Combining Behavioral and Structural Predictors of Violent Civil Conflict: Getting Scholars and Policymakers to Talk to Each Other

Douglas M Gibler*

October 1, 2015

Abstract

This paper uses conflict narratives from the International Crisis Group's *CrisisWatch* publications to cross-validate structural analyses of civil conflict and confirm the mechanisms that lead to outbreaks of violence in conflict-prone countries. I correct for selection bias in the narrative data with an underlying model of conflict, and I find that several indicators thought to be causally related to civil conflict do indeed continue to have an effect after selection. I also find a tendency in the narrative data to over-emphasize the importance of several low-intensity, separatist conflicts within developed democracies and the potential for conflict among oil-rich states. Overall, the analyses highlight the importance of combining structural, large-N analyses of structure with qualitative assessments of government, citizen, and international community behavior and describe a state-capacity explanation of conflict behavior.

Keywords

conflict narratives, structural causes of civil war, selection bias

*Professor, Department of Political Science, University of Alabama, Tuscaloosa, AL; dmgibler@ua.edu.

This paper provides a unique set of analyses of the causes of civil conflict. I leverage in-country assessments of public, government, and international community behaviors to assess the ability of structural models of civil conflict to predict dangerous states. I then combine these approaches to understand what behaviors are associated with continued civil conflict and escalation, and, just as importantly, which behaviors are not.

There are several advantages of this approach. First and foremost is the added information afforded to both types of analysis. As I detail in the next section, structural predictors are most often rough approximations of key concepts we believe are associated with potentially distressed states. We use mountainous terrain as a proxy for rebels, for example. Moreover these variables provide actual values that are often invariant or move very slowly across time, and the analyses provide little prediction about when conflicts are likely to occur across these environments.

Conversely, a focus on behaviors in conflicts misses the key point that certain states are going to suffer because they do not possess the political, economic, or even geographic conditions that make good governance likely. Thus, the factors naively associated with conflicts may be a product of structural conditions and not at all associated with conflict themselves. In this way analysts need an understanding of the selection effects that make certain types of events likely.

Of course there are advantages and disadvantages to any approach, but, by combining in-country reports with large-N analyses of conflict likelihoods, I am able to cross-validate the predictions of each method based on real-world events. I find that the structural variables commonly associated with failed states and poor state capacity continue to predict conflict well, and this remains true even after corrections for selection. Missed in the structural models, however, is the tendency for several developed democracies to have low-intensity, separatist movements with continual conflict. Structural models tend to ignore these types of incidents.

The paper proceeds as follows. First, I provide a general outline of how large-N analyses of civil conflict have focused on the predictors of unstable states and the disadvantages of this type of approach. I then discuss how narrative data on the events in particular countries can be used to inform these structural models. Cross-validation of the two approaches uncovers potentially biases, especially with regard to regime. However, the analyses that combine the data confirm the importance of several predictors of intrastate conflicts. I argue that the findings confirm a state-capacity explanation of conflict and close with a discussion of the implications of these findings.

Research on the structural predictors of civil conflict

Existing research on civil conflict is often unable to provide much guidance as to the specific triggers leading to large-scale violence. The focus has instead been on a broad range of conditions that make the occurrence of civil war more likely. Since Fearon and Laitin's (2003) landmark study of civil wars, such factors typically include low levels of economic development, low state capacity (broadly defined), countries with weak military forces, countries with larger populations, the presence of ethnic and/or religion divides, weak or absent democratic political institutions,

the presence of natural resources, extreme variations in rainfall, and the presence of mountainous terrain.[1]

These factors represent general conditions that are correlated with the occurrence of civil war. Low levels of economic development may lead to the outbreak of civil war through the creation of widespread grievances resulting from a scarcity of resources. In such environments, competition over access to those resources may be more intense—and more likely to turn violent—than would be the case in wealthier countries. Similarly, countries with an abundance of mountainous terrain have been found to be at a higher risk of experiencing civil wars. Rugged, mountainous terrain can facilitate conflict by providing safe-havens for rebel groups, and they can also stymie government efforts to monitor and pursue such groups.

However, though factors such as poor economic development and mountainous terrain have been linked to conflict, we still see significant variation in the actual occurrence of civil conflict. Many poor countries do not experience civil war while others do, and the severity of violent conflicts also varies across countries as well. What explains this variation? Factors like low economic development contribute to an environment in which conflict is *more likely* but do not themselves *cause* conflict. Similarly, mountainous terrain can provide safe-havens that enable violent, anti-government insurgency campaigns, but the mere presence of such safe-havens does not actually create the rebel or insurgent groups. Accordingly, these indicators give us a rough approximation of where countries fall in terms of relative risk, but they cannot account for the actual onset or escalation of violent conflict within a given country.

The structural predictors of civil conflict are also largely static, or at least slow-to-change, and this means that it is often difficult to predict *when* the various dangerous environment will erupt, how long the conflict will last, and how severe it will become. The emphasis on static measures also limits the ability of research to affect policy outcomes. Economic development, political institutions, demographic composition, and terrain are not easily manipulable by foreign or domestic policies. This makes existing research unable to provide much guidance for interventions designed to prevent civil conflict before it even begins.[2]

In order to improve our ability to predict the onset of violent conflict we also need to refine how we conceptualize and measure some of the key mechanisms that are expected to cause conflict. Theory testing is often hampered by the need to rely on relatively crude measures as proxies for concepts such as rebels and insurgencies or economic and ethnic grievances. These groups and grievances are simply assumed to be present based on the underlying conditions of the state. The problem with these assumptions, though, is that they can serve as proxies for multiple theoretical mechanisms and concepts, making it difficult to accurately assess the specific causes of violence. For example, does low economic development affect conflict through the creation of poverty and grievances, or does it facilitate conflict through its effects on government revenues and economic resources? In either case, there are several steps that link economic development to the onset of conflict that are assumed away, just as must be done with all other structural factors.

[1] For more information on these issues, and for examples, see the following: Hegre et al. (2001), Collier and Hoeffler (2004), Miguel, Satyanath and Sergenti (2004*a*), Regan and Bell (2010), Hendrix and Salehyan (2012), and Young (2012).

