Integrating Qualitative and Quantitative Methods

James D. Fearon
David D. Laitin

Prepared for the Oxford Handbook of Political Methodology

Almost by definition, a single case study is a poor method for establishing whether or what empirical regularities exist across cases.[1] To ascertain whether some interesting pattern, or relationship between variables, obtains, the best approach is normally to identify the largest feasible sample of cases relevant to the hypothesis or research question, then to code cases on the variables of interest, and then to assess whether and what sort of patterns or associations appear in the data.

However, with observational data (data not generated by random assignment of presumed causal factors), mere patterns or associations are not enough to allow us to draw inferences about what causes what. The standard approach in large-N statistical work in political and other social sciences is to accompany the presentation of associations (often in the form of regression results) with arguments about (a) why the reader should believe that the variation in an independent variable could cause variation in the dependent variable, and (b) why the reader should believe that the association observed in the data is not due to the independent variable happening to vary with some other, actually causal factor. The latter is usually done by adding "control" variables to the regression model, and arguing that one has not omitted important factors that are correlated with the independent variables of interest.[2]

The arguments for (a) and (b) amount to a sort of *story* the researcher tells about the associations observed in the regression results. (The arguments, particularly those for (a), are often referred to as a "theory."[3]) To some extent these stories can be evaluated as to whether they are deductively valid, that is, whether the conclusions do indeed follow from the premises, and whether the arguments are consistent. For example, it may be that

---

[1] Of course, a single "discrepant" case can disprove a hypothesis asserting a deterministic relationship between two things. But deterministic relationships between variables of interest are at best rare in social science. Eckstein (1975) argues that case studies are, under some conditions, the most efficient method for theory development. But his conditions require theories that make deterministic (i.e. not probabilistic) predictions.

[2] With regression analysis, there may be additional assumptions that should be justified in order to warrant a causal inference, such as correct functional form and additional assumptions about error variances.

[3] As Wagner (2007) observes, partisans of quantitative social science sometimes seem to mistake a regression equation itself for a theory.

the argument for why one independent variable matters contradicts the argument made on behalf of some other variable. Or it may be that an argument for a particular independent variable is internally inconsistent, confused, or doesn't follow from the premises on closer inspection. However, while a good analysis tells a valid and internally consistent story about the observed correlations, there may be multiple possible consistent stories about particular independent variables, and in general the reader may not know how much weight to put on the researcher's interpretation. Is the researcher's story capturing "what is really going on" in the cases to generate the observed patterns, or is something else driving the results?

At this point case studies can be extremely useful as a method for assessing whether arguments proposed to explain empirical regularities are plausible. One selects particular cases and examines them in greater depth than was required to code values on the outcome and explanatory variables of interest. When the "cases" are sequences of events in different countries or regions, or years in a particular country, or bills proposed in a legislature, and so on, the case study will entail a narrative account of what led to the outcome, including an assessment of what role the proposed causal factors played. In these narratives one typically uses additional data about the beliefs, intentions, considerations, and reasoning of the people who made the choices that produced the outcome, in order to test whether the "higher level" general story told about many cases is discernible in particular, concrete cases. One can also ask if there were important factors omitted from the large-N analysis that might in fact be driving the results. Finally, one can use the finer-grained analysis possible in a narrative account to ask about the validity and accuracy of measures being used in the large-N analysis.

For these several reasons, so-called "multi-method" research has become increasingly popular in political science in recent years, especially in Comparative Politics (Laitin 1998; Mares 2003; Iversen 1999; Boix 1998) and International Relations (Huth 1998; Martin 1994; Schultz 2001; Goemans 2000; Stone 2002; Walters 2001; Fortna 2004; Mansfield and Snyder 2006; Collier and Sambanis 2005; and Doyle and Sambanis 2006).[4] Done well, multi-method research combines the strength of large-N designs for identifying empirical regularities and patterns, and the strength of case studies for revealing the causal mechanisms that give rise to political outcomes of interest.

An important but neglected problem for this research approach is the question of how to choose the cases for deeper investigation. Most work in this vein adopts the implicit criterion of choosing cases that support (or can be argued to support) the researcher's interpretation of the regression results. This criterion need not yield worthless results, since knowing that there are at least *some* cases that show good evidence of the causal mechanisms proposed by the researcher is something. But "cherry picking" by the researcher, or even the perception of cherry picking when it did not

---

[4] . To our knowledge survey researchers in American politics do not normally do in-depth, unstructured interviews with or ethnographies of particular respondents to assess their theories based on interpreting regression coefficients. In Congressional research, ethnographies produced by scholars such as Richard Fenno inspire quantitative work but rarely combine methods with the goal of making better causal inferences.

occur, will tend to undermine a reader's confidence that the case study part of the design demonstrates that the researcher's causal story is on target.

In this article we propose that choosing cases for closer study *at random* is a compelling complement in multi-method research to large-N statistical methods in its ability to assess regularities and specify causal mechanisms. We discuss the advantages of random selection (or random selection within strata) for case studies, as well as problems with other possible criteria. These include choosing cases "on the regression line" that appear to fit the researcher's theory well; cases "off the regression line" that do not; "hard" or "critical" cases that are allegedly "tough tests" for the theory to pass; and choosing cases that have particular values on the explanatory factor of interest.

We illustrate what we will call the "random narratives" approach with work in progress we have been doing on the causes of civil war.[5] In Fearon and Laitin 2003, we report the main results of a cross-national statistical analysis of factors that distinguish countries that have had civil war onsets in the period 1945 to 1999. On the basis of these findings, theoretical arguments, and prior, unsystematic reading of diverse cases, we proposed a story about how to interpret why certain factors (such as low per capita income and high country population) are strongly related to civil war risk, whereas other factors (such as ethnic diversity, autocracy and broad grievances) are not, once one controls for level of economic development. In order to assess this story we randomly selected 25 countries, stratified by region and whether or not the country had at least one civil war, and undertook narrative accounts of these countries' civil war experience (or lack thereof) using secondary sources.

