INSURGENCY OUTCOME AND DURATION

INFORMING A GRAND STRATEGY FOR COIN WARFARE

A Thesis
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Master of Arts
in Security Studies

By

Harsh R. Pandya, B.S.
Table of Contents

I. Introduction ..................................................................................................................... 1

II. Research Design.............................................................................................................4
   Case Selection ................................................................................................................. 5
   Selection Bias .............................................................................................................. 6
   Dependent Variables ....................................................................................................... 7
   Defining Duration and Termination ............................................................................ 9
   Explanatory Variables ................................................................................................... 10
   Control Variables .......................................................................................................... 19

III. Empirical Analysis...................................................................................................... 20
   Part One: Outcome ........................................................................................................ 20
   Part Two: Duration ........................................................................................................ 23

IV. Additional Robustness Tests....................................................................................... 28

V. Conclusion ...................................................................................................................3 0
   Policy Implications........................................................................................................ 30
   Future Research............................................................................................................ 33

Appendix A: Codebook .................................................................................................... 37
   Dependent Variables ..................................................................................................... 37
   Explanatory Variables ................................................................................................... 37
   Control Variables .......................................................................................................... 39

Appendix B: Summary Statistics ...................................................................................... 41

References......................................................................................................................... 42
I. Introduction

A decade ago, only a handful of researchers had examined the duration of civil war. During the Cold War, the realist outlook prevailed, limiting research by its assumptions of the state as a unitary actor. The big questions were about collective security, democratic peace, and deterrence; it was as if every conflict could be mapped out by the game theory approach. But by the end of the Cold War, a worrisome trend emerged. Despite the end of bipolar hegemony and proxy wars, intrastate conflicts began to emerge (or more appropriately, began to gain publicity). From Somalia to Kosovo, state failure and rebel insurgency were rampant. Some scholars tried to adjust the precepts of systemic theories to internal conflict, but with confusing results. It was now less crucial to understand the causes of interstate conflict than it was to understand those of domestic conflict, particularly insurgencies and civil wars.

Figure 1: Percent of COIN wars won by incumbents by half-decade.
Outside academia, the problem was very real to policymakers as well. As shown in Figure 1, over time incumbents have been less successful at fighting insurgencies. Whether this is because of global democratization, improvements in technology, or some other not readily observable effect is yet to be seen, and is the focus of this analysis. Given the proportion of conflicts in which the state authority loses outright or makes significant concessions to the insurgent group, there is a particular urgency today for a more complete understanding of insurgencies—not only their causes (though onset analysis may prove useful in understanding the more critical need), but also more importantly their longevity and end. Under what conditions do insurgencies end? What factors explain the duration of insurgent conflict? More specifically, what is the role of any one factor in the duration and termination of insurgencies? The answer to the latter question gives policymakers an opportunity to cut the problem off at its root and come to a peaceful solution. In order to answer that question, I conduct a broad quantitative analysis that draws from Lyall and Wilson’s (2009) insurgency dataset and research design.

Almost all of the early analyses relating to intrastate conflict had been devoted to modeling the onset of civil war.¹ These studies broadly looked at two hypotheses: that civil war was a product of political grievance, or that civil war was a product of

¹ At this time, no distinction was made amongst insurgency (non-state organizations employing guerilla tactics), civil war (two competing centers of power with credible claims to sovereignty), and ethnic violence. This was probably due to the unavailability of such datasets.
opportunity.\(^2\) Now, as political scientists have parted with the wisdom of the Cold War, more and more scholars have applied new models to the study of the timing and duration of events such as civil war. Much of the footwork here was done by Fearon and Laitin (2003). They determined that from a policy perspective, ethnic differences, ethnic minority status, and group grievances are difficult – if not impossible – to eliminate if they are the primary causal factors of civil war. Attributing these factors to conflict as reasons for their onset does little for analysts in the sense that there is no remedy short of genocide and ethnic cleansing (and some leaders have, unfortunately, chosen that route). Moreover, more civil conflicts are characterized by insurgency than they are by ethnic or nationalist conflict, further putting into question the utility of differentiating any particular conflict as based on group differences or grievances (Fearon and Laitin 2003, 79).\(^3\)

There are some shortcomings to the early civil conflict analysis. For one thing, they pool insurgencies fought with guerilla tactics along with coup attempts and unconventional wars.\(^4\) This is an important criticism, because if cases are not uniform, it becomes much more difficult to deduce a uniform causal process. All internal wars are not alike. The reason Fearon and Laitin (2003) has been selected as an anchor text, however, is that it sheds light on the type of methodology that seems most effective to

\(^2\) The first study to clearly posit these two competing hypotheses was Collier and Hoeffler (2002).
\(^3\) In Fearon and Laitin (2003, 79), insurgency is defined as “a technology of military conflict characterized by small, lightly armed bands practicing guerilla warfare from rural base areas.”
assess the correlates of insurgency to determine causal variables. Moreover, authors who speak directly to insurgency termination refer to this study as well.

The remainder of this paper will be organized as follows. First, I will set up the research design for this analysis by providing a justification for the selection of cases and individual variables, with reference to the literature on civil conflict, outcome, and duration. The empirical analysis is divided into two parts – an outcome analysis and a survival analysis for 135 counterinsurgency wars since the end of World War II. I then conduct several robustness tests to bolster the analysis. In the last section, I offer concrete policy implications for the results and determine several profitable avenues for future research.

II. Research Design

I adopt a two-stage research design, followed by a series of robustness tests to enhance the accuracy of my analysis. After determining proxy measures for a variety of explanatory variables, I run a regression analysis against the two dependent variables – outcome and duration – to check for statistically significant variables. Outcome captures whether the incumbent won the counterinsurgency war, and duration measures the length of conflict. Jones and Libicki (2008) do a similar analysis with terrorist groups built on a point system – i.e. groups that did not last long received few points, and groups that survived for decades or entered the political process received more points (Jones and Libicki 2008, 197-8).
Case Selection

For this project, I used the Lyall and Wilson (2009) dataset, which covers 286 insurgent conflicts beginning in the 19th century. Lyall and Wilson (2009) draw cases from four datasets: the Correlates of War (COW) Intrastate War Dataset (version 3.0); the Fearon and Laitin civil war dataset (2003); the Uppsala Conflict Data Program (UCDP) Uppsala Armed Conflict Dataset (version 3); and the Political Instability Task Force’s (PITF) Internal Wars Dataset. As these are some of the most commonly cited and rigorously maintained conflict datasets, I feel that Lyall and Wilson (2009) is a reliable dataset on which to model my analysis. Then, the experiment can be replicated using any one of these datasets, granted that the definition of insurgencies stays consistent.

Another key advantage of the Lyall dataset is the binary variable “Outcome” that has been assigned to every insurgency on the list. It opens up the possibility to not only determining what factors bring insurgencies to an end, but also what factors promote a positive outcome. Admittedly, this is dicey territory, because many of the insurgent conflicts on the list can be seen as sparked by legitimate separatist claims, so a “positive” outcome is not very telling in all cases. However, for the purposes of the analysis, there is certainly value in judging on the whole whether a state security apparatus is effective in dealing with an insurgency under particular conditions.

