A Framework for Measuring Leaders' Willingness to Use Force*

Jeff Carter
Department of Political Science
The University of Mississippi
jcarter3@olemiss.edu

Charles E. Smith, Jr.
Department of Political Science
The University of Mississippi
pssmith@olemiss.edu

Abstract

Political leaders' relative willingness to use military force is central to most theoretical models of interstate conflict. Unfortunately, existing measures of this important concept offer limited spatial or temporal coverage or are derived partially from leaders' participation in interstate conflicts. This paper outlines a strategy for combining data on leaders' personal attributes, political orientations, and psychological traits in a Bayesian latent variable framework with informative priors to construct measures of leaders' underlying willingness to use force for all national political executives between 1875 and 2004. Our framework produces indicators that offer multiple advantages over existing proxies used in the literature and yields measures of uncertainty that can be used to test theories about the relationship between uncertainty over leaders' preferences and interstate conflict. An application demonstrates that measures incorporating data on leaders' personal attributes, political orientation, and psychological willingness to challenge constraints outperforms measures based exclusively on leaders' personal attributes in models of interstate conflict initiation.

*Giacomo Chiozza is thanked for comments on a previous draft. We thank Michael Horowitz, Allan Stam, Kali Ellis, Katsunori Seki, Laron Williams, and Jonathan Keller for making their data publicly available and Scott Gartner for the suggestion to consider measuring leaders' preferences using latent variable techniques. We will make all data and code needed to replicate all analysis associated with this project publicly available upon publication.

1 Introduction

The increased theoretical and empirical focus on the role of political leaders arguably represents the most important development in international relations scholarship during the last twenty years. Central to understanding the behavior of any decision-maker is identifying his or her preferences over potential outcomes (Frieden 1999). Where the initial wave of leader-centric research focused on how political incumbents' desire to remain in power drove the policies they pursued (among many others, Bueno de Mesquita et al. 1999, Chiozza and Goemans 2004, Debs and Goemans 2010, Croco 2011), scholars increasingly consider how variation in leaders' personal attributes, beliefs, and decision-making processes influence patterns of foreign policy and interstate conflict (e.g., Horowitz, Stam and Ellis 2015, Colgan and Weeks 2015, Kertzer 2016). This research program has led to an accumulation of knowledge about the effects of particular leader attributes and characteristics, but does not explicitly analyze the influence of leaders' preferences regarding the use of force on interstate conflict processes because no explicit indicator of this key concept exists. This manuscript develops measures of leaders' underlying willingness to use military force based on their personal attributes, political orientation, and psychological characteristics.

Scholars commonly face two challenges when constructing measures of actors' underlying preferences over any issue. First, preferences are not directly observable. Second, empirical proxies of preferences often suffer from missing data and/or come in very different forms. Together, these challenges present a question of how to go about structuring measurement models that attempt to use all of the available information about a concept. Some covariates of a latent variable might best be reduced via principal components analysis, while others may present empirically and theoretically such that the data contain an IRT model-like structure, and are thus better reduced via a modern variant of the classic Rasch model (Rasch 1960). Political scientists have long relied on variants of the latter strategy (prominent examples include Poole and Rosenthal (1991), Martin and Quinn (2002), and Voeten (2000)), and we do so here.[1] With respect to missing data, we use conventional

[1] We are not suggesting that all of the data we have on hand are properly reduced in this way, and indeed do not in this paper use all of it effectively precisely because we have not yet fully explored PCA options. In the end, we expect PCA to be an essential feature of our framework.

imputation methods where possible, but notably introduce a novel method for incorporating information on small fractions of the overall data into a Rasch-like model via assignment of Bayesian priors to subsets of the overall data where the extra information exists. We then estimate from the data what (if any) contribution the added information makes to the underlying latent variable and use the estimates to structure the information's contribution to a one-dimensional measure of the latent trait.

The result is a framework that we use to construct measures of leaders' latent willingness to use military force. While still a work in progress both conceptually and empirically, the basic ideas underneath the framework are sufficiently mature to animate a preliminary paper (this one), and we have some initial (and quite promising) results. Our measures are based on publicly available data sets of leaders' personal attributes and background experiences (Ellis, Horowitz and Stam 2015), political orientations (Seki and Williams 2014), and psychological traits (Keller 2005, Hermann 2005). Our strategy is to use the information contained in each underlying data set (e.g., the spatial and temporal coverage of data on leaders' background experiences and nuanced measures of their stated policy positions and psychological willingness to challenge environmental constraints) for data reduction to a one-dimensional scale. Our results generate measures of leaders' latent willingness to use military force for 2,965 national political executives between 1875 and 2004. They are noisy and imperfect. Moving forward, our goal is to make them less noisy and fashion a data and code environment wherein we and others can change the underlying data and our assumptions about how the data best fit together (e.g., distributional assumptions) to produce measures for their own purpose.

The preliminary measures yielded by our framework offer four advantages over indicators of leaders' willingness to use force used in the literature. First, such measures should be better predictors of interstate conflict initiation than the underlying data on which they are based. Second, they offer greater spatial and temporal coverage than indicators based on leaders' psychological traits or political orientation. Third, they do not structurally depend upon data that indicate whether certain leaders have initiated conflicts, as this indicator must stand as the discriminator on the question of noise reduction in entropy-based tests of predictive validity. Fourth, they include

estimates of the uncertainty around leaders' latent preferences for conflict. This is important for many reasons, perhaps most notably because it provides analysts with a variable that can be used to test theoretical claims about how uncertainty over decision-makers' preferences influence crisis bargaining and interstate conflict.

The remainder of the paper contains our preliminary ideas about how to initiate a measurement development program that accomplishes these objectives, and more importantly, establishes a conceptual framework for building the aforementioned environment wherein we and others can continually better accomplish them. The next section provides an overview of existing research that analyzes leaders' willingness to use military force. The third and fourth sections describe the data we use and provide more details about our approach and four preliminary measures. The fifth presents an application of our measures in models of interstate conflict initiation. The paper concludes with a brief summary.