In this chapter, we first summarize the findings of our statistical analysis of civil war onsets. We then in section 2 look more carefully at different criteria for choosing which narratives to tell. In section 3, we discuss a method for structuring narratives that is complementary to the statistical work. In section 4 we illustrate in light of our narrative findings the incompleteness of the statistical models we initially ran. In section 5, we highlight one narrative as an example of its potential yield. In the conclusion, we underline some surprises and advantages of the random narrative approach.

1. Statistical Results

Cross-national statistically-based research by us and several other researchers has tended to find little or no support for two well-entrenched theories of civil war onset. First, our data show that by most measures of broad societal grievance – for example, lack of democracy, lack of religious or linguistic rights, or economic inequality – knowing the *level* of grievances in a country does not help differentiate countries susceptible to a civil war from those that are not.[6] Second, our data show that measures of

---

[5] . Two examples of the random narratives are available in James Fearon and David Laitin (2005).

[6] . This is *not* to say that if a state increases the level of grievance for a set of its citizens it can't provoke an insurgency. It may be that some states aggrieve minority groups as much as they can get away with, but some groups will tolerate higher levels of abuse (perhaps due to their weakness). Therefore increasing

cultural divides (the level of ethnic heterogeneity or the degree of cultural distance) do not help differentiate countries susceptible to a civil war from those that are not, once one controls for level of economic development.

In their stead, we have advanced an argument that points to the conditions that favor insurgency, a technology of military conflict characterized by small, lightly armed bands practicing guerrilla warfare from rural base areas. This perspective ties together a set of variables that correlate with civil war onset. Our interpretation of all of them is that they point to the relative incapacity of a state to quell insurgencies, which may begin at random and be "selected" for growth in countries with favorable conditions, or may be actively encouraged by signs of state weakness. The key variables that are significant and robust in our statistical models are listed below.

* Per capita income – we argue that low per capita income matters primarily because it marks states lacking in financial, bureaucratic, military and police capability;

* Mountainous terrain – we interpret high degrees of mountainous terrain in a country as a tactical advantage to potential insurgents for hiding from government forces;

* Population – large populations require more layers of principals and agents to govern, making it harder for the regime to competently monitor, police, and respond to security threats at the village level;

* Oil – oil increases per capita income and government revenues and so works against insurgency, but controlling for the level of income we expect that oil producers have weaker governmental institutions, inasmuch as oil revenues make it unnecessary to develop intrusive tax bureaus that need to track individual citizens. Oil revenues can also increase the "prize" value of capturing the state or a region;

* Instability – we interpret rapid shifts in the regime type (a two or more shift in a single year in the Polity score for democracy) as a proxy for weak or weakening central state institutions;

* New state – in the immediate aftermath of independence, due to withdrawal of colonial forces before indigenous institutions have taken root, states are especially fragile. This provides an opportunity for challengers. They may fear that the leaders of the new state cannot commit not to exploit their region or group in the future after the government consolidates. Or they may see that the government cannot commit to provide state benefits in the future worth as much as their short-run expected value for trying to seize power by force now. Unable to hold as credible such commitments, insurgents can take advantage of a window of opportunity to seek secession or capture of the state before the state develops its national army;

---

grievances can lead to insurgency even if levels of grievance across countries vary without implications for civil war onsets.

* Anocracy – Regimes that mix autocracy with some democratic features (such as a legislature or partly competitive national elections) suggest the presence of political conflict that weakens its ability to counter an internal threat.

A multivariate analysis thereby helped us to address the question of what factors correlate with higher likelihood of civil war onset. Our selection of variables was motivated by reading the literature on civil war, reading about specific cases, and thinking theoretically about the literature and cases using game theoretic tools (see for example Fearon 1998, Laitin 1995, Fearon and Laitin 1999, Fearon 2004, Fearon 2008). We continue to work on formal models of civil war-related interactions in an effort to clarify and deepen the informal arguments we have proposed (e.g., Fearon and Laitin 2007).

Lacking in both approaches, however, is a clear empirical answer as to whether the variables in our statistical and theoretical arguments are actually "doing the work" in raising a country's susceptibility to a civil war onset. This is where case narratives can play an especially valuable role.

2. Choosing Narratives

But which narratives to tell? In statistical work, there are some methodological standards for case selection and analysis. In formal and particularly in game-theoretic analysis, there are well-developed rules and standards for what constitutes a well-posed model and proper analysis of it. There is no intellectual consensus, however, on how to choose cases for narrative exposition in a research design that combines statistical and case study evidence.[7] We consider several criteria that have been used in practice or that might be argued to be desirable.

(a) "Good cases" for the researcher's theory

In practice, perhaps the most common approach is for the researcher to choose "good cases" that nicely illustrate the causal mechanisms that the researcher proposed to explain the broader empirical regularities found in the statistical analysis. This is not an entirely worthless procedure in terms of producing evidence. It tells us that there exist at least some cases showing plausible evidence of the causal mechanisms hypothesized by the researcher. But selection bias is obviously a problem here. "Good cases" have been cherry-picked, so it is not clear that they convey much information about the importance of the proposed causal mechanism in explaining the observed patterns. Even worse, what if these are the very cases that led the researcher to propose the causal mechanisms that she hypothesized to explain the statistical patterns in the first place?

---

[7] Nor is there a consensus on what, given the choice of a particular case, makes for a methodologically strong narrative. Still, most would agree that factual errors and tendentious interpretation make for a bad narrative or case study. Alex George has been a trail blazer in setting criteria for drawing causal inferences from narrative accounts. For his latest (and alas, his last) statement on this, see George and Bennett (2005).

In one version of "good case" selection, the researcher tells the reader that one or more of the cases selected for narrative exposition are, in fact, "hard cases" for her theory. It is then argued that received theories predict that the causal mechanism advanced by the author would be especially unlikely to be found in these "hard cases." Invariably, however, the historical narrative shows that the author's proposed causal mechanism is at play or that it trumps other factors stressed by existing theories ("surprisingly").[8] Selection bias is again a major concern. If the reader's prior belief put a great deal of weight on the existing theories, then there may be some evidentiary value (actual surprise) in learning that in at least one case, the author's favored causal mechanism can be argued to have been as or more important. But the reader can also be confident that the researcher would not present a developed "hard case" narrative unless it could be rendered in a way that seemed to support the claims of the researcher's preferred causal mechanism. So readers may be forgiven for thinking that talk of "hard cases" may be as much a rhetorical as a defensible methodological strategy.

