....Accordingly, results suggesting variable-specific non-proportionality should not be ignored in the face of a residual indicating global proportionality." I therefore identify suspected non-proportional variables using the covariate-specific test statistics, adopting a 15% cut-off level for significance. I then re-estimate Model 6, introducing interaction terms between each suspect variable and time (see Collett 2003:146-8).[17] Table B.8 reports the results of the non-proportional version of the model.

The results of the non-proportional hazards estimation support the robustness of the original results reported in Section B.2.1, above. All variables that were statistically significant in the proportional hazards model and for which no interaction with a function of time was required in the non-proportional hazards model continue to be so at the same level of statistical significance. Of the four variables interacted with a function of time, only *area* drops below statistical significance at the 5 percent level. *Cold War* is not statistically significant in the original proportional hazards model and continues not to achieve statistical significance in the non-proportional hazards model. *Prior civil war* is significant only in its interaction with a function of time. *Ethnic polarization*, which was not statistically

the covariate-specific test statistics point to four variables as likely violating the non-proportional hazards assumption: *area* (at the 10% level), *ethnic polarization* (at the 10% level), *previous war* (at the 1% level), and *Cold War* (at the 5% level).

[16]The global test statistics for Model 6 is $\chi^2(16) = 14.41$ (p-value $= 0.5683$).

[17]It is also possible to account for possible non-linear effects of the interaction by using a function of time in the interaction, for example the log or square root. Collett (2003) recommends using the log if the units of time take on large values. This is not a concern here, since time is measured in years.

Table B.7: Test of Proportional Hazards Assumption

Variable	ρ	χ^2	$\Pr > \chi^2$
rough terrain	−0.025	0.12	0.731
area	−0.120	2.94	0.086
gdp per capita	−0.096	1.47	0.226
gdp per capita, squared	0.079	1.03	0.311
ethnic polarization	0.126	3.46	0.063
military personnel	0.066	0.76	0.382
military expenditures	0.015	0.04	0.833
age of polity	0.063	0.70	0.402
executive constraints	−0.092	1.41	0.235
prior civil war	−0.186	9.15	0.003
secession	0.090	1.22	0.269
borders	0.034	0.18	0.675
neighboring civil war	−0.016	0.04	0.841
superpower sphere of influence	0.051	0.36	0.546
hydrocarbon exporter	0.050	0.44	0.508
Cold War	−0.143	5.76	0.016
Global test		14.41	0.568

Test uses scaled Schoenfeld residuals from estimation of Model 6, Table B.4
χ^2 statistics have one d.f. except for the global test (16 d.f.)

Table B.8: Proportional and Non-Proportional Hazards Models of Duration Compared

	Model 6 (Original)	Model 6 (NPH)
rough terrain	−0.128	−0.218
	(0.223)	(0.248)
area	−0.237∗	−0.224
	(0.100)	(0.128)
gdp per capita	−2.190∗	−2.356∗
	(0.944)	(0.948)
gdp per capita, squared	1.039∗	1.114∗∗
	(0.411)	(0.420)
ethnic polarization	−0.725	−1.613∗
	(0.534)	(0.705)
military personnel	0.834∗∗	0.868∗∗
	(0.312)	(0.335)
military expenditures	0.098	0.083
	(0.090)	(0.092)
age of polity	−0.663∗∗	−0.625∗∗
	(0.218)	(0.210)
executive constraints	−0.220∗∗∗	−0.209∗∗∗
	(0.054)	(0.054)
prior civil war	−0.590∗	0.158
	(0.257)	(0.361)
secession	−0.120	−0.004
	(0.196)	(0.197)
borders	−0.069	−0.086
	(0.220)	(0.205)
neighboring civil war	−0.162	−0.201
	(0.181)	(0.181)
superpower sphere of influence	0.312	0.330
	(0.249)	(0.255)
hydrocarbon exporter	0.246	0.319
	(0.253)	(0.287)
Cold War	−0.302	0.387
	(0.285)	(0.398)
TVCs		
area		−0.012
		(0.012)
ethnic polarization		0.132∗
		(0.067)
prior civil war		−0.115∗∗
		(0.043)
Cold War		−0.160
		(0.086)
Obs.	140	140

* significant at 5%; ** significant at 1%; *** significant at 0.1%

significant in the proportional hazards model is now statistically significant at the 5 percent level and has the same sign, but its effect appears to decrease in size over time.

Ties

The estimation procedure used in Section B.2.1 treated tied failure times in the data as the result of imprecise measurement and used Efron's (1977) approximation of the exact partial likelihood. If two or more civil wars in the dataset lasted the same number of years, the assumption was that in fact there was an order to the true length of time each lasted, but that this was observed imperfectly, leading to the tie. Two alternatives to this approach may be recommended based on the literature. The first addresses the potential for the approximation of the exact partial likelihood to bias estimated coefficients towards zero if there are many ties in the data. As there are a relatively large number of ties in the civil war duration data, particularly given the use here of a yearly metric, it is important to verify that the results are not sensitive to changes in the method of calculating the partial likelihood. The second addresses the sensitivity of the results to the conceptual frame applied to the data. If it is truly the case that the duration of civil wars is continuous, the methods already considered likely deal appropriately with tied failure times. If, however, it is possible for the length of one civil war to be truly identical to the length of another, the exact discrete partial likelihood method of handling ties should be used. Civil wars are a social phenomenon, making the perceptions of the actors involved key to understanding them. The length of a war may be just as likely to be experienced by the warring parties as discrete rather than as continuous (i.e., the war could be said to have been ongoing in one year, but not the

next, without the ability to identify a specific end date). I therefore also employ the exact discrete partial likelihood method to verify that the results presented earlier are not sensitive to the conceptualization of time in the data.

The results of re-estimating Model 6 using both the exact marginal likelihood method and the exact partial (discrete) likelihood method for handling ties, not reported here, are equivalent to those presented for Model 6 in Section B.2.1, Table B.4 above. Compared to the models using the exact methods, the original model does show a downward bias towards zero for some coefficients. The estimated coefficients from the discrete version are larger than those using the exact method based on continuous time.