2 Existing Research

All rational choice theories of politics assume that decision-makers' valuation of possible outcomes influence the choices they make (Frieden 1999). Variation in decision-makers' willingness to use force is typically incorporated into theoretical models of interstate crises by assuming "hawkish" leaders pay a lower subjective cost for participating in a conflict than "dovish" leaders (Bueno de Mesquita and Lalman 1992, Schultz 2005, Wolford 2007). This implies, among other things, that political leaders' underlying willingness to use military force should influence the conditions under which they are willing to initiate interstate conflicts. Indeed, empirical research from multiple traditions conclude leaders with hawkish characteristics are more likely to initiate conflicts than those with dovish characteristics. The most prominent of these literatures connect patterns of interstate conflict to leaders' psychological characteristics, political orientation towards the use of force, or personal attributes and background experiences.[2]

While mainstream international conflict research focused nearly exclusively on system- and

[2]See Carter and Chiozza (N.d.) for a longer review of this literature.

state-level analyses in the decades following the publication of Waltz (1959), scholars working in the foreign policy analysis tradition located the source of states' behavior in the psychological traits and beliefs of political leaders.[3] Research on political executives' psychological traits and "operational code" directly speak to leaders' underlying willingness to use military force. Analyses of leaders' psychological traits and their implications for interstate conflict are closely associated with the work of Margaret Hermann. Hermann argues leaders differ along a number of dimensions (e.g., nationalism, cognitive complexity, and need for power) that shape variation in their underlying willingness to challenge political constraints, openness to new information, and motivation for pursuing policies (among others, Hermann 1980, Kaarbo and Hermann 1998, Hermann et al. 2001). To test these arguments, Hermann developed cross-national data on the leadership styles and psychological traits of political executives and elites during the post-World War II era based on their political speeches (Hermann 1987, Hermann 2005). Empirically, leaders who are more willing to challenge political constraints, have lower levels of cognitive complexity, or have higher levels of distrust, are more likely to pursue aggressive foreign policies (Kowert and Hermann 1997, Keller 2005, Keller and Foster 2012, Foster and Keller 2014).

A related strand of research focuses on political leaders' operational codes. Leaders' operational codes refer to their 1) philosophical beliefs that shape their views of the world and political goals and 2) instrumental beliefs regarding the methods that will best allow them to achieve their goals (George 1969, Walker 1983, Renshon 2008). Scholars argue that a leader's operational code influences her decisions and her state's foreign policies. For example, leaders who believe the world is fundamentally one of competition and goals are best pursued with aggression are more likely to engage in conflictual behavior (Schafer and Walker 2006). Most empirical operational code scholarship tends to focus on the specific beliefs of a single leader or a small group of leaders. Articles by Walker (1995) and Walker, Schafer and Young (1998), for example, consider how the operational codes of U.S. President Woodrow Wilson and U.S. President Bill Clinton and U.K. Prime Minister Tony Blair influence their decision-making. Unfortunately, the idiographic nature of these studies makes systematic, cross-national analyses of the relationship between leaders' operational codes

[3] Hudson (2005) provides a nice account of the relationship between foreign policy analysis scholarship and mainstream international relations research.

and their states' participation in interstate conflicts rare.

A second research tradition uses characteristics of a leader's government or party to identify her relative willingness to use military force. In particular, scholarship on differences among democratic states' foreign policy and conflict behavior commonly argues that leaders of right-governments and parties are relatively more hawkish and pro-military than leaders of governments and parties on the left (Schultz 2001, Palmer, London and Regan 2004). Scholars consistently find empirical support for the implications of this argument using indicators of decision-makers' policy preferences and left-right orientation using data derived from political parties' manifestos. For example, governments run by hawkish/right-wing leaders are more likely to initiate interstate conflicts (Palmer, London and Regan 2004, Arena and Palmer 2009, Clare 2010, Heffington 2016), fight longer conflicts (Koch 2009, Koch and Sullivan 2010), and have higher levels of military spending (Whitten and Williams 2011) than governments run by dovish/left-wing leaders. Taken together, these findings suggest democratic leaders affiliated with hawkish and right-wing parties are more willing to use military force than are democratic leaders associated with dovish and left-wing parties.

A third research program locates the source of variation in leaders' willingness to use military force in their personal attributes and background experiences. There is strong evidence from psychology, sociology, and political science that formative experiences influence individuals' attitudes, beliefs, and preferences (among others, Matthews 1954, Serpell 1981, Jennings and Niemi 2014). In general, early experiences are more likely to influence a leader's willingness to use force regarding an issue when they provided the leader with experiences relevant to conflict and/or reward risk taking (Horowitz, Stam and Ellis 2015, especially Chapters 1 and 2). As one might suspect, then, leaders who previously served in the military, on average, are more likely to initiate interstate conflicts than leaders with a civilian background (Horowitz and Stam 2014, Horowitz, Stam and Ellis 2015, Carter and Nordstrom Forthcoming). However, not all military service is the same. Leaders who saw combat, and thus more directly experienced the costs of war, are less likely to initiate conflicts than those leaders with a military background that did not involve combat (Horowitz and Stam 2014, Horowitz, Stam and Ellis 2015). Beyond national military service, leaders who participated in armed rebellion against the state are more likely to initiate interstate conflicts than

other political leaders (Horowitz and Stam 2014, Horowitz, Stam and Ellis 2015) .

Personal attributes and background experiences beyond prior military service or participation in a rebellion also have been linked to leaders' foreign policy decision-making. Political executives who led a revolutionary movement prior to obtaining power (Colgan 2013) or obtained office through irregular means (Debs and Goemans 2010) are disproportionately likely to initiate conflicts. Older leaders are more likely to initiate conflicts than younger leaders in most scenarios (Horowitz, McDermott and Stam 2005), although this relationship appears to be conditional on both regime type and how long a leader has been in power (Horowitz, Stam and Ellis 2015, Calin and Prins 2015). Perhaps less obvious from a theoretical perspective, Horowitz, Stam and Ellis (2015, pg. 67) report that leaders who had a job in the medical field prior to obtaining office or were considered "illegitimate" as a child were relative less likely to initiate a conflict than other political executives while those who had a creative occupation before assuming power were more likely to start an interstate conflict than other leaders.

An important variant of this approach uses multiple personal attributes and background experiences to construct measures of leaders' general preferences and orientation towards the use of force. The most prominent example of this strategy is Horowitz, Stam and Ellis's (2015) *Leader Risk Index*. Their *Leader Risk Index* was created by estimating interstate conflict initiation as a function of thirty-three leader attributes and background experiences and calculating a given leader's average annual probability of initiating a conflict over his or her time in office (Horowitz, Stam and Ellis 2015, pgs. 67, 73-74).[4] One downside to this strategy is that the resulting measure should not be used to in statistical models attempting to estimate interstate conflict initiation or onset because the measure is based on whether or not a leader initiated a conflict. Taking a different approach, Carter and Nordstrom (Forthcoming) proxy leaders' underlying hawkishness using indices based on multiple personal attributes and background experiences (e.g., prior military service, a military education). Empirically, both Horowitz, Stam and Ellis's *Leader Risk Index* and Carter

[4]The description here is of the monadic version of the *Leader Risk Index*. They also construct a version of the *Leader Risk Index* based on directed-dyadic data, which is described on pages 118-119 of Horowitz, Stam and Ellis (2015). The data and code required to recreate both versions of the *Leader Risk Index* is provided in the replication materials available at `https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/ZK3FYV`.

and Nordstrom's indices of background experiences are significant predictors of interstate conflict.