They chose two rules for identifying cases:

5 All of these datasets are publicly available.
• (1) A minimum threshold of 1000 battle deaths, with at least 100 casualties suffered on each side.

• (2) The non-state actor must have adopted a guerrilla warfare strategy.

Guerrilla warfare is defined as a strategy of armed resistance that:

• (1) uses small, mobile groups to inflict punishment through hit-and-run strikes while avoiding direct battle when possible and

• (2) seeks to win the allegiance of at least some portion of the noncombatant population.

Thus, insurgencies as defined here are not necessarily synonymous with civil wars since these can be fought conventionally between opposing armies, with guerrilla tactics, or through non-violence.

135 counterinsurgency war cases are adapted from Lyall and Wilson (2009). These wars are all from the post-1945 era, for which the data collection is much more uniform, accurate, and comprehensive. For instance, language data only exists for this period (adapted from Fearon and Laitin 2003). For future analysis, this is also the period for which data collection for other causal factors, such as income inequality or police per capita, would be easiest to collect.

Selection Bias
I go into this analysis with a healthy skepticism in the quality of the data points that I have adapted. I did not inspect each individual case to ensure the accuracy of the coding, but I did randomly investigate several cases to gauge data quality. It is necessary to spend some time discussing the conflicts to which this analysis cannot apply. Conflicts with less than 1,000 battle deaths cannot be accounted for in this study, though they may have been persistent nuisances in many countries. More importantly, those small-scale conflicts may be able to enhance our knowledge of COIN operations, as many nascent conflicts may have been suppressed before reaching 1,000 battle-related deaths.

In order to minimize selection bias otherwise, I account for ongoing counterinsurgency wars (i.e. wars that have not terminated yet) with nine right-censored observations in the data set. Notable examples include the Mexican drug war, the US war in Afghanistan, the insurgency in Iraq, and the Baluchi insurgency in Pakistan. The number of ongoing conflicts is subjective, of course, since often casualty numbers cannot be immediately determined nor easily agreed upon. Moreover, it is difficult to determine whether civil conflict is dominated by guerilla warfare when the conflict has not ended yet and the dynamics of it could change unpredictably over time. That said, the right-censored observations that are included in the data set all meet the casualty threshold and the guerilla warfare strategy requirement.

Dependent Variables

Analyzing the conditions under which civil conflicts endure and terminate is useful, but perhaps more salient is the question of which side the outcome favors. Lyall
and Wilson (2009) address this question in detail, with particular attention to insurgencies. Termination of insurgencies can take place in one of two ways – the incumbent government can win, or the insurgent organization can extract concessions or even defeat the incumbent politically. Lyall and Wilson (2009) distinguish between these two possible outcomes of insurgency and investigate the factors that tip the balance in either direction. The outcome can be a win, a draw (which counts as a win for the non-state insurgency organization because they have forced concessions from the government), or loss, each with regard to the incumbent government (Lyall and Wilson 2009, 71). The study notably does not rely on military indicators such as recruitment or loss-exchange ratios.

Following that design, war outcome is coded from the incumbent’s perspective here. A win occurs when the insurgency is militarily defeated and its organization destroyed or the war ends without any political concessions granted to insurgent forces. A loss occurs when the incumbent unilaterally concedes to all, or nearly all, insurgent demands, including the granting of independence or the deposition of current leaders. A draw occurs when an incumbent is forced to concede to some, but not all, insurgent demands, and neither side obtains its maximal aims. From Lyall and Wilson (2009).

---

6 An insurgency is defined as “a protracted violent struggle by non-state actors to obtain their political objectives – often independence, greater autonomy, or subversion of existing authorities – against the current political authority (the incumbent)” (Lyall and Wilson 2009).
7 For instance, if an insurgent group demobilizes with the condition that it can be recognized as a political party, it is determined a draw, but the government is forced into a concession that previously did not exist.
**War duration** measures the length of the conflict in months (logged). Data are taken from the UCDP-PRIO Armed Conflict Dataset (Version 4). War start dates are coded from the first large-scale insurgent attack, open declaration of hostilities, or the date of occupation. War termination dates are identified by several means, including the last day of war-related fatalities and the date of a peace settlement.

**Defining Duration and Termination**

An interjection must be made here about modeling civil conflict duration and termination. Analysis of the duration of intrastate conflict has become more sophisticated over the last decade. Many of the scholars cited here have adapted to use proxies more effectively to estimate abstract information. But there are some more fundamental problems that relate to measurement error in most of these studies; the data that has been used to estimate the duration of wars is especially problematic. Three problems stand out: 1.) the casualty threshold used to define a civil conflict; 2.) accounting for repeated events in the same country and contemporaneous events occurring in a single country; and 3.) the precision of the measurement of time. These issues have not been addressed by most scholars to satisfaction. Without a new, resource-intensive data collection effort to correct these selection issues, it will be difficult to resolve the measurement errors. This analysis therefore cannot estimate duration as would be ideal, but the author is very much aware of these concerns.

---

8 See Gates and Strand (2004).
Explanatory Variables

There is a host of explanatory (independent) variables I intend to look at for this study. First, I test the thesis that democracies are particularly vulnerable to defeat in counterinsurgency wars. Scholars have not come to a consensus on this point. Democracies may be especially constrained in COIN wars by their commitment to international norms and laws governing the use of military force. Unlike autocratic regimes, democratic leaders are often thought to be risk-averse in terms of the reputations costs associated with violating these norms, particularly widely adopted treaties like the Geneva Convention. Indeed, the prospect of suffering audience costs at home and abroad has emerged as a reality in the post-1945 era of global media that transmits both information on compliance with treaty obligations and on insurgent attacks aimed at disillusioning their democratic audience. “Beware the scripted enemy,” Kilcullen (2006) writes, “who plays to a global audience and seeks to defeat you in the court of global public opinion.” We should therefore expect varying outcomes across regime types, especially as the ranks of democracies swell over time and as global media brings the reality of war home to democratic publics. The democratic systems place significant constraints on state behavior and ability to conduct effective COIN operations (Kydd and Walter 2006, 61).

Others have argued that democracy is associated with less discrimination and
discriminatory

9 Downes (2006, 161–70) argues that democracies will violate these laws if they are desperate and believe that these actions will end the war swiftly.
10 Also see Pape (2005: 44).
repression along cultural or other lines, since it endows citizens with a political power—the vote—that they do not enjoy in dictatorships. Even more directly, how well states observe civil rights—such as freedom of association, expression, and due process—should be associated with lower grievances (Jones and Libicki 2008, 17).

