APPENDIX A

Methods and Data

This appendix describes the overall methodological approach employed in this research, the historical COIN cases informing the analyses and how they were selected, and the specific methods used in the analyses. Our goal was to test the validity and range of applicability of the 24 COIN concepts described in Chapter Three against substantial historical evidence. How have COIN forces that have adhered to the tenets of the various concepts fared historically? How can these lessons inform preparations for contemporary and future COIN contingencies?

Charles Ragin's Qualitative Comparative Analysis

Early in our planning for the original research in this series (documented in *Victory Has a Thousand Fathers: Sources of Success in Counterinsurgency*), we remembered a previous encounter with sociologist Charles Ragin's work on case-based comparative historical analysis using QCA, a tool designed to assess configurations of case similarities and differences using simple logical rules.[1] We carefully considered the application of his methods to this problem and concluded that QCA was an ideal match. We structured our data collection and analysis to allow us to employ Ragin's QCA approach in the original study, and we retained a similar structure for this extension of that research.

[1] See Ragin, 1987.

Through the use of "truth tables," QCA provides a holistic approach to qualitative historical comparison by viewing cases in terms of combinations of binary (present or absent) factors.[2] Using computer algorithms first developed for the simplification of switching circuits, researchers are able to compare a large number of cases as configurations—many more than they could possibly "hold in their heads" using traditional case-oriented comparative methods. This case-based method for analytic aggregation allows for the quantification of otherwise voluminous amounts of qualitative data. As such, it compels researchers to be explicit about outcomes of interest and proposed causal relations, including necessary or sufficient causes and conditional or contributing causes.

QCA relies on the application of Boolean algebra to a truth table, in which selected factors are scored as present or absent (1 or 0) for all selected cases.[3] The truth table has as many rows as there are logically possible combinations of values for the selected factors. (For example, including four binary factors in the analysis would result in $2^4 = 2 \times 2 \times 2 \times 2 = 16$ rows.) Rows are first reduced by removing patterns of factors that do not occur in the data—that is, any row that does not correspond to one or more actual cases. Boolean algebra then allows further reduction of the combinatorial matrix to expose simplified patterns of relationships and determine the prime implicants.

---

2 "Binary" indicates that a factor can take on only one of two values. In our case, that is *present* or *absent*, always represented by 1 and 0, respectively. A truth table, then, is a collection of rows of 1s and 0s that represent every pattern of presence and absence of the factors of interest that appear in the data.

3 Boolean algebra was developed in 1954 by George Boole. See George Boole, *An Investigation of the Laws of Thought*, Amherst, N.Y.: Prometheus Books, 2003. Boolean algebra differs from standard high school algebra in two ways. First, values are logical instead of numerical values. These are *true* or *false*, *present* or *absent*, and are represented as 1 or 0. Second, logical values dictate slightly different mathematical operations obeying slightly different mathematical laws. Many readers will be familiar with Boolean search operators, such as *and*, *or*, and *not*, as they can be used in some search engines. The application of Boolean algebra here has two implications: It requires us to structure our data with logical values (true or false, or, in our case, present or absent), and it allows complex patterns of data to be reduced to the minimum set of factors necessary to determine a pattern, called *prime implicants*.

Prime implicants are the minimally sufficient patterns of factors that fully describe the pattern of outcomes of a set of cases. In our analysis, the prime implicants are concepts (or patterns of factors representing concepts) that describe the patterns of success or failure (the outcomes) in our cases and thus received strong support.

Though the prime implicants are determined mathematically, once they are identified, the analysis can turn back to the qualitative nuances of the individual cases. Cases with surprising patterns, or patterns that usually result in success but did not, can be singled out for more detailed case-study analysis. This can lead to further inductive theory development. Imagine a situation in which the presence of three factors leads to a COIN force win in all cases except one. Thorough and careful examination of the details of that exceptional case could reveal many different things, any of which would be informative. It could be that one or more of the three critical factors are not really present in the exceptional case but they were evaluated as present based on a superficial reading of the history. Or it could be that the three critical factors are very much present, but a detailed exploration of the case reveals a narrative showing that the impact of the three factors was thwarted by the presence of a fourth factor, which proves to be absent in the other cases containing the original three factors of concern. In this event, the addition of a fourth factor perfects the set of prime implicants. (Now, the presence of three factors plus the absence of the new fourth perfectly predicts COIN force victory.) Discerning what exactly is exceptional about the exceptional case leads to a better understanding of that case and the other cases as well.

This method is particularly well suited to our research effort because it allows mathematical principles to be applied to fundamentally qualitative data without in any way compromising the qualitative nuance necessary to identify and resolve exceptions. Boolean reduction allows us to identify and evidence factors and interactions between factors that have historically led to successful COIN outcomes. Thus, we can test the concepts associated with these factors.

In many cases, the intention to apply QCA drove how we structured our data and the collection of those data. For a more in-depth explanation of how QCA was actually applied to the data, see the section "Additional Details on the Use of Ragin's Qualitative Comparative Analysis," later in this appendix.

