Peaceful Neighborhoods and Democratic Differences*

Mark David Nieman† Douglas M. Gibler‡

Forthcoming in *The Journal of Politics*

Abstract

Democracies are thought to behave differently than other states, particularly when cooperating in international institutions, such as alliances. We contend, however, that these democratic differences largely depend upon geopolitical environments that make cooperation possible. Though studies have demonstrated endogeneity between democracy and peace, few analyze the effects of this joint relationship on democratic differences in cooperative foreign policy behavior. We explore this argument using the alliance literature. We argue that the empirical finding that democracies are more reliable alliance partners is driven by the tendency of democracies to cluster in peaceful environments. Alliances are more likely to be "scraps of paper" when found in more dangerous environments. By jointly modeling regime type and political environment using data on alliance termination from 1920–2001, we show that alliance reliability is a function of the latter rather than the former. Our argument has important ramifications for a host of literatures focused on regime type, as well as current debates over the effectiveness of democratic deterrence.

*Gibler thanks the National Science Foundation (Awards #0923406, #1260492, and #1729244) for their generous support on research related to this project. Nieman thanks the National Science Foundation (Award #1728395) for their support for research. For questions and comments, please contact dmgibler@ua.edu.

†Department of Political Science, University of Toronto, Toronto, ON.

‡Institute for Social Science Research, University of Alabama, Tuscaloosa, AL.

Since the establishment of the empirical finding that democracies rarely fight one another, many studies have sought confirmation of democratic differences in other types of relationships. Democracies are thought to trade more with other democracies, are more likely to form and cooperate in intergovernmental organizations, and are more likely to ally with each other and be reliable partners. These are just a few examples of the larger democratic peace research program.

The explanations, however, often fail to model that democracies themselves tend to cluster in mostly peaceful geographical regions. As democratization is more likely in peaceful environments, analyses examining any type of democratic differences must be careful to separate the independent causal effect of democracy on political outcomes from that of the political environment. Since democracy is itself at least partially determined by the political environment, a failure to model both the direct and indirect effect of the political environment on policy outcomes may incorrectly attribute the effect of the political environment to democratic institutions. This, in turn, makes it easier to find statistically significant differences in foreign policy behaviors across regime types. We argue that once the political environment is accounted for—and this source of bias properly modeled—differences in foreign policy behavior between regime types are no longer evident.

We focus our analysis on democratic differences in alliance behavior, and build on a recent study finding democracies to be more reliable partners.[1] An often overlooked aspect of the alliance reliability argument is that alliances exist within specific geopolitical circumstances and that their commitments necessarily vary in salience across these conditions. The degree to which an alliance faces salient security threats, we contend, directly correlates with the degree that it is honored or maintained. Alliances that exist in high-threat environments are also the most "at-risk" of having their terms violated.[2] As democracies make up a

[1] As we demonstrate below in section *Beyond alliances*, our argument equally holds for other examples of democratic differences, such as international trade.

[2] We define alliance violation as any action that results in the termination of an alliance prior to its scheduled expiration, other than renegotiation or extension. Note this definition permits a violation to occur without an invocation, i.e. a state may preemptively end a commitment. We use the terms *alliance violation*, *termination*, and *abrogation* interchangeably.

disproportionate number of alliances that are unlikely to face security challenges—due to democracies arising and perpetrating in more peaceful conditions—a failure to appropriately model threat environment results in a conflation of the effect of regime type with that of the geopolitical environment.

We model these relationships by using a split-population logit with an instrumental variable to estimate the likelihood of conflict during the duration of an alliance, and introduce this likelihood into the study of democratic reliability. The split-population model allows us to statistically account for the risk of alliance termination that is due to the political environment. We build an instrumental variable and include it in the first equation to identify a state's threat level while also accounting for possible endogeneity between alliance reliability and threat environment. We find that the association between democracy and alliance reliability is not evident once political environment is modeled. Further, our analysis suggests the combined, threat-and-reliability model outperforms the reliability model alone.

Our argument also applies to other ancillary findings suggesting democratic differences, and we demonstrate this with a brief application to democratic trading partners. Any time peace affects cooperation—whether it is trade, institutions, or similar types of cooperation—the endogeneity we document will pose problems for confirming that democracies behave differently. We begin our argument in the next section with a brief review of the democracy and alliance literature.

Alliance behavior

Traditional alliance theory is replete with arguments that threats to the state cause alliance-making in order to deter aggressors. Morgenthau (1960) calls it external balancing: faced with threat and unable to respond quickly enough with an increase in internal capability, leaders seek partners in other states to help them avoid, or survive, against external threats to their sovereignty. Alliance-making in this manner forms a key component of traditional

realist theory (see also, Waltz 1979; Walt 1985), and most empirical studies find that threats do matter in determining whether alliances form (Siverson and Emmons 1991; Lai and Reiter 2000; Johnson 2017). The implication of these arguments and findings is that alliances covary with threat, and, when threat diminishes, the need for the alliance does as well. Alliances are, as the famous phrase puts it, "scraps of paper" to be torn as situations change.

Forming and maintaining an effective alliance to counter a threat is costly and diverts resources from other goals. A well-functioning alliance, in particular, requires commitment of budget and personnel, information-sharing, joint military exercises, etc (Morrow 1994; Fearon 1997). Alliances may also restrict the foreign policy actions of their members, as they often include precise conditions and actions required of each state (Leeds, Long and Mitchell 2000; Leeds et al. 2002).

Alliances do not, however, function only as responses to threat. Instead, alliances may also be used to facilitate a number of different tasks (Schroeder 1976; Altfeld 1984). These tasks range from resolving contentious issues (Weitsman 2004; Mattes and Vonnahme 2010), to facilitating diplomatic functions and policy-alignment (Lake 2009; McManus and Nieman 2019). Gibler (1996, 1997), for example, identifies a number of alliances that, rather than aggregate power, serve instead to resolve outstanding territorial disputes. Powers (2004, 2006) demonstrates that alliances sometimes serve economic purposes. Alliances may also be used as a method of gaining influence over smaller states (Morrow 1991; Lake 1996; Johnson 2015). Finally, alliances may be used to signal policy-alignment, specifically when bandwagoning with hegemonic powers (McDonald 2015; Nieman 2016; Mousseau 2019).

Some have tried classifying alliances into types, such as whether they include territorial settlement or economic clauses (Gibler 1997; Powers 2004), or the symmetry of members' capabilities (Morrow 1991), but classifying alliances this way often ignores that alliances can serve multiple purposes. The former requires that alliances which resolve issues, such as the 1887 pact between Prussia and Russia or the 1960 USSR and China alliance, are not

also, at least in part, power aggregating.[3] The latter forces an assumption that some trade-off between capabilities and autonomy is the primary reason for why major powers would partner with minor powers that offer little in the way of additional military capabilities. While asymmetric alliances may be an effort to "buy influence" from the perspective of major powers, they can also provide access and basing rights necessary to confront distant adversaries (Nieman et al. 2020). From the perspective of minor powers, these arrangements do supplement their security (McManus 2018) and increase the minor power's bargaining position vis-à-vis rivals in disputes (Langlois 2012). We build on this literature and argue that the functions of alliances are often varied, even within the same treaty, and are conditioned by the current political threat within their environment.

Democracies in alliance

Related to the research noting variation in the emphasis on security versus non-security concerns has been the growth of studies associating democracies with alliance behavior that is quite different from traditional alliance theory expectations. Democracies may engage in deterrent alliances, for example, but their commitments are seemingly not scraps of paper. Their commitments are more likely to deter other states and, when threatened, to be honored by the democracies involved.

The theoretical argument for this has focused on the idea that cooperation is more likely among similar types of states. Leeds (1999), for example, develops a model where cooperation is essentially a method of policy coordination, and leaders consider the likelihood of agreement fulfillment—foreign policy changes—when forming or proposing cooperation. Without the likelihood of fulfillment by the other actor, then there is little incentive to alter state policies when it will not be reciprocated. These audience costs seemingly make it more likely that democracies make better alliances and have longer-lasting cooperation.

A number of studies have empirically analyzed whether pairs of democracies tended to

[3]The Prussia-Russia alliance addressed disputes in the Balkans and the Dardanelles, while the USSR-China alliance resolved a border dispute between the two parties (Gibler 1997).