Another version of "good case" selection occurs when the researcher selects "easy cases" for the preferred theory. For example, Schultz (2001) chooses for narrative several international crises involving Britain, in part because Britain's Westminster system most closely approximates the way that his theoretical model of crisis bargaining represents domestic politics, and from which Schultz drew his hypotheses about the role of opposition signaling in international disputes. This makes it easier to evaluate whether the causal mechanism at work in the model is at work in the cases. One downside, however, is that we have less information about whether Schultz's mechanism is more general than the sample of cases that led him to his model in the first place.

(b) Convenience samples

Another common approach in practice is to choose cases for narrative exposition that are relatively easy for the analyst to research, due to language skills or data availability, in what is sometimes referred to as a convenience sample. This procedure may be justified in some circumstances. The researcher will have to adjudicate trade offs between the risk of selecting non-representative cases, the accuracy of the narratives (for example, the validity and reliability of the measurement of the variables in question), and the number of cases that can be studied in depth. Representativeness will often be a significant problem for this approach, however, since the cases that are easy for the analyst to research will often be systematically unrepresentative on important variables. For example, in cross-national studies it is normally easier to find the sort of detailed information necessary for a good political narrative for wealthier countries, or for poor countries that happen to have attracted a lot of OECD attention and money for strategic or other reasons.

(c) Selecting on variation in the dependent or independent variables

---

[8] Though they may exist, we are not aware on any published paper or book in which the author asserts that X is a hard case for the preferred theory, and then finds that the case does not support the theory.

In principle one might select cases for narrative on the basis of values of the dependent variable, on values of an important independent variable, or on some kind of joint variation in both. For example, if the thing to be explained is the occurrence or absence of a phenomenon like democracy, war, rapid economic growth, or post-conflict peace, the researcher might select for close study a few cases where the phenomenon occurred and a few where it did not. Once again, though, we run into a problem of representativeness. If one is selecting a few cases from a larger set, why this one and not another? Why shouldn't the reader be suspicious about selection of "good cases" if no explanation is given for the choice? If an explanation is given and it amounts to convenience sampling, don't we still need to worry about representativeness? We can probably learn something about the proposed theoretical story from these cases and the contrast between them, but are we maximizing the amount? The same concerns apply for a small sample of cases selected on variation in some independent variable.

Another possibility here would be to select cases that show some particular combination on both the dependent variable and on independent variables of interest. For example, in their large-N analysis Mansfield and Snyder (2005) find that democratizing states with "concentrated" political power have been more likely to initiate interstate wars. They explain this with an argument about the incentives for threatened authoritarian elites to use nationalism and conflict for diversionary purposes. For narrative accounts they choose cases of democratizing states that initiated wars, which they call "easiest" for their theory (p. 169), presumably because these exhibit war and democratization together. As they argue, if they were not able to make a plausible case for the operation of their preferred mechanism in these cases, this would strongly undermine their theoretical claims.

Mansfield and Snyder look at all ten countries that were coded as democratizing when they initiated one or more wars, which protects them against the concern that they might have intentionally or unintentionally cherry-picked within this set. And intuitively, it seems plausible that if one's mechanism links the presence of X to phenomenon Y via such-and-such steps, one can learn more empirically about whether the mechanism matters as hypothesized from cases where X and Y occurred than from cases where one or the other did not. If the empirical question is "How often did democratization lead to war by the specific sequence of events that we have proposed?", then these are obviously the only cases one needs to consult. So here is a strong rationale for selecting cases that are not only "on the regression line," but that show a particular combination of values on Y and X.

If a causal mechanism implies specific sequences of events for more than one value of an independent variable, then the same reasoning leads to the suggestion that cases be sampled that are "on the regression line." For example, we proposed that per capita income proxies for several aspects of a state's capability to conduct effective counterinsurgency, relative to insurgent groups' ability to survive. Thus we would expect to find that in rich states nascent insurgent groups are detected and easily crushed by police (or they stay at the level of small, not very effective terrorist groups), while in poor states we should find would-be insurgent groups surviving and growing due to the states'

incompetence (e.g., indiscriminate counterinsurgency) and virtual absence from parts of their territory. One could try to evaluate how much of the empirical relationship between income and civil war this mechanism explains by selecting for narratives poor countries that fell into civil war and rich countries that did not.

Of course, selecting on Y and X in this manner still faces the difficulty of *which* cases among those on the regression line. If it is not feasible to write narratives of all such cases, as Mansfield and Snyder did, we still face the problem of cherry-picking, or appearance of selection bias.

Moreover, there are good reasons to think that cases that are *off* the regression line might hold important, if different, information about the mechanisms in question as well. First, recall that the fundamental threat to causal inference in non-experimental settings is the risk that there are other causes of Y that happen to be correlated with the proposed cause X. Cases off the regression line are more likely to show these other causal mechanisms at work. Seeing what they are and how they work should increase the ability of the researcher to say whether the observed relationship with X is the result of omitted variable bias. In addition, identifying other causes of the phenomenon in particular cases can lead to new hypotheses about general patterns and explanations, to be evaluated in subsequent rounds of research.

Second, narratives of cases off the regression line improve the researchers' chances of understanding why the proposed causal mechanism sometimes fails to operate as theoretically expected. For example, democratization does not always (or in fact, all that often) lead to interstate war – why not? Were authoritarian elites insufficiently threatened in these cases? Were they threatened but judged that there was no good opportunity for diversionary war? What additional variables determine each of these steps? If we accounted for these as well, would the original relationship hold up? In Snyder and Mansfield's study, answering such questions would require narratives of cases of democratization with concentrated political power that failed to produce interstate war.

Third, cases may fall off the regression not only because of the influence of unmeasured factors, but due to measurement error in Y or X. Particularly with cross-national data, large-N statistical analyses must often employ crude indicators that can be coded for a large number of cases in a reasonable amount of time. Case narratives can then be used to make rough estimates of the validity and reliability of the large-N indicators, and to estimate how often measurement errors are masking instances where the proposed causal mechanism actually did work as predicted, or where it did not work as predicted but mistakenly got credit in the statistical analysis.