Case Selection

QCA is potentially applicable across any set of cases. As is true with any inferential analyses, findings are generalizable only across cases that can be argued to be comparable with the sampled cases. Since our sponsor's interest was in preparing U.S. forces for success in contemporary and future operations, we sought historical cases that were likely to be as representative as possible of the contemporary state of the art in insurgency and COIN. In an effort to be contemporary yet comprehensive, we elected to study all insurgencies worldwide begun and completed between WWII and 2010. We chose completed cases because we were interested in factors that contributed to the outcomes, which are impossible to assess if the outcome is not yet determined. Once we had compiled a list of the world's resolved insurgencies in the post-WWII era, we sought to narrow down our data set using an agreed-upon collection of distinguishing characteristics.

Identifying and enumerating historical insurgencies worldwide is a nontrivial undertaking. There have been many insurgencies in the course of human history and many other similar conflicts from which they must be distinguished.[4] RAND's Martin Libicki recently prepared a list of 20th- and 21st-century insurgencies.[5] He began with a list of 127 insurgencies started by 1999 that was developed by other scholars.[6] These 127 cases met three criteria:

---

[4] Insurgency is a centuries-old form of conflict that pits the weak against the strong. Indeed, writing between 400 and 300 B.C., with an emphasis on intelligence, hit-and-run tactics, and adaptability, Chinese strategist Sun Tzu essentially laid out the basis for guerrilla warfare in his timeless classic *The Art of War*. Ancient Rome also provided fertile ground for insurgency in such places as Gaul and Judaea.

[5] Libicki, 2008.

[6] The base list comes from Fearon and Laitin, 2003.

- They involved fighting between states and nonstates seeking to take control of a government or region or that used violence to attempt to change government policies.
- The conflict killed at least 1,000 people over its course, with a yearly average of at least 100.
- At least 100 people were killed on each side (including civilians attacked by rebels).

Starting with this list, Libicki first excluded cases that could be classified as coups, countercoups, or insurrections. (There were 51 such cases; subtracted from 127, this leaves 76.) He then added 11 insurgencies that began (or crossed the threshold of 1,000 deaths) after the 1999 cutoff of the foundational list (so, 87 cases). Finally, careful consideration led two conflicts that had previously been excluded to be returned to the list. This left 89 insurgencies covering the period from 1934 to 2010.

To extend Libicki's list, we added four cases from a list prepared by the Center for Army Analysis and the Dupuy Institute that were missing but appeared to meet Libicki's criteria, for a total of 93 cases.[7] Of the 93 total cases, we excluded 17 conflicts still considered ongoing or unresolved, which included not only conflicts listed as unresolved on Libicki's list but also two conflicts listed as resolved whose resolution our analysts disputed: Burma (1948–2006) and the Moro Islamic Liberation Front insurgency in the Philippines (1977–2006). We then excluded one conflict that began before WWII: China (1934–1950); two conflicts that were not clear-cut cases of insurgency but insurrections followed by massive superpower interventions, Lebanon (1958–1959) and the Dominican Republic (1965–1966); one case that was more akin to a "police action," Congo/Katanga (1960–1965); and one case that was less an insurgency and more of a coup (and, thus, should have been excluded by Libicki), the Biafran secession in Nigeria (1967–1970). These reductions left 71 cases, 30 of which were examined in the *Victory Has a Thousand Fathers* research. The new set of 71 cases

7 See C. Lawrence, 2008.

includes all insurgencies worldwide begun and completed between WWII and 2010.

This set of cases has several attractive features from an analytical perspective. First, it is exhaustive over the period under examination, so it constitutes the universe of insurgencies begun and resolved between WWII and 2010. This is not a sample of insurgencies over this period—this is the whole population. No statistics are necessary to make inferences about the extent to which these data represent a larger population; the data are perfectly representative of the past 65 years of completed COIN operations. Second, they represent many different regions, with cases in South America, Central Asia, Africa, and the Far East. If regional differences in the conduct or context of COIN were to significantly affect the performance of various COIN concepts, these data would reflect them. Third, there is significant variation among COIN forces—from world superpowers (United States, Soviet Union) to near-peer nations (Turkey, United Kingdom) and non-peer nations (Rwanda, Tajikistan)—and insurgent forces, which span the spectrum from highly advanced (Lebanese Hizballah, LTTE) to less advanced (Revolutionary United Front, MNLF) and everywhere in between.

The Exclusion of Colombia (La Violencia) from the 59 Core Cases

As discussed in Chapter Three, Colombia (La Violencia) was excluded from the 59 core comparative cases because its outcome was deemed essentially indeterminate. The Colombia case was highly distinctive in its general narrative but most distinctive in its outcome. The COIN force did not "lose" in the traditional sense. By the end of the conflict, COIN forces and the associated political party had become so disenchanted with the current president, who was excessively corrupt and under whom the economy had suffered, that they were willing to negotiate a power-sharing agreement that would remove him from office. The government party retained substantial political rights under the power-sharing agreement, including a turn-taking process that would have the presidency alternate between parties, beginning with the leftists (the side of the insurgents). And there is the rub: Had the first turn been taken by the rightists (the party that controlled the government for most of the conflict), we likely would have scored this case as "mixed, favoring COIN" instead of "mixed, favoring insurgents." However, following the procedures laid out in Figure 2.2, we classified La Violencia as favoring the insurgents. Our procedure for making black and white out of gray worked so well that it allowed us to assign a clear outcome to a case that perhaps we should not have, a case whose outcome (such as it was) hinged on a specific personality and quirks of negotiation for power-sharing about which side would share the power first. Had the outcome favored the COIN force, this case still would not be a good example of good COIN practices; given how marginally the outcome favored the insurgents, it is definitely not a ringing condemnation of the COIN practices that were followed. We decided to consider the case as a poor learning example due to its minimally determinate outcome and thus flagged it for exclusion.