Some scholars have dismissed the effect of regime type on state capacity altogether. Engelhardt (1992), for example, compared ten countries—five of which were democracies—in 25 post-1945 cases and concluded that, despite authoritarian states’ greater willingness to sustain casualties, victory was “elusive” for all regime types. More recently, Abrahms (2007) contrasted 20 democracies with 20 authoritarian states, all of which had been targeted by terrorists, and employed simple bivariate analysis to conclude that democracies, by virtue of their greater restraint and low tolerance for civilian casualties, made better counterterrorists. By contrast, Maria Stephan and Erica Chenoweth (2008) studied 267 cases of opposition to state authorities and found that as regimes become more democratic, they also become more likely to offer concessions when confronted with either violent or non-violent campaigns.

Thus our first explanatory variable is regime, and it lends itself to two hypotheses.

**H1:** Political democracy and the presence of civil liberties will be associated with shorter insurgency duration.

**H2:** Insurgencies against democratic regimes are more likely to terminate in favor of the insurgents.
A regime’s sensitivity to international pressure arising from normative commitments may also hinder an incumbent by removing more effective COIN strategies from the realm of the politically possible. Since it is difficult to measure normative pressure directly, we can assume that the greater the state’s exposure to trade flows, the more vulnerable it will be to reputational and economic costs for breaching existing norms. Hence I include the variable *trade* in this study. The variable is coded by Lyall and Wilson (2009) and draws trade data from the International Monetary Fund Direction of Trade data set, while GDP is taken from the Banks Cross-National Time Series Dataset for the 1946–2005 era. Measures are taken in the year prior to the war.

**H3:** The number of military personnel in the incumbent’s army will be positively correlated with incumbent victory, especially in lengthy conflicts.

Next, we look at state capacity. A simple measure is the number of personnel in a state’s military. Per capita police would be a more ideal measure for state capacity, but it is not exactly an easy variable to find or approximate. *Personnel* is used as a best guess variable to reflect a state authority’s presence, though significant differences between presence in rural areas and urban population centers may exist. It is also foreseeable that the larger the number of personnel in a state’s army, the longer it could sustain a counterinsurgency war.

**H4:** The number of military personnel in the incumbent’s army will be positively correlated with incumbent victory, especially in lengthy conflicts.
Lyall and Wilson (2009) found that the trend of incumbent losses could be ascribed to the force-structures of modern (post-World War I) militaries. That is, the advent of *mechanization* had deleterious consequences on modern militaries’ capacity to engage in counterinsurgency operations (Lyall and Wilson 2009, 72). Prior to mechanization, soldiers were forced to interact with local populations; armies consisted of infantry organized around the principle of “foraging,” where the bulk of military supplies were obtained by coercing civilians. This dynamic of extensive interaction with locals generated high volumes of information. Mechanized armies that are isolated from society, therefore, suffer from “information starvation,” which inhibits their ability to recruit reliable collaborators among local populations (Lyall and Wilson 2009, 74-5).

**H5: The level of mechanization of a state’s army negatively affects a state’s capacity to fight a COIN war.**

After Fearon and Laitin (2003), some scholars saw the need to understand the duration of civil conflict because they acknowledged that there was more variation in war duration than in the rate of war onset. This implied that means to alter the chances for war termination were more readily available than measures for conflict prevention. Instead of investigating which factors are systematically associated with onset of civil conflict, they sought to understand why civil wars endure and how they end. Some of the recurring variables that are tested in these texts are discussed below.

Here, there is a dearth of analysis. Most scholars seem to be focused on conditions that are determined *a priori*, and do not include counterinsurgency policy change in their
analyses. Therefore, for instance, instead of factors like policing, they look at the
efficiency of the bureaucracy. Oddly enough, counterinsurgency measures do not have to
be military-based to be effective. DeRouen and Sobek (2004, 311) show that an efficient
bureaucracy increases the chance of reaching negotiated outcomes (although it does not
reduce the duration of civil wars).\textsuperscript{11}

The great waves of state creation after the post-World War II imperial collapses
may have populated the dataset with weak states that are more likely to experience
insurgent challenges precisely because they are unable to defeat them. Fearon and Laitin
(2003) have noted that states with low gross domestic product (GDP) per capita incomes
record the highest levels of civil war onset (Fearon and Laitin 2003, 80). Another variable
to look at is how insurgencies are financed. Often, insurgents are able to maintain a
guerilla war through contraband finance. Collier and Hoeffler (2002) find that a high
(around 33 percent) share of GDP derived from primary commodity exports is correlated
with civil war onset (Collier and Hoeffler 2002, 574).

Collier and Hoeffler’s (2002) work on why civil war starts sparked a vivacious
debate on whether rebel groups most commonly are motivated by ‘grievance’ or ‘greed’
– whether the aim of the rebels is to generate public goods such as equal political rights
or redistribution, or to yield private gains for the rebels.\textsuperscript{12} They argue that the latter is the

\textsuperscript{11} The model in Olsson & Fors (2004: 332) also indicates how improvement of the rebels’ institutional
environment may reduce conflict intensity.
\textsuperscript{12} Multiple iterations of this article have been released since then, but this one has received the most critical
attention and therefore I use it as a primary reference in this literature review.
case more often, partly because of the collective action problems involved in bringing about public goods, and partly because their empirical analysis indicates that civil wars occur where the opportunities for armed conflict are present, rather than where social or political injustice is particularly prevalent.

Collier, Hoeffler and Söderbom (2004) spell out the implications of this argument for the duration of civil conflict. They restate the two conceptualizations as “rebellion-as-investment” and “rebellion-as-business” (Collier, Hoeffler and Söderbom 2004, 254-5). If most rebellions are best seen as investments where the payoff is attainment of some political goal in the future, wars are expected to last longer the larger are the contested stakes, as long as there is no funding or credit constraint. For instance, if the goal is freedom from repression, wars should be longer depending on how severe the pre-war repression was. Stakes will also be high if control of the government gives access to large revenues from natural resource extraction.13 A different implication would follow from conceptualizing rebellion as business (or as greed), where only the profitability of the ongoing rebellion should affect duration, not the gains that can be reaped only upon victory.

Collier, Hoeffler and Söderbom (2004) conclude from their empirical analysis that the rebellion-as-investment (or as grievance) conceptualization sits less well with the data than the rebellion-as-business one. Declines in commodity prices that may hurt the

13 See Olsson and Fors (2004).
profitability of the rebellion, on the other hand, are found to increase the chances of conflict termination.

Fearon (2004) chooses to code all of the conflicts in his sample into different types. These types are to some extent based on rebel group motivations; he finds that wars emerging from coups, revolutions, and anti-colonial struggles are relatively brief. On the other hand, ‘sons of the soil’ conflicts – or “land or natural resource conflicts between a peripheral ethnic minority and state-supported migrants of a dominant ethnic group” (Fearon 2004, 277)—and wars where the rebel group is funded by contraband—such as drugs or diamonds—tend to be long. However, Fearon explains the variance in civil conflict duration less with reference to the motivations of the rebels than to the commitment problems involved in reaching a negotiated solution.14 Thus Fearon (2004) suggests that different motivations are likely to lead to different outcomes for the conflicts.