(d) Selecting Randomly

These several considerations and the general concern about selection bias lead us to propose that *randomly choosing a (relatively) small number of cases for narrative analysis will often be a desirable strategy in "multi-method" research.*

With random selection, the investigator is asked to write narratives for cases that were chosen *for* him or her by a random number generator. The investigator goes from the first case on a list ordered by a random series to some number down the list, in each case using the narrative to ask about what connects or fails to connect the statistically significant independent variables to the coded value on the dependent variable. Most importantly, the researcher is protected against the risks caused by (known or unknown) systematic bias in case selection. In addition, with cases both on and off the regression line, the researcher can ask both about whether the proposed causal mechanisms operated in the cases on the line and about why and whether they failed to operate in cases off the line. If there were missing variables previously unexamined that would have improved the initial statistical analysis, they are probably more likely to be found in cases forced upon the investigator than in cases she or he chose, and the investigator has an unbiased, if small, sample of them. Finally, the researcher has the opportunity to estimate the impact of measurement error both in cases that are well predicted by the statistical model and those that are not.

Of course, there are downsides to random selection as well, most resulting from possible trade offs against the quality and total number of narratives that can be produced. The researcher may be able to write more and better quality narratives for cases that she already knows well or for which she has language or other skills. On the other hand, this may itself argue against selecting such cases, since it is more likely that they were consciously or unconsciously already used to generate the theoretical arguments that are being tested. Thus, random cases can have the virtue of serving as out-of-sample tests. In addition, case studies of unfamiliar or obscure events may gain from a fresh reading of the standard literature about a country with an eye to how much mileage can be gotten in understanding outcomes through a special focus on significant independent variables validated from cross-national analysis.

Rather than choosing purely at random, it may be more efficient to stratify on particular variables. For example, if one is interested in evaluating whether one's theoretical account for a particular variable in a cross-sectional study is plausible, then it makes sense to sample cases that have a range of values on this particular variable. Or, as we discuss below, it makes sense to stratify on certain variables, for instance, region in a cross-national study, to avoid random selection yielding lots of (say) Eastern European cases but almost no Latin American cases. Still, short of doing case studies for every data point, random selection within strata or values of an independent variable will be warranted for the same reasons just advanced.

To construct the sample for our narrative dataset, we took a random sample of countries stratified by region and by whether the country had experienced a civil war in the period under study (1945-1999). The rationale for stratifying by region was to ensure an even distribution across a factor that is correlated with common historical experience, culture, religion, and level of economic development.[9] We distinguished between "war"

---

[9] If we had not stratified by region, there was a reasonable probability that at least one region would have been significantly underrepresented, and another overrepresented. Since there are so many common and

9

and "no war" countries for a different reason. We initially expected that there was more to be learned by studying a country that had an outbreak of war at some time than one that never did, because a "war country" has periods of both peace and war (variation), whereas a "no war country" has only peace. There is certainly information in the "no war" cases, and we thought it would be wrong to exclude them entirely. But we wanted to make possible the over sampling of countries that experienced a transition from peace to war, as this provides within country variation on the dependent variable in which, in effect, a great many country-specific factors are controlled for.

While this expectation was to some extent borne out, as we note below we were surprised by how theoretically informative and empirically interesting were the narratives for countries that never had a civil war in the period under study.[10]

3. Structuring the Narratives

From our random selection of cases, we created a graph of predicted probabilities of civil onset by year for all chosen countries. The predicted probabilities were generated using the following (logit) model, using coefficients from the estimations using the data discussed in Fearon and Laitin (2003):

Log odds of onset at time t = b0 + b1*Prior war + b2*Per cap Income_t-1 + b3*log(Population_t-1) + b4*log(%Mountainous) + b5*Noncontiguous + b6*Oil + b7*New state + b8*Instability in prior 3 years + b9*Anocracy_t-1.

In other words, we include variables found to be statistically and substantively significant based on that analysis.[11] In generating the predicted probabilities for a given country, we estimate the model *without the data for that country*, so that the experience of the country in question is not being used to shape the predictions for that country.[12] In generating the predicted probabilities, we set "prior war" (which is "1" if there was a civil war in progress in the prior year and zero otherwise) to zero for every year in the country's history, since we did not want to use actual war experience to help predict subsequent war experience. We place a tic on the *x* axis if in fact there was an ongoing civil war in that country for the given year.

---

distinguishing features of the politics and economics of states with "regions" as conventionally described, we wanted to have a better chance of distinguishing the impact of our independent variables from unmeasured, region-specific factors.

[10] . For countries such as Japan and the United States – both in the dataset – we looked for proto-insurgencies (e.g. the Zengakuren protests against the Narita airport in Japan; the Aryan Nation militias at Ruby Ridge) to explain in terms of the model how and why they were successfully marginalized.

[11] . Noncontiguous territory – not fully justified statistically -- was added to the model for the country graphs. Its effects were minor.

[12] We drop all observations for the country, which can be up to 55, depending on how many years the country has been independent since 1945. Changes in predicted probabilities by this procedure were small, giving us confidence that our results did not turn on any single country. Of course, the country's experience has very indirect influence, since it was used in the earlier data analyses that led to this particular model specification. We are merely trying to avoid saying, in effect, "wow the model does great here" if part of the reason is that the model is reflecting the experience of a particularly influential case.

The graph below (Figure 1) illustrates the case of Algeria. Accompanying this graph (Table 1) are data on the key variables in comparison to the mean values of the region and the world. Throughout its independent history, Algeria has had a higher susceptibility (.040) than the mean country of the world (.017) and the region as well (.016). We can see that there were two civil war onsets in years that our model predicted heightened susceptibility. Consistent with many other cases, the first civil war coincided with Algeria being a new state and the second civil war coincided with a period of anocracy and instability. Further, there is only one apparent "false positive" in the sense of a sharp rise in the estimated probability of civil war when no war occurred, in 1996-99; and this is fairly excusable given that it occurs during a war in progress, so inclusion of the variable "prior war" in the predictive model would have greatly reduced this "spike."