Factor Generation, Evaluation, and Scoring

For each case, we completed a case narrative and collected data on roughly 289 specific factors.[8] Selecting the factors to evaluate was, itself, a methodologically interesting process.

Crisp-set QCA requires binary data for reduction to prime implicants using Boolean algebra. Given the difficulty of trying to quantify many of the concepts that we sought to test (e.g., security, democracy, legitimacy) in any discrete, scaled, or even ordinal way, binary (present/absent, or 1/0) scoring was eminently suitable.

The identification and refinement of these binary factors was an inductive and iterative process. We began with the 79 factors scored for the original *Victory Has a Thousand Fathers* data set, which were based on an extensive review of the literature on strategic communication and COIN. We added several hundred additional factors as refinements and expansions, based on questions we asked ourselves about

[8] While 289 is the number of factors used in our analyses and listed in Appendix E, we actually collected (or attempted to collect) data on several additional factors. Some of these factors are included in the data and set off from their main factor with a subordinate number (for example, factor 155a), and some factors are not included because they proved impossible to reliably ascertain in many of the cases (factor 139, for example).

the cases and data, questions raised during briefings and discussions related to *Victory Has a Thousand Fathers*, or emergent questions in the ongoing and evolving literature on COIN.

Once we identified the practices advocated by the various COIN concepts and laid them out as measurable factors, we engaged in vigorous debate over whether the factors truly represented what we intended for them to capture. We revisited the factor list repeatedly as data collection and analysis progressed. This process of refinement spanned much of the project and relied on examples and experiences from the individual case studies whenever possible. Factors were adjusted (or eliminated) due to the difficulty assessing them with the available historical data, because of the nuance necessary for specific cases or to better capture the tenets of the concepts as they played out in real cases. Whenever a factor or its criteria changed, all previously scored data on that factor were reviewed for consistency across all phases.

For example, several of our preliminary factors were dropped as being too difficult to measure against the historical record. These included "messages consistent (or at least progressive) over time" in the realm of strategic communication, and "COIN force employed ID cards/checkpoints for population control." Other factors were changed subtly to make them either easier to assess or more representative of the tenets. For example, "Leaders selected in a manner considered just and fair by the majority of the population" became "Government leaders selected in a manner considered just and fair by the majority of the population in the area of conflict." "COIN forces attempted to secure border(s)" became "Flow of cross-border insurgent support significantly decreased or remained dramatically reduced or largely absent."

In addition to factors derived from specific COIN concepts and inductively revised based on experience with the actual data, we included factors induced from the cases. As we conducted the case studies, the preliminary narratives revealed other factors that appeared to make important contributions to determining case outcomes. After some discussion, we added these inductive factors to our factor list if they could not be explained away through reference to other factors.

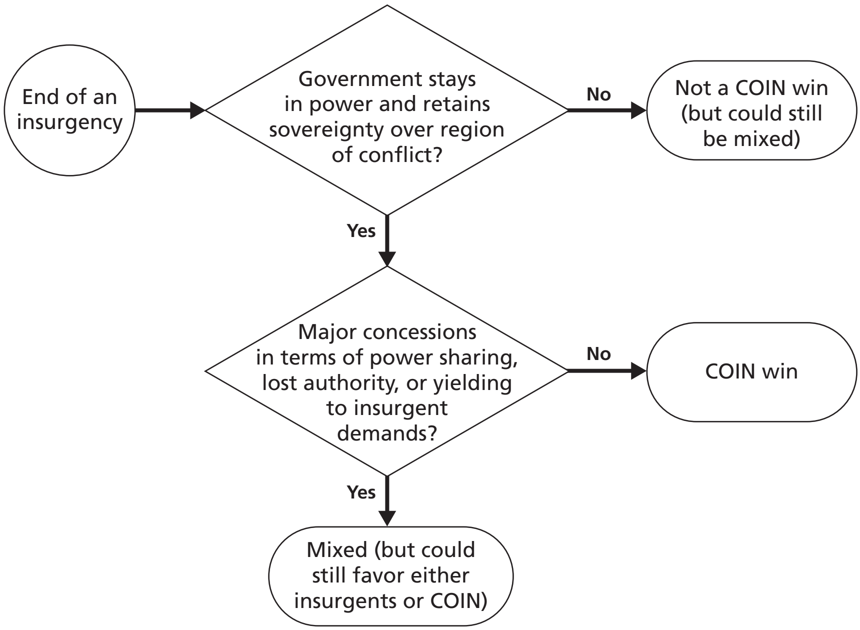
All factors were scored as present or absent (1 or 0) for each case based on the best assessment of the analyst responsible for that case (unless the factor was specified as categorical, in which case the analyst used his or her best judgment to assign the phase to the proper category). To ensure consistency in criteria for evaluating the presence or absence of each factor, the research team met regularly to discuss factor assignments. Each project team member was responsible for a subset of the cases. Each analyst worked on at least ten cases concurrently, so all had ample examples on which to draw to illustrate a point, highlight a challenge to discrimination, or test candidate criteria language. We discussed factors and criteria to ensure shared understanding, and we collectively examined the details of difficult or borderline cases for certain factors. This exchange of concrete examples and counterexamples resulted in either new consensus and understanding of existing criteria or revision to the factor's wording or criteria. A full list of factors scored for each phase of each case appears in Appendix E.