[2] There are exceptions to the sole focus on structural conditions as explanations for conflict. Noting the invariance of many structural conditions, several studies have explored the role of economic shocks (Miguel, Satyanath and Sergenti 2004*b*), food or commodity-price changes (Dube and Vargas 2013, Weinberg and Bakker 2014), and climate change (Hendrix and Glaser 2007) as means for predicting variation in the outbreak of conflict. I believe the focus on "trigger events" and change models in these studies highlights the lack of agency in most of our best models of conflict states.

International Crisis Group's CrisisWatch narratives and the causes of civil conflict

The divide between current scholarship and the policymaking community hints at some of the problems in both fields. For example, rather than relying on the structural determinants of civil conflicts, many policymakers focus instead on in-country narratives and reports of conflict hotspots and potential hotspots. Consider the work of the International Crisis Group (ICG) which is one of the premier and most comprehensive sources of conflict narratives for policymakers available. Founded by high-ranking government officials and diplomats, the ICG is a non-governmental organization whose primary purpose is to monitor events on the ground in many of the most conflict-prone countries around the globe. Since 2003 the ICG covers approximately 70 countries, with field analysts located in roughly 30 of those countries at any given time. Through this network of field analysts, and by drawing on global and regional news reports, the ICG releases a closely-watched, twelve-page monthly *CrisisWatch* bulletin that contains updates on the 70+ countries that it covers. These detailed textual accounts group countries into three general "trend" categories according to the state of violent conflict within that country: unchanged, deteriorated, and improved. Each monthly report also provides a "watchlist" listing countries that are deemed to be at an especially high risk of conflict, as well as countries that are viewed as having an increased opportunity for conflict resolution. Figure 1 gives an example of several country reports from August 2014.

West Africa

➡ **Burkina Faso** Amid continued tensions over possible referendum to amend constitution and remove 2-term limit, President Compaoré 13 July said undecided about running in 2015 elections. U.S. 9 July stressed need to respect democratic process; 17 July pledged $1mn to support free, fair and transparent elections.

- "Blaise Compaoré: 'Il faudra bien partir un jour…'", *Jeune Afrique*, 21 July 2014.

➡ **Côte d'Ivoire** Opposition Ivorian Popular Front (FPI) leader and presidential hopeful Pascal Alli N'Guessan 4 July restructured party leadership ahead of 2015 presidential elections, attempted to remove General Secretary and pro-Gbagbo hardliner Laurent Akoun; Akoun 15 July called decision illegal, reiterated calls for Gbagbo's release and role in party. Ongoing tensions between govt and FPI: FPI 3 July said some 1,000 supporters still detained, many tortured; 9 July rejected new Independent Electoral Commission (IEC). Fearing Ebola outbreak, border authorities 11 July blocked 400 Ivorian refugees returning from Liberia, UNHCR called decision "unacceptable". French President Francois Hollande 17 July visited Abidjan, said Côte d'Ivoire now "inspires confidence", discussed economic and security cooperation.

- "Le FPI dans la tourmente", RFI, 16 July 2014.

➡ **Guinea** Restart of govt-opposition talks on electoral process and opposition return to parliament 4 July; verbal agreement reached 5 July quickly vaporised as opposition 11 July accused govt of omitting key verbal deals in concluding statement; opposition 12 July submitted own version, signature pending. As Ebola killed hundreds, tensions continued between authorities and health organisations; Red Cross, WHO, MSF reported denied access to villages around Guéckédou. Guinea was declared in full conformity with Extractive Industry Transparency Initiative (EITI) 2 July.

- "Processus électoral bloqué", Nouvelle Tribune, 13 July 2014.

Figure 1: Sample narratives from ICG's August 2014 CrisisWatch.

The ICG is dedicated to providing policymakers with accurate and up-to-date information on developments in at-risk countries, so it is an incredibly useful source for regular, systematic

reporting on events that directly pertain to the status of violent conflict within its sample of the most dangerous conflict environments. Further, because the ICG was created in large part to supply information to policymakers, many of the factors on which the ICG reporting focuses are likely to be manipulable, and are less likely to focus on broad indicators like economic development.

The problem with these narratives is that, of course, the coverage only includes a highly-selected sample of countries already in conflict or at risk of conflict. This selection effect may instill substantial bias in any links between the various narratives and conflict outbreak or escalation. So, for example, any trend in the narratives between riots or protests in various countries and subsequent civil conflict may be a function of the structural conditions that make riots and conflict more likely and not some causal mechanism suggesting riots cause civil wars. Riots could be taking place in many of the peaceful countries not reviewed by the narratives, and, without this baseline, inferences become almost useless.

Correcting selection bias is likely to yield insights into both the narrative-based data as well as the structural predictors of conflict. Because the ICG is an organization that operates with limited resources, and because it needs to maintain its relevance insofar as policymakers and donors are concerned, the ICG seeks to identify the states that are most at risk of seeing new or recurring civil conflict. They approach this task in two ways: ICG analysts identify specific states believed to be at an especially high risk of conflict in their monthly *Crisis Watch* bulletins and then they deploy their limited staff to monitor those areas that are believed to be high-risk environments. This pattern of behavior is useful because it provides an opportunity to cross-validate models predicting the outbreak of conflict from political science with the predictions and coverage decisions made by policymakers and policy analysts.

An anonymous reviewer pointed out, rightly, that the division between structural and behavioral accounts of civil conflict may be less pronounced than it first appears. Large-N models have strong predictions about where conflicts are likely to occur, and several of these findings began as counter-intuitive relationships that have been embraced over time. That these may now often be intuitive selection devices for policymakers underscores the policy relevance of the structural predictors of conflict. Of course, differences remain between the two approaches, and those differences can provide much information about the causes of conflict.

Using a simple cross-tabulation, Figure 2 illustrates this basic idea. When both ICG and political science predictions match—where both predict no conflict or where both predict the outbreak or escalation of a conflict—we can assume that there are certain factors that clearly indicate a strong likelihood of conflict. When this occurs, the narratives and the structural approaches to conflict cross-validate each other, and the narratives provide agency and the causal mechanisms that link underlying conditions to the outbreak of conflict. However, the cases in which the ICG and political science predictions diverge demonstrate the potential for measurement error, measurement bias, or selection effects in either or both approaches.

