Research Note

Journal of Conflict Resolution
1-27

DOI: 10.1177/0022002715595700
jcr.sagepub.com

Estimating Dynamic State Preferences from United Nations Voting Data

Michael A. Bailey[1,2], Anton Strezhnev[3], and Erik Voeten[1,4]

Abstract

United Nations (UN) General Assembly votes have become the standard data source for measures of states preferences over foreign policy. Most papers use dyadic indicators of voting similarity between states. We propose a dynamic ordinal spatial model to estimate state ideal points from 1946 to 2012 on a single dimension that reflects state positions toward the US-led liberal order. We use information about the content of the UN's agenda to make estimates comparable across time. Compared to existing measures, our estimates better separate signal from noise in identifying foreign policy shifts, have greater face validity, allow for better inter-temporal comparisons, are less sensitive to shifts in the UN' agenda, and are strongly correlated with measures of liberalism. We show that the choice of preference measures affects conclusions about the democratic peace.

Keywords

international cooperation, international law, international organization, international institutions

[1]Department of Government, Georgetown University, Washington DC, USA
[2]McCourt School of Public Policy, Georgetown University, Washington, DC, USA
[3]Department of Government, Harvard University, Cambridge, MA, USA
[4]Edmund A. Walsh School of Foreign Service, Georgetown University, Washington, DC, USA

Corresponding Author:
Erik Voeten, Department of Government, Georgetown University, 301 ICC, Washington, DC 20052, USA.
Email: ev42@georgetown.edu

Foreign policy preferences play a crucial role in international relations theory. Yet, they are not directly observed and thus need to be inferred from observable behavior. Votes in the United Nations General Assembly (UNGA) have become the standard data source for constructing measures of state preferences, as they are comparable and observable actions taken by many countries at set points in time. Unfortunately, with few exceptions, scholars have ignored methodological advances in the roll call literature on how to estimate latent preferences from observed votes. Moreover, existing measures fail to separate changes in the UN's agenda from shifts in state preferences. Consequently, these measures frequently suggest that states change their foreign policy orientations when all that changes is the content of votes.

We develop a state-of-the-art ideal point model to estimate dynamic national ideal points along a single dimension from 1946 to 2012. We use resolutions that were identical across years to serve as bridge observations to help make the preference estimates comparable over time. The resulting estimates consistently capture the position of states vis-à-vis a US-led liberal order. During the Cold War, this was the East–West conflict between communist and capitalist states. Since the end of the Cold War, the non-Western pole has been occupied by a motley crew of states that have little ideological cohesion other than their opposition to the Western liberal order.

Our estimates have several advantages over dyadic similarity indicators such as affinity or S scores (Gartzke 1998, Signorino and Ritter 1999). First, our estimates allow for more valid intertemporal comparisons. For example, according to S scores, Russia and the United States have more conflictual relations in the mid-2000s than the Union of Soviet Socialist Republics (USSR) and the United States ever had during the Cold War. By separating agenda changes from changes in preferences, we offer more plausible estimates of long-term shifts in this and other relationships. This matters given the frequent usage of UN-based preference measures in dynamic analyses.

Second, our estimates are better able to distinguish signal from noise in identifying meaningful shifts in foreign policy orientations. For example, ideal points analysis of our ideal point estimates consistently indicates that left-wing regimes in Latin America were systematically less favorable to the United States than right-wing regimes. We do not find this pattern with S scores.

Finally, dyadic indicators can only reveal shifts in preference similarity between two states but not which state has shifted. Ideal point estimates allow us to assess, for instance, whether the Soviet Union or the United States was responsible for the two states moving closer or further apart. This allows researchers to examine foreign policy orientations in monadic analyses. Moreover, we offer assessments of uncertainty in our point estimates.

We first discuss how UN votes have been used in the literature. We then explain our estimation method and data. The following section evaluates the validity of our estimates. Finally, we replicate an application to the study of conflict (Gartzke 2000