"flock together". Siverson and Emmons (1991) found some evidence that democracies were more likely to form alliances with other democracies, but there were strong period effects. The finding was consistent in the post-World War II era data, but not prior (see also Kimball 2006). Lai and Reiter (2000) revisited this empirical claim and found a strong relationship for joint threats to the dyad. Jointly-democratic dyads were more likely to be involved in defense pacts than mixed or non-democratic dyads, and regime similarity systematically predicted both defense pacts and other types of commitments in the dyad.[4]

Alliances composed of jointly democratic states also appear to last longer than those comprised of other types of states. Gaubatz (1996) found that jointly democratic alliance dyads lasted twice as long as other alliance dyads. There was no empirical difference, however, between mixed-regime and non-democratic dyads. This led Gaubatz (1996, 135) to conclude, "democracy by itself does not appear to either increase or decrease the ability of a state to make commitments to nondemocracies." Bennett (1997), using the same data, took the average number of liberal regimes in an alliance and found a positive, statistically significant effect on alliance duration. The substantive effects were especially strong, since all-liberal alliances increase the duration of an average alliance by almost fifteen years.[5]

Finally, democratic states have also been found to be more reliable alliance partners than nondemocracies. When confronted with threats to the alliance, democracies are more likely to honor the provisions of their alliances because their leaders risk sanction by their publics when reneging. Thus, among all the states in alliances, democracies are expected to be better partners—more reliable, less likely to terminate their alliances, and less likely to violate the terms of the agreement (Leeds 2003; Leeds and Savun 2007; Leeds, Mattes and Vogel 2009).

[4]McManus and Yarhi-Milo (2017) suggest that while democracies are more likely to engage in public acts of support, their cooperation with autocracies is often less public in order to avoid domestic backlash.

[5]It is worth noting that only a handful of alliances (less than 1% of the data) were comprised solely of democratic states at the time of alliance formation.

Do political institutions explain these differences?

Some evidence suggests that democratic differences in alliance behavior may be less well understood than it appears. For example, regarding alliance formation, Gibler and Wolford (2006) questioned the research design used by both Siverson and Emmons (1991) and Lai and Reiter (2000). Gibler and Wolford argued that these studies were not technically examining alliance formation but were instead identifying whether dyads were allied. By switching the analysis from whether a dyad was allied in a given year to focusing on a dyad at the time of alliance formation, Gibler and Wolford demonstrated that democracies were not more likely to form alliances; instead, states were becoming democratic *after* having formed an alliance.

Gibler and Wolford's (2006) analysis also showed that the peace provided from the deterring effect of large, regional defense pacts promoted the development of democracy. In fact, over 90% of jointly-democratic alliance-dyads exist within three broad, regional defense pacts: NATO (55%), OAS (29%), and the WEU (7%). The regional clustering associated with these regional defense pacts confirmed a more complicated relationship between democracies and alliance-making. It also hints that the distribution of democracies in alliance is at least partially determined by something within their political environment.

This finding raises questions about the reliability of other alliance outcomes associated with democracy. For example, given the logic of cooperation among similar regime types outlined by Leeds (1999) and Lai and Reiter (2000), it is clear why democracies may be more reliable allies with other democracies, but it is less clear why democracies would unilaterally restrict their options when interacting with non-democracies given the latter's expected higher degree of defection. While democratic states may simply be less willing than non-democracies to break their international agreements, this commitment is not evident in other policy areas. Democracies do not, for example, honor their monetary commitments (Simmons 2000), nor their territorial treaties (Chyzh 2014), more than other regime types. Moreover, Gartzke and Gleditsch (2004) looked at whether alliance partners intervene in response to their obligations and found that democracies were actually less reliable than other

states. Taken together, these additional results suggest that differences in alliance behavior often attributed to domestic institutions may instead be driven by some omitted factor.

Peaceful environments and democracy

A possible omitted factor is the political environment around a state: specifically, how threatening a political environment is may shape a state's foreign policy behavior (Vasquez 2009). For example, alliance formation is often motivated by the goal counteracting a threat (Kimball 2006; Johnson 2017). Yet, peace—or the lack of threat—encourages or even causes democracy (Gleditsch and Ward 2006; Gleditsch 2009; Gibler 2012). This creates a puzzle in terms of alliance behavior because, without a threat, democratic states should have no need for alliances. Nevertheless, democracies do make and maintain alliances.

Peace causing democracy is not a new argument, of course, and has developed over time and been integrated into the larger democratic peace project. Russett and Oneal (2001, 37), for example, contend in their foundational work that "[d]emocracy is easier to sustain in a peaceful environment," and "external threats become reasons or justifications for suspending normal civil liberties, elections, and constitutional government." Their model of a Kantian peace recognizes the endogenous "feedback loops" from peace to democracy, trade, and international organization, so there is an explicit recognition that peace at least partially causes democracy even among some of the staunchest democratic peace advocates.

The problem for those who study democratic differences is that, if unmodeled, *any* degree of endogeneity between the political environment and democracy will introduce bias into statistical estimates. Even a weakly-endogenous relationship, such as that suggested by Russett and Oneal (2001), will bias estimates of the coefficients, misattributing the effect of the political environment to that of democracy. Depending on the degree of the correlation, it may also lead to inflation/deflation of standard errors and, subsequently, altogether incorrect inferences. Given the large amount of effort that has been used to determine that democracies

are different from other states in their relations, this implies far-reaching concerns for studies examining the ancillary properties of the democratic peace.

If peace causes democracy (see, e.g., arguments by Thompson (1996) or Gibler (2012)), then democratic differences in a particular variable may simply underscore a more pervasive sample-selection process that made these cases observable. Alliances among democracies would, almost by definition, be more likely to be those formed or continuing after a threat has subsided. Thus, if peace is causally related to both democracy and cooperation, studies that fail to explicitly model this when looking at the effect of democracy on cooperation will suffer from a specific form of omitted variable bias: functional form misspecification. The omitted variables represent nonlinearities—such as those introduced by selection processes—between the dependent and independent variables (Heckman 1979; Signorino and Yilmaz 2003).

This is a problem facing many studies of the associated effects of democracy, such as the finding that democracies are more reliable alliance partners. Democracies exist within more peaceful political environments than other states (Gleditsch and Ward 2000; Ward and Gleditsch 2002; Gibler 2012). Without a clear, immediate threat, alliances that democratic states participate in have no need to deter potential enemies, and their security clauses are rarely invoked. This implies that countering threats is a less important goal for these alliances.

Contemporaneous security challenges are likely to affect whether an alliance is at risk of abrogation. States in dangerous or threatening environments are more likely to face threats, and these alliances face a greater expectation that their security clauses will be invoked. As the risk of invocation increases, current leadership in a member-state may re-evaluate the costs and benefits of maintaining the commitment versus terminating the alliance. France, for example, failed to support Czechoslovakia, instead reaching a deal with Germany for its annexation in 1938. Nearby territorial threats may also prompt termination, as leadership may want to focus on more pressing matters, rather than maintain a relationship that is less

relevant for immediate security. Egypt, for example, ended its alliance with Yemen in 1967 as a result of the Six-Day War and its aftermath.

A member-state may choose to terminate an alliance even without the alliance being formally invoked, as a previous commitment may risk entanglement, exacerbate an existing security threat, or antagonize a more pertinent ally. Afghanistan terminated its alliance with Turkey in 1950, as each turned towards different poles in the Cold War. Poland, Bulgaria, and other Warsaw Pact members ended relations with Yugoslavia, following that latter's expulsion from the Cominform. In some cases an ally may even be the source of danger as in 1915, when Italy invaded its ally, Austria-Hungary, in order to resolve longstanding territorial disputes while Austria-Hungary was occupied with other conflicts. In 1977, Somalia terminated its alliance with the USSR because the latter was supplying arms to Ethiopia.

Conversely, alliances in more peaceful environments are at a lower risk of termination, as their members do not face immediate security concerns. That democratic states are also likely to exist in peaceful political environments has profound implications for evaluating whether regime type affects alliance reliability. By implication, alliances among democratic states are less likely to be abrogated, more likely to last longer, and more likely to be institutionalized over time. Since states democratize in these same peaceful environments, ignoring the role of the peaceful political environment leads to the misattribution for the role of democratic institutions and alliance reliability. When we observe democracy correlated with reliable, durable, and institutionalized alliance partnerships, we are really observing the effects of peaceful environments on both democracies and alliance behavior.[6] This argument leads to two hypotheses:

H1 (Political environment): States in peaceful political environments are less likely to abrogate their alliance commitments.