Better understanding the error and biases in our approaches to predicting the onset and escalation of conflict is valuable. To the extent that political science models of conflict are failing to predict cases that the ICG succeeds in predicting, we can use this information to ask what factors policymakers are focusing on that political scientists are neglecting. Alternatively, looking at cases where political science models successfully predict conflict, but the ICG does not, also allows us to cross-reference cases to better understand the factors that policymakers may be

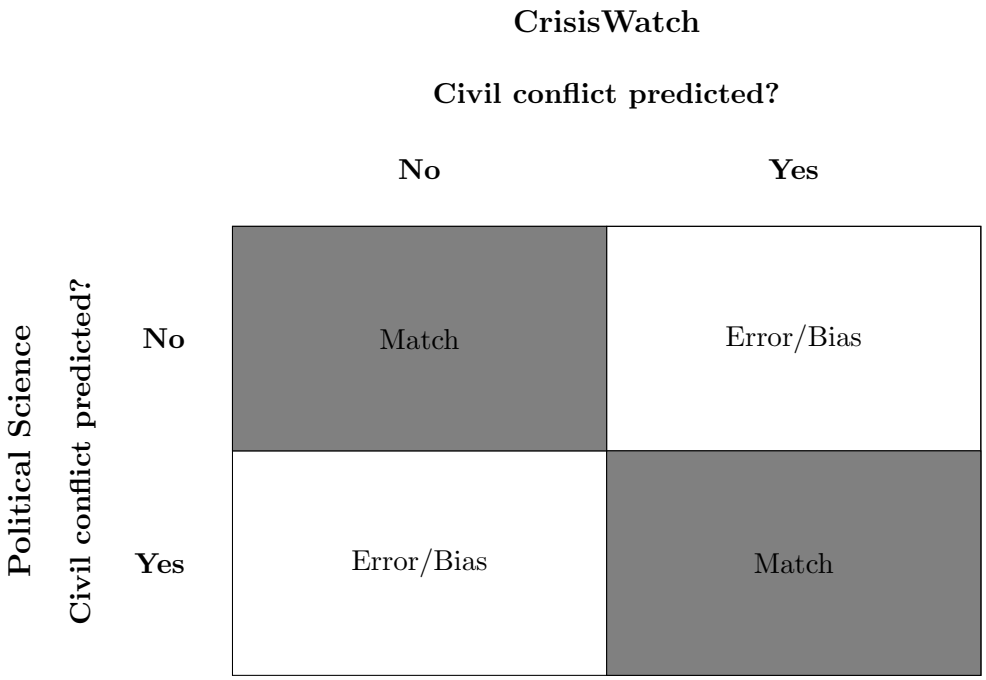

Figure 2: Cross-validation and potential bias in approaches to predicting civil conflict.

overlooking, or what factors might be leading them to emphasize some cases over others. A combined approach facilitates a more refined study of conflict causes as well as the development of evidence-based policy prescriptions and interventions by policymakers. Since conflicting predictions may represent tough cases, cross-validation also improves our understanding of where and why specific predictions prove inadequate.

To make the ICG narratives useful for systematic study, I first coded each individual narrative for three to six keywords or phrases that best described the situation. The narratives are each one to two paragraphs long and highlight major incidents taking place in each potential conflict zone in a particular month. I chose the keywords that would best describe the paragraph-long narratives; these keywords or phrases were most often associated with shorter narratives, while six keywords or phrases were sometimes used for longer, more complicated narratives.[3] The keywords most often included such phrases as protest, riots, government repression, arrests of key figures, terrorist incidents, etc. From this list of keywords and phrases I then created dummy variables for the presence of each indicator using a simple search function. Finally, I confirmed that every mention of an indicator was correctly described in the narrative as relating to the conflict taking place in the country. This method privileges the elimination of Type 1 (false positive) over Type 2 (false negative) errors.[4]

The theoretical choice behind each keyword or phrase can really be traced back to what the ICG chose to report as important. The keywords were chosen to reflect the concepts described in the narratives, so I again confirmed that the dummy variables matched the case narratives well for each of the most numerously mentioned keywords or phrases. Each mention had to also be

[3] All *CrisisWatch* publications are available at http://www.crisisgroup.org/en/publication-type/crisiswatch.aspx.
[4] Inter-coder reliability tests using three separate random samples of approximately 20 narratives correlated at 95% or higher from keyword selection for each narrative through the second step of confirmation.

temporally proximate to the crisis environment and not describe long-past events; the keyword also had to refer to the country covered. So, for example, ethnic tensions across the border that were described in a narrative would only be coded as positive for this concept if the narrative explicitly mentioned that the tensions had also enflamed ethnic issues in the crisis country. Similarly, a narrative that mentions an election from ten years prior that divided on ethnic cleavages would not be coded as positive for this concept unless it was somehow directly related to the ongoing potential for crisis. Table 1 provides summary statistics for each of the bolded keywords in the ICG *CrisisWatch* narratives from 2003-2011. The keywords are also embedded within longer questions that better describe each concept.

Table 1: Frequency of keyword and phrase description of CrisisWatch narratives

		Structural Likelihood of Conflict in State†
Popular unrest		
Did **Protests** occur?	1,248 mentions	-8.95%
Were there **Riots**?	98 mentions	-8.42%
Government behavior		
Were there **Arrests** of political figures?	849 mentions	-8.42%
Were the incidents of **Repression**?	224 mentions	-7.08%
Were **Elections** imminent or discussed?	224 mentions	+7.46%
Were there **Coups or coup attempts**?	918 mentions	-6.24%
Conflict		
Were there one or more **Terrorist incident(s)**?	780 mentions	+10.19%
Were anti-government **Rebels** mentioned?	2,028 mentions	+49.43%
Was **Ethnicity or ethnic issues** mentioned in the narrative?	318 mentions	+37.15%
International reaction		
Was the state under **Sanctions** or threatened by them?	152 mentions	+7.54%
Was conflict **Mediation** mentioned in the narrative?	266 mentions	-10.94%
Was **United Nations** involvement mentioned?	616 mentions	+63.34%
More specifically, was **United Nations** mediation discussed?	587 mentions	+66.54%
Was the presence of **Peacekeepers** discussed?	185 mentions	+63.89%

†Difference in average probability of civil conflict in state-year for states with each behavior versus the entire sample average. Sample includes 8,076 narratives analyzed, 2003-2011

In the remaining sections I describe first the construction of a set of predicted probabilities from a structural-condition-based model of civil conflict. Using common indicators of conflict provides a baseline for further analyses. I then use those predictions for comparisons to the narrative-based keywords that describe the situations in conflict-likely countries. Finally, I correct for selection bias and merge the two sets of analyses to examine the underlying causes and mechanisms that are likely to generate civil conflicts.