Central to the arguments of Collier and Hoeffler (2002) and Fearon and Laitin (2003) is the question of opportunities available for rebel groups to organize an army, recruit soldiers, and obtain the funds to cover running costs. Collier and Hoeffler (2002) focus in particular on the importance of natural resources as sources of rebel group finances. Scholars seem to have reached an overwhelming consensus here. Fearon (2004)

-------------------------------

14 The problem Fearon (2004, 294) discusses is that a government will not commit credibly to autonomy deals signed in periods when it is relatively weak. When the government regains strength, nothing will stop it from overturning the agreement. This commitment problem is more severe when the government’s stakes in the conflict are high.
shows that wars where the rebels have access to contraband finance are long. Collier, Hoeffler and Söderbom (2004) present results indicating that increases in the prices of primary commodities tend to prolong war, and DeRouen and Sobek (2004, 314) show that dependence on primary commodities increases the chance that civil wars will end in truces rather than outright victories.

If the consensus that these civil war scholars have come up with is correct, we should observe a correlation between state strength and insurgent defeat. This argument lends itself to the following hypothesis.

**H₆:** The relative strength of the incumbent will have a direct relationship with incumbent’s chances of victory and a higher risk of war termination (i.e. shorter duration).

Two proxies to measure relative strength are chosen for this study. The first is *power*, defined as a country’s share of global military (army size, military spending) and economic (iron production, energy consumption, and population) power. This variable is coded by the Correlates of War’s Composite Index of National Capabilities (CINC) and logged. Measures are taken in the year prior to the war to minimize endogeneity. *Energy* indicates a state’s per capita energy consumption (logged); the variable is more tied to a country’s ability to sustain military power than GDP per capita.

Some scholars examine external military intervention as another means to alter the outcome of civil conflict. How intervention affects the duration of war depends on which side the intervener aligns himself with, and on the initial distribution of power.
between the parties. Logically, an intervention that enables one side to achieve outright military victory is likely to shorten wars. Practically, it would be difficult to test this with cross-national studies, since military strength is only one factor determining duration and it is hard to measure relative power. This is not to say that scholars have not tried to include the variable in their models. Two types of external assistance—safe havens in neighboring states and the provision of military or economic aid—have been deemed especially crucial for insurgent success (Lyall and Wilson 2009, 82).

**H7:** External intervention on behalf of the insurgent group in the form of safe havens and/or military and economic aid will shorten insurgency duration and increase the insurgent group’s chances of victory.

Finally, it is possible that if an incumbent regime is seen as an external occupier, the insurgent group will perform better due to some *ex ante* nationalist sentiment (the “grievance” argument). To capture this, a variable *occupy* is introduced.

**H8:** External occupiers are more likely to lose counterinsurgency wars.

The last variable, and the most difficult to operationalize, addresses the goals and ideologies of the insurgent movements. Jones and Libicki (2008) provide us with a fairly robust classification for both. The goals of a group may be to maintain the status quo, to influence a policy change, to make some sort of territorial gain/bid for autonomy, to overthrow the incumbent regime, or to seek an empire. Ideologies may range from left-wing to right-wing, nationalist to religious. Operationalizing goals and ideologies, however, is very difficult when often, insurgent groups adopt many different goals and
elements of different ideologies. Nevertheless, insurgent ideology could have significant effects on war outcome and duration. In order to resolve the issue, I coded ideology as a binary variable indicating whether the insurgent group’s objectives and rhetoric indicated only one ideological element or if it was tainted with multiple ideological elements. The assumption here is that if a group adopts elements of different ideologies, it can communicate to a wider constituency and draw support from the population base – which is essential to sustaining an insurgency. Moreover, a mixed ideology can decrease the risk of war termination because it may mean that the group is changing its ideology over time to survive in a long war.

**H9:** When facing insurgent groups with a mixed ideology, a state is more likely to lose a COIN war.

**H10:** Mixed ideology decreases the risk of war termination.

*Control Variables*

Four control variables are included in this analysis. *Elevation* records the roughness of terrain – often cited as favoring insurgent success (Fearon and Laitin 2003 and Collier and Hoeffler 2002) – using the average of five altitude measures (in meters) in the conflict area itself. Second, we might anticipate that as the distance increases between the state’s capital and the war zone, the odds of state victory decrease as logistical difficulties multiply. *Distance* therefore measures the distance between the capital city and the principal fighting area in logged kilometers. Following Fearon and Laitin (2003, 81, 84-85), *new state* was included as a variable to denote whether war
onset occurred during the first two years of a new state’s post-colonial existence. It is likely that the combination of colonialism’s de-legitimization in the 1950s and surging nationalism rendered wavering colonial states and weak successor governments especially likely to be challenged and to suffer defeat. Finally, states may also face increased difficulties in defeating insurgents in linguistically complicated areas due to challenges associated with acquiring information. *Language* records the number of languages spoken in each conflict era for all post-1945 wars. Data are drawn from Fearon and Laitin (2003).

**III. Empirical Analysis**

*Part One: Outcome*

The first part of the analysis requires running an ordered logistic regression against the *outcome* variable (Model 1), which is a trichotomous categorical variable that captures wins, draws, and losses from the perspective of the incumbent. The full set of variables, including the control variables, is included in this model. Table 1 (below) shows the results. *Regime* type achieves statistical significance at $p \leq 0.10$ and is negatively correlated with victory – i.e. more democratic states are more likely to lose counterinsurgency wars. This would seem to corroborate the “vulnerable democracy” thesis.

Mechanization is also of significance and negatively correlated with incumbent victory. Lyall and Wilson (2009, 73) argue that the advent of mechanization had deleterious consequences for a military’s ability to wage counterinsurgency wars.
Foraging armies, often quite rudimentary in their level of technological sophistication, were forced to interact extensively with local populations to acquire their provisions – mostly food, water, and fodder. Frequent soldier-population interaction generated high volumes of information that enabled foraging armies to be more selective in their application of rewards and punishments. Highly complex modern armies, by contrast, are isolated from local populations since conflict zones cannot provide either the type or quantity of needed supplies. Mechanized militaries therefore suffer a kind of “information starvation” that inhibits their ability to seek out the insurgents operating from within the population.

Foreign support for insurgents is also negatively correlated with incumbent victory. Perhaps the diffusion of the nation-state system through colonialism may have created greater opportunities for insurgents to seek sanctuary in neighboring states, thereby hiding behind a shield of sovereignty that prevents incumbents from crossing national borders. In turn, the surge of newly independent states over time may well have increased the incentives for patrons or state sponsors to fight proxy wars through the provision of cheap arms to insurgents, in the hopes of destabilizing their enemy domestically. External occupation is also significant and explains a great deal of the model—occupiers are much less likely to win counterinsurgency wars than non-occupiers.

Of note are the variables that do not achieve statistical significance, namely personnel (which predicts almost none of the model), power, and energy – all indicators
of state capacity. Of the three, the result for personnel is most interesting. It suggests one of two scenarios – either large militaries systematically do not involve a considerable amount of military personnel in “low-intensity” conflicts like counterinsurgency wars, or that the number of boots on the ground does not necessarily reflect a winning strategy.