Existing research makes clear that leaders' personal attributes and experiences, political orientation, and psychological traits influence patterns of interstate conflict. Taking a step back, the idea that leaders' relative willingness to use force influences their states' involvement in conflict underlies research in each of these traditions. This implies indicators of leaders' political orientations, psychological traits, and background experiences at least partially reflect their unobserved willingness to use force in an interstate crisis. Empirical measures from each of these research traditions have relative strengths and weaknesses with respect to what they can tell us about leaders' general willingness to use military force. For example, variables capturing leaders' psychological traits and political orientation are more nuanced and often have a stronger theoretical connection to leaders' underlying preferences than the dichotomous indicators of leaders' background experiences. Further, arguments about how personal attributes and background experiences influence leaders' willingness to initiate conflict are often about how prior experiences influence or reflect psychological traits and decision-making biases.[5] This suggests psychological and orientation-based measures might more precisely capture the underlying relationships and effects of interest than indicators of leaders' attributes and experiences prior to assuming office. Data on leaders' personal attributes and background experiences, though, offer far greater spatial and temporal coverage than measures of leaders' psychological traits and political orientation. Most notably, the Leader Experience and Attribute Data (LEAD) project (Ellis, Horowitz and Stam 2015) includes data on leaders' personal attributes and background experiences for the universe of 2,965 national political executives between 1875 and 2004. In contrast, measures of leaders' psychological traits are largely drawn from the post-World War II period (Hermann 2005) while data on leaders' hawk-dove and left-right orientations are limited to democracies after 1950 and concentrated among developed, parliamentary systems in Europe (Seki and Williams 2014, Heffington 2016).

The preceding discussion suggests using data on leaders' political orientation, psychological traits, *and* background experiences should yield a more comprehensive measure of leaders' general

[5]For example, participation in a rebel movement is thought to indicate a leader is relative risk-acceptant and, therefore, more willing to use military force than leaders who were not rebels (Horowitz, Stam and Ellis 2015).

willingness to use force than an indicator based exclusively on any one of the three approaches. There are two primary reasons for this. First, if leaders' willingness to use force is a function of their experiences, psychological traits, and political orientation, then indicators based on any one of these types of characteristics necessarily contain less information about the underlying concept than a measure based on all three characteristics. Second, a measure derived from data on leaders' background experiences, psychological traits, and political orientation could potentially leverage the advantages associated with indicators from each tradition while minimizing their less desirable qualities. For example, measures that offer the spatial and temporal coverage associated with data on leaders' personal attributes and background experiences while accounting for variation induced by their psychological traits and political orientation would incorporate the best features of indicators from each research tradition while minimizing their respective downsides. The next two sections describe the data and latent variable framework we use to construct such a measure.

3 Data

Our measures are built around data from the LEAD project (Ellis, Horowitz and Stam 2015). This initiative produced data on a range of personal attributes and experiences for all national leaders between 1875 and 2004. There are 2,965 political executives included in the leader-level version of LEAD.[6] We use thirty-six variables from the LEAD project. These variables cover leaders' military service, involvement in a rebel movement, sex, family life, occupations, and other experiences and are reported in Table 1.[7] The spatial and temporal coverage of the variables in the LEAD project is impressive, but, unfortunately, most of the indicators of leaders' personal attributes are missing observations. To address missing data concerns, we multiply imputed the values of the

[6]The leader-level LEAD data is available at `https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/ZK3FYV`.

[7]A description of how the variables in Table 1 are coded can be found in the codebook associated with the LEAD project (available at the link in footnote 1). Variables were recoded to produce dichotomous indicators and ensure higher values corresponded to more hawkish preferences/positive relationship with conflict initiation. The following variables were recoded to yield dichotomous measures: age, childtotal, and spousesinlife. The following variables from the LEAD data were recoded based on bivariate correlations with MID initiation: law, medicine, careerpolitician, business, aristocratlandowner, police, totalspouses, married, marriedinpower, divorced, childtotal, illegit, royalty, and orphanbinary.

missing observations using Amelia II (Honaker, King and Blackwell 2007). As one might expect, many of the variables in Table 1 are highly collinear. Importantly, some of this collinearity is due to the construction of the variables. For example, a leader cannot have served in combat if he or she was not in the military at some point prior to obtaining office. Not coincidentally, the bivariate correlation between *Military Service* and *Military Service-Combat* is 0.8. We therefore included a ridge prior of 5% to aid with the convergence of the imputation model.[8] The imputation model yielded five imputed data sets that were then combined to produce a single complete data set of thirty-six personal attributes and experiences for 2,965 national leaders.

Table 1: Variables from LEAD

Military Service	Older Leader	Journalism
Military Service - Non-Combat	Education	Law
Military Service - Combat	Number of Spouses	Medical
Win War	Married	Religion
Lose War	Married in Power	Activist
Military Career	Divorced	Career Politician
Military Education	Number of Children	Creative Occupation
Rebel	Parental Status	Businessman
Rebel - Win War	Legitimate Child	Aristocrat/Landowner
Rebel - Lose War	Royalty	Police
Irregular Entry	Orphan	Science/Engineer
Male	Teacher	Blue Collar

We incorporate information on leaders' political orientation towards the use of military force into our measures using data from Seki and Williams (2014).[9] These data are derived from the prime minister/presidential versions of the *rile*, *intpeace*, and *hawk* measures from the Seki-Williams Annual Government Partisanship data set. The measures, respectively, proxy a leader's general political orientation (*Right-Left*), support for peaceful international relations in general or with respect to specific countries (*International Peace*), and net support for military engagement with other countries (*Hawk*) based on data from the Manifesto Data Project (Volkens et al. 2013, version 16a). Notably, each variable is coded such that higher values suggest relatively more hawkish policy positions. For our purposes, Seki and Williams's (2014) measures are preferable to other indicators

[8]See pg. 22 of the Amelia II manual on this point: https://r.iq.harvard.edu/docs/amelia/amelia.pdf.
[9]Replication material available at `http://faculty.missouri.edu/williamslaro/govtdata.html`.

used to proxy democratic leaders' political orientations or hawkish/dovish preferences because they offer greater spatial and temporal coverage. The Seki and Williams data cover thirty-seven democracies between 1944 and 2014. In contrast, Heffington's (2016) data on leaders' hawkishness contains information on 25 democracies between 1951 and 2000. Ultimately, converting Seki and Williams's (2014) annual measures to the leader-level produces data on 398 political executives included in the LEAD data set (Ellis, Horowitz and Stam 2015).

Incorporating systematic data on leaders' psychological traits into a general measure of leaders' latent willingness to use force with substantial spatial and temporal coverage is challenging. This is because, as discussed above, most of the psychology-based research on political executives and conflict tends to analyze a small number of leaders in a given study. We use a measure of leaders' psychological willingness to challenge constraints from Keller (2005).[10] Keller's indicator is derived from Hermann's (2005) leadership trait data and covers 42 national leaders between 1937 and 1998. This measure, and the underlying data from Hermann, are based on an analysis of at least 50 speeches by national political leaders. In this case, speeches were coded for what they revealed about leaders' underlying "need for power," "task emphasis," "distrust of others," and "nationalism." Keller standardized and combined leaders' scores on these four indicators to create a single index (*Constraint Challenger*) that represents leaders' "willingness to challenge potential pacifying constraints in the pursuit of aggressive foreign policy behavior" (Keller 2005, pg. 211-212). Following Keller, we assume leaders who are relatively willing to challenge environmental constraints are relatively more willing to use military force than are leaders who are relatively more willing to accept their environmental constraints.