Analyzing the structural determinants of civil conflict

I begin the analyses with a simple model of intrastate conflict derived from the civil war literature. I control for the economic situation in the state with a measure of its gross domestic

product, using Gleditsch's (2002) expanded trade data with imputations for missing values. The population of the state is based on estimates from the Correlates of War Project (Singer, Bremer and Stuckey 1972). Non-contiguous territory, a substantial presence of oil in the state, and ethnic and religious fractionalization are each coded consistently with the analyses in Fearon and Laitin (2003). Unified democracy scores provide a proxy for the level of democracy in the state, and I also include its square to identify any curvilinear effects in the relationship between government type and civil conflict (Pemstein, Meserve and Melton 2010). Finally, my dependent variable is the presence of a civil conflict as identified by the UCDP/PRIO Armed Conflict Dataset (Themnér and Wallensteen 2014), and in the model I control for duration dependence using the years since observing the dependent variable, its square, and its cube (Carter and Signorino 2010). The first two columns of Table 2 present estimates of these variables on the effects of civil conflict between the years 1946 and 2008. The third column substitutes coverage of the state by the ICG as the dependent variable; as I describe below, this allows a direct comparison between a structural model and the ICG's choice of coverage.[5]

The structural measures of conflict generally perform as expected. Wealthy states and less populous states are not as likely to experience intrastate conflicts. This is true during the entire post-World War II period of the data and also in the constrained temporal sample that corresponds to the ICG narrative data.

The Fearon and Laitin (2003) argument of insurgency holds in both samples as well as mountainous terrain and large populations—where rebels can hide in the geography or among the population—are both related to an increased risk of civil conflict in each sample. Meanwhile, democracy and its square reduce predict fewer civil conflicts, so too does increased religious fractionalization. With the UCDP dependent variable ethnic fractionalization also now has an effect, predicting conflict in both samples. Non-contiguous territory, the presence of large exports, and political instability are not associated with conflict in either of the samples. Finally, after controlling for all other predictors, the presence of peace in a state in one year makes it more likely that the next year will be peaceful as well.

Overall, these results describe civil conflicts as unlikely in stable, older democracies but more often present in states with rugged terrain and an ethnically diverse, large population. These relationships suggest an association of civil conflict with the inability of governments to find and quash rebellions from state authority.

The ICG also tends to cover less developed, more populous, and more ethnically homogenous states. Substantial differences arise, however, when considering the other predictors of coverage. For example, oil states are much more likely to be covered by the ICG, even though the structural model predicts no relationship with internal conflict. Too, while both types of fractionalization are associated with a reduced likelihood of conflict, the ICG tends to over-emphasize ethnic fractionalization while ignoring religious fractionalization. The differences are more nuanced for regime type as the structural model demonstrates that democracy pacifies once a certain threshold is reached; meanwhile, the effects remain linear in the ICG-coverage model. Finally, the ICG reports seem to be mobile and responsive to changes across the sample since I found no duration dependence in the data analyzed.

I also compared the state-year predicted probabilities of conflict in the second model to the predicted probability of coverage by the ICG for every state-year in the third model. The

[5]Data use note, anonymized.

Table 2: Structural predictors of civil conflict

DV:	*UCDP intrastate conflict* 1946-2008	2000-2008	*ICG Coverage* 2003-2011
GDP (1yr lag)	-0.646***	-0.662***	-0.776***
	(0.130)	(0.145)	(0.163)
Population (ln, 1yr lag)	0.779***	0.905***	0.628***
	(0.080)	(0.103)	(0.097)
Mountains (ln)	0.257**	0.221*	0.195
	(0.090)	(0.099)	(0.113)
Non-contiguous territory	-0.058	0.509	-1.073**
	(0.299)	(0.337)	(0.392)
Oil state	0.461	-0.035	1.903**
	(0.306)	(0.327)	(0.638)
Democracy (UDS mean, 1yr lag)	0.199	-0.173	-0.603*
	(0.231)	(0.256)	(0.288)
Democracy squared (1yr lag)	-1.441***	-0.917**	0.095
	(0.292)	(0.304)	(0.205)
Political instability	-0.497	0.150	1.632*
	(0.360)	(0.419)	(0.680)
Ethnic fractionalization	-7.654***	-3.385***	-7.039***
	(0.758)	(0.555)	(0.605)
Religious fractionalization	-2.176***	-1.932***	1.297
	(0.516)	(0.543)	(0.674)
Peace years	-0.225***	-0.123*	
	(0.045)	(0.051)	
Peace years (square)	0.007***	0.003	
	(0.002)	(0.002)	
Peace years (cube)	-0.000*	-0.000	
	(0.000)	(0.000)	
Years since ICG			-16.97
			(251.0)
Years since ICG (square)			15.51
			(376.5)
Years since ICG (cube)			-3.932
			(125.5)
Constant	-3.141*	-4.342**	2.160
	(1.303)	(1.353)	(1.443)
N	7,980	1,889	1,214

Standard errors in parentheses
Symobls: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

correlation between predicted probabilities was surprisingly low at .43 for the full sample. This reinforces the argument that there are indeed differences between what countries ICG decides to cover and where a structural model predicts conflict.[6]

Model Diagnostics: Comparing structural explanations and ICG's CrisisWatch narratives

I use this section as a bridge between the structural conditions that are associated with civil war and the narratives that describe the behaviors associated with conflict. For example, the first part of Table 3 provides a cross-tabulation of model predictions from Table 2 and the presence of state coverage in CrisisWatch. I dichotomize the predictions and code likely conflicts as those state-years with predicted probabilities of civil war that are twice the system-wide average for the given year; predictions of peace are those cases with predicted probabilities less than half the system-wide average. The narrative data is monthly, but I assume that any coverage in a given year serves as a match for that state-year's prediction from the structural model.