H2 (Democratic institutions): Once the political environment is accounted for, democratic states are no more reliable alliance partners than other states.

[6]In quantitative studies, this means that, unless explicitly modeled, democracies are attributed the indirect effect of peaceful environments, *even if they include a measure for peaceful environment.*

Research design

We test the effect of a state's conflict environment and regime type on its propensity to violate alliance agreements using data from the Alliance Treaty Obligations and Provisions dataset (Leeds et al. 2002; Leeds and Mattes 2007). As we are interested in assessing characteristics of states that terminate or violate the terms of an alliance, our unit of analysis is the directed alliance member-year. To better compare to existing studies of alliance reliability, we use the same data and follow the same coding decisions as Leeds, Mattes and Vogel (2009). Leeds, Mattes and Vogel's (2009) sample includes all bilateral alliances formed between 1919–1989 and traces these for violations between 1919–2001, yielding a total of 234 bilarteral alliances.[7] We use a split-population logit estimator to probabilistically identify and separate alliances that exist in peaceful political environments—i.e. those at low risk of violation—from those alliances in threatening political environments.[8] We adopt an instrumental variable approach to account for any endogeneity between alliance reliability and militarized conflict.

Methodology

We expect alliances to face varying risk of invocation or termination based on each state's external threat environment. States in more threatening environments are at a greater risk of having their alliances invoked, which provides more opportunities to violate the alliance's terms, while states in safer environments have fewer opportunities to commit violations. Even without formal invocation, states in threatening environments are more likely to reevaluate the costs and benefits of their alliances, as they may face other more pressing security concerns. If our argument is correct, then estimates of the predictors of alliance violation,

[7]Leeds, Mattes and Vogel's (2009) sample excludes alliances that were formed after 1989 as there may not have been enough time to test their reliability.

[8]By limiting our sample to alliances, rather than modeling alliance formation, we treat alliances as weakly exogenous—leaders often have short tenures and short time horizons, thus treating alliances as a sunk cost. Any correlation between alliance formation and alliance violations should be negative, as states that intend to break their alliances are less likely to find states willing to form an alliance with them and, thus, opt out of the sample of observed alliances. We also perform a series of robustness checks that show no evidence of selection effects due to alliance formation in the Appendix.

which ignore these different types of environments, will recover biased estimates.

Ignoring the conditioning effect of threat environments, and treating all alliance observations as equally at-risk of entering the sample that may violate their alliance commitments—which is true for traditional binary-choice estimators, such as logit or probit—is a type of model misspecification (Heckman 1979; Signorino and Yilmaz 2003). Unfortunately, we cannot definitively know *ex ante* which alliances face high risk of invocation or termination; leaders do not often reveal information on whether they considered violating their alliance obligations. Instead, we only have data on whether an alliance was violated, but not direct data on the degree that the alliance is "at-risk."

To address these data limitations, we use a split-population logistic regression model (Xiang 2010; see also Beger et al. 2011). A split-population logit is a type of mixture model, in which an outcome variable is a function of two processes.[9] The logic of the estimator is that there are two populations in the data, and entry into each population can be estimated *probabilistically*. Though the structure of the alliance data does not let us directly observe which cases are actually in the "at-risk" pool, the sub-sample can be estimated. The estimator does this by using two equations: one equation that functions as the selector, identifying *relevant* observations to include in the at-risk sample, and a second equation that estimates the outcome of interest on these relevant observations. The *relevance* equation affects the *outcome* equation probabilistically: some cases are treated as more "at-risk" than others, and this probability conditions the estimates of the outcome equation.[10]

[9] All selection and zero-inflated models are types of mixture models, with the familiar censored probit-types of selection models (e.g., Heckman 1979; Sartori 2003) including data on the outcome of the first stage, and more recent extensions modeling selection when there are not data available on the outcome for the first stage (Xiang 2010; Nieman 2015, 2018; Bagozzi 2016; Bagozzi and Mukherjee 2012). The relationship between selection and zero-inflated models is such that the probit variant of the split-population model is mathematically equivalent to Poirier's (1980) bivariate probit with partial observability (Xiang 2010, 488).

[10] Partial observability models, including split-population and other zero-inflated models, have been shown to correctly recover the sign and significance for parameters, even if variables are specified in the wrong equation, permitting accurate hypothesis testing (Nieman 2015, 438-439), though there are some criticisms of their overall reliability (Rainey and Jackson 2017). In this particular case, however, the alternative to estimating a partially observed model is to simply ignore the assumed bias induced by the political environment, which is likely to result in inaccurate hypothesis testing and substantive effects (Xiang 2010). To address

More formally, the estimator treats the outcome variable as a function of two processes:

$$Y_i = 0 \text{ with probability } (1 - R_i) + (R_i)(1 - V_i) \quad (1)$$

$$Y_i = 1 \text{ with probability } R_i V_i \quad (2)$$

where R and V are cumulative distribution functions of a binary choice model (see Xiang 2010, 487-488). R_i represents the probability that a case is *relevant* to the sample—that the observation should be in the outcome equation, i.e. an "at-risk" state[11]—and conditions V_i, which represents the probability of violating an alliance. These probabilities can be specified as:

$$\Pr(Y_i = 0) = [1 - \Lambda(Z_i\gamma)] + [(\Lambda(Z_i\gamma))(1 - \Lambda(X_i\beta))] \quad (3)$$

$$\Pr(Y_i = 1) = (\Lambda(Z_i\gamma))(\Lambda(X_i\beta) \quad (4)$$

where Z_i and X_i are vectors of covariates associated with the relevancy and outcome equations, respectively, γ and β the accompanying parameter estimates, and Λ is the logistic link function. Equation 3 can, of course, be simplified as $\Pr(Y_i = 0) = [1 - (\Lambda(Z_i\gamma))(\Lambda(X_i\beta))]$. The likelihood function of the split-population logit is written as:

$$\mathcal{L} = \prod_{i=1}^{n} [(\Lambda(Z_i\gamma))(\Lambda(X_i\beta))]^{y_i} [1 - (\Lambda(Z_i\gamma))(\Lambda(X_i\beta))]^{1-y_i} \quad (5)$$

and estimates of β and γ are recovered via maximum likelihood estimation.[12]

The estimator treats cases where $Y = 0$ in the data as being the outcome of either (1) not at-risk, or (2) being at-risk, but not terminated or abrogated, whereas $Y = 1$ is the outcome of being both at-risk and terminated/abrogated. This modeling approach allows

concerns of reliability, we formally assess model fit and conduct robustness checks in the Robustness and Validity Checks section of the Appendix.

[11] This implies, of course, that the inverse, $1 - R$, is the probability of an observation selecting out of the "at-risk" subsample—i.e. being identified as "not at-risk."

[12] As the estimator is not included as an 'off-the-shelf' option with most statistical software, we include sample stata code in section 4 of the Appendix for interested readers.

us to statistically separate alliances that exist in non-threatening environments, and whose terms are unlikely to become salient, from alliances for which threats to member states increase the possibility of invocation or termination. As an example, suppose an alliance is formed during a relatively high-threat time period in which the likelihood of conflict in that dyad-year is 35%. The split-population logit would then assign 35% of the estimation to the relevance equation since it is part of the at-risk population of alliances. The remaining percentage of the estimation would be considered not at-risk and would be grouped with the alliances formed during more peaceful periods. The result of this weighting is analogous, in a sense, to including an interaction term, since the model corrects for the conditional effect of the sample-selection process. However, rather than interacting two variables, the interaction is between the full set of variables from the outcome and relevance equations.

Though this process does weigh each observation by its political threat environment, the test remains quite conservative. The likelihood of threat in any given dyad-year is often much smaller than the 35% figure used in our example, so for each alliance we are only assigning a small portion of its effect on the overall model to the relevance equation. Conversely, the current standard approach within this research tradition is to conflate such cases as peace settlements or trade pacts that have alliance clauses with the offensive and defensive pacts formed in the years prior to major wars. As we demonstrate below, this standard assumption significantly affects whether several key variables predict alliance reliability or failure.