Outcome Assessment

The step that was most critical to the results of the analysis was the assessment of the outcome of each case. Unsurprisingly, since we do not live in a dichotomous world, some of the case outcomes were somewhat ambiguous. Libicki, in the 89 cases from which we started our case selection, had provisional outcomes for each case as assessed by his research team, and many of them were "mixed." While we retained "mixed" outcome as a factor in the data, we knew we wanted a discrete binary outcome for our core analyses. In other words, "mixed" was not good enough. For each case with a mixed outcome, the case analyst made a determination of "mixed, favoring the COIN force" or "mixed, favoring the insurgents." In no case was the outcome so truly ambiguous that the result could not be clearly identified as favoring one party or the other. However, as described in Chapter Three, the outcome for Colombia (La Violencia), though identified as "mixed, favoring insurgents," was determined to be so thoroughly mixed as to render it effectively indeterminate with regard to its utility as a comparative case. It is the only case excluded from the core 59 cases based on its outcome.

To adjudicate unclear case outcomes, we followed the logic illustrated in Figure A.1. First, for each case, we asked whether the government against which the insurgency arose had stayed in power through

Figure A.1
Logic for Assignment of Case Outcomes



SOURCE: Paul, Clarke, and Grill, 2010b, p. xiv, Figure S.2.
RAND *RR291/1-A.1*

the end of the conflict and whether it retained sovereignty over the region of conflict. If insurgents either deposed (or otherwise led to the fall of) the government or won de facto control of a separatist region, then the COIN force did *not* win. If the government remained in power and the country intact, then we further considered whether the government had been forced (or chose) to make major concessions to the insurgents, such as power sharing in government or loss of territory or other sovereign control, or whether it was otherwise forced to yield to insurgent demands. If the government stayed in power, the country remained intact, and no major concessions were granted to the insurgents, then the COIN force unambiguously won. If, however, major concessions were made, then the outcome was mixed. In all cases, what constituted a "major" concession and who (the COIN force or the insurgents) had the better of a mixed outcome was decided at the discretion of the individual case analyst based on the distinct narrative of that case.

p and (1 – *p*)

As noted, virtually all factors were scored as present or absent, 1 or 0, for each case. Some of the factors are described as negations; for example, one factor is "COIN force *not* viewed as an occupying force in the area of conflict." If this factor is scored present (1) for a case, that means that the COIN force was not viewed as an occupying force in the area of conflict. This follows standard practice for dummy or indicator variables and also adheres to English-language conventions regarding double negatives.

Some of the analyses focused on the presence of certain factors, while others focused on the absence of those factors. (Specifically, our analysis of good COIN practices focused on the presence of those good practices, while our analysis of detrimental COIN practices usually identified a poor practice as the absence of an otherwise positive factor.) While leaving the underlying data intact, we avoid double negatives throughout the discussion and presentation of the findings to the extent possible. We do this by invoking the relationship between a probability p and $(1 - p)$. Consider factors in which p is either 1 or 0 (as is the case for all our factor scores): $(1 - p)$ will always be the other of 1 or 0. So, if a case is scored 0 for "COIN force *not* viewed as an occupying force in the area of conflict," that means that it is not *not* viewed as an occupier, which means that it *is* viewed as an occupier. We avoid awkward double negatives by describing the obverse factor rather than the negation of the factor; in this example, we would simply say, "COIN force viewed as an occupying force in the area of conflict," if that were the relationship of interest.

Data Collection

Data for the case studies (both narrative and factor evaluation) came from secondary sources. The analyst assigned to each case thoroughly reviewed the available English-language history and secondary analysis of the conflict for that case. Documentation proved voluminous for some cases (particularly those in Central and South America, but also cases in which Russian or Soviet forces were involved); it was much more sparse for other cases (particularly those in Africa). In all cases, available information was sufficient to meet our data needs.

Phased Data

We initially set out to score factors for the decisive phase of each case. Many of these cases lasted ten or more years and saw many different strategies employed by the government and the insurgents, as well as significant wholesale changes in exogenous factors that could be relevant to the outcome. By focusing on the factors present or absent at or immediately prior to the decisive point in the case, we hoped to capture the conditions that led to the observed outcome. Throughout this discussion, *case data* refers to the data for the decisive phase of the case.

We intentionally sought data for the decisive phase rather than the terminal phase because the two did not match in all cases. In three of the 71 cases, the decisive phase preceded the terminal phase: Baluchistan, Western Sahara, and Nagorno-Karabakh (see details in the accompanying case-study volume, *Paths to Victory: Detailed Insurgency Case Studies*, and that for the previous effort, *Victory Has a Thousand Fathers: Detailed Counterinsurgency Case Studies*).[9] A single example is instructive. The insurgency in Nagorno-Karabakh followed an interesting path. In the initial phase, the Karabakh Armenian insurgency made modest headway against the government. In the second phase, the Russians provided heavy weapons to both sides, but the more disciplined insurgents took advantage of political discord in the government to seize the initiative and occupy and control the majority of the territory in their declared separatist region. In the third and final phase, the COIN force reorganized and put significant pressure on the insurgents, beginning to roll them back with a series of stinging victories. However, before the government could press its advantage, the

[9] Paul, Clarke, Grill, and Dunigan, 2013; Paul, Clarke, and Grill, 2010a.