Table 3a shows that only about half of the CrisisWatch's coverage would be predicted by the structural model. Of 236 state-years covered by the publication, 101 were predicted to be peaceful by the structural conditions in the state. However, ICG covers the majority of conflict cases, as over 78% of the conflict predictions are described in their narratives. This makes sense, of course, since the ICG is often sending their analysts to conflict areas *post hoc*, after events warrant their coverage. Nevertheless, the high number of incorrect predictions from the structural model, even with the high threshold of double the average probability, leads to some questions regarding its efficacy as a predictor of civil strife.

To further assess these differences, the remainder of Table 3 provides two separate examinations of the cases that are contained in the summary cross-tabulation. Table 3b lists countries that the structural model predicts as likely candidates for civil conflict, and, on the face of it, these are pretty decent choices for dangerous environments. In fact, only three of the cases (Cameroon, Iran, and Malawi) were not in the narrative data during the entire 2003 to 2008 time period. The remaining cases were represented in the data, but the years were not correctly predicted by the structural model.

The cases of peaceful predictions that were not in Table 3c present an interesting difference from the rest of the data. Over 65% of these cases were democracies with low-level, often separatist, conflicts that had been ongoing for some time. However, given the capabilities of many of these governments—in Britain, France, and Spain, for examples—these conflicts had little or no chance of escalating to large-scale conflicts and civil wars. In most of the narratives, too, there is little mention of violence, much less casualties or deaths, and several of these cases had been coded as ending in the UCDP data (Northern Ireland, for example). These cases may be over-emphasized by the ICG for a variety reasons—more Western news coverage or closer proximity to donors; the conflicts may also endure in the minds of policymakers long after the cases cool. Regardless, their coverage is not consistent with the other conflict cases in the narrative data or the UCDP data with its strict definition of case inclusion.

[6] The sample sizes for these separate analyses are different, of course, but constraining the models to the same set of observations changes none of the conclusions above in any way.

Table 3: Differences between predicted conflict and ICG CrisisWatch coverage

Table 3a: Summary differences between conflict predictions and ICG CrisisWatch coverage
Positive predictions of conflict are those probabilities that are twice the average prediction of conflict in a system-year; predictions of peace are half the same average.

	In ICG's CrisisWatch?		
	No	Yes	Total
Conflict			
No	437 (359.2)	101 (178.8)	538
Yes	37 (114.8)	135 (57.2)	172
Total	474	236	710

Pearson $\chi^2 = 209.43$, ($p < 0.000$)
Expected counts in parentheses.

Table 3b: States predicted to be conflict-prone that were not
These are states predicted to be conflict-prone by a structural model that did not have ICG coverage. States listed here have a probability of conflict more than double the average probability of conflict for that year among states in the international system.

State Name:	Conflict Predictions Based on Probabilities of Table 1					
Bhutan				2006	2007	2008
Cameroon	2003	2004	2005	2006	2007	2008
Ethiopia	2003	2004				
Iran	2003	2004	2005	2006	2007	2008
Kenya	2003	2004				
Madagascar	2003		2005			2008
Malawi	2003	2004	2005	2006	2007	2008
Mali	2003					
Morocco			2005			
Niger		2004	2005			
Tanzania	2003	2004			2007	2008
Zambia			2005			

Table 3c: States predicted to be peaceful that were not
These are states predicted to be peaceful by a structural model but still had conflict according to ICG. States listed here have a probability of conflict less than half the average probability of conflict for that year among states in the international system.

Predicted conflicts:	Years of ICG Coverage (Deteriorated Status in Red)					
Albania		2004	2005	2006	2007	2008
Armenia		2004	2005	2006	2007	2008
Bahrain		2004	2005	2006	2007	2008
Belarus			2005	2006	2007	2008
Croatia				2006		
Cyprus	2003	2004	2005	2006	2007	2008
Equatorial Guinea		2004				
France (Corsica)	2003	2004			2007	
Israel	2003	2004	2005	2006	2007	2008
Lebanon	2003	2004	2005	2006	2007	2008
Macedonia	2003	2004	2005	2006	2007	2008
Maldives	2003	2004	2005	2006	2007	
Moldova		2004	2005	2006	2007	2008
Mongolia						2008
Sao Tome and Principe	2003					
Saudi Arabia	2003	2004	2005			
Serbia				2006	2007	2008
Solomon Islands		2004		2006		
Spain (Basques)	2003	2004	2005	2006	2007	2008
Swaziland	2003	2004	2005	2006	2007	2008
Taiwan	2003	2004	2005	2006	2007	2008
Tonga				2006	2007	
Turkmenistan	2003	2004	2005	2006	2007	
UK (Northern Ireland)	2003	2004	2005	2006	2007	2008
Vanuatu					2007	

Behavioral indicators of civil conflict

The ICG CrisisWatch journals report basic narratives of one or two paragraphs that summarize what in-country analysts believe to be significant potential causes of conflict and instability. As I mentioned before, my dataset includes keyword-coded summaries of these narratives. I then used a two-part process to associate many of the keywords with various behaviors thought to provoke conflict in the civil war literature. These include measures of popular unrest, government and antigovernment behavior, conflict-related activities, and international reactions to states at risk.

The problem with naively analyzing these keywords is that ICG includes only a subset of all countries in their publications, and that subset is intentionally selected based upon a high likelihood of civil instability and conflict. This selection effect will bias any attempts to make inferences when analyzing conflict or conflict-related dependent variables.