Model 2 is a logistic regression that includes the full set of variables, but the outcome variable has been collapsed into a dichotomous variable combining draws and losses into a single defeat variable. The results are shown in Table 1. Regime, mechanization, and support all have a negative impact on an incumbent’s likelihood of winning a counterinsurgency war. Again, this supports the democratic vulnerability hypothesis. Information-starved mechanized armies may be forced to offer more concessions to terminate and insurgency. Being a new state also raises the likelihood of defeat. Personnel, again, explains no part of the model.

Table 1: War Outcomes

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1 (W/D/L)</th>
<th></th>
<th>Model 2 (W/L)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Standard Error</td>
<td>P&gt;</td>
<td></td>
</tr>
<tr>
<td>REGIME</td>
<td>-0.05*</td>
<td>0.03</td>
<td>0.07</td>
<td>-0.06**</td>
</tr>
<tr>
<td>MECHANIZATION</td>
<td>-0.31*</td>
<td>0.17</td>
<td>0.06</td>
<td>-0.35*</td>
</tr>
<tr>
<td>PERSONNEL</td>
<td>0.00</td>
<td>0.00</td>
<td>0.35</td>
<td>0.00</td>
</tr>
<tr>
<td>POWER</td>
<td>0.22</td>
<td>0.17</td>
<td>0.20</td>
<td>0.16</td>
</tr>
</tbody>
</table>

*Significant at 10%; **Significant at 5%; ***Significant at 1%.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient 1</th>
<th>Coefficient 2</th>
<th>Coefficient 3</th>
<th>Coefficient 4</th>
<th>Coefficient 5</th>
<th>Coefficient 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENERGY</td>
<td>-0.03</td>
<td>0.09</td>
<td>0.73</td>
<td>-0.08</td>
<td>0.11</td>
<td>0.44</td>
</tr>
<tr>
<td>TRADE</td>
<td>0.44*</td>
<td>0.27</td>
<td>0.10</td>
<td>0.26</td>
<td>0.25</td>
<td>0.30</td>
</tr>
<tr>
<td>OCCUPY</td>
<td>-2.03***</td>
<td>0.62</td>
<td>0.00</td>
<td>-0.95</td>
<td>0.66</td>
<td>0.15</td>
</tr>
<tr>
<td>SUPPORT</td>
<td>-0.67***</td>
<td>0.21</td>
<td>0.00</td>
<td>-0.74***</td>
<td>0.28</td>
<td>0.01</td>
</tr>
<tr>
<td>IDEOLOGY</td>
<td>0.64*</td>
<td>0.35</td>
<td>0.07</td>
<td>0.65</td>
<td>0.46</td>
<td>0.16</td>
</tr>
<tr>
<td>ELEVATION</td>
<td>0.20*</td>
<td>0.11</td>
<td>0.06</td>
<td>0.32**</td>
<td>0.14</td>
<td>0.02</td>
</tr>
<tr>
<td>DISTANCE</td>
<td>0.00</td>
<td>0.07</td>
<td>0.96</td>
<td>-0.03</td>
<td>0.10</td>
<td>0.75</td>
</tr>
<tr>
<td>LANGUAGE</td>
<td>-0.05**</td>
<td>0.03</td>
<td>0.08</td>
<td>-0.05</td>
<td>0.04</td>
<td>0.19</td>
</tr>
<tr>
<td>NEW STATE</td>
<td>0.78</td>
<td>0.59</td>
<td>0.19</td>
<td>1.28**</td>
<td>0.59</td>
<td>0.03</td>
</tr>
</tbody>
</table>

**Cutpoints/Constant**

-3.07

**Observations**

135

**Wald chi2**

46.95*** 35.28***

**Prob > chi2**

0.00 0.00

**Log Likelihood**

-126.89 -72.02

**$R^2$**

0.14 0.16

*Part Two: Duration*

The next part of the analysis tested the explanatory variables against the *duration* variable. In order to decide on a strong model for testing against *duration*, I first conducted a Kaplan-Meier survival analysis (KMSA) to get a general sense of the shape of the distribution. The estimate is visualized in Figure 1. Fortunately, KMSA can estimate survival functions even in the presence of estimated cases. It is limited, however, in that it models survival based only on time dependence, without covariate effects.
Since the KMSA curve shows a relatively smoothly increasing hazard rate over time, a parametric Weibull hazard regression model was employed for the duration analysis in this study, and the regressions were run with shared frailty terms centered on country codes to account for unobserved heterogeneity that may lead certain states to “fail” at a higher rate than others. Table 2 reports hazard ratios for these Weibull regressions. Hazard ratios are interpreted relative to the baseline ratio of 1, where ratios higher than 1 indicate that a variable increases the risk of the failure event over time (i.e.

16 The Weibull parametric model allows a flat, smoothly decreasing, or smoothly increasing hazard rate. Other survival analysis models, such as the Cox proportional hazard model, have different assumptions about the shape of the hazard curve.
the war is shortened), while a ratio lower than 1 indicates that a variable reduces the risk of these events.

Model 3 encompassed the full set of variables, and three explanatory variables came up significant at $p \leq 0.01$ – trade, occupy and support. Trade and occupy increase the risk of war termination, whereas support reduces the risk of shorter wars. A country integrated into the global economy cannot sustain a long civil war without being vilified by some contingent of international opinion. These results are not totally intuitive, however. That external support for insurgents typically lengthens wars is contested by Collier, Hoeffler and Söderbom (2004, 267), who find military interventions on the rebel side to be significant in decreasing the duration of civil war. The implication is that interventions in favor of governments are likely to increase the chance of stalemated conflicts, and their results confirm this expectation. The difference may be explained by the operationalization of the variable support – this analysis takes into account material, economic, and military aid, as well as the ability of an insurgent to use a neighboring country as a sanctuary. If an insurgent group can take refuge in another state’s territory, it limits the ability of the incumbent to defeat it.

Though occupy increases the risk of war termination, it indicates an extremely high risk of state defeat as well (Model 4). These two results are complicated even more by the control variable distance, which is statistically significant in reducing the risk of war termination (i.e. lengthening it) but also reduces the risk of state defeat. The results may be explained by the fact that more powerful states with the capacity to fight and win
counterinsurgency wars are more likely to project their forces a greater distance and achieve victory – i.e. they self-select themselves into distant conflicts. In the post-Cold War era, many conflicts have been defined by local, separatist conflicts, and even though an occupation might be adjacent to the incumbent’s territory, the incumbent loses because of weak state capacity to begin with. The nations that broke away from the Federal Republic of Yugoslavia and fought counterinsurgency wars—e.g. Bosnia and Croatia—epitomize this trend, where the incumbent did not have to project a force far away to occupy. Spain and France occupied nearby Morocco and Algeria, respectively, and China only had to cross the Taiwan Strait.

Regime does achieve conventional levels of significance in the duration analysis, and this is somewhat puzzling. The prevailing wisdom says that democratic publics are impatient and casualty-averse; hence, constituents of democratic regimes would withdraw support from protracted conflicts (Reiter and Stam 2002, 168-72). Lyall (forthcoming) remains skeptical of this consensus, arguing that nearly every study of democracies fighting counterinsurgency campaigns rests on a no-variance research design – “audience costs” may be found in authoritarian regimes too, but no one has studied the two types together.