To sum up, we have thirty-six binary indicators of 2,965 leaders' personal attributes and background experiences from the LEAD project (Ellis, Horowitz and Stam 2015), three variables that tap into the political orientations of 398 leaders from Seki and Williams (2014), and one indicator that measures 42 leaders' psychological willingness to challenge environmental constraints from (Keller 2005). The next section describes how we use these data to construct measures of leaders'

[10] Replication data are available from `https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/10053`.

latent willingness to use military force.

4 Measures and Models

Our central concern is the formulation of a data reduction strategy to combine information from the three data sources described above. Our starting point is the LEAD data. A variety of data reduction techniques might seem ideal. For example, as previously mentioned, we have briefly experimented with principal components analysis and examined (not reported) a one-dimensional solution over the thirty-six indicators. It is highly correlated (.77) with with the latent variable in the first dimension-reduction model described below, and we intend to return to this strategy in the near future. Here, though, we begin with a Bayesian version of a Rasch (1960) model of the form:

$$Pr(Y_{ij} = 1) = logit^{-1}(\theta_i - \alpha_j) \tag{1}$$

where $Pr(Y_{ij} = 1)$ is the probability that the ith leader ($n = 2965$) has the jth characteristic ($J = 36$).

This is a very common model for student test score data (questions scored as correct (1) or incorrect (0)) in educational testing research. The $logit^{-1}$ term represents the inverse of the logistic function. In the education tradition, θ_i represents the "ability" of the student, and the α_j terms form cutpoints on the "ability" dimension around which the θ_i terms float. Again in testing terminology, the α_j terms should increase in value according to the increasing "difficulty" of the questions. In our circumstance, to the extent that the data fit the model, we should expect characteristic vectors containing many ones and few zeroes to be associated with low α_j values. An example in our data is the sex of a given leader. Approximately 99% of the leaders in our data set are male (scored as one); a leader's sex is thus an "easy question," and should be associated with a low α_j. On the other extreme, our data contains an indicator scored as "one" when and only when a leader joined a rebellion that subsequently won a war. This vector is of course dominated by zeroes (approximately

96%), and thus by analogy is associated with a "hard question" that should in turn be associated with a higher α_j estimate.

We do not expect all of the data from Table 1 to be a good fit to this model, as we suspect some attributes and experiences are more relevant than others in shaping leaders' preferences regarding interstate conflict. Still, this is a starting point. Using a Bayesian framework, we seek to estimate posterior distributions and associated moments for the parameters. Absent restrictions in the form of priors/hyperpriors, the model described in Equation 1 is not identified (see Gelman and Hill (2007, pp. 315-316)). For this paper we estimated four versions of the model, all with bland but weakly informative priors for the α_j and θ_i terms.[11] We used the probabilistic programming language Stan (Carpenter et al. 2016), a relative newcomer in the toolbox of Bayesian analysis, to estimate our models.[12] Stan is distinguishable from the familiar BUGS lineage (e.g. WINBUGS, JAGS, and OPENBUGS) of Bayesian programming/software in a variety of ways, but two stand out: (1) it utilizes Hamiltonian rather than Markovian mechanics as its formal model of probabilistic state transitions over Monte Carlo trials,[13] and (2) relatedly samples probability space using Hoffman and Gelman's (2014) "No-U-Turn" sampler. NUTs has properties that relieve the user of many sensitive tuning tasks that have slowed the application of Hamiltonian Monte Carlo to a variety of tasks to which it is well-suited. These features coupled with well-tuned defaults and excellent diagnostic documentation (credit to the Stan Development Team (2017)) make Stan efficient and user-friendly.[14]

The four versions of the model differ from each other in the following ways. A pair of base models contain only information from the LEAD data and produce estimates according the equation above – one using all thirty-six characteristics and one using a theoretically-motivated subset of these characteristics. Then we create a new version of each of these that includes four weighted (β) covariates that fashion the θ terms in combination with a unit normal error (υ_i), as in:

[11] The θ_i terms from the Rasch models/data are unit normal. The α_j terms are centered at zero and confined to the interval [-5,5].

[12] The appendix contains the Stan code to estimate one of our models.

[13] See Betancourt, Byrne, Livingstone, and Girolami (2014) and Betancourt (2017).

[14] Stan fit objects (we used the R implementation of Stan, rstan) contain a wealth of diagnostic information, all of which in the case of results described in this paper point to clean convergence. In all four runs, we used four HMC chains with 2000 iterations, half dedicated to burnin.

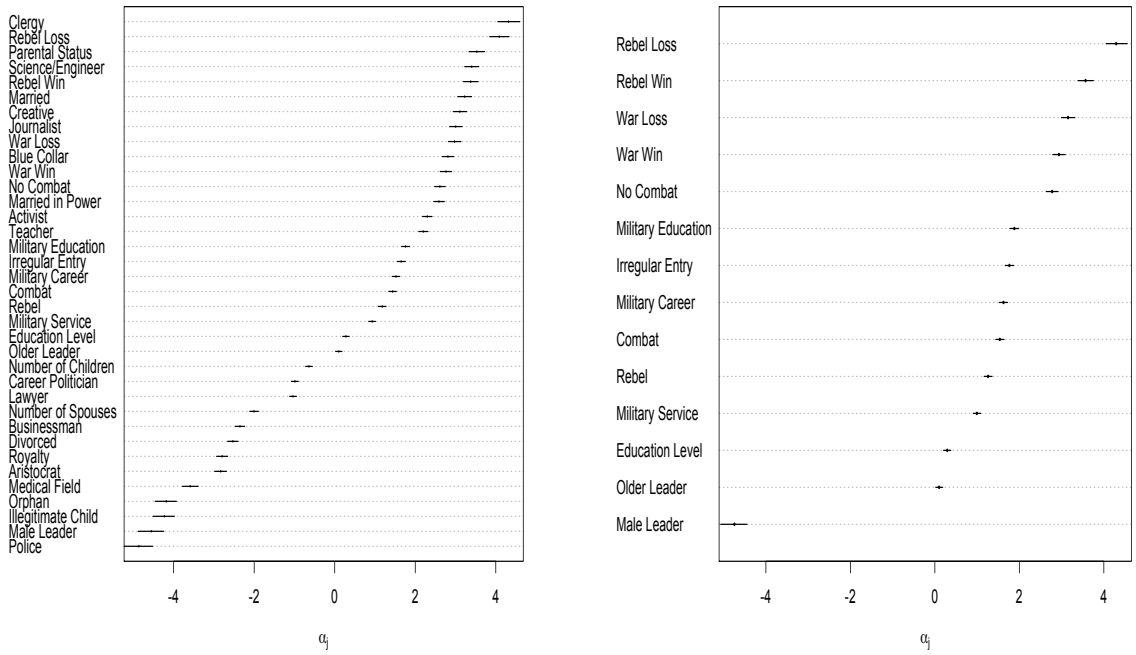
$$\theta_i = X_i\beta + \upsilon_i \tag{2}$$

, where υ_i is the latent trait estimated from the LEAD data. The four X vectors contain zeroes in every cell where the associated covariate is either missing or observed at its mean (they are all scaled unit normal). The β vector is estimated from the data. So $\theta_i = \upsilon_i$ if either the β vector or the X matrix is null. Otherwise, the variance of θ will be larger than the variance of υ, with the difference being a weighted (β) function of the variances on the X vectors. This set-up allows us to incorporate leaders' political orientations and psychological traits into our measures despite the fact that we have information about these characteristics for such a small percentage of the leaders included in our data.