Data

Our dependent variable is *alliance violation*, which captures whether a state abrogates its alliance commitments. We follow Leeds, Mattes and Vogel (2009, 469-470) and code *alliance violation* as 1 if state A violated the terms of an alliance. They code an alliance as abrogated if (1) a major provision is violated and governments do not agree to continue with the alliance or (2) one government unilaterally ends the alliance prior to its terms.[13] There are

[13]This definition implies that alliances can be terminated without necessarily being invoked to assist against an aggressor. For example, a state may preemptively terminate an agreement as the threat of war to

74 instances of *alliance violation* among the 234 alliances in the data set, roughly 32% of all alliances.[14]

Next, we specify the relevancy and outcome equations of the split-population logit (Xiang 2010). Following our theory, we specify the *relevance equation* with predictors related to a state's geopolitical threat environment. We expect that the absence of a threatening environment is associated with fewer reasons for a state to terminate its alliances, as the risk of invocation is lower. Alliances formed and maintained in more threatening environments, however, are more likely to be invoked and/or terminated. We specify the *outcome equation*, with predictors that have previously been identified to affect alliance reliability. We build on Leeds, Mattes and Vogel (2009) and include variables such as joint democracy and whether a state has had a change in their leader's societal coalitions, as well as dyadic- and alliance-specific features.

Relevance equation

Our primary measure of a state's threat environment is *territorial threat*. We conceptualize *territorial threat* as a threat to occupy and hold territory, as opposed to any militarized threat. We expect territorial threat to be the most important determinant of whether a state is at-risk of an alliance violation due to the targeted nature of this type of threats, as opposed to more diffuse system-wide threats, such as the Cold War. We operationalize *territorial threat* as the maximum predicted probability of a fatal militarized interstate dispute for state A for all contiguous neighbors. This value provides a continuous, latent measure of *territorial threat*. We construct a time-varying measure of *territorial threat* for each observation in the data set.

The predicted probability of a fatal militarized interstate dispute is estimated based on a

either party increases, prior to invocation, in order not to be dragged into a conflict, or a leader may decided to discontinue an agreement signed by their predecessor during peacetime.

[14] After accounting for missing data, there are 70 violations in the exact replication of Leeds, Mattes and Vogel (2009) reported in Table 1 Model 1, and 68 violations in the samples reported in Models 2 and 3. See Table A.12 in the Appendix for the complete list of abrogated alliances.

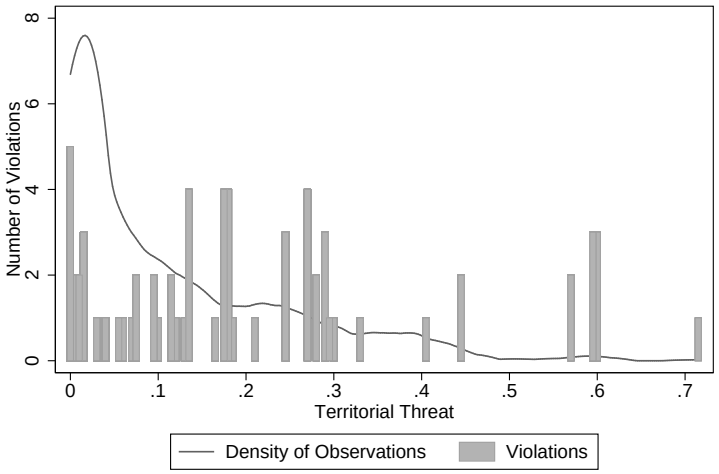
model from Gibler and Tir (2014, Table 1). Gibler and Tir emphasize *territorial* predictors of conflict among contiguous neighbors, such as previous peaceful and violent transfers of territory within a dyad, the highest level of militarization of a state's neighbors, previous territorial MIDs within a dyad, and the age of the dyad's border. They include controls for whether there is a shared colonizer, a civil war in either state within a dyad, or defense pacts with neighbors. The results for the logit model used to construct *territorial threat* are presented in Table A.1 of the Appendix.

We use this measure for two reasons: first, it best captures the theoretical concept of a dangerous neighborhood. While we include other possible sources of a dangerous political environment, we expect *territorial threat* to be the best identifier of "at-risk" alliance observations. There is strong evidence that territorial disputes are the single best predictor of militarized conflict (Bremer 1992; Vasquez 2009; Reed and Chiba 2010). Yet, it is the threat of conflict—rather than just its realization—that create dangerous political environments. Neighbors with territorial disputes are likely to appear as more salient threats than more distant adversaries, as neighboring states are usually able to quickly project power to their borders. Thus, we expect territorial threats to be the primary driver of whether alliances are "at-risk."

Second, using an instrumental variable—the predicted probability of conflict, *territorial threat*—rather than looking at observed militarized conflict is advantageous methodologically, as it helps avoid issues related to endogeneity between alliance reliability and militarized conflict.[15] We account for uncertainty in our estimate of the instrumental variable by taking 10 draws from the estimated distribution of the maximum predicted territorial threat, and use these to calculate point estimates and standard errors, following Rubin's (1987) formula for multiple imputation (see Boehmke, Chyzh and Thies 2016, for a similar

[15] An instrument of territorial threat allows us to attribute any correlation between territorial threat and violations solely to the effect of territorial threat, while ruling out any correlation/endogeneity between threat and alliance formation. Our instrument appears to be strong; the difference in the F-statistic between nested logits is >26, well above the threshold of 10 used to indicate that it does not suffer from weak instrument bias (Stock and Watson 2011). For the F-test, we estimate a logit with *alliance violation* treated as a function of the independent variables from the relevance equation.

Figure 1: Density of Territorial Threat and Frequency of Violations.



Note: The frequency of violations is overlayed with the kernel density of Territorial Threat within the sample from Table 1 Model 3.

re-sampling approach).[16]

As initial evidence of a relationship, we find that 57 of the 74 alliance violations ($\approx 77\%$) have a *territorial threat* above the median territorial threat for all allied states. Figure 1 displays the kernel density of *territorial threat* for the observations within the sample. It also reports the frequency of *alliance violations* at differing threat levels. The figure shows that at low threat levels, alliances are terminated less than expected by chance, while at high threat levels there are more terminations than would be expected.

We also control other factors that may influence the hostility of state A's political environment. We include the *number of borders* and *proportion of democratic borders*. States with more boarders have more opportunities for conflict (Vasquez 2009), though this is mitigated as a greater proportion of a state's neighborhood is democratic (Kadera, Crescenzi and Shannon 2003). These variables are based on Stinnett et al. (2002) and Marshall and Jaggers (2014). We control for whether state A is involved in a *rivalry* with any other state,

[16]The point estimate for each parameter is the mean of the 10 draws, or $\frac{1}{10}\sum_k^{10}\beta_k$, while the standard error is the average of the estimated variances within the datasets plus the variance in the point estimates across datasets, or $\sqrt{\frac{1}{10}\sum_k^{10}s_k^2+(1+\frac{1}{10})\sigma_\beta^2}$, where s_k^2 is the standard error for dataset k and σ_k^2 is the variance in β between datasets. See Rubin (1987). As few as 5 draws from the estimated distribution is sufficient to incorporate uncertainty (Mislevy 1991).

as this indicates already heightened tensions (Diehl and Goertz 2000; Rasler and Thompson 2006), as well as whether it is a *major power*, as these are more generally more active and attractive alliance partners (Chiba, Martinez Machain and Reed 2014). Data on rivalries and major powers are obtained from Klein, Goertz and Diehl (2006) and the Correlates of War Project (2016). We include an indicator variable for the *Cold War* to account for any systemic effects of a bipolar system. Economically developed states, often clustered geographically, may have complex economies that create norms that constrain conflictual behavior (Mousseau 2003). We operationalize *economic development* as the log of energy consumption per capita (Singer, Bremer and Stuckey 1972).[17] Finally, we control for whether a state is an *oil producer*, as such states are more conflict-prone (Colgan 2013).

Outcome equation

We rely on Leeds, Mattes and Vogel (2009, Table 1, Model 1) to specify our outcome equation. Leeds, Mattes and Vogel (2009) find that changes in a leader's core constituency and (the absence of) democratic institutions are highly correlated with alliance violations. Their conclusion affirms previous studies that find democratic governments to be seemingly more reliable alliance partners. *Democracy* is coded 1 if state A has a score ≥ 6 on the -10 to 10 Polity IV index (Marshall and Jaggers 2014), and *change in leader's societal coalition* is measured as a binary variable that is 1 if there is a change in the core domestic supporting coalition of state A in a year.