Normally, a selection model could be estimated to control for this bias (Heckman 1979), but there are two problems with this type of approach. First, my temporal domain for my structural predictors of conflict ranges from 1946 to 2008, but the narrative data spans 2003 to 2011. Second, the narrative data is monthly while the structural data is yearly. Since I am only interested in the effects of the behavioral predictors of conflict from the narrative data, I use the structural model from Table 2 as the selection equation in a heckman probit. I introduce a three-year lag of this model to correspond with the narrative data temporal domain. The implicit assumption with this lag is that the structural conditions do not vary substantially over time,[7] and the lag allows analysis of the full narrative dataset while also controlling for the conditions in the country-year. Also, the selection model is estimated twelve times for each country-year in the model since the narrative data is monthly. This is not a problem for inferences regarding the narrative data since I am only trying to infer the effects of the variables in the outcome model.[8]

The dependent variable for the selection equations is coverage by the ICG narratives. As discussed above, these should be consistent with the UCDP civil conflict data but are limited to high-risk cases. The dependent variable in the outcome equations is the month-year observation of a UCDP conflict. According to the start-day precision variable from the dataset, there are several cases in which the start-month was unknown and assigned; however, all but one of these conflicts—Myanmar in 2005 and 2011—began prior to and lasted throughout my temporal domain. The results below include Myanmar for the entire year in 2005 and 2011, but results that eliminate this case do not affect the results. Therefore, missing month data is not really a concern

[7] This is an empirically true assumption since predictions of conflict correlate with three-year lags in the same state at a rate over .97.

[8] The repeated observations in the selection portion of the equation may artificially lower the standard errors of those coefficients. I re-estimated the three models using bootstrapped standard errors (200 iterations), but there were no identifiable effects on the coefficients or their standard errors in the outcome model. Another way to think about this, though, is that the structural model is actually time-invariant across the months in the outcome model. GDP, mountainous terrain, the presence of oil, non-contiguous territory, fractionalization, etc, do not really change from month to month in these models, so it makes theoretical sense to have the repeated observations. The only likely changes—and these are rare in the sample—result from changes in regime-type, its square, and the political instability variable. I re-estimated the Heckman models without these three variables, and the results were completely consistent with those reported here. Finally, since the ρ is statistically significant, the correlation across errors could affect interpretations of the outcome coefficients. Therefore, I also re-estimated the models using randomly selected months, dropping the other eleven months from the analyses. Again, the results remained stable across these models.

in this sample. Similarly, I consulted the episode-end data for the UCDP cases, but there were no cases with unknown end dates in the temporal domain of the analyses that follow.

Table 4 provides estimates of the effects of various behaviors based on the keyword indicators in the ICG CrisisWatch reports. The six columns in the table correspond to three separate types of monthly lags in each of the measures—the first two columns have one-month lags of each independent variable, followed by two columns each of two-month and three-month lags. Each set of two columns presents estimates for an unconditioned (logistic) analysis of the civil conflict measure and the outcome model of a conditioned (heckman probit) analysis. I label the latter conditioned since the coefficients in the conflict equation also incorporate the selection process.[9] Table 5 provides substantive effects estimates for the statistically significant variables in the one-month lag models.

Several interesting patterns emerge in the estimates, and many of the null results are surprising as well. First, protests *reduce* the likelihood of civil conflict in each model. Conditioning the case by its underlying likelihood of conflict changes its substantive effect by half, but in all cases the implication is that protests are associated with fewer conflicts. This could indicate that protests have a longer-term effect on conflict, with protests indicating some sort of grievance that is eventually followed by war. However, the statistically significant, negative coefficient is probably better interpreted as a function of overall state stability and capacity for handling contentious issues. Thus, a selection effect may exist in which unstable governments cannot allow protests to occur for fear that they will devolve into open conflict. Meanwhile, riots have no demonstrable effect on the likelihood of conflict in any of the models—naive or selection-corrected. Riots are more immediate manifestations of public grievances. Thus, it is unlikely that there is a longer-term or more complicated association with civil conflict.

Conditioning the cases affects how arrests are associated with civil conflict. Arrests are more likely in certain types of states, and, after taking this into account, the presence of arrests dampens the likelihood of civil conflict in the state. Nevertheless, arrests are the only government-related behavior analyzed that actually affects the likelihood of civil conflict in the following month. Repression, discussion of elections and election-related issues, or coups and coup attempts each has no effect in any of the models.

Among the conflict variables, the presence of rebels predicts conflict in every model. Fearon and Laitin (2003) argued that the structural determinants of conflict—mountainous terrain, low GDP, etc—are good predictors of the presence of rebels since these conditions allow primitive supplies for the groups as well as places to hide. However, the analyses of the CrisisWatch summaries suggests the relationship is even stronger than would initially be suspected. Even after controlling for the effects of selection with terrain, the potential for ethnic division, large populations, and poor development, the presence of rebels has a strong added effect on conflict likelihoods. It doubles the state's chances of intrastate conflict according to the estimates in Table 5.

This finding associating the existence of a rebel group with civil conflict is important because it provides the first direct evidence of a mechanism that was previously only suspected by most large-N studies. According to the sample I use here, structural variables are correlated with the

[9]The number of observations vary some across the two model sets due to missing data in the selection equations, but these do not substantially affect the estimates that are presented. Estimates with similar case sets are nearly identical. I omit presentation of the selection equation to save space, but the results are entirely consistent with those reported in the structural model of Tabel 2.

outgrowth of rebel groups, but the correlation is weak. Only thirty percent of high-mountainous terrain states[10] have rebel groups within their borders; therefore, there remains an exogenous variable that encourages rebel group formation within these states and that variable has much higher effect on conflict than we would predict with our current structural models.

The analyses of ethnic issues also confirm previous suspicions. The structural model predicts well which states are likely to have ethnic issues; still, after controlling for that effect, the presence of ethnicity-related issues in the narratives predicts future conflict well. The variable has one of the largest substantive effects in the models.

Perhaps surprising is that the presence of terrorism is not associated with civil conflict in any of the models. Indeed, the standard error of the measure is greater than the coefficient in all but one of the estimates. The measure uses terrorism or terror mentions, but even analyses restricted to only terrorist incidents still provide no statistically significant effects.