As noted above, we estimated four models that yielded measures of leaders' underlying willingness to use military force for this paper. The models differ in the indicators of leaders' personal attributes and background experiences that underlie the models. The first two models are based on all thirty-six attributes from the LEAD project (Ellis, Horowitz and Stam 2015) listed in Table 1. The first measure, θ_1, is based exclusively on the thirty-sex personal attributes while the second, θ_2, incorporates the political orientation (Seki and Williams 2014) and psychological measures (Keller 2005). The third and fourth models are based on fourteen theoretically-motivated personal attributes and background experiences from the LEAD data that are thought to influence leaders' willingness to use force. The fourteen indicators are related to leaders' previous military experiences, involvement in a rebellion, manner of obtaining office, sex, age, and education.[15] The third model, θ_3, is based exclusively on the fourteen personal attributes while the fourth, θ_4, incorporates data on leaders' political orientation and psychological traits.

Figure 1 reports the estimates of α_j (with 95% credible intervals) for each of the four models. Recall that the values of α_j reflect how rare an attribute is in the data underlying each model, with more positive values reflecting attributes that are rarely observed in the data. Thus, as discussed

[15] More specifically, this model uses the following indicators of leaders' personal attributes and background experiences: *Military Service*, *Military Service - Non-Combat*, *Military Service - Combat*, *Win War*, *Lose War*, *Military Career*, *Military Education*, *Rebel*, *Rebel - Win War*, *Rebel - Lose War*, *Irregular Entry*, *Male*, *Older Leader*, and *Education Level*.



(a) Models 1 and 2

(b) Models 3 and 4

Figure 1: Relative Prevalence of Leader Attributes and Experiences (α_j) in Data for Each Model. Positive values imply attribute is less common.

above, it is very common to have male leaders (2nd most common trait in Models 1 and 2 and the most common trait in Models 3 and 4), but it is relatively rare that national leaders have participated in a rebellion that defeated the state (5th least common trait in Models 1 and 2 and 2nd least common trait in Models 3 and 4).

Our four measures allow us to make two substantively important inferences. First, and most importantly from our perspective, comparing the performance of measures based on the same set of personal attributes and experiences (i.e., θ_1 vs. θ_2 and θ_3 vs. θ_4) allows us to directly assess whether using data from multiple scholarly traditions yields more predictive measures of leaders' latent willingness to use force based exclusively on their personal attributes, political orientation, or psychological characteristics. Second, comparing measures yielded from models with analogous specifications in terms of $X_i\beta$ but different sets of personal attributes (i.e., θ_1 vs. θ_3 and θ_2 vs. θ_4) allows us to assess whether measures based on a wide-range of personal attributes outperform measures based on previous experiences thought to be more directly relevant to a leader's decision

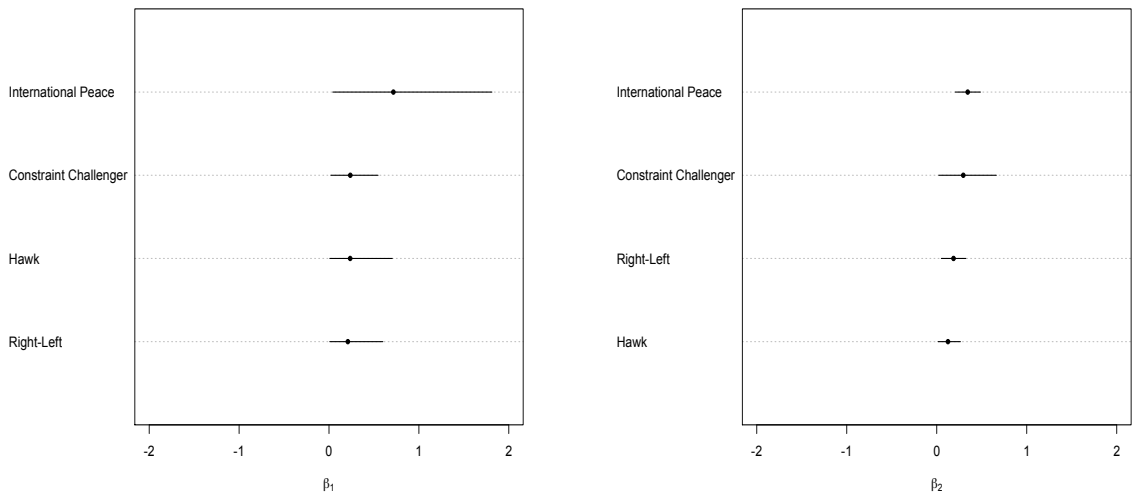
to use force. While limiting the indicators included in the model to theoretically-motivated personal attributes makes it more likely that we are omitting an empirically relevant factor, it also minimizes the losses in efficiency associated with including irrelevant variables in our model. Comparing the predictive power of analogous models based on different underlying data, therefore, identifies whether the gains in efficiency associated with a more limited set of personal attributes outweigh the consequences of omitting background experiences that influence leaders' willingness to use force in less obvious ways.

Table 2 reports the bivariate correlations between our four measures. Two patterns immediately emerge from Table 2. First, all of our measures are highly correlated. The lowest correlation between any two of our four measures is 0.787 (θ_1 and θ_4). Second, measures derived from models estimated on the same sets of personal attributes are more highly correlated than measures derived from models estimated on different sets of personal attributes. Indeed, these bivariate correlations are essentially one (0.997 for θ_1 and θ_2 and 0.987 for θ_3 and θ_4). Substantively, this suggests incorporating information about leaders' political orientation and psychological characteristics does little to change our measures of their latent willingness to use military force based exclusively on their personal attributes and background experiences. This is the case despite the positive effects our indicators of leaders' political orientations and willingness to challenge constraints has on our measures of leaders' latent hawkishness, as demonstrated in Figure 2.

Table 2: Correlations between Measures of Latent Hawkishness for All Leaders

	θ_1	θ_2	θ_3	θ_4
θ_1	1.000			
θ_2	0.997	1.000		
θ_3	0.789	0.789	1.000	
θ_4	0.787	0.791	0.987	1.000

Figure 2 reports the effects of *Right-Left*, *Hawk*, *International Peace*, and *Constraint Challenger* (with 95% credible intervals) on θ_1 (Panel A) and θ_2 (Panel B). The positive parameter estimates imply that, all else equal, leaders who score higher on these variables are relatively more willing to use military force. The results in Figure 2, therefore, suggest that leaders' political orientations and psychological characteristics influence their relative hawkishness. In particular, democratic



(a) β_1 (b) β_2

Figure 2: Estimated Effects of Political Orientation and Psychological Variables on Latent Hawkishness with 95% Credible Intervals.

leaders whose party platforms placed a low emphasis on international peace and those leaders who are relatively willing to challenge environmental constraints have a greater latent willingness to use military force.