4. The Incompleteness of Statistical Models

How to interpret this graph and how to bring narrative evidence to bear? A first potential concern is that the predicted probabilities from the model are typically very small. Indeed, some 90% of the predicted probabilities fall between .0004 and .05. This is probably as it should be – or as it has to be – for two reasons. First, measured by year of onset, civil war starts are extremely rare. We have only about 127 onsets in nearly seven thousand country years. So, not conditioning on any other factors, the probability of a civil war starting in a randomly selected country year in this period was only about .017. Predicting civil war onset in a given country year from factors that can be coded across a large sample of countries and years is a bit like trying to find a needle in a haystack.

Second, it is virtually impossible to identify factors that are codable for a broad cross-section of countries and that can do an excellent job of predicting in *which year* a civil war will break out, if any. The random narratives reinforced our prior belief that a great deal of essentially random historical contingency is involved in determining whether and exactly when a country will "get" a civil war. Bad luck, idiosyncratic choices by leaders, complicated and ex ante unpredictable political and social interactions may all factor into explaining why a particular civil war started at a particular time, or even at all, in a particular country. It is a historian's project, and an admirable project at that, to try to understand such particularities and idiosyncrasies for particular cases. Our social science project, implausible though it may be, is to try to identify some factors and mechanisms that do "travel" across countries and years, raising the risk of civil war onset in a consistent and appreciable manner. This is a difficult task and we do not expect that even in the best case it will be possible to come anywhere close to perfect ex ante prediction. (Arguably, we could only do this if the participants themselves could do even better, given that they always have access to much more relevant case-specific information. But many examples suggest that it is quite rare for people and even politicians in a country to be able to forecast with confidence, a full year in advance, that a civil war will begin.)

Several of the key explanatory variables in our statistical model do not change much over time within countries. These variables are mountainousness, per capita income, population, and to some extent oil production. This means that while they may contribute a lot to explaining variation in the propensity for civil war outbreaks *across* countries, they are seriously handicapped for explaining in which year a war will start. The four variables that do (or can) change sharply from year to year – Prior war, New state, Instability, and Anocracy – are all quite crude measures. They are relatively easily coded across a broad cross-section of countries and years, but they do not condition in any sophisticated way on political circumstances or events occurring in the country. While they are statistically and substantively significant as predictors of onset, none of them is diagnostic in the sense that they have empirically been followed by onset with a near certainty. Thus, even when several of these time-varying variables are "on," the predicted probability of civil war breaking out in the very next year may only reach .2 or .3.

So, if one had to bet on civil war starting in a particular year using the model's predicted probabilities, even the country years with some of the highest estimated chances of onset are more likely than not to remain peaceful. The highest probability in the dataset is .51 for Indonesia in 1949 and 1950. After that is Pakistan in 1947 and 1948, with a probability of onset at .35. For all but two cases in the entire dataset, for any given country/year, a betting person should bet against a civil war onset occurring that year.

How then to interpret year-to-year changes in predicted probabilities for a given country, like those seen in Figure 1? Relatively large jumps or falls in the graph (as seen, for example, for 1961-62, 1989-90, or 1996-97) correspond to changes on one or more of the sharply time varying explanatory factors. For each such case we can ask (1) if the change coincides with or shortly precedes a civil war onset, do we see a causal link from the independent variable to the outcome?; (2) if there is no onset, do we see signs of strife or other indications of increased conflict that might have become a civil war (and which appear causally related to the change on the independent variable)? And if so, why didn't the violence escalate to the level of civil war?; and (3) if there is no onset and no sign of increased violent conflict, why not? For all these cases we can and should also ask if we have miscoded (measured) the dependent variable or an independent variable, giving too much or too little credit to the model and the theory.

Similarly, for years in which we see an onset but no change in predicted probabilities, what occasioned the onset? Is there some new explanatory factor evident that could be coded across cases? Are there miscodings of variables in the model? If the predicted probability of conflict from the model is high on average (relative to the mean for the region or the world), do we see evidence linking any of the slow-moving variables like income or population and the outcome?

In more statistical terms, one way to interpret the graph of predicted onset probabilities for a given country is to interpret the statistical model as a model of an underlying, unobserved *propensity* for civil war in a given country year. The logit transformation scales the estimated propensity index into the [0, 1] interval as predicted

probabilities. With narrative analysis we can try to go beyond assessing the fit of the model by looking at 0/1 outcome predictions, as in the standard quantitative approach to binary outcomes. Instead, to some extent we may be able to assess the model's fit by comparing changes in the predicted propensity of civil war to changes in the actual propensity as judged from the narrative evidence.

Despite difficulties in interpreting blunt variables in case analysis, our narratives of particular countries provide a useful complement and comparison to the statistical model. Before summarizing why they are useful, we will draw from the Algerian narrative to give a specific example of this approach.

5. Learning from the Narratives

In the case of Algeria's two civil wars, practitioners in the field of comparative politics committed only to the statistical approach might see the nice fit between model and the real world would and then let sleeping dogs lie. Algeria, with its poverty, its oil, its large population, and its mountains was a likely candidate for civil war. This was especially the case for two periods, 1962-63 and 1991-1992, when political factors (being a new state in 1962; political instability and the movement toward anocracy beginning in 1990) pushed the expected probability of civil war well above the world average. And in fact, the onsets of civil war took place precisely when our models showed that Algeria was especially prone to such violence.

Why examine with a fine-toothed comb cases for which no explanation is needed? In our method, however, Algeria was chosen through random selection, and so we could not let this sleeping dog lie. Waking him up proved rewarding.[13]

The civil war onset when Algeria was a new state indeed involved a commitment problem like that we theorized in proposing why the variable New State associates with civil war. Berbers who had prospered under French rule feared a loss of status to Arabs who would be the majority group in the new Algerian state. Furthermore, the local guerrilla units that bore the greatest brunt of the fighting for independence feared marginalization once the regular army entered Algeria from Morocco and Tunisia. If these regional militias did not fight for power in the summer of 1962, they feared marginalization as they did not trust the the new leadership to give them as much as they might be able to take by fighting. Indeed some of these units and at least one from the Berber region took part in the 1962 rebellion.