Russians put irresistible pressure on both sides for an immediate settlement, "freezing" the conflict with the insurgents still in de facto control of much of the territory they sought. Because of this peculiar close to the terminal phase, the second phase became the decisive phase; the factors changed in the third phase and did not have any effect on the outcome.

Due to the kinds of complexity that the Nagorno-Karabakh example illustrates, we ultimately separated each COIN case into one to five phases. While our core analyses still focus on the decisive phase, collecting data for all phases helped us avoid several pitfalls.

First, it seemed like a critical omission to summarize a case in a single row, with factors scored as present or absent that had not been present or absent for the majority of the conflict but were at the point of resolution. Second, those of us with backgrounds in comparative historical narrative research understood the possible importance of sequence in historical outcomes, a possibility we were ignoring by reducing our cases to a single row. The phased record for the whole case accurately reflects the condition of all factors throughout the conflict, not just in the decisive phase.

Identifying phase durations and break points proved to be at least as much art as science. Phases are not uniform in duration. A new phase was declared when the case analyst recognized a significant shift in the COIN approach, in the approach of the insurgents, or in the exogenous conditions of the case that caused changes in the assessment of several factors. Phases were *not* intended to capture microchanges or tight cycles of adaptation and counteradaptation between the insurgents and the COIN force; rather, these were macro-level and sea-change phases. Case analysts had discretion regarding the number of phases and the number of factors that needed to change to constitute a phase change. As with the individual factors, phase breaks were discussed during team meetings to ensure comparability across cases. Secondary analysis of the cases often helped, as other analysts would include periods or phases in their narratives. Similarly, elections resulting in a change in government, or the entrance or exit of an important external participant in the conflict, were often clear indicators of phase change.

Analyses

Using these data, we conducted four different types of analysis. The first was a narrative for each case, presented in the companion volume of 41 new case studies, *Paths to Victory: Detailed Insurgency Case Studies*, and the previously published volume, *Victory Has a Thousand Fathers: Detailed Counterinsurgency Case Studies*, for the 30 cases studied earlier.[10] Full data for all factors for every phase of all 71 cases can be found in the accompanying Microsoft Excel® spreadsheet. The second type was a bivariate analysis of factors or concepts employed in cases or phases. Results from these analyses are presented in Chapter Four. The third was QCA, as described at the beginning of this appendix. The QCA results are presented in Chapter Five, and a detailed presentation of that analysis can be found in Appendix B. Fourth, and finally, was survival analysis of factors that increased or decreased the duration of insurgencies, as well as those that increased or decreased the duration of postconflict peace intervals. Results for these analyses are presented in Chapter Five, with further details in Appendix C.

Narratives

To give context to the raw phased factor data, we developed a brief narrative for each case. Each narrative includes a short summary of the case, a brief summary of each phase, a discussion of the conventional explanations of the case offered in the existing secondary analysis, and a list of distinct factors that were either uncommon but present in that case or wholly unique to that case.

Beyond this, we offer no separate analysis of the individual cases. In fact, one of the most striking findings of this research is that we do not need to discuss any of the distinct features or narrative peculiarities of the individual cases to wholly explain the outcomes. Unlike other research efforts, here, we are not relying on narrative historical methods to reach our conclusions.[11] In fact, our analysis supports the idea

---

10 Paul, Clarke, Grill, and Dunigan, 2013; Paul, Clarke, and Grill, 2010a.

11 For various discussions of narrative historical methods, see Andrew Abbott, "Conceptions of Time and Events in Social Science Methods: Causal and Narrative Approaches,"

that it is a mistake to learn too many "lessons" from a single case, as the peculiarities and distinctions of a single case may obfuscate otherwise critical and enduring relationships between COIN practices and outcomes.

Bivariate Relationships

Our quantitative analysis began by identifying simple bivariate relationships between the various factors and the outcome of the case (or phase).[12] We computed bivariate correlations for all factors and case outcomes and also created 2×2 tables for each factor and the case outcome. We computed these bivariate relationships for all 71 cases, for the 59 core cases, and for the various subsamples identified in Chapter Three. For the reasons discussed in Chapter Three, the core analyses presented address the 59 core cases rather than all 71 cases. These 2×2 tables provided particularly interesting results, especially when the "diagonal" cells contained small values or were 0s, indicating a very strong degree of association between the factor and the outcome.

Table A.1 shows, for example, that in all 11 cases in which the government reduced corruption and or increased good governance, the government prevailed.

As is always the case with bivariate displays, no effort is made to control for the presence or absence of other factors. Thus, while Table A.1 suggests that reducing corruption is a good COIN practice, it tells us nothing about the other things those victorious governments and COIN forces were or were not doing.

Because our cases fully represent our population (we have the complete set of resolved insurgencies from 1944 through 2010), we do not compute inference statistics (e.g., χ^2 tests, p-values) for any of our

---

Historical Methods, Vol. 23, No. 4, 1990; Ronald Aminzade, "Historical Sociology and Time," *Sociological Methods and Research*, Vol. 20, No. 4, 1992; and Robin Stryker, "Beyond History Versus Theory: Strategic Narrative and Sociological Explanation," *Sociological Methods and Research*, Vol. 24, No. 3, 1996.