Figure 2 and Table 2 seemingly contradict one another: variation in political orientations and psychological traits are associated with variation in leaders' latent willingness to use force (Figure 2) yet measures that are informed by these variables do not meaningfully differ from measures based exclusively on leaders' personal attributes and background experiences (Table 2). This tension is explained by the relative lack of data on leaders' political orientations and psychological traits. Our measures of leaders' latent willingness to use force are informed by leaders' political orientation or psychological willingness to challenge constraints in roughly 14% of cases (426/2965) and by political orientation and psychological traits in only 0.4% of cases (14/2965).[16] Tables 3 and 4, respectively, report the bivariate correlations between our measures for observations when a leader's latent hawkishness is informed by personal attributes and either political orientation or psychological indicators and all three types of data.

[16]Those fourteen cases are Eisenhower, Kennedy, Johnson, Reagan, Bush and Clinton from the United States; G. Eyskens from Belgium; De Gaulle from France; Adenauer from Germany; A. Papandreou from Greece; Inonu and Demirel from Turkey; and Ben Gurion and Eshkol from Israel.

Table 3: Correlations between Measures of Latent Hawkishness for Leaders with Personal Attribute and Political Orientation or Psychological Indicators

	θ_1	θ_2	θ_3	θ_4
θ_1	1.000			
θ_2	0.968	1.000		
θ_3	0.687	0.694	1.000	
θ_4	0.692	0.734	0.886	1.000

Table 4: Correlations between Measures of Latent Hawkishness for Leaders with Personal Attribute, Political Orientation, and Psychological Indicators

	θ_1	θ_2	θ_3	θ_4
θ_1	1.000			
θ_2	0.684	1.000		
θ_3	0.731	0.781	1.000	
θ_4	0.467	0.855	0.889	1.000

The correlations reported in Tables 3 and 4 indicate a greater divergence in measures of leaders' latent willingness to use military force when estimates are informed by multiple types of data. This suggests two things. First, including information about leaders' political orientation and psychological characteristics changes our estimates of their latent hawkishness. Second, and as suggested above, the limited effect of leaders' political orientation and psychological willingness to challenge constraints on our measures reflected in Table 2 is due to the relatively small number of cases these data inform. The next section assesses the relative ability of our measures to predict interstate conflict initiation.

5 An Application

Table 5 reports a set of models that estimate the initiation of a militarized interstate dispute (MID) (Ghosn, Palmer and Bremer 2004, version 3.1) on a directed-leader-dyad-year data set that covers the period from 1875 to 2001.[17] Model 1 estimates MID initiation as a function of an index based on the thirty-six personal attributes and background characteristics from the LEAD project listed in Table 1 and serves as a baseline to compare the performance of our four measures. Models

[17]The data set was taken from the replication material associated with (Horowitz, Stam and Ellis 2015) and is available at `https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/ZK3FYV`.

2-5 estimate MID initiation as a function of θ_1, θ_2, θ_3, and θ_4, respectively. The control variables included in each model reflect the specification used by Horowitz, Stam and Ellis (2015) and are commonly employed in quantitative models of interstate conflict initiation or onset.

Table 5 yields three noteworthy results. First, the coefficient estimates associated with θ_1, θ_2, θ_3, and θ_4 indicate that each of our measures is a statistically significant predictor of MID initiation. While perhaps a low bar to cross, this is important because it lends face validity to our approach for measuring leaders' underlying willingness to use military force.

Second, Table 5 indicates measures that incorporate information about leaders' political orientation and psychological traits outperform measures based exclusively on information about leaders' personal attributes and background experiences. This inference is based on the model fit statistics (AIC, BIC, and log-likelihood) associated with each model in Table 5. In terms of apple-to-apple comparisons, θ_2 and θ_4 have lower AIC, BIC, and log-likelihood statistics than θ_1 and θ_3, respectively. Further, both θ_2 and θ_4 outperform the baseline measure *Background Index* in terms of AIC, BIC, and log-likelihood statistics. Consistent with our claims above, then, our results suggest measures of leaders' willingness to use military force that incorporate information from multiple research traditions do a better job at predicting conflict initiation than measures derived from any single research tradition.

Third, our measures based on theoretically-relevant personal attributes and background experiences outperform measures that include a broader range of personal characteristics. The relevant comparisons here are between Model 2 (θ_1) and Model 4 (θ_3) and Model 3 (θ_2) and ¡Model ; 5 (θ_4). While θ_1 and θ_2 are based on more data from the LEAD project than θ_3 and θ_4 (36 personal characteristics vs. 14 personal characteristics), Models 4 and 5 yield lower AIC, BIC, and log-likelihood statistics than Models 2 and 3, respectively. This suggests that the gains in efficiency that follow from focusing on theoretically-relevant personal attributes outweigh the potential advantages of modeling leaders' underlying willingness to use force using data on as many personal characteristics as possible.

Table 5: Leaders' Latent Willingness to Use Force and Interstate Conflict Initiation, 1875-2001

	Model 1	Model 2	Model 3	Model 4	Model 5
Background Index	0.09*	–	–	–	–
	(0.01)	–	–	–	–
θ_1	–	0.38*	–	–	–
	–	(0.04)	–	–	–
θ_2	–	–	0.41*	–	–
	–	–	(0.04)	–	–
θ_3	–	–	–	0.31*	–
	–	–	–	(0.03)	–
θ_4	–	–	–	–	0.35*
	–	–	–	–	(0.03)
Democracy_1	0.37*	0.36*	0.37*	0.44*	0.45*
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Democracy_2	0.50*	0.50*	0.50*	0.51*	0.50*
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Joint Democracy	-0.77*	-0.77*	-0.77*	-0.77*	-0.76*
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Relative Capabilities	0.23*	0.24*	0.23*	0.23*	0.21*
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Defense Pact	0.82*	0.83*	0.81*	0.80*	0.78*
	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)
Contiguous States	3.50*	3.50*	3.50*	3.50*	3.50*
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
System Concentration	1.42*	1.52*	1.52*	1.22*	1.17
	(0.61)	(0.61)	(0.61)	(0.62)	(0.62)
Satisfied Dyad	-1.46*	-1.45*	-1.44*	-1.43*	-1.40*
	(0.25)	(0.25)	(0.25)	(0.25)	(0.25)
Peace Years	0.00*	0.00*	0.00*	0.00*	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Peace Years^2	-0.00*	-0.00*	-0.00*	-0.00*	-0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Peace Years^3	0.00*	0.00*	0.00*	0.00*	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
(Intercept)	-7.61*	-6.33*	-6.34*	-6.29*	-6.29*
	(0.27)	(0.23)	(0.22)	(0.22)	(0.22)
Observations	931799	931799	931799	931799	931799
AIC	16800.67	16805.59	16788.88	16803.08	16776.18
BIC	17411.41	17416.32	17399.61	17413.81	17386.92
$\log L$	-8348.34	-8350.80	-8342.44	-8349.54	-8336.09

Standard errors in parentheses

* indicates significance at $p < 0.05$

6 Conclusion

Political leaders' relative willingness to use military force lie at the center of nearly every formal and informal model of crisis bargaining and interstate conflict, yet the field does not have a comprehensive measure of this important theoretical concept. This paper introduces a framework for measuring leaders' underlying willingness to use force. More specifically, we use Bayesian Rasch models to estimate leaders' latent hawkishness based on their personal attributes, political orientation, and psychological willingness to challenge constraints. The resulting measures cover the universe of national political executives between 1875 and 2004 and have advantages over indicators used in the literature to proxy leaders' underlying willingness to use force on multiple dimensions.