Leeds, Mattes and Vogel (2009) include several dyadic measures expected to decrease the reliability of international commitments. *Change in international power* is a binary variable coded 1 if there is a change of >20% in either state since the alliance was formed. *Change in political institutions* is a dichotomous variable coded as 1 if either state experiences a change in political institutions since the alliance was formed. *Change in external threat* is a binary variable coded as 1 if the level of external threat between the current year and the start year

[17]Economic consumption per capita is available for a broader time frame than GDP/capita and is highly correlated.

of the alliance changed by 30%.[18] *Formation of new outside alliance* is a binary measure coded as 1 if state A formed a new alliance.

Leeds, Mattes and Vogel (2009) also include four alliance-specific variables in their analysis. Each of these are expected to reduce the risk that an alliance is abrogated. *Asymmetry* is a dichotomous variable equal to 1 if an alliance includes a major and minor power. *Non-military cooperation* is a binary variable coded as 1 if an alliance has provisions linking nonmilitary issues to the alliance. *Ratification* is a dichotomous variable coded as 1 if an alliance was formally ratified. *Military cooperation* is a binary variable coded as 1 if an alliance includes provisions related to peacetime military cooperation. Lastly, cubic polynomials are included to account for temporal dependence (Carter and Signorino 2010).

Empirical analysis

Table 1 presents the results comparing *reliability* and *threat-and-reliability* models. We estimate three models: the first two are the *reliability* models, and are estimated with a traditional logit, while the third is the joint *threat-and-reliability* model, estimated with a split-population logit. Model 1 provides an exact replication of Leeds, Mattes and Vogel (2009) for estimating alliance reliability. Model 2 re-estimates Leeds, Mattes and Vogel, but restricts the sample to only those observations that are also in the split-sample logit, to make the models comparable. Finally, Model 3 reports the estimates of the *threat-and-reliability* model using the split-population logit, which includes both *relevance* and *outcome* equations.[19] The top of the table reports the *outcome* (violation) equation, and the bottom of the table reports the *relevance* (the degree an observation is "at-risk") equation.

[18] This measure differs substantially from our measure of *territorial threat* in terms of both composition and by focusing on whether there is a *change* in environment. The measure used by Leeds, Mattes and Vogel is based on a variable from Leeds and Savun (2007, 1127), which represents the sum of the capabilities (Correlates of War CINC scores) for politically relevant states (neighbors and major powers) that do not share an alliance and have a foreign policy affinity score (S score, a similarity score based on alliance portfolio) below the median value in their sample (median = .775).

[19] To ensure model robustness to likelihood optimizer choice and starting values, we estimated our results using several different optimization algorithms, and followed Veall's (1990) procedure in evaluate the log-likelihood function at a large number of randomly selected starting values.

The results are interpreted in a relatively straightforward way: positive coefficients indicate that increases in a variable make the outcome for that equation more likely. Hence, positive coefficients for variables in the *relevance* equation indicate an *increase* in the probability of being in the "at-risk" sub-sample of violating an alliance, while negative coefficients indicate a *decreased* likelihood of being in the "at-risk" sub-sample.[20] Similarly, positive coefficients for variables in the *outcome* equation indicate an increased likelihood of alliance termination.

Again, the first model is an exact replication of Leeds, Mattes and Vogel (2009, Table 1, Model 1): all coefficients and standard errors are the same as in the original study. Model 2 restricts the Leeds, Mattes and Vogel (2009) sample of cases to only those observations included in both the exact replication and the full split-population model. All of the parameter estimates are approximately the same, and all relationships are in the same direction and have the same level of significance as the original analysis. Model 2 thus provides a baseline from which to compare the *threat-and-reliability* model.

Model 3 estimates a split-population model where territorial threat and other factors related to the political environment are treated as part of the *relevance* equation, which identifies and assigns probabilistic weights that observations select into the pool of cases "at-risk" of terminating their alliances, estimated in the *outcome* equation. As expected, the coefficient on *territorial threat* is positive and statistically significant in the *relevance* equation. This result, consistent with H1, indicates that states are more likely to enter the at-risk population for alliance abrogation when they face greater territorial threats.

Turning to the *outcome* equation, we see that, after accounting for the underlying sample-selection process, the sign on *democracy* is now positive, though statistically insignificant. This suggests the negative and statistically significant effect associated with *democracy* in the *reliability* models (Model1 1 and 2) may, in fact, have arisen due to omitting the key role

[20]We focus on "opting into," rather than "opting out of," the at-risk sample. Our focus on observations being treated as "at-risk" or "opting in" to the outcome equation, of course, is the mathematical inverse of identifying the "zero-inflated" observations that "opt out." Reporting our results this way is consistent with previous studies using this estimation technique (e.g., Xiang 2010, 2017; Bagozzi 2016; Nieman 2015, 2018).

Table 1: Political Environment, Democracy, and Alliance Violations.

	Reliability (Replication)	Reduced Reliability	Threat and Reliability
Outcome Equation			
Democracy	-1.322*	-1.341*	1.358
	(0.393)	(0.401)	(1.924)
Change in Leader's Societal Coalition	0.889*	0.910*	5.504
	(0.436)	(0.444)	(5.507)
Change in International Power	0.803*	0.877*	2.996
	(0.330)	(0.340)	(2.251)
Change in Political Institutions	0.131	0.195	1.699
	(0.308)	(0.308)	(1.276)
Change in External Threat	0.421	0.425	0.763
	(0.270)	(0.278)	(1.255)
Formation of New Outside Alliance	1.070*	1.064*	2.447
	(0.253)	(0.261)	(1.626)
Asymmetry	-0.408	-0.503	-1.620
	(0.259)	(0.265)	(0.908)
Non-military Cooperation	-0.746*	-0.746*	-2.652*
	(0.257)	(0.261)	(1.198)
Ratification	-0.083	-0.134	1.525
	(0.354)	(0.355)	(2.387)
Military Cooperation	0.557*	0.575*	4.949*
	(0.180)	(0.186)	(2.434)
Time	-0.086	-0.069	-0.123
	(0.077)	(0.076)	(0.273)
Time Squared	0.001	-0.001	-0.003
	(0.004)	(0.004)	(0.010)
Time Cubed	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)
Constant	-4.448*	-4.464*	-3.226*
	(0.440)	(0.455)	(1.520)
Relevance (At-risk) Equation			
Territorial Threat			4.062*
			(1.025)
Proportion of Democratic Borders			-0.292*
			(0.123)
Number of Borders			0.122
			(0.063)
Major Power			0.417
			(0.516)
Rivalry			0.007
			(0.398)
Cold War			0.446
			(0.312)
Economic Development			-0.153
			(0.083)
Oil Producer			0.670
			(0.512)
Constant			-4.979*
			(0.558)
Log-likelihood	-352.394	-339.988	-306.909
Observations (Alliances)	6612 (223)	6395 (223)	6395 (223)

Note: *p <0.05, two-tailed. Standard errors in parentheses. Point estimates and standard errors in model 3 were calculated from 10 draws using Rubin's (1987) formula for multiple imputation to account for uncertainty in the *territorial threat* instrumental variable.

of the political environment. As democratic alliance members tend to experience lower threat levels than non-democracies, their alliances are less likely to have military provisions invoked

or those provisions are never salient.[21] Similarly, *changes in leader's societal coalition* also fails to reach any traditional level of statistical significance.

To more formally assess whether *democracy* exerts a null effect once the political environment is accounted for, we use a technique introduced to political science by Rainey (2014). The idea is to identify the smallest 'meaningful effect', and then determine whether the estimated quantity of interest meets this threshold. If the estimate (and its 95% confidence interval) fails to meet the threshold identified as the smallest 'meaningful effect', then the effect is marginal, and there is statistical evidence that the variable has little or no effect on the outcome of interest. If the estimate (and its 95% confidence interval) is equal to or surpasses the threshold, then the effect is appears to be non-marginal.

Given that estimated parameters in logit-based models are difficult to interpret directly, we follow Rainey's advice and focus on assessing whether the independent variable affects the predicted probability of the outcome of interest. In our case, we evaluate whether *democracy* exerts a meaningful effect by seeing if it reduces the likelihood of alliance violations by at least one-half of one percent (0.5%); i.e. the effect size should be less than -0.005. To do this, we take the first difference of the parameter of interest, holding all other parameters at their mean or modal values. The 95% confidence interval *democracy* [-0.001, 0.010] is *greater* than -0.005, indicating that the effect of *democracy* on reducing alliance termination is negligible. This result is consistent with H2 and indicates that the influence of political institutions may actually be attributable to the previously omitted political environment rather than the institutions themselves.