[12] *Bivariate analysis* denotes consideration of the relationship between two variables. In these analyses, there is always some factor (or stack of factors representing a concept's factors combined into a single factor) considered in relationship to the outcome of the phase or case.

Table A.1
Sample 2x2 Table: Government Corruption Reduced Versus Case Outcome for the 59 Core Cases

		Case Outcome	
		COIN Loss	COIN Win
Government corruption reduced/good governance increased since onset of conflict	Yes	0	11
	No	31	17

analyses. The relationships observed are perfectly representative of the relationships in this population of cases.

Factor Stacks

We also sought to examine the bivariate relationships between the 24 COIN concepts presented in Chapter Four and the phase and case outcomes. Because each concept is represented by more than one factor (see Chapter Four for the detailed breakdown of the factors for each concept), we faced a challenging question: How many of the factors associated with a given COIN concept must have been present in a case before the COIN force is considered to have implemented that concept? Rather than attempting to answer this question in an abstract or arbitrary way, we let the data speak and sought the best empirical cut point for each concept.

For each COIN concept, we created a new factor or variable that was the sum of all the factors tied to that concept and present in a given phase or case. We then chose a threshold value for that sum that maximized the number of COIN wins associated with the concept and minimized the number of COIN losses. Here is a concrete example: Legitimacy of the use of force as a COIN concept is represented in the data by six discrete factors (listed in Chapter Four in the section "Legitimacy"). For each case, we summed these six factors, creating a new variable, "sum of legitimacy of the use force factors." The results are shown in Table A.2.

Table A.2
Sum of Legitimacy of the Use of Force Factors Versus Case Outcome (empirical cut point in red)

		Case Outcome	
		COIN Loss	COIN Win
Sum of legitimacy of the use of force factors	6	0	7
	5	0	3
	4	1	4
	3	2	1
	2	9	3
	1	7	5
	0	12	5

Here, the empirical cut point was identified to be at four or more. Having at least four legitimacy of the use of force factors captures 14 of the COIN wins and excludes all but one of the COIN losses. Thus, we created a single factor to represent legitimacy of the use of force in the analysis: "at least four legitimacy of the use of force factors," which was evaluated as present or absent in each phase of each case, just like all the other factors in the analysis.

We created a "factor stack" for each of the 24 concepts that we tested. These single factor stacks were used to represent each of the concepts in both the bivariate and qualitative comparative analyses. We also used the intermediate stage, the sum of factors, to combine and compare "good" practices and factors with "bad" factors (see Table 5.1 in Chapter Five).

Our decision to let the data speak and identify thresholds for satisfaction criteria for the adherence to certain concepts based on empirically observed cut points (see the discussion in Chapter Four) is open to criticism. One might argue that we should have set a theoretically based standard, either across all concepts (e.g., a threshold of 50 percent or even of 100 percent of an concept's factors must be present to qualify) or based on individual concepts (e.g., How many of these factors or practices do the proponents of an concept suggest are necessary in order to prevail?).

Our decision to use empirically observed cut points is not a conservative one; it shows each concept in its best possible light by maximizing the ability of the factors to predict COIN success versus failure. We do not present the sums of factors for each concept, though they were part of our preliminary analysis. In our defense, for all concepts receiving strong support in our analysis (as listed in Table 4.33), choosing a higher threshold would only increase the predictive power of the single-factor expression of the theory and the outcome. That is, for each supported concept, higher thresholds would exclude case losses, ultimately to the point of perfection. Consider, for example, Table A.2.

Imagine if we had used a higher threshold. For the sake of argument, suppose we had insisted on the presence of more than two-thirds of a concept's factors before considering it to be implemented. For a six-factor stack like the one for legitimacy of the use of force, this would require five or six of the factors to be present. If we used that threshold, then we would conclude that legitimacy of the use of force was present in only ten of the 59 core cases. However, we would also conclude that it perfectly predicted a win every time it was employed. Similar patterns would be observed for all the supported concepts: Fewer cases would get credit for implementing each concept, but each concept would be shown to be even more successful as a predictor of outcome.

Additional Details on the Use of Ragin's Qualitative Comparative Analysis

As indicated at the beginning of this appendix, we structured our data to facilitate the application of Ragin's QCA approach. The construction of crisp-set truth tables requires that all data be binary, hence our efforts to reduce all factors and concepts to present or absent (1 or 0). For the actual analysis, we used Ragin's fsQCA (fuzzy-set QCA) software.[13]

[13] Charles C. Ragin, Kriss A. Drass, and Sean Davey, *Fuzzy-Set/Qualitative Comparative Analysis 2.0*, Tucson, Ariz.: Department of Sociology, University of Arizona, 2006. See also Charles C. Ragin, *User's Guide to Fuzzy-Set/Qualitative Comparative Analysis 2.0*, Tucson, Ariz.: Department of Sociology, University of Arizona, 2006.

We used the crisp-set option in fsQCA to analyze our data.[14] A wholly atheoretical data-mining approach would have encouraged us to take all 289 of our factors, enter them into a truth table, and allow Ragin's software to reduce them to prime implicants using Boolean algebra. While this might have exposed unexpected and interesting patterns in the data, it also would have increased our vulnerability to Type I error.[15] In any event, this proved impossible. A truth table has a number of possible rows equal to 2^n, where n is equal to the number of factors included. Including all 289 factors would have required a table with 2^{289} rows, a mind-blowing matrix size, but, more importantly, a computer-blowing one as well. The current software limited us to the inclusion of no more than 11 factors at a time (so, a truth table of 2^{11} possible combinations, or 2,048 unique rows).