Surprisingly, ideology does for once achieve a conventional level of significance in Model 4 at the $p \leq 0.01$-level, reducing the risk of state defeat. This statistic suggests that a mixed ideology is beneficial to insurgents and likely will lead to incumbent defeat. The longer a war continues, the harder it becomes for the purists to communicate to a
wider constituency that is so essential to sustaining an insurgency. Groups that tend to adopt other ideological elements, at least in their rhetoric, may be better able to maintain the support of the indigenous population.

Table 2: War Duration

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 3 (Full Set)</th>
<th>Model 4 (Defeats Only)</th>
<th>Model 5 Wins Only)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hazard Ratio</td>
<td>Std. Error</td>
<td>P&gt;</td>
</tr>
<tr>
<td>REGIME</td>
<td>0.98</td>
<td>0.02</td>
<td>0.24</td>
</tr>
<tr>
<td>MECH</td>
<td>0.91</td>
<td>0.10</td>
<td>0.37</td>
</tr>
<tr>
<td>PERSONNEL</td>
<td>1.00</td>
<td>0.00</td>
<td>0.17</td>
</tr>
<tr>
<td>POWER</td>
<td>1.11</td>
<td>0.10</td>
<td>0.21</td>
</tr>
<tr>
<td>ENERGY</td>
<td>0.97</td>
<td>0.06</td>
<td>0.59</td>
</tr>
<tr>
<td>TRADE</td>
<td>1.54***</td>
<td>0.19</td>
<td>0.00</td>
</tr>
<tr>
<td>OCCUPY</td>
<td>1.70*</td>
<td>0.52</td>
<td>0.08</td>
</tr>
<tr>
<td>SUPPORT</td>
<td>0.63***</td>
<td>0.08</td>
<td>0.00</td>
</tr>
<tr>
<td>IDEOLOGY</td>
<td>1.10</td>
<td>0.22</td>
<td>0.64</td>
</tr>
<tr>
<td>ELEVATION</td>
<td>0.81***</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>DISTANCE</td>
<td>0.89***</td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>LANGUAGE</td>
<td>0.98</td>
<td>0.01</td>
<td>0.23</td>
</tr>
<tr>
<td>NEW STATE</td>
<td>1.26</td>
<td>0.35</td>
<td>0.41</td>
</tr>
<tr>
<td>Observations</td>
<td>144</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wald chi2</td>
<td>128.02***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

17 *Significant at 10%; **Significant at 5%; ***Significant at 1%.
IV. Additional Robustness Tests

The data was then tested to include state-fixed and time-fixed effects to see if there were any significant difference by country or by time period. State-fixed effects were tested for in order to see whether the results were being driven by the performance of specific outliers. I created a dummy variable for every incumbent with multiple observations whose mean outcome was significantly higher or lower than the mean outcome. Significance was assigned to mean country outcomes that were more than half a standard deviation (0.41) above or below the mean outcome (0.98).18 The countries that performed significantly well in multiple counterinsurgency wars were Algeria, Iran, the Philippines, Yemen and Zimbabwe. The set of countries that performed poorly consisted of Afghanistan, Cambodia, Chad, Liberia, Portugal, South Africa, South Vietnam, and the United States.

It is also imaginable that unobservable factors in specific time periods (e.g. a technological innovation) might also be affecting the results. Time-fixed effects were tested by generating dummy variables for every decade; for instance, the first period was

---

18 See summary statistics for OUTCOME in Appendix B. $x \leq 0.57$ or $x \geq 1.39$ are significant.
1945-1954, and so on. Decade 6 (1995-2004) was dropped because the data is right-censored – i.e. there are not enough data points for wars that began in that decade and have already terminated. Both state- and time-fixed effects were tested with the full model against the trichotomous *outcome* variable. The results are shown in Model 6, below.

It turns out that both state-fixed effects (higher and lower) are statistically significant in the model, as well as several decade-level effects. However, the addition of these variables does not seriously affect the impact of the explanatory variables that were significant in the first model. The measures for regime type, external support for insurgents, occupation, and ideology still remain significant (though *trade* and *mechanization* notably do not).

Table 3: State- and Time-Fixed Effects

| Variables   | Coef. | Std. Error | P>|z| | Variables   | Coef. | Std. Error | P>|z|
|-------------|-------|------------|-----|-------------|-------|------------|-----|
| REGIME      | -0.06*| 0.03       | 0.07| HIGHER      | 1.28**| 0.61       | 0.04|
| MECH        | -0.24 | 0.16       | 0.14| LOWER       | -1.97***| 0.46       | 0.00|
| PERSONNEL   | 0.00  | 0.00       | 0.39| DECADE1     | 1.82*  | 1.12       | 0.10|
| POWER       | 0.19  | 0.20       | 0.34| DECADE2     | 1.80   | 1.17       | 0.12|
| ENERGY      | -0.03 | 0.11       | 0.79| DECADE3     | 1.99** | 0.92       | 0.03|
| TRADE       | 0.45  | 0.27       | 0.09| DECADE4     | 1.42   | 0.88       | 0.11|
| OCCUPY      | -1.89***| 0.69       | 0.01| DECADE5     | 1.57*  | 0.86       | 0.07|
| SUPPORT     | -0.47**| 0.21       | 0.03| Cutpoints   | -1.72  |            |     |
|IDEOLOGY    | 0.59* | 0.35       | 0.09| Observations| 0.21   |            |     |
|ELEVATION    | 0.12  | 0.12       | 0.35|             | 135    |            |     |

19 *Significant at 10%; **Significant at 5%; ***Significant at 1%.
V. Conclusion

Policy Implications

In general, state capacity variables (personnel, global share of military and economic power, and energy use per capita) notably did not achieve significance. Clearly then, winning a counterinsurgency war is not necessarily a matter of having the material resources and human capital. On the other hand, trade flows, mechanization, external support for the insurgents, and external occupation all came up as significant. The *trade* results in Models 1 and 3 suggest that in order to develop a winning strategy, a country may need to discount international opinion against strategy options that may lead to a swift defeat of the insurgents.

Mechanization speaks to the force structure of the army units fighting an insurgency. It may be prudent for armies that find themselves fighting a disproportionate amount of counterinsurgency wars or for whom the cost of COIN warfare is incredibly high, to consider restructuring their units to exploit the useful elements of the older “foraging” units – namely interaction with the local population, military policing, more infantry, and less dependencies and logistical ties with the homeland. Self-sustaining infantry and policing units are arguably the best way to organize a military’s forces in
order to win a counterinsurgency war. But in the end, mechanization never reached strong levels of statistical significance in this analysis, bringing Lyall and Wilson’s (2009) conclusions into question.