Applying the same test to *changes in leader's societal coalition*, we find different results. We expect a meaningful effect to increase the likelihood of alliance violations by at least one-half of one percent when there is a change in a leader's societal coalition. In this case, the 95% confidence interval [-0.001, 0.024] includes 0.005, indicating that, though it is not

[21] The mean *territorial threat* for democracies 0.059 with a standard deviation of 0.086 and N=1841, while the mean for non-democracies is 0.129 with a standard deviation of 0.135 and N=4554. A difference of means between the two samples is statistically significant with $p<0.001$.

statistically significant, we cannot rule out a meaningful effect for *changes in leader's societal coalition.*

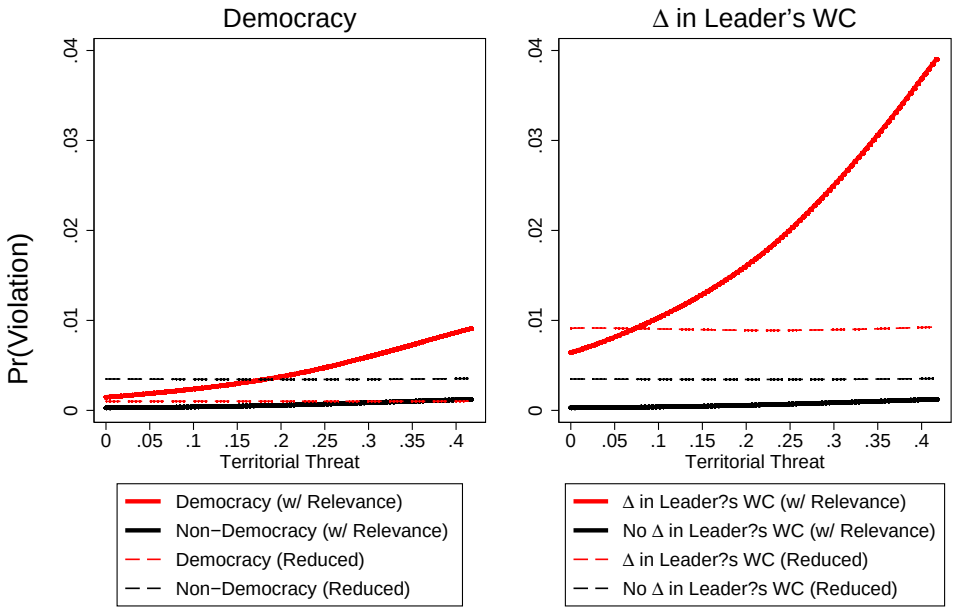
We formally compare and evaluate model fit statistics for the reduced and split-population models using the Clarke (2003, 2007) distribution-free test, in Table A.2 in the Appendix. The test indicates a strong preference for the *threat-and-reliability* model over the *reliability* model, even after penalizing the split-population model for its inclusion of additional parameters. We also assess the reliability and robustness of the estimates from the split-population logit in Section 3 of the Appendix, with the results being consistent across models.

Overall, the results are consistent with our theoretical expectations. Threatening environments affect the underlying propensity of states to enter the sample at-risk of violating/terminating their alliance terms. Moreover, once the political environment is accounted for, democratic institutions do not appear to exert a significant impact on whether alliance termination occurs.

Substantive effects

To illustrate the substantive effect of territorial threat and political institutions on alliance abrogation, we report predicted probabilities of alliance violations in Figures 2. In the first subfigure, we report predicted probabilities for two conditions: for democracies [red line] and for non-democracies [black line], after accounting for the level of territorial threat affecting the state [solid lines]. As a point of reference, we also compare these predicted probabilities to those from the reduced model [dashed lines], which does not account for the effect of territorial threat on an observation's probability of being part of the at-risk subsample. In the second subfigure, we repeat this procedure, reporting a change in the leader's winning coalition [red line] and when there is no change in the leader's winning coalition [black line], after accounting for territorial threat [solid lines]. The reduced model is again provided as a reference [dashed lines]. To make the substantive results more realistic, and to ensure that outliers are not skewing our interpretation, we visualize predicted probabilities of alliance

Figure 2: Predicted Probabilities of an Alliance Violation, Democracy, and Territorial Threat.



Note: Estimates from Table 1, Models 2 and 3. All variables held at mean or median. Figure reports the predicted probabilities for the middle 95% of values of Territorial Threat from the samples.

abrogation for the middle 95% of values of *territorial threat* from the estimated sample.

Figure 2(a) shows that, while democracy is associated with a lower likelihood of alliance termination in the reduced model [red dashed line is below the black dashed line], in the full model democracies are associated with an increased probability of an alliance violation [red solid line above black solid line]. Moreover, increases in territorial threat raise the probability of an alliance violation regardless of whether regimes are democratic [both solid lines increase], suggesting that the threat environment is driving the change. It is worth keeping in mind, of course, that the difference between democracies and non-democracies were shown to be negligible (or in the wrong direction), according to the test suggested by Rainey (2014).

Figure 2(b) looks at the impact of a change in the a leader's societal coalition. The figure demonstrates that a change is associated with increases in the probability of an alliance

violation [red lines are above corresponding black lines] and also that a change in the winning coalition increases the likelihood of termination as the degree of territorial threat rises [solid red line]. Territorial threat also increase the probability of alliance termination when there is no change in a leader's winning coalition [solid black line], though this effect is much smaller.[22] The figure shows that threat environment significantly influences the effect of changes in the leader's winning coalition on the probability of an alliance violation.

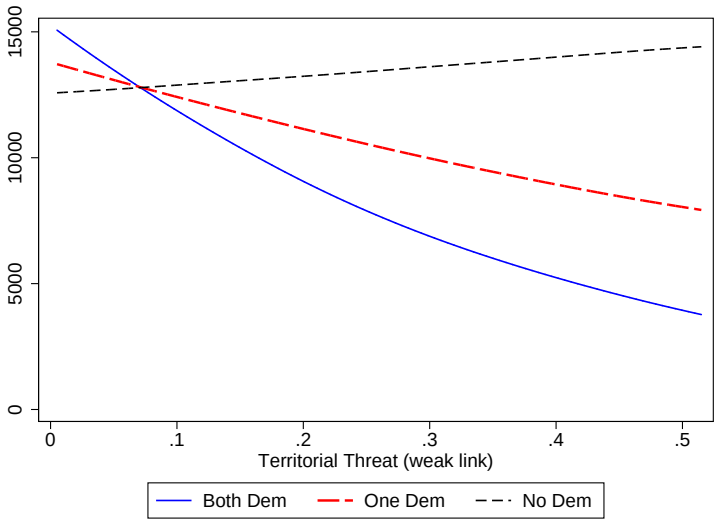
Notably, territorial threat matters at all points along the spectrum for both subfigures: in low threat environments, the presence of a democracy exerts only a small risk of alliance violation. Similarly, in the absence of an external threat, a change in a leader's winning coalition has very little effect on alliance abrogation. Instead, it appears abrogation becomes much more likely for democracies as the level of territorial threat increases. The same holds for a change in the leader's winning coalition: alliance violation is more likely as territorial threat increases. An implication is that ignoring threat environment significantly *overestimates* the effect of both democracy and changes in a leader's winning coalition at low levels of threat and significantly *underestimates* these effects at high levels of threat. Accounting for territorial threat and the political environment more broadly improves our understanding of the roles of political institutions and leadership changes in substantively meaningful ways.

Beyond alliances

Our core argument—that the empirical results of the virtues of democratic cooperation may be, in part, driven by the unmodeled causal relationship between peace and democracy—holds beyond the application to international alliances. To illustrate, we perform a secondary set of analyses, in which we specify and test the theory's application to the relationship be-

[22]The predicted probability of alliance violation in a democratic state, which experiences a change in their leader's winning coalition, is almost identical to the probability of a violation when there is a change in the leader's winning coalition in a non-democracy, at every territorial threat level. This suggests that the interaction of the two variables exerts little substantive impact, once territorial threat is accounted for.

Figure 3: Predicted Level of Trade by Regime Type at Varying Levels of Territorial Threat.



tween peace, democracy, and international trade.[23] In parallel to the theoretical discussion above, international trade and democracy both tend to grow and prosper in peaceful environments. In addition to stifling democratization, threatening political environments also dampen trade by disrupting trade routes, creating uncertainty and raising transaction costs, and even leading to the imposition of sanctions, trade barriers, and termination of trade agreements. As a result, the effect of democracy on trade may be conditional on the level of threat within the political environment.