Based on our preliminary bivariate analyses of the case data, we entered composite factors representing the 18 COIN concepts that received strong support at the bivariate level into fsQCA.[16] To ensure that we identified as many of the prime implicant patterns of these 18 composite factors as possible, we ran fsQCA analyses repeatedly, iteratively removing and replacing a factor each time we identified a prime implicant pattern. We iterated through composite elements of each set of prime implicants, slowly removing factors whose role as part of a prime implicant pattern had been explored, until the remaining factors were unable to fully explain the data. Details and results from this analysis can be found in Appendix B.

---

[14] On the distinction between fuzzy sets and crisp sets, see Charles C. Ragin, *Fuzzy-Set Social Science*, Chicago, Ill.: University of Chicago Press, 2000.

[15] A Type I error is rejecting the null hypothesis when the null hypothesis is true—that is, asserting a finding when, in fact, what you have found is strictly the result of chance. This is a frequent problem in data mining. See Egon S. Pearson and Jerzy Neyman, "On the Problem of Two Samples," in Jerzy Neyman and Egon S. Pearson, *Joint Statistical Papers*, Cambridge, UK: Cambridge University Press, [1930] 1967.

[16] Chapter Four reveals that 17 concepts received strong support. One of them, legitimacy, was broken into two different factor stacks to test different aspects of legitimacy. That division was retained for this analysis.

Survival Analysis

Survival analysis is a statistical technique originally developed to answer questions in engineering about time to failure in mechanical systems (usually called "reliability analysis" in engineering) and in epidemiological studies to determine the impact of different treatments for (usually terminal) diseases or afflictions. The technique was broadened in its application in sociology (in which it is called, simply, "event history analysis"), and it is mathematically applicable to any situation in which *Y*, the dependent variable, is *time to event*, whether "event" is the failure of a ball-bearing, the death of a patient, or the end of an insurgency.

We use survival analysis to answer two questions in Chapter Five. First, what factors extend or reduce the duration of insurgencies? Second, once an insurgency has been resolved, what factors extend or reduce the subsequent period of peace (the "peace interval")?

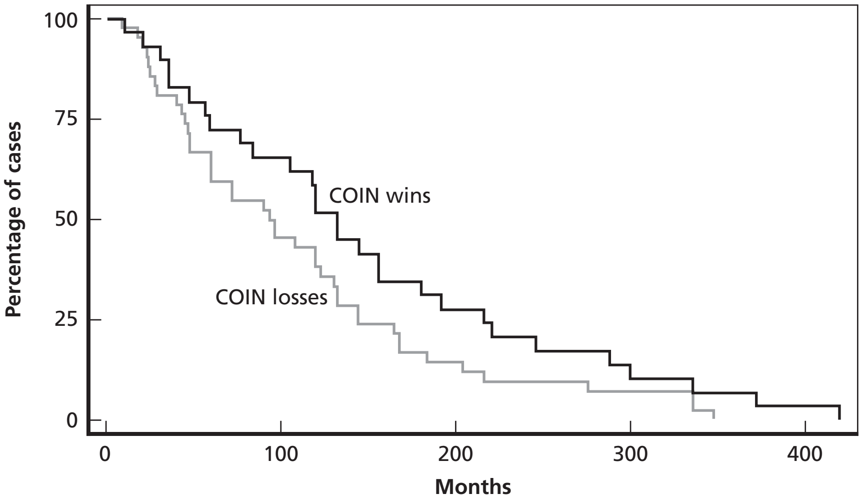
This discussion tries to remain at the general audience level and will not digress into formulae or advanced mathematical discussion. The reader who is interested in greater detail is referred to one of the many textbooks available on survival analysis or event history analysis.[17] The core of the calculations for survival analysis is the *survival function*. The survival function is the probability that the time of event (traditionally, death or failure, but in our analyses the end of an insurgency or the end of the peace interval following an insurgency) is later than some specified time *t* and can be calculated for a population based on an appropriate sample of cases, as is true for most regression-based statistical techniques. Survival functions are most interesting in comparison. Consider, for example, the two survival functions depicted in Figure A.2.

Figure A.2 shows, for all 71 cases, the survival functions for COIN force wins and COIN force losses. The vertical axis of the case reports the proportion of cases remaining in the analysis ("surviving," or, in this instance, still having an insurgency), and the horizontal axis reports the time in months from the beginning of each insurgency.[18]

[17] See, for example StatSoft, "Survival/Failure Time Analysis," web page, undated.

[18] Note that the beginning of each insurgency is time = 0 for that case, regardless of the actual historical year in which the cases began. When we say that two or more insurgencies

Figure A.2
Survival Functions for COIN Wins and Losses (n = 71)



RAND *RR291/1-A.2*

Points on the curve can be interpreted as follows. Visually find a point on the lower curve (labeled "lost cases"), perhaps the point where the curve reads "50 percent" on the vertical axis and is a few millimeters shy of "100" on the horizontal axis. The vertical axis reading of 50 percent indicates that half of the lost cases (in the full 71 cases, 42 cases were COIN losses, so 21 cases) survived to almost 100 months (actually 96 months, or eight years).