Case studies might show that mechanization fosters a particular kind of operational culture that removes the soldier from the population. Nevertheless, dismounting is a choice, and when the stakes are high enough, regardless of complicated logistical dependencies on the homeland and the ratio of armed personnel to vehicles, a soldier will interact with the population and get the information he or she needs when necessary.20 States should not worry about how much their forces rely on machines when fighting an insurgency, but should focus on engendering a foraging soldier’s instinct and intelligence gather capability in a mechanized soldier who can choose to dismount.

Foreign support for insurgencies, whether sanctuary or financial/military aid, is ruinous for an incumbent. Support for insurgents is strongly correlated with a negative outcome for the incumbent, protracts conflicts, and reduces the chances of a quick victory. If an insurgent group is able to obtain territorial safe haven in a neighboring country, or if they are receiving aid from a foreign patron—be it military or economic—a policymaker may choose to go after the patron first and cutting off support to the insurgents, increasing its likelihood of winning the war. This strategy can be pursued concomitantly to fighting the insurgents directly, but the earlier the support is severed, the faster the war should terminate—an important consideration in conflicts where the ________________________

20 Thanks to Doug Schaffer for this bit of insight.
leaders are reluctant to sanction or coerce state sponsors to minimize political risk. A cost-benefit calculation must be made in those cases.

External occupation is hands down the most significant variable in its affect on both outcome and duration. In simple terms, military occupation is typically a losing proposition. Edelstein (2008, 5) finds that occupation has only succeed in 7 out of 26 occasions studied. Occupation also has substantial impact on lengthening war duration. Of the 33 counterinsurgency wars in which an incumbent was an external occupier, 19 (57%) were fought by democratic regimes (score ≥ 7). Hence, democratic regimes are more vulnerable to being sucked into a counterinsurgency war on foreign soil, and when they occupy, they will likely lose. The causal process needs to be studied, but it could range from democracies’ attraction to humanitarian intervention to force projection over great distances to the inability of democratic publics and policymakers to understand the perils of occupation. Whatever the case may be, policymakers should exercise extreme caution before making the decision to occupy a foreign country to quell an insurgency.

Group ideology came up as significant in Model 4, and if policymakers are convinced of that relationship between long wars and mixed ideologies, he or she might consider ways in which the counterinsurgency effort can undermine the communication of the insurgent group with its core and peripheral constituencies – especially the latter. The more grievances a group can lay claim to, even if it is reflected only in its rhetoric and not in its actions, the wider the base of support that the group achieves.
Counterinsurgency is warfare of the most brutal kind. It can only be hoped that this analysis has yielded valuable results to inform policymakers in their efforts to minimize the amount of blood and treasure that goes into fighting an insurgency. Executing counterinsurgency operations in Iraq and Afghanistan (and perhaps in the future, Pakistan) has required an immense number of troops and massive expenditures. Policymakers would do well with some historical background on what works and what does not in COIN operations. In an economic crisis, it becomes more imperative to spend resources wisely.

Future Research

My findings suggest several profitable avenues for future research. First, more work can be done to develop variables associated with insurgent groups—group size, breadth of goals, and strategy, to begin with. It is plausible that outcomes vary according to the extent and nature of group demands. Similarly, the number of insurgent groups active in a conflict likely affects the difficulties an incumbent faces. Moreover, better state efficacy variables can be generated, including police per capita. Advances could also be made by exploring draws, which represents a growing share of outcomes since the 1960s. There certainly must be variation in the types of concessions made by incumbents. Low-level concessions to end a conflict might still be considered a victory by the incumbent, though it would still be coded here as a draw.

While findings in my analysis focus on COIN outcome and duration, they nonetheless suggest new avenues of research for the study of COIN war initiation among
the subset of combatants that fought as external occupiers. It remains unclear, for example, why powerful democracies represent such a disproportionate share of this subset and why their leaders have apparently proved incapable of “selecting” into wars with a higher probability of success. It is difficult to tell whether this is because democracies do not invest in force structures suitable for COIN wars in anticipation of conventional wars, or if they do not have the political will to pass the necessary reforms to increase odds of victory in a long war. Prior to drawing broad policy conclusions, however, we require quantitative analyses that not only vary regime type, but also the very decision to intervene. It may be ideal for future studies to include cases of non-intervention, enabling us to examine whether different regime types are driven by the same motives and policy goals when choosing to intervene militarily.

Though better measures of democracy may lead to different conclusions, it appears that regime type has little analytical utility for explaining COIN war outcomes and duration. Analysis that focuses on regime type is not very informative anymore, especially because the results vary so drastically on case selection—hence the disagreement in the academic community on the democratic vulnerability thesis. Instead, emphasizing battlefield dynamics—that is, the strategies and tactics used by an incumbent—may prove a better line of research than focusing on regime-specific variables. In other words, instead of focusing on who is fighting, we might be better off examining how they are fighting. Any study of COIN strategy, however, will have to distinguish between the proximate strategy that wins a war and the evolution of strategy
from directly before onset through to termination. Jones and Libicki (2008) look at proximate cause of termination, but often this could be a fluke. Almond and Genco (1977, 496-7) write, “In their fascination with powerful regularities and uniformities that have the properties of causal necessity […] social scientists have overlooked the fact that much of social and political change has to be explained neither by strong […] nor weak regularities, but by accidental conjunctions – by events that had a low probability of occurring.” Proximate strategies, be it military intervention or policy change, may just be an “accidental conjunctions” that lead to war termination. Case studies and process tracing may be better suited to answer the question of strategy than quantitative analysis, though they may never achieve parsimony.

A variable that I have not come across in my research but still seems logically worth inclusion is income inequality within a state, as that would be a strong indicator of grievance. But this data is hard to come by. Although the World Bank has published GINI coefficients for 138 countries over the period form 1890-1998, the dataset did not include many of the country-year dyads that were necessary for this analysis. Even if data measuring income inequality for a large number of states over a long period of time could be collected, there would still be some doubt as to whether it would adequately tap the intended theoretical concept. As Weede (1981, 641) points out, “Only if there is a discrepancy between what people get and what they feel rightly entitled to is there relative deprivation. As long as an untouchable street sweeper in Calcutta accepts his low status and extremely meager income as deserved (that is, earned by bad deeds in a
previous life), he will not feel deprived relative to his standards of justice and expectations, however absolutely deprived he may seem by Western standards.”

Finally, Fearon and Laitin’s (2003) analysis gives us a good starting point in suggesting that ethnic differences and group grievance are not causal factors for the onset of civil war. However, it may well be that ethnic differences and group grievance help perpetuate civil war, either because those issues are exacerbated in a protracted conflict or because of the development of myths by leaders to sustain the ideological appeal of the conflict. Fearon and Laitin (2003, 88) propose “intense grievances are produced by civil war.” My ideology variable could not capture this dynamic.
Appendix A: Codebook

Dependent Variables

OUTCOME: War outcomes are coded from the incumbent’s perspective. A win occurs when the insurgency is militarily defeated and its organization destroyed or the war ends without any political concessions granted to insurgent forces. A loss occurs when the incumbent unilaterally concedes to all, or nearly all, insurgent demands, including the granting of independence or the deposition of current leaders. A draw occurs when an incumbent is forced to concede to some, but not all, insurgent demands, and neither side obtains its maximal aims. From Lyall and Wilson (2009).