Several variables also assess the effects the international community can have on at-risk states, but, once again, some of these findings may seem counter-intuitive. First, sanctions predict conflict in two of the unconditioned models, but controlling for state selection suggests this association is most likely due to chance. Consistent across all models is the association of conflict with both mediation and calls for mediation (so long as the mediation involves the United Nations) and discussions of peacekeeping. Both variables *increase* the likelihood of observing conflict the following month. Of course, there is likely to be a selection effect in these models since mediation and peacekeeping are implemented and discussed when conflicts are ongoing. The Heckman model only controls for the structural determinants of likely conflict states, not the underlying tensions prior to conflict that may occur in the state. Nevertheless, the interpretation of the findings is straightforward: the presence of both variables increases—not decreases—the likelihood that the conflicts will still be ongoing one, two, and three months later, and this is true even after including a lag of the dependent variable which controls for the conflict environment the state.

The ρ in each of the conditioned models is statistically significant and negative. This implies that any unobservable variables correlated with the selection equation is a negative predictor of the outcome equation of a UCDP conflict. This makes sense when we consider the ICG's tendency to over-report low-level, latent conflicts in democracies such as the Basques in Spain or Northern Ireland after 2000. These are not represented in the UCDP data, and the variable guiding ICG to cover these cases remains unobserved.

Table 5 provides a nice comparison of the overall substantive effect differences between naive estimates of the *CrisisWatch* narratives and the narratives conditioned by underlying conflict likelihoods. In only a couple of cases do the rankings of the predicted probabilities change between the two models—and then only slightly. In most cases, too, the effects of each variable are dampened somewhat by the conditioning of selection, but that is not the case for the ethnicity and peacekeeper variables. The differences for the latter are quite small. Ethnic issues, though, have substantially larger effect in the conditioned estimates.

Overall, these results provide added support for a state-capacity explanation of civil conflict. Cohesive rebel groups emerge when the government is weak and unable to counter their formation

[10] Average mountainous terrain is only slightly higher in states with rebel groups. The log of mountainous terrain is 2.68 in states with rebel groups and 2.51 in states without.

Table 4: Behavioral influences on the likelihood of civil conflict

Lag of explanatory variables:	1-month		2-month		3-month	
DV: UCDP conflict	Unconditioned	Conditioned	Unconditioned	Conditioned	Unconditioned	Conditioned
UCDP (lagged one month)	0.889***	0.459***	0.864***	0.447***	0.852***	0.443***
	(0.079)	(0.048)	(0.083)	(0.050)	(0.088)	(0.053)
Popular unrest						
Protests	-0.453***	-0.250***	-0.425***	-0.240***	-0.483***	-0.281***
	(0.114)	(0.067)	(0.120)	(0.071)	(0.128)	(0.075)
Riots	-0.178	-0.200	-0.192	-0.226	-0.116	-0.186
	(0.359)	(0.222)	(0.380)	(0.239)	(0.387)	(0.247)
Government behavior						
Arrests	-0.245*	-0.143*	-0.270*	-0.155*	-0.266*	-0.164*
	(0.118)	(0.071)	(0.126)	(0.075)	(0.134)	(0.080)
Repression	-0.030	-0.149	0.041	-0.137	0.239	-0.041
	(0.258)	(0.171)	(0.268)	(0.181)	(0.274)	(0.189)
Election issues	0.049	0.039	0.011	0.009	0.049	0.018
	(0.109)	(0.065)	(0.116)	(0.069)	(0.121)	(0.073)
Coups/coup attempts	-0.250	-0.145	-0.083	-0.039	-0.158	-0.084
	(0.361)	(0.203)	(0.368)	(0.211)	(0.410)	(0.238)
Conflict						
Terrorism	-0.141	-0.056	-0.120	-0.029	-0.072	-0.009
	(0.109)	(0.066)	(0.114)	(0.069)	(0.121)	(0.073)
Rebels	1.018***	0.588***	1.033***	0.603***	1.024***	0.609***
	(0.079)	(0.048)	(0.083)	(0.050)	(0.088)	(0.054)
Ethnicity issues	0.656***	0.478***	0.625***	0.433***	0.669***	0.449***
	(0.155)	(0.098)	(0.165)	(0.104)	(0.175)	(0.111)
International reaction						
Sanctions	0.520*	0.253	0.516*	0.259	0.466	0.220
	(0.217)	(0.133)	(0.231)	(0.142)	(0.244)	(0.151)
Non-UN Mediation	0.155	0.129	0.070	0.068	0.091	0.063
	(0.195)	(0.115)	(0.217)	(0.127)	(0.232)	(0.136)
UN Mediation	0.910***	0.527***	0.942***	0.555***	0.994***	0.584***
	(0.119)	(0.074)	(0.124)	(0.078)	(0.133)	(0.083)
Peacekeeping	0.509**	0.314**	0.528**	0.319**	0.501*	0.335*
	(0.193)	(0.117)	(0.201)	(0.123)	(0.218)	(0.134)
Constant	-2.305***	-1.255***	-2.284***	-1.244***	-2.270***	-1.244***
	(0.065)	(0.042)	(0.069)	(0.045)	(0.073)	(0.049)
ρ		-0.196***		-0.189***		-0.166***
		(0.048)		(0.049)		(0.050)
N *ICG reports*	5,838	5,112	5,126	4,491	4,489	3,929
N *unconstrained*		12,623		12,002		11,440

Unit of analysis is the country-month (2003-2011), including all countries reported by the ICG for their CrisisWatch publication. Columns with *unconditioned* results are estimated using logistic regression. Columns with *conditioned* results are estimated with a Heckman-type model with structural controls for selection effects associated with civil conflict. Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Predicted probabilities of civil conflict the following month

Substantive effect of...	Unconditioned	Conditioned
Is there a cohesive group of rebels in the country?	+115.5%	+106.5%
Is UN mediation taking place or being talked about?	+85.0%	+79.0%
Was there a UCDP conflict in the previous month?	+94.1%	+75.4%
Are ethnic issues prominent?	+56.1%	+72.5%
Are there peacekeepers or is a mission being considered?	+42.0%	+44.3%
Are sanctions in place or being considered?	+42.6%	+33.1%
Are there popular protests?	-28.7%	-27.6%
Have there been arrests by the government?	-16.5%	-16.1%
Base probability of civil conflict in next month of ICG-covered country: 17.2%		

(Fearon and Laitin 2003), and ethnicity issues are more likely in divided states that lack a cohesive national identity (see, Author XXXX). Similarly, it takes a strong government to be able to allow popular protests without being threatened (Kitschelt 1986, Weiss 2013, Ritter 2014). Arrests of agitators are important, or, more nefariously, a strong government can arrest opposition leaders without threats of backlash among the population. The presence of peacekeepers, mediation, and sanctions are all likely to increase the chances of future conflict, and all are likely when the government itself is unable to police the state and establish its sovereignty (see, for example, Hultman, Kathman and Shannon 2014)

Finally, notice, too, that no real information is provided by the narratives on either greed or grievance or any of the possible underlying causes of conflict. Instead, the policymaker focus tends to be on events ongoing in the state and what behaviors are related to conflict escalation and expansion. These provide agency to the structural conditions arguments, but they still do not answer why conflicts occur.