To illustrate, we estimate a model of bilateral trade for 178 countries from 1948–1999, using data from Rose's (2004) canonical study (see full description of variables, tables, and figures in the Appendix). The results support our argument. Figure 3 displays the predicted values of (unlogged) bilateral trade for dyads consisting of two democracies, one democracy, and no democracies, holding all control variables are their median values. The predicted values highlight the degree to which trade among democratic states is conditioned by their political environment. Democratic pairs are expected to see a loss of approximately 40%

[23]For a full discussion of this application, see section *Additional Application: International Trade* in the Appendix.

of bilateral trade by increasing the threat level from the around 0 (no threat) to .2 (high threat). Dyads with one democratic members experience much less dramatic decreases, but still observe a loss of approximately 20% over the same range.

Conclusion

We began this paper by pointing out that endogeneity between peace and democracy may bias statistical estimates of the effects of many other democracy-related arguments, and we found evidence consistent with this in regard to international alliances and international trade. Democracies in alliances have been thought to be more reliable, but we demonstrate that this result is likely to be spurious. Democracy is more likely to take hold in peaceful international environments—environments that seldom produce security challenges that trigger alliance termination. Thus, democratic alliances are different from other types of alliances, but this has little to do with regime type.

Our argument and results have implications beyond the alliance literature, raising concerns regarding a number of second-order findings associated with the democratic peace research program more broadly. Current scholarship suggests that democratic states trade more often with other democracies, and democracies may also be more active in international governance. Each of these literatures, however, tend to pool samples without regard to threat environment, potentially biasing estimates by attributing sole explanatory power to political institutions rather than the underlying causal processes produced by peaceful political environments. This criticism extends to almost all studies that find some type of democratic difference in state behavior.

Also noteworthy is our finding that, under some conditions, traditional alliance theories may be correct. Quantitative analyses of alliances and conflict generally pool the sample of all cases to assess conflict-proneness and reliability. Our findings imply, however, that the nature of alliances may vary both across cases and over time, as alliances formed for security

may continue as tools for economic and other cooperation as threats subside. Future research may focus on developing more accurate measures of the context-specific purposes of alliances. Accounting for the threat environment may also help explain why alliances are correlated with peace in some periods and with conflict, or even the diffusion of conflict, in others (e.g., Levy 1981; Kadera 1998; Senese and Vasquez 2008).

Ultimately, democratic institutions may still affect state behavior once peaceful environments take hold, but more evidence is needed. Instead, our findings present an important set of questions for these long-accepted relationships: without controlling for the effect of dangerous environments, current estimates of the effect of democracy on behavior are biased and may be spurious.

References

Altfeld, Michael F. 1984. "The Decision to Ally: A Theory and Test." *Western Political Quarterly* 37(4):523–544.

Bagozzi, Benjamin E. 2016. "The Baseline-inflated Multinomial Logit Model for International Relations Research." *Conflict Management and Peace Science* 33(2):174–197.

Bagozzi, Benjamin E. and Bumba Mukherjee. 2012. "A Mixture model for Middle Category Inflation in Ordered Survey Responses." *Political Analysis* 20(3):369–386.

Beger, Andreas, Jacqueline H.R. DeMeritt, Wonjae Hwang and Will H. Moore. 2011. "The Split Population Logit (SPopLogit): Modeling Measurement Bias in Binary Data." SSRN Working Paper. http://ssrn.com/abstract=1773594.

Bennett, D Scott. 1997. "Testing alternative models of alliance duration, 1816-1984." *American Journal of Political Science* 41(3):846–878.

Boehmke, Frederick J, Olga Chyzh and Cameron G Thies. 2016. "Addressing Endogeneity in Actor-Specific Network Measures." *Political Science Research and Methods* 4(1):123–149.

Bremer, Stuart A. 1992. "Dangerous Dyads: Conditions Affecting the Likelihood of Interstate War, 1816-1965." *Journal of Conflict Resolution* 36(2):309–341.

Carter, David B. and Curtis S. Signorino. 2010. "Back to the Future: Modeling Time Dependence in Binary Data." *Political Analysis* 18(3):271–292.

Chiba, Daina, Carla Martinez Machain and William Reed. 2014. "Major powers and militarized conflict." *Journal of Conflict Resolution* 58(6):976–1002.

Chyzh, Olga. 2014. "Can You Trust a Dictator: A Strategic Model of Authoritarian Regimes' Signing and Compliance with International Treaties." *Conflict Management and Peace Science* 31(1):3–27.

Clarke, Kevin A. 2003. "Nonparametric Model Discrimination in International Relations." *Journal of Conflict Resolution* 47(1):72–93.

Clarke, Kevin A. 2007. "A Simple Distribution-free Test for Nonnested Model Selection." *Political Analysis* 15(3):347–363.

Colgan, Jeff D. 2013. *Petro-aggression: When Oil Causes War*. Cambridge University Press.

Correlates of War Project. 2016. "State System Membership List, v2016.". url: http://correlatesofwar.org.

Diehl, Paul and Gary Goertz. 2000. *War and Peace in International Relations*. Ann Arbor: University of Michigan Press.

Fearon, James D. 1997. "Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs." *Journal of Conflict Resolution* 41(1):68–90.

Gartzke, Erik and Kristian Skrede Gleditsch. 2004. "Why Democracies May Actually Be Less Reliable Allies." *American Journal of Political Science* 48(4):775–595.

Gaubatz, Kurt Taylor. 1996. "Democratic states and commitment in international relations." *International Organization* 50(01):109–139.

Gibler, Douglas M. 1996. "Alliances that never balance: The territorial settlement treaty." *Conflict Management and Peace Science* 15(1):75–97.

Gibler, Douglas M. 1997. "Control the Issues, Control the Conflict: The Effects of Alliances that Settle Territorial Issues on Interstate Rivalries." *International Interactions* 22(4):341–368.

Gibler, Douglas M. 2012. *The Territorial Peace: Borders, State Development, and International Conflict*. New York, NY: Cambridge University Press.

Gibler, Douglas M. and Jaroslav Tir. 2014. "Territorial Peace and Democratic Clustering." *Journal of Politics* 76(1):27–40.

Gibler, Douglas M. and Scott Wolford. 2006. "Alliances, Then Democracy: An Examination of the Relationship between Regime Type and Alliance Formation." *Journal of Conflict Resolution* 50(1):129–153.

Gleditsch, Kristian Skrede. 2009. *All International Politics is Local: The Diffusion of Conflict, Integration, and Democratization*. University of Michigan Press.

Gleditsch, Kristian Skrede and Michael D. Ward. 2000. "War and Peace in Space and Time: The Role of Democratization." *International Studies Quarterly* 44(1):1–29.

Gleditsch, Kristian Skrede and Michael D. Ward. 2006. "Diffusion and the International Context of Democratization." *International Organization* 60(4):911–933.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1):153–162.

Johnson, Jesse C. 2015. "The cost of security: Foreign policy concessions and military alliances." *Journal of Peace Research* 52(5):665–679.

Johnson, Jesse C. 2017. "External Threat and Alliance Formation." *International Studies Quarterly* 61(3):736–745.

Kadera, Kelly M. 1998. "Transmission, Barriers, and Constraints: A Dynamic Model of the Spread of War." *Journal of Conflict Research* 42(3):367–387.

Kadera, Kelly M., Mark J. C. Crescenzi and Megan L. Shannon. 2003. "Democratic Survival, Peace, and War in the International System." *American Journal of Political Science* 47(2):234–247.

Kimball, Anessa L. 2006. "Alliance formation and conflict initiation: The missing link." *Journal of Peace Research* 43(4):371–389.

Klein, James P., Gary Goertz and Paul F. Diehl. 2006. "The New Rivalry Dataset: Procedures and Patterns." *Journal of Peace Research* 43(3):331–348.

Lai, Brian and Dan Reiter. 2000. "Democracy, Political Similarity, and International Alliances, 1816-1992." *The Journal of Conflict Resolution* 44(2):203–227.

Lake, David A. 1996. "Anarchy, Hierarchy, and the Variety of International Relations." *International Organization* 50(1):1–33.

Lake, David A. 2009. *Hierarchy in International Relations*. Ithaca, NY: Cornell University Press.

Langlois, Catherine C. 2012. "Power and Deterrence in Alliance Relationships: The Ally's Decision to Renege." *Conflict Management and Peace Science* 29(2):148–169.