DURATION: War duration measures the length of the conflict in months (logged). Data are taken from the UCDP-PRIO Armed Conflict Dataset (Version 4). War start dates are coded from the first large-scale insurgent attack, open declaration of hostilities, or the date of occupation. War termination dates are identified by several means, including the last day of war-related fatalities and the date of a peace settlement.

Explanatory Variables

REGIME: Each country’s regime type is coded using Polity2 values from the PolityIV dataset. Polity2 is a 21-point scaled composite index of regime type that ranges from highly autocratic (-10) to highly democratic (+10). Values are from one year prior to the conflict.
MECH: Drawing on data from Lyall and Wilson (2009), mechanization is defined as the ratio of soldiers per motorized vehicle, is measured one year prior to war, and takes five values, with 1 representing the lowest level of mechanization and a 4 the highest.

PERSONNEL: Drawn from the Correlates of War by Lyall and Wilson (2009), this is the estimated number of personnel in a state’s military (in thousands). Measured in the year prior to war’s outbreak.

POWER: Recorded by the Correlates of War’s Composite Index of National Capabilities (CINC), this variable denotes a country’s share of global military (army size, military spending) and economic (iron production, energy consumption, and population) power. Measures are logged and taken in the year prior to the war.

TRADE: Trade data are drawn from the International Monetary Fund Direction of Trade data set, while GDP is taken from the Banks Cross-National Time Series Dataset. Measures are taken in the year prior to the war.

ENERGY: A state’s per capita energy consumption (logged). This has the advantage of extending across a longer time period than the standard measure of capacity—GDP per capita—and is more tied to a country’s ability to sustain military power. Measures are taken one year prior to war onset. As coded by Correlates of War.

OCCUPY: An estimation of the depth of ex ante nationalist sentiment. Binary variable that denotes whether the state authority was an external occupier. A country is coded as
an external occupier (1) if its military forces cross an internationally recognized border in order to suppress an insurgency. From Lyall and Wilson (2009).

SUPPORT: A variable that measures whether insurgents received two critical types of assistance: material, economic and military aid, and the ability to use a neighboring country as a sanctuary (with or without the permission of the neighboring state). A conflict scores a 2 if the insurgent group received both types of assistance; a 1 if only one type was granted; and a 0 if neither aid nor sanctuary was received by a particular state. From Lyall and Wilson (2009).

IDEOLOGY: A binary variable indicating whether the insurgent group’s objectives and rhetoric indicated only one ideological element (1) or if it was tainted with multiple ideological elements (0). Drawn from various sources.

**Control Variables**

ELEVATION: Quantitative and qualitative studies converge on the finding that rough terrain, specifically mountains and deep forests, increase the likelihood of insurgent victory by providing refuge from incumbent power. The roughness of terrain is estimated by taking the average of five altitude measures (in meters) of the conflict area itself.

DISTANCE: It can also be anticipated that as the distance between the incumbent’s capital and the conflict area increases, the probability of incumbent victory decreases as logistical difficulties multiply. Therefore, the distance in kilometers (logged) between the incumbent’s capital and the principal conflict area is included in the analysis. Wars that
are primarily fought in or near a capital city are assigned a nominal one kilometer in distance.

LANGUAGE: Number of languages spoken in the conflict areas. Drawn from Fearon and Laitin (2003).

NEW STATE: Denotes whether war onset occurred during the first two years of a new state’s post-colonial existence.
### Appendix B: Summary Statistics

<table>
<thead>
<tr>
<th>Variables</th>
<th>Num</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>OUTCOME</td>
<td>135</td>
<td>0.98</td>
<td>0.82</td>
<td>0</td>
<td>2</td>
<td>Categorical win-draw-loss variable</td>
</tr>
<tr>
<td>DURATION</td>
<td>135</td>
<td>3.83</td>
<td>1.47</td>
<td>0</td>
<td>6.32</td>
<td>Length of conflict in months (logged)</td>
</tr>
<tr>
<td>EXPLANATORY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Explanatory Variables</td>
</tr>
<tr>
<td>REGIME</td>
<td>135</td>
<td>-0.33</td>
<td>7.02</td>
<td>-10</td>
<td>10</td>
<td>Incumbent’s Polity2 value from the PolityIV data set</td>
</tr>
<tr>
<td>MECHANIZATION</td>
<td>135</td>
<td>2.56</td>
<td>1.14</td>
<td>1</td>
<td>4</td>
<td>Ratio of soldiers per vehicle</td>
</tr>
<tr>
<td>PERSONNEL</td>
<td>135</td>
<td>619.36</td>
<td>1747.39</td>
<td>1</td>
<td>12500</td>
<td>Number of armed personnel (in thousands)</td>
</tr>
<tr>
<td>ENERGY</td>
<td>135</td>
<td>-1.30</td>
<td>2.38</td>
<td>-8.07</td>
<td>2.78</td>
<td>Incumbent’s per capita energy consumption (logged)</td>
</tr>
<tr>
<td>TRADE</td>
<td>135</td>
<td>-3.67</td>
<td>1.20</td>
<td>-9.28</td>
<td>-0.46</td>
<td>Distribution of incumbent’s exports and imports (from IMF Direction of Trade Statistics)</td>
</tr>
<tr>
<td>OCCUPY</td>
<td>135</td>
<td>0.24</td>
<td>0.43</td>
<td>0</td>
<td>1</td>
<td>Denotes whether or not incumbent is an external occupier</td>
</tr>
<tr>
<td>SUPPORT</td>
<td>135</td>
<td>1.01</td>
<td>0.85</td>
<td>0</td>
<td>2</td>
<td>2 if sanctuary and support, 1 if only sanctuary or support; and 0 if none.</td>
</tr>
<tr>
<td>IDEOLOGY</td>
<td>135</td>
<td>0.61</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
<td>1 if pure, 0 if mixed</td>
</tr>
<tr>
<td>CONTROL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Control Variables</td>
</tr>
<tr>
<td>ELEVATION</td>
<td>135</td>
<td>6.00</td>
<td>1.34</td>
<td>0</td>
<td>8.50</td>
<td>Average of five elevation readings in conflict area (logged)</td>
</tr>
<tr>
<td>DISTANCE</td>
<td>135</td>
<td>4.18</td>
<td>3.51</td>
<td>0</td>
<td>9.54</td>
<td>Distance from incumbent’s capital to conflict area (km, logged)</td>
</tr>
<tr>
<td>LANGUAGE</td>
<td>135</td>
<td>8.04</td>
<td>7.39</td>
<td>1</td>
<td>30</td>
<td>Number of languages spoken in conflict area</td>
</tr>
<tr>
<td>NEW STATE</td>
<td>135</td>
<td>0.16</td>
<td>0.37</td>
<td>0</td>
<td>1</td>
<td>Denotes whether conflict broke out within 2 years of independence</td>
</tr>
</tbody>
</table>
References