This is very much a work-in-progress. We have multiple ideas for moving forward with this project. We briefly discuss three here. First, we plan to incorporate more data on leaders' political orientation and psychological traits into our models. In particular, we plan to add data on leaders' political orientations from the new Ideology of Heads of Government data set (Brambor, Lindvall and Stjernquist 2017)[18] and hope to gain access to data that offers greater coverage of political leaders' psychological traits and/or operational codes. Second, we plan to investigate whether performing logistic principle component analysis on leaders' personal attributes can improve the resulting measures of their latent willingness to use force. Extracting the most relevant factor(s) in the personal attribute data in this way would allow us to reduce the noise in these data and, hopefully, improve our measures. Third, we plan to conduct at least two additional applications with our measures. The first is a leader-year analysis of interstate conflict initiation and/or onset. While dyadic analyses are standard in the quantitative conflict literature, a leader-year unit-of-analysis represents a more straightforward and, arguably, more appropriate way to demonstrate the utility of our measures. The second application we plan to conduct will make use of the uncertainty around our measures of leaders' underlying willingness to use military force. Uncertainty about decision-makers' subjective costs of conflict figures prominently in theoretical accounts of interstate conflict onset and escalation (among many others, Fearon 1995, Wolford 2007). Scholars rarely directly test

[18] We thank Giacomo Chiozza for bringing this data set to our attention.

these arguments, though, because systematic measures of uncertainty about leaders' preferences have not existed. Thus, we plan to demonstrate how our measures of uncertainty can be used to test existing theoretical claims in the field.

While the our results are preliminary, we believe we have demonstrated the usefulness of our framework for measuring leaders' latent willingness to use military force and think there are multiple ways for us to improve upon the measures described here.

References

Arena, Philip and Glenn Palmer. 2009. "Is it Politics or the Economy? Domestic Correlates of Dispute Involvement in Parliamentary Systems." *International Studies Quarterly* 53(4):955–975.

Brambor, Thomas, Johannes Lindvall and Annika Stjernquist. 2017. "The Ideology of Heads of Government, 18702012. Version 1.5." .

Bueno de Mesquita, Bruce and David Lalman. 1992. *War and Reason: Domestic and International Imperatives*. New Haven, CT: Yale University Press.

Bueno de Mesquita, Bruce, James Morrow, Randolph M. Siverson and Alastair Smith. 1999. "An Institutional Explanation of the Democratic Peace." *American Political Science Review* 93(4):791–808.

Calin, Costel and Brandon Prins. 2015. "The Sources of Presidential Foreign Policy Decision Making: Executive Experience and Militarized Interstate Conflicts." *International Journal of Peace Studies* 20(1).

Carpenter, Bob, Andrew Gelman, Matt Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Michael A Brubaker, Jiqiang Guo, Peter Li and Allen Riddell. 2016. "Stan: A probabilistic programming language." *J Stat Softw* .

Carter, Jeff and Giacomo Chiozza. N.d. State Leaders and Foreign Policy. In *Oxford Encyclopedia of Foreign Policy Analysis*.

Carter, Jeff and Timothy Nordstrom. Forthcoming. "Term Limits, Leader Preferences, and Interstate Conflict." *International Studies Quarterly* .

Chiozza, Giacomo and H.E. Goemans. 2004. "International Conflict and the Tenure of Leaders: Is War Still *Ex Post* Inefficient?" *American Journal of Political Science* 48(3):604–619.

Clare, Joe. 2010. "Ideological Fractionalization and the International Conflict Behavior of Parliamentary Democracies." *International Studies Quarterly* 54(4):965–987.

Colgan, Jeff D. 2013. "Domestic Revolutionary Leaders and International Conflict." *World Politics* 65(04):656–690.

Colgan, Jeff D and Jessica LP Weeks. 2015. "Revolution, Personalist Dictatorships, and International Conflict." *International Organization* 69(01):163–194.

Croco, Sarah E. 2011. "The Deciders Dilemma: Leader Culpability, War Outcomes, and Domestic Punishment." *American Political Science Review* 105(3):457–477.

Debs, Alexandre and H.E. Goemans. 2010. "Regime Type, the Fate of Leaders and War." *American Political Science Review* 104(3):430–445.

Ellis, Cali Mortenson, Michael C Horowitz and Allan C Stam. 2015. "Introducing the LEAD Data Set." *International Interactions* 41(4):718–741.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.

Foster, Dennis M and Jonathan W Keller. 2014. "Leaders' Cognitive Complexity, Distrust, and the Diversionary Use of Force." *Foreign Policy Analysis* 10(3):205–223.

Frieden, Jeffry A. 1999. "Actors and Preferences in International Relations". In *Strategic Choice and International Relations*, ed. David Lake and Robert Powell. Princeton University Press Princeton, NJ pp. 39–76.

George, Alexander L. 1969. "The "Operational Code": A Neglected Approach to the Study of Political Leaders and Decision-Making." *International Studies Quarterly* 13(2):190–222.

Ghosn, Faten, Glenn Palmer and Stuart Bremer. 2004. "The MID3 Data Set, 1993-2001: Procedures, Coding Rules, and Description." *Conflict Management and Peace Science* 21(3):133–154.

Heffington, Colton. 2016. "Do Hawks and Doves Deliver? The Words and Deeds of Foreign Policy in Democracies." *Foreign Policy Analysis* .

Hermann, Margaret G. 1980. "Explaining Foreign Policy Behavior Using the Personal Characteristics of Political Leaders." *International Studies Quarterly* 24(1):7–46.

Hermann, Margaret G. 1987. "Handbook for Assessing Personal Characteristics and Foreign Policy Orientations of Political Leaders, Columbus, OH: Mershon Center, Ohio State University.".

Hermann, Margaret G. 2005. Assessing Leadership Style: A Trait Analysis. In *The psychological assessment of political leaders*, ed. Jerrold M. Post. The University of Michigan Press An Arbor, MI pp. 178–212.

Hermann, Margaret G, Thomas Preston, Baghat Korany and Timothy M Shaw. 2001. "Who Leads Matters: The Effects of Powerful Individuals." *International Studies Review* 3(2):83–131.