The narrative also made clear that this "weak state" commitment problem stemmed as well from a factor not previously considered, namely, the ineffectiveness of France's transfer of power. France left the scene when the independence movement was divided in several ways, with no clearly dominant force and no general commitment to any constitutional mechanism to decide among them or guarantee future, peaceful chances at power. Independence to a new state without a credible commitment by the former metropole to support the leadership to which it transfers power yielded a vacuum

---

[13] . The full Algeria narrative is available at http://www.stanford.edu/group/ethnic/.

that drew in insurgents. It was France's inability to commit to Prime Minister Ben Bella rather than just Ben Bella's inability to commit to the future security of minorities that accounted for post-independence violence in 1962.

The civil war coded as beginning in 1992, as our theoretical arguments concerning anocracy and political instability expected, emerged out of the political opening granted by the authoritarian regime in 1990. The opening indicated division within the governing elite and a general sense of weakening in the face of public dissatisfaction and pressure. Clerics in the FIS (Islamic Salvation Front) were emboldened to exploit an economic crisis to challenge the regime in the name of fundamentalist ideals. The commitment problem again appears as critical in the explanation for why there was violence rather than a negotiated or democratic solution – the military regime feared that the first national election lost to the Islamists would be the last election held anytime soon. This is consistent with our theoretical account linking anocracy and political instability (indicated by changed Polity scores) to civil war.

However, a careful look at the civil war that ensued brought into question our interpretation of the impact of per capita income. We have argued that on average, country poverty proxies for a weak central government administration and coercive apparatus, unable to collect information on its own population or to use information strategically to root out insurgents. In the Algerian case, we find a very strong army, one that that had learned much from the experience in fighting the French during the long war of independence. The army was well taken care of and it had the resources and will to develop sophisticated counter-insurgency units. So Algeria's moderately low per capita income (versus the regional or world averages) is in this case a poor measure of state coercive capabilities, and civil war occurred despite relatively strong coercive capability.

The narrative suggested, however, another quite different mechanism linking low per capita income and civil war. Area experts have pointed to the earlier closing of the migration outlet to Algerian youth as a causal factor in the onset of the war. Indeed, the rebellion came shortly after France cut off the immigration spigot. Instead of further inciting the anti-immigration program of M. Le Pen and his Front National in France, young Algerians that would formerly have been unemployed or sending back remittances from France were being recruited into the FIS. And so, the low GDP and weak economy in Algeria worked through a second mechanism (available recruits) rather than the first (weak military) to translate high likelihood to actual onset (consistent in this case with Collier and Hoeffler's (2004) interpretation of the role of income).

Since the FIS was a religious mobilization, the narrative almost invited us to ask whether the civil war of 1992 can be explained by some religious factor, even if religion played no role in our statistical analysis. There can be little doubt that Islamic symbols had a powerful emotional impact on the population. In the late 1970s, Muslim activists engaged in isolated and relatively small-scale assertions of fundamentalist principles: harassing women whom they felt were inappropriately dressed, smashing establishments that served alcohol, and evicting official imams from their mosques. The Islamists escalated their actions in 1982, when they called for the abrogation of the National

Charter and for the formation of an Islamic government. Amidst an increasing number of violent incidents on campuses, Islamists killed one student. After police arrested 400 Islamists, about 100,000 demonstrators thronged to Friday prayers at a university mosque. Islamists were also able to mobilize large numbers of supporters successfully to demand that the government abrogate rights given to women in the colonial period. And of course, the Islamist political party shocked and awed the military authorities in their impressive first round electoral victory in December 1991 (Metz 1993). Fundamentalism was popular!

However, it is not clear why Islamic fundamentalists confronted the FLN. Algerian nationalism, consequent on the French denial of citizenship to Muslims in the 1870 Cremieux Decree, was always "Islamic" in sentiment. The FLN was never considered, as many in the army command considered themselves, secular and perhaps even anti-Islam. Some FLN leaders were Islamists. The FIS did not represent a deep cultural cleavage in Algeria. In fact, there is a popular pun among Algerians, "le FIS est le fils du FLN" (Quandt 1998, 96-7). The trump in the FIS hand was not its religious devotion or its sole identification with Islam.

Furthermore, a careful examination of the FIS reveals little about Islam as the source for the Algerian rebellion. For one, the clerics followed the urban proletariat into war rather than led them. There is evidence that in fact the clerics sought in the late 1980s to calm the riots in the streets instigated by the unemployed youth (Pelletiere 1992, 6). To be sure, the GIA (the leading insurgent militia) relied on fundamentalist ideology in order to finance the war through the "Islamic rent" paid by Middle East states (Martinez 2000, pp. 198-206, 240). But this was a strategy of raising funds more than a sign of Islamic devotion.

The narrative suggests that it was not Islamic fundamentalism, but rather state strategies in regard to religion that played a vital role in driving the insurgency. After independence, the Algerian government asserted state control over religious activities for purposes of national consolidation and political control. Islam became the religion of the state in the new constitution and the publicly displayed religion of its leaders. No laws could be enacted that would be contrary to Islamic tenets or that would in any way undermine Islamic beliefs and principles. The state monopolized the building of mosques, and the Ministry of Religious Affairs controlled an estimated 5,000 public mosques by the mid-1980s. Imams were trained, appointed, and paid by the state, and the Friday *khutba*, or sermon, was issued to them by the Ministry of Religious Affairs. That ministry also administered religious property, provided for religious education and training in schools, and created special institutes for Islamic learning.

What is the implication of state control over religion? Our statistical analysis of the whole sample found that religious discrimination could not distinguish countries that experienced civil wars from those that did not. But the narrative suggests a different mechanism. The very act of authorizing and subsidizing religious organizations and leaders automatically politicized religious protest. As religious entrepreneurs sought to capture new audiences for their local mosques, they were in fact challenging state

authority. Through its subsidization and promotion of Islam, the Algerian authorities opened themselves up to forms of symbolic attack they could not easily repel. By seeking to suppress religious experimentation, the FLN found itself more vulnerable to attack than if it kept entirely out of religious affairs.

In sum, rather than some deep religious message of FIS that articulated with the religious sentiments of the people, it was the situation in which the state, by seeking to co-opt religious opposition, that gave that opposition a chance to articulate a clear anti-regime message through renegade mosques. State sponsorship of religion backfired grievously. The Algerian case thus suggests that state *sponsorship* of religion (rather than *discrimination* against it) raises the probability of a successful insurgency.