Moving Forward

These analyses have demonstrated that a convergence of policy-based analysis and large-N approaches to studying the underlying dimensions of conflict likelihood can be quite useful. Using conflict narratives from the ICG's *CrisisWatch* publication, I was able to demonstrate that several causal mechanisms associated with conflict in the large-N literature do indeed play a large part in determining whether a country experiences civil violence. This is true even after controlling for the selection effects inherent in moving from all countries to the interesting cases covered by policy analysts.

The factors provoking conflict include likely suspects such as the presence of rebels, ethnic issues, and UN mediation and peacekeepers, and conditioning the ICG cases by selection effect rarely changed these associations. Interestingly, many behaviors often associated with conflict episodes—such as riots, terrorist activities, coups and coup plots, and elections—have no effect in any of the models analyzed. As a whole, the findings support the argument that the capacity of the state to deal with conflict-related behaviors is of primary importance in understanding when violence erupts. This is true regardless of the political, geographic, and economic pre-conditions of the state, and these findings affirm that large-N structural models do quite well when predicting the states likely to breakdown into civil conflicts.

What is notably missing from the structural models is a fair treatment of the effects of regime type. Structural conditions do not do well when guessing which states the ICG will cover, and a

majority of these cases are due to low-intensity, separatist movements in strong democratic states. This affirms the connection between state capacity and the prevention of large-scale civil conflict, but it underscores the need to think more about how democracies and near-democracies deal with minority groups.

Finally, the analyses should provide some caution for those who examine only the events on the ground without understanding first the structural conditions constraining the population and government. Several behavioral variables often assumed linked to conflict were, in fact, not related to civil disruptions at all, and a naive reading of the statistically significant associations found using only narrative data would overestimate the predictors of conflict in most cases. Of course, the structural conditions will never be able to provide an answer for when conflict is likely to occur in the dangerous-environment states or what actions are likely to escalate the conflict. Combined inferences are important.

References

Carter, D.B. and C.S. Signorino. 2010. "Back to the future: Modeling time dependence in binary data." *Political Analysis* 18(3):271.

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

Dube, Oeindrila and Juan F Vargas. 2013. "Commodity price shocks and civil conflict: Evidence from Colombia." *The Review of Economic Studies* 80(4):1384–1421.

Fearon, J.D. and D.D. Laitin. 2003. "Ethnicity, insurgency, and civil war." *American Political Science Review* 97(01):75–90.

Gibler, Douglas M and Steven V Miller. 2014. "External territorial threat, state capacity, and civil war." *Journal of Peace Research* p. 0022343314531003.

Gleditsch, K.S. 2002. "Expanded trade and GDP data." *Journal of Conflict Resolution* 46(5):712–724.

Heckman, James J. 1979. "Sample selection bias as a specification error." *Econometrica: Journal of the econometric society* pp. 153–161.

Hegre, Håvard, Tanja Ellingsen, Scott Gates and Nils Petter Gleditsch. 2001. "Toward a Democratic Civil Peace? Democracy, Political Change, and Civil War, 1816–1992." *American Political Science Review* 95(1):33–48.

Hendrix, Cullen S. and Idean Salehyan. 2012. "Climate Change, Rainfall, and Social Conflict in Africa." *Journal of Peace Research* 49(1):35–50.

Hendrix, Cullen S and Sarah M Glaser. 2007. "Trends and triggers: Climate, climate change and civil conflict in Sub-Saharan Africa." *Political geography* 26(6):695–715.

Hultman, Lisa, Jacob Kathman and Megan Shannon. 2014. "Beyond keeping peace: United Nations effectiveness in the midst of fighting." *American Political Science Review* 108(04):737–753.

Kitschelt, Herbert P. 1986. "Political opportunity structures and political protest: Anti-nuclear movements in four democracies." *British journal of political science* 16(01):57–85.

Miguel, Edward, Shanker Satyanath and Ernest Sergenti. 2004*a*. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy* 112(4):725–753.

Miguel, Edward, Shanker Satyanath and Ernest Sergenti. 2004*b*. "Economic shocks and civil conflict: An instrumental variables approach." *Journal of political Economy* 112(4):725–753.

Pemstein, Daniel, Stephen A Meserve and James Melton. 2010. "Democratic compromise: A latent variable analysis of ten measures of regime type." *Political Analysis* 18(4):426–449.

Regan, Patrick M. and Sam R. Bell. 2010. "Changing Lanes or Stuck in the Middle? Why are Anocracies More Prone to Civil Wars?" *Political Research Quarterly* 63(4):747–759.

Ritter, Emily Hencken. 2014. "Policy Disputes, Political Survival, and the Onset and Severity of State Repression." *Journal of Conflict Resolution* 58(1):143–168.

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

Themnér, Lotta and Peter Wallensteen. 2014. "Armed conflicts, 1946–2013." *Journal of Peace Research* 51(4):541–554.

Weinberg, Joe and Ryan Bakker. 2014. "Let them eat cake: Food prices, domestic policy and social unrest." *Conflict Management and Peace Science* p. 0738894214532411.

Weiss, Jessica Chen. 2013. "Authoritarian signaling, mass audiences, and nationalist protest in China." *International Organization* 67(01):1–35.

Young, Joseph K. 2012. "Repression, Dissent, and the Onset of Civil War." *Political Research Quarterly* 66(3):516–532.