Honaker, James, Gary King and Matthew Blackwell. 2007. "Amelia II: A Program for Missing Data." http://gking.harvard.edu/amelia/.

Horowitz, Michael C, Allan C Stam and Cali M Ellis. 2015. *Why Leaders Fight.* Cambridge University Press.

Horowitz, Michael C. and Allan C. Stam. 2014. "How Prior Military Experience Influences the Future Militarized Behavior of Leaders." *International Organization* 68:527–559.

Horowitz, Michael, Rose McDermott and Allan C Stam. 2005. "Leader Age, Regime Type, and Violent International Relations." *Journal of Conflict Resolution* 49(5):661–685.

Hudson, Valerie M. 2005. "Foreign Policy Analysis: Actor-Specific Theory and the Ground of International Relations." *Foreign policy analysis* 1(1):1–30.

Jennings, M Kent and Richard G Niemi. 2014. *Generations and Politics: A Panel Study of Young Adults and Their Parents*. Princeton University Press.

Kaarbo, Juliet and Margaret G Hermann. 1998. "Leadership Styles of Prime Ministers: How Individual Differences affect the Foreign Policymaking Process." *The Leadership Quarterly* 9(3):243–263.

Keller, Jonathan W. 2005. "Leadership Style, Regime Type, and Foreign Policy Crisis Behavior: A Contingent Monadic Peace?" *International Studies Quarterly* 49(2):205–232.

Keller, Jonathan W and Dennis M Foster. 2012. "Presidential Leadership Style and the Political Use of Force." *Political Psychology* 33(5):581–598.

Kertzer, Joshua D. 2016. *Resolve in International Politics*. Princeton University Press.

Koch, Michael T. 2009. "Governments, Partisanship, and Foreign Policy: The Case of Dispute Duration." *Journal of Peace Research* 46(6):799–817.

Koch, Michael T and Patricia Sullivan. 2010. "Should I Stay or Should I Go Now? Partisanship, Approval, and the Duration of Major Power Democratic Military Interventions." *The Journal of Politics* 72(03):616–629.

Kowert, Paul A and Margaret G Hermann. 1997. "Who Takes Risks? Daring and Caution in Foreign Policy Making." *Journal of Conflict Resolution* 41(5):611–637.

Martin, Andrew D and Kevin M Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the US Supreme Court, 1953–1999." *Political Analysis* 10(2):134–153.

Matthews, Donald R. 1954. *The Social Background of Political Decision-makers*. Vol. 8 Garden City, NY: Doubleday.

Palmer, Glenn, Tamar R. London and Patrick M. Regan. 2004. "What's Stopping You? The Sources of Political Constraints on International Conflict Behavior in Parliamentary Democracies." *International Interactions* 30(1):1–24.

Poole, Keith T and Howard Rosenthal. 1991. "Patterns of Congressional Voting." *American Journal of Political Science* pp. 228–278.

Rasch, Georg. 1960. "Probabilistic models for some intelligence and achievement tests." *Copenhagen: Danish Institute for Educational Research* .

Renshon, Jonathan. 2008. "Stability and Change in Belief Systems: The Operational Code of George W. Bush from Governor to Second-term President." *Journal of Conflict Resolution* .

Schafer, Mark and Stephen G Walker. 2006. "Democratic Leaders and the Democratic Peace: The Operational Codes of Tony Blair and Bill Clinton." *International Studies Quarterly* 50(3):561–583.

Schultz, Kenneth A. 2001. "Hawks and Doves: Estimating Military Policy Positions from Election Platforms.".

Schultz, Kenneth A. 2005. "The Politics of Risking Peace: Do Hawks or Doves Deliver the Olive Branch?" *International Organization* 59(01):1–38.

Seki, Katsunori and Laron K Williams. 2014. "Updating the Party Government data set." *Electoral Studies* 34:270–279.

Serpell, James A. 1981. "Childhood Pets and Their Influence on Adults' Attitudes." *Psychological Reports* 49(2):651–654.

Voeten, Erik. 2000. "Clashes in the Assembly." *International organization* pp. 185–215.

Volkens, Andrea, Pola Lehmann, Nicolas Merz, Sven Regel, Annika Werner, Onawa Promise Lacewell and Henrike Schultze. 2013. "Comparative Manifesto Project (MRG/CMP/MARPOR).".

Walker, Stephen G. 1983. "The Motivational Foundations of Political Belief Systems: A Re-analysis of the Operational Code Construct." *International Studies Quarterly* 27(2):179–202.

Walker, Stephen G. 1995. "Psychodynamic Processes and Framing Effects in Foreign Policy Decision-Making: Woodrow Wilson's Operational Code." *Political Psychology* pp. 697–717.

Walker, Stephen G, Mark Schafer and Michael D Young. 1998. "Systematic Procedures for Operational Code Analysis: Measuring and Modeling Jimmy Carter's Operational Code." *International Studies Quarterly* 42(1):175–189.

Waltz, Kenneth N. 1959. *Man, the State, and War: A Theoretical Analysis*. New York: Columbia University Press.

Whitten, Guy D. and Laron K. Williams. 2011. "Buttery Guns and Welfare Hawks: The Politics of Defense Spending in Advanced Industrial Democracies." *American Journal of Political Science* 55(1):117–134.

Wolford, Scott. 2007. "The Turnover Trap: New Leaders, Reputation, and International Conflict." *American Journal of Political Science* 51(4):772–788.

7 Appendix: Stan Example

```
data {
  int<lower=1> K;               // # characteristics
  int<lower=1> J;               // # leaders
  int<lower=1> N;               // # observations
  int<lower=1, upper=K> kk[N];  // characteristic for n
  int<lower=1, upper=J> jj[N];  // leader for n
  int<lower=0, upper=1> y[N];   // presence of characteristic for n
  real x1[J];                   // Keller score for leader j
  real x2[J];                   // Seki-Williams pm_hawk score for leader j
  real x3[J];                   // Seki-Williams pm_rile score for leader j
  real x4[J];                   // Seki-Williams pm_intpeace score for leader j
}
parameters {
  vector[K] alpha4;              // position on line for characteristic k
  real<lower=0> beta4a;          // regression coefficient of x1
  real<lower=0> beta4b;          // regression coefficient of x2
  real<lower=0> beta4c;          // regression coefficient of x3
  real<lower=0> beta4d;          // regression coefficient of x4
  vector[J] error4;              // error term in the regression model
}
model {
  vector[N] holdervec4;
  vector[J] theta4;              // POTUOF for leader j
  alpha4 ~ normal(0,10);
  error4 ~ normal(0,1);
  for (j in 1:J)
    theta4[j] <- (beta4a * x1[j]) +
                 (beta4b * x2[j]) +
                 (beta4c * x3[j]) +
                 (beta4d * x4[j]) +
                  error4[j];
  for (n in 1:N)
    holdervec4[n] <- theta4[jj[n]] - alpha4[kk[n]];
  y ~ bernoulli_logit(holdervec4);
}
```