6. Conclusion

Despite some claims to the contrary in the qualitative methods literature, case studies are not designed to discover or confirm empirical regularities. However they can be quite useful – indeed, essential – for ascertaining and assessing the causal mechanisms that give rise to empirical regularities in politics. We have argued that random selection of cases for narrative development is a principled and productive criterion in studies that mix statistical and case-study methods, using the former for identifying regularities, and the latter to assess (or to develop new) explanations of these.

Using the Algerian example, the narratives suggest a return to large-N analysis with several new ideas. First, the variable "new state" might be productively interacted with the capacity of the metropole to commit to the transitional leadership. The expectation is that a strong metropole would better be able to protect the leaders to whom it transferred authority, and thereby deter (at least for a time) potential insurgents. France in the wake of occupation in World War II, the loss of the colonial war in Vietnam, the collapse of the Fourth Republic, and the long war for Algerian independence, was not in a position to manage the transition to the new leadership in Algiers.

This insight emerging from the narrative of a case that was "on the regression line" illuminated a not-so-obvious pattern in vulnerability of new states to civil war onsets. Many countries that received independence in the immediate postwar era when metropoles were devastated (such as in Indonesia, Vietnam, South Asia, and the Palestine Mandate) fell quickly into civil war. Those countries that became new states when the Soviet metropole disintegrated (Azerbaijan, Georgia and Moldova are examples) were also highly susceptible to civil war onsets. However, those countries that received independence in the 1960s and 70s in Africa when the metropoles were strong (except for Belgium and Portugal that could not manage the transitions to new leadership in their colonies) were less likely to suffer immediately insurgent success. In several of these cases, "commitment problem" wars started several years later, after the colonial power really did stand back. "New state" is more dangerous the weaker the metropole that grants it independence.

Second, the Algeria narrative suggests that we might develop a coding rule for extent of migration by young men to more productive economies for work. The expectation would be that in countries where young men can relatively easily escape unemployment through migration to industrialized countries, insurgent recruitment will be more difficult than otherwise. (Subsequent narratives reported on the relationship of blocked migration opportunities and civil war in Haiti and near civil war in Jamaica; meanwhile open migration opportunities may have helped save Portugal and Dominican Republic from joining the list of onsets under revolutionary conditions).

Third, the Algerian narrative suggested a new way to think about the religious sources of insurgency. Instead of modeling hatreds between people of different religions, or of state discrimination against minority religions, it might be more productive to model the relationship of dominant religious authority and the state. The more the state seeks to regulate the dominant religious organization, the more it is setting up a recruitment base against the state within the religious organization. Preliminary data analysis for our large-N dataset gives support to this narrative-inspired conjecture.

There are several more general lessons as well to be learned from the random narrative exercise. Through narrative, it is possible to point to interactions among individual variables that may not matter in a consistent way by themselves, but that jointly may make for civil war susceptibility. It may then be possible to specify more sharply the conditions when a variable will have some theorized effect. It is possible to point to micro-factors for future coding such as tactical decisions by states and by insurgents that are usually ignored in large-N data collection exercises.

As well, the random narrative method allows us to estimate measurement error for variables that are hard to code across reliably across large numbers of cases. In the set of narratives we examined through random selection, we found not insubstantial error in the coding of civil war onset, our dependent variable. To give but one example, northern Thailand has been held by area experts to be a zone of peace compared to the mountainous rebellions in neighboring Burma and Laos. A number of civil war lists have thus ignored the northern troubles in Thailand as a possible civil war. As a result of the research that went into the random narrative, however, we found that the northern insurgency clearly passed the death threshold that our scheme determines as a civil war. In general, we estimate that as many as five percent of our initial codings on the dependent variable were erroneous. Statistically, if these errors are random, in a logit analysis this will tend to bias effect estimates towards zero. Of course, the errors may not be random – they are surely more likely for relatively low level civil wars close to whatever death threshold is employed – so a direct advantage of combining narrative analysis with statistical analysis is better measurement and more accurate effect estimates.

These random narratives, in sum, have already proven both troubling and useful as a complement to, or extension of, a large-N analysis of civil war onsets. They suggest a natural way that qualitative work might be integrated into a research program as a complement to rather than as a rival or substitute for quantitative analysis.

References:

Bates, Robert et al. 1998. <u>Analytic Narratives</u> (Princeton: Princeton University Press)

Boix, Carles 1998. Political Parties, Growth and Equality (Cambridge: Cambridge University Press)

Collier, Paul and Nicholas Sambanis. 2005. Understanding Civil War, 2 volumes. (Washington, D.C.: World Bank)

Doyle, Michael and Nicholas Sambanis. 2006. Making War and Building Peace (Princeton: Princeton University Press)

Eckstein, Harry. 1975. Case Study and Theory in Political Science. In <u>Handbook of Political Science</u>, eds. Fred Greenstein and Nelson Polsby. (Reading, MA: Addison Wesley), pp. 94-137

Elster, Jon (1998) "A Plea for Mechanisms". In P. Hedstrom and R. Swedberg (eds.), <u>Social Mechanisms: An Analytical Approach to Social Theory</u> (Cambridge: Cambridge University Press)

Fearon, James D. (1998) "Commitment Problems and the Spread of Ethnic Conflict. In David Lake and Donald Rothchild, eds., <u>The International Spread of Ethnic Conflict: Fear, Diffusion, and Escalation</u>. Princeton: Princeton University Press, 1998.

Fearon, James D. (2004) "Why Do Some Civil Wars Last So Much Longer Than Others? <u>Journal of Peace Research</u> 41, 3 (May 2004), 275-301.

Fearon, James D. (2008) Economic Development, Insurgency, and Civil War. In Elhanan Helpman, ed., <u>Institutions and Economic Performance</u> (Cambridge, MA: Harvard University Press, forthcoming).

Fearon, James D. and David D. Laitin (1999) "Weak States, Rough Terrain, and Large-Scale Ethnic Violence since 1945." Presented at the Annual Meetings of the American Political Science Association, Atlanta, GA, 2-5 September 1999.