Leeds, Brett Ashley. 1999. "Domestic political institutions, credible commitments, and international cooperation." *American Journal of Political Science* pp. 979–1002.

Leeds, Brett Ashley. 2003. "Alliance Reliability in Times of War: Explaining State Decisions to Violate Treaties." *International Organization* 57(4):801–827.

Leeds, Brett Ashley, Andrew G. Long and Sara McLaughlin Mitchell. 2000. "Reevaluating Alliance Reliability: Specific Threats, Specific Promises." *Journal of Conflict Resolution* 44(5):686–699.

Leeds, Brett Ashley and Burcu Savun. 2007. "Terminating Alliances: Why do States Abrogate Agreements?" *Journal of Politics* 69(4):1118–1132.

Leeds, Brett Ashley, Jeffrey M. Ritter, Sara McLaughlin Mitchell and Andrew G. Long. 2002. "Alliance Treaty Obligations and Provisions, 1815-1944." *International Interactions* 28(3):261–284.

Leeds, Brett Ashley and Michaela Mattes. 2007. "Alliance Politics during the Cold War: Aberration, New World Order, or Continuation of History?" *Conflict Management and Peace Science* 24(3):183–199.

Leeds, Brett Ashley, Michaela Mattes and Jeremy S. Vogel. 2009. "Interests, Institutions, and Reliability of International Commitments." *American Journal of Political Science* 53(2):461–476.

Levy, Jack S. 1981. "Alliance formation and war behavior: An analysis of the great powers, 1495-1975." *Journal of Conflict Resolution* pp. 581–613.

Marshall, Monty G. and Keith Jaggers. 2014. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2013." http://www.systemicpeace.org/polity/polity4.htm.

Mattes, Michaela and Greg Vonnahme. 2010. "Contracting for Peace: Do Nonaggression Pacts Reduce Conflict?" *Journal of Politics* 72(4):925–938.

McDonald, Patrick J. 2015. "Great Powers, Hierarchy, and Endogenous Regimes: Rethinking the Domestic Causes of Peace." *International Organization* 69(3):557–588.

McManus, Roseanne W. 2018. "Making it personal: The role of leader-specific signals in extended deterrence." *Journal of Politics* 80(3):982–995.

McManus, Roseanne W and Keren Yarhi-Milo. 2017. "The Logic of "Offstage" Signaling: Domestic Politics, Regime Type, and Major Power-Protégé Relations." *International Organization* 71(4):701–733.

McManus, Roseanne W. and Mark David Nieman. 2019. "Identifying the Level of Major Power Support Signaled for Protégés: A Latent Measure Approach." *Journal of Peace Research* 56(3):364–378.

Mislevy, Robert. 1991. "Randomization-based Inference about Latent Variables from Complex Samples." *Psychometrika* 56(2):177–196.

Morgenthau, Hans. 1960. *Politics Among Nations*. 3rd ed. New York, NY: Knopf. Originally published in 1948.

Morrow, James. 1991. "Alliances and Asymmetry: An Alternative to the Capability Aggregation Model of Alliances." *American Journal of Political Science* 35(4):904–933.

Morrow, James. 1994. "Alliances, Credibility, and Peacetime Costs." *Journal of Conflict Resolution* 38(2):270–297.

Mousseau, Michael. 2003. "The Nexus of Market Society, Liberal Preferences, and Democratic Peace: Interdisciplinary Theory and Evidence." *International Studies Quarterly* 47(4):483–510.

Mousseau, Michael. 2019. "The End of War: How a Robust Marketplace and Liberal Hegemony Are Leading to Perpetual World Peace." *International Security* 44(1):160–196.

Nieman, Mark David. 2015. "Statistical Analysis of Strategic Interaction with Unobserved Player Actions: Introducing a Strategic Probit with Partial Observability." *Political Analysis* 23(3):429–448.

Nieman, Mark David. 2016. "The Return on Social Bonds: Social Hierarchy and International Conflict." *Journal of Peace Research* 53(5):665–679.

Nieman, Mark David. 2018. "Strategic Binary Choice Models with Partial Observability." *Statistica Sinica* 28(4):2089–2105.

Nieman, Mark David, Carla Martinez Machain, Olga Chyzh and Sam Bell. 2020. "An International Game of Risk: Troop Placement and Major Power Competition." *Journal of Politics* . Forthing.

Poirier, Dale J. 1980. "Partial Observability in Bivariate Probit Models." *Journal of Econometrics* 12(2):209–217.

Powers, Kathy L. 2004. "Regional Trade Agreements as Military Alliances." *International Interactions* 30(4):373–395.

Powers, Kathy L. 2006. "Dispute Initiation and Alliance Obligations in Regional Economic Institutions." *Journal of Peace Research* 43(4):453–471.

Rainey, Carlisle. 2014. "Arguing for a Negligible Effect." *American Journal of Political Science* 58(4):1083–1091.

Rainey, Carlisle and Robert A. Jackson. 2017. "Unreliable Inferences About Unobserved Processes: A Critique of Partial Observability Models." *Political Science Research and Methods* 6(2):381–391.

Rasler, Karen A. and William Thompson. 2006. "Contested Territory, Strategic Rivalries, and Conflict Escalation." *International Studies Quarterly* 50(1):145–167.

Reed, William and Daina Chiba. 2010. "Decomposing the Relationship Between Contiguity and Militarized Conflict." *American Journal of Political Science* 54(1):61–73.

Rose, Andrew K. 2004. "Do We Really Know that the WTO Increases Trade?" *American Economic Review* 94(1):98–114.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.

Russett, Bruce M. and John R. Oneal. 2001. *Triangulating Peace: Democracy, Interdependence, and International Organizations*. New York, NY: WW Norton & Company.

Sartori, Anne E. 2003. "An Estimator for Some Binary-Outcome Selection Models without Exclusion Restrictions." *Political Analysis* 11(2):111–138.

Schroeder, Paul W. 1976. *Alliances, 1815-1945: Weapons of power and tools of management*. Lawrence, KS: University of Kansas Press pp. 227–62.

Senese, Paul D. and John A. Vasquez. 2008. *The Steps to War: An Empirical Study*. Princeton, NJ: Princeton University Press.

Signorino, Curtis S. and Kuzey Yilmaz. 2003. "Strategic Misspecification in Regression Models." *American Journal of Political Science* 47(3):551–566.

Simmons, Beth A. 2000. "International Law and State Behavior: Commitment and Compliance in International Monetary Affairs." *American Political Science Association* 94(4):819–835.

Singer, J. David, Stuart 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.

Siverson, Randolph M. and Juliann Emmons. 1991. "Birds of a Feather: Democratic Political Systems and Alliance Choices in the Twentieth Century." *The Journal of Conflict Resolution* 35(2):285–306.

Stinnett, Douglas M., Jaroslav Tir, Philip Schafer, Paul F. Diehl and Charles Gochman. 2002. "The Correlates of War Project Direct Contiguity Data, Version 3." *Conflict Management and Peace Science* 19(2):58–66.

Stock, James H. and Mark W. Watson. 2011. *Introduction to Econometrics, 3rd international edition*. Upper Saddle River, NJ: Prentice Hall.

Thompson, William. 1996. "Democracy and Peace: Putting the Cart before the Horse?" *International Organization* 50(1):141–174.

Vasquez, John A. 2009. *The War Puzzle Revisited*. New York, NY: Cambridge University Press.

Veall, Michael R. 1990. "Testing for a Global Maximum in an Econometric Context." *Econometrica* 58(6):1459–1465.

Walt, Stephen M. 1985. "Alliance Formation and the Balance of World Power." *International Security* 9(4):3–43.

Waltz, Kenneth N. 1979. *Theory of international politics*. Waveland Press.

Ward, Michael D and Kristian Skrede Gleditsch. 2002. "Location, location, location: An MCMC approach to modeling the spatial context of war and peace." *Political analysis* 10(3):244–260.

Weitsman, Patricia. 2004. *Dangerous Alliances: Proponents of Peace, Weapons of War*. Stanford, CA: Stanford University Press.

Xiang, Jun. 2010. "Relevance as a Latent Variable in Dyadic Analysis of Conflict." *Journal of Politics* 72(2):484–498.

Xiang, Jun. 2017. "Dyadic Effects, Relevance, and the Empirical Assessment of the Kantian Peace." *International Interactions* 43(2):248–271.