The curves representing the survival functions follow a "stepped" pattern because of the relatively modest number of discrete cases represented. Each step "down" represents the exact duration of one or more insurgencies (if more than one, a bigger step down, it means that multiple insurgencies had the same duration).

Overall, Figure A.2 shows several interesting patterns. First, on average, COIN wins had longer durations than COIN losses. This can be seen by the fact that the won cases' curve is always above the lost

---

ended at the same time, we mean, for example, that they all ended after 72 months, not that they all ended on December 3, 1971. All times are duration times and are relative to the start times of the individual cases.

cases' curve. (It would also be interesting if they crossed; that would mean that one type of case tended to last longer up to a certain point. We might see these curves cross like that if, for example, time really did favor the insurgents.) Second, Figure A.2 shows us that, for wins and losses, the distribution of durations is not uniform. If the distribution were uniform, the curves would be straighter and more closely aligned with the diagonal on the figure. Instead, both curves are below the diagonal, and both curves have relatively longer (vertical) steps toward the lower right of the figure, indicating that some cases lasted a great deal longer than others, disproportionately longer. Of course, Figure 5.1 in Chapter Five also contains that information.

Comparing survival functions for different groups is at the heart of survival analysis, but looking at graphs of the survival functions is not always the most informative way to make these comparisons. Often more useful is the *hazard ratio*. The hazard function is slightly more complex than, though derivative of, the survival function. The hazard function is the event rate (as determined by the survival function) at a give time *t*, conditional on survival to that time *t*. (So, basically, the hazard is the calculated risk of experiencing the event after any point in time, assuming you've "survived" at least that long.) The *hazard ratio* is an overall comparison of the hazard functions of two groups, usually a group defined by the presence of one or more factors or variable as compared to the rest of the data. The hazard ratio is reported as a ratio, so if the hazard ratio is equal to 1, it means that the two conditions have equal hazard of experiencing the event (of the insurgency ending); if the hazard ratio for a group is positive, it means that it is at greater risk of the event (the insurgency is likely to end sooner, so duration is likely to be shorter); and if the hazard ratio is negative, it means that the group is at less risk of the event (so duration is likely to be longer). For survival analysis of duration, we sought factors with positive hazard ratios, as they are correlated with decreased durations. When examining peace intervals, however, we sought factors with negative hazard ratios—that is, those that decrease the likelihood of experiencing the event (in this instance, the end of peace), relative to cases without the factor—because longer peace intervals are preferable.

Data for survival analysis can be structured in a variety of ways. An analyst can calculate a hazard function for cases in single rows of data, where a single event time is noted and all other factors in the model are assumed to be constant from time = 0 to time of event, or the model can be calculated with multiple rows representing a single case, where only one of those rows ends with the event and the other rows represent blocks of time (or, in our study, phases) in which the event of interest did not occur but various other factors might have changed (perhaps the administration of some kind of treatment, or the presence or absence of one or more COIN concepts). Because our phased data allow us to identify blocks of time in which certain factors were present or absent, we used survival analysis techniques that were appropriate for individual cases with multiple sets of conditions prior to event. Note that while we report the total duration of each phase in months, we do not report the within-phase start time for each factor that changed its state (went from present to absent or absent to present) during a phase. For all survival analyses, we assumed that factors changed state right at the start of a phase and held only one value (present or absent) for the entire duration of the phase. This slightly decreased the precision of our results and weakened the relationships of beneficial factors to duration,[19] so it is a conservative assumption.

One of the analytical challenges that survival analysis often faces is referred to as "right-censoring," the inclusion of cases that do not experience the event during the period recorded by the data. This could be a patient who survives past the end of the study, a machine that continues to operate through the entire observation period, or an insurgency that is not resolved. In our analyses of duration, there is no issue with right-censored data. To be included in our data set, a case must

[19] Imagine a factor that strongly increases the hazard of a conflict ending when present. It became present at some point during a 24-month phase. Regardless of when it actually became present, we assumed that it was present from the beginning. Now, imagine that it actually became present and began to exert its strong influence toward ending the conflict, at month 12 of the phase. That would mean that we had inaccurately attributed the factor as present for 12 months during which no such influence was being exerted, diminishing the calculated strength we attributed to the factor. If we still find that it had a strong impact on hazard, then it must be strong indeed.

have concluded, so when the event of interest is "end of insurgency," that event always occurs in the data. However, the data on peace intervals are frequently right-censored. A peace interval is the time between the end of an insurgency and the start of the next one; many of our cases were not followed by subsequent insurgencies, so they have long peace-intervals that are right-censored; we do not know when (or if) those peace intervals will end. Right-censored data are a common and understood challenge in survival analysis, and the software we used to conduct this analysis is appropriate for the right-censored data.

All survival analyses conducted as part of this research were produced using STATA® and report estimates from Cox proportional hazard models calculated using the robust variance estimator (to correct for the fact that some countries hosted multiple insurgency cases and are thus not fully independent of each other).[20]

---

[20] On Cox regression, see D. R. Cox, "Regression Models and Life-Tables," *Journal of the Royal Statistical Society*, Series B, Vol. 34, No. 2, 1972. On robust variance correction, see D. Y. Lin and L. J. Wei, "The Robust Inference for the Cox Proportional Hazards Model," *Journal of the American Statistical Association*, Vol. 84, No. 408, December 1989.