Fearon, James D. and David D. Laitin 2003. "Ethnicity, Insurgency and Civil War". <u>American Political Science Review</u>. 97 (1) (February): 75-90

Fearon, James and David Laitin (June 2005) "Civil War Narratives" Estudio/Working Paper 2005/218, Centro de Estudios Avanzados en Ciencias Sociales, Instituto Juan March de Estudios e Investigaciones.

Fearon, James D. and David D. Laitin (2007) "Civil War Terminations". Paper presented at the 103rd Annaul Meeting of the American Political Science Association, Chicago, IL., August 30-September 2.

Fortna, Virginia Page. 2004. <u>Peace Time: Cease-Fire Agreements and the Durability of Peace</u>. (Princeton: Princeton University Press)

George, Alexander and Andrew Bennett. 2005. <u>Case Studies and Theory Development in the Social Sciences</u>. (Cambridge, MA: MIT Press)

Gerring, John. 2006. <u>Case Study Research: Principles and Practices</u>. (Cambridge: Cambridge University Press).

Goemans, Hein. 2000. War and Punishment. (Princeton: Princeton University Press)

Goldthorpe, John. 2000. On Sociology (Oxford: Oxford University Press)

Huth, Paul. 1998. Standing Your Ground: Territorial Disputes and International Conflict. (Ann Arbor: University of Michigan Press)

Iversen, Torben. 1999. Contested Economic Institutions. (Cambridge: Cambridge University Press)

Laitin, David. 1995. "National Revivals and Violence" Archives Européennes de Sociologie (Spring)

Laitin, David. 1998. Identity in Formation (Ithaca: Cornell University Press)

Mansfield and Snyder 2005. Electing to Fight: Why Emerging Democracies Go to War (Cambridge, MA: MIT Press).

Mares, Isabela. 2003. The Politics of Social Risk. (Cambridge: Cambridge University Press)

Martin, Lisa. 1994. Coercive Cooperation (Princeton: Princeton University Press)

Martinez, Luis. 2000. The Algerian Civil War 1990-1998 (London: Hurst), trans. Jonathan Derrick

Metz, Helen ed. 1993. Algeria: A Country Study Federal Research Division, Library of Congress (http://memory.loc.gov/cgi-bin/query2/r?frd/cstdy:@field(DOCID+dz0000), section on Chadli Bendjedid and Afterward)

Pelletiere, Stephen C. 1992. Mass Action and Islamic Fundamentalism: The Revolt of the Brooms (Carlisle Barracks, PA: US Army War College)

Quandt, William B. 1998. Between Ballots and Bullets (Washington, D.C.: Brookings), pp. 96-7.

Schultz, Kenneth. 2001. Democracy and Coercive Diplomacy (Cambridge: Cambridge University Press)

Stone, Randall. 2002. Lending Credibility (Princeton: Princeton University Press)

Van Evera, Stephen. 1997. Guide to Methods for Students of Political Science (Ithaca: Cornell University Press)

Wagner, R. Harrison (2007) War and the State (Ann Arbor, MI: University of Michigan Press)

Walters, Barbara. 2001. Committing to Peace (Princeton: Princeton University Press)

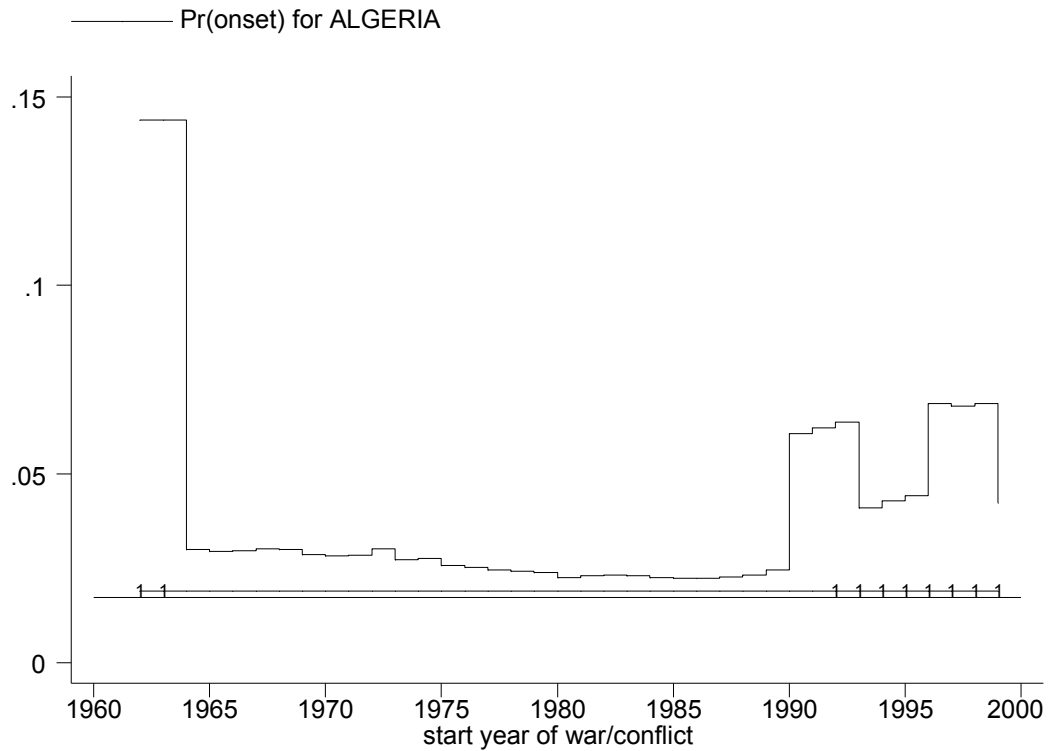Figure 1: Probability Graph for Civil War Onsets in Algeria

Table 1: Key Variables for Algeria, and in Comparison to Regional and World Means

Variable	Algerian mean	Regional Mean	World Mean
Onset Probability	.040	.016	.017
GDP/capita (in 1985 USD)	2340	5430	3651
Population (in millions)	19.411	11.482	31.787
Mountainous Area (as a percent of total area in country)	15.7	18.6	18.1
Oil (dummy for extensive exports as percentage of total exports)	1	.49	.13
Instability (dummy)	.24	.13	.15
Anocracy (dummy)	.18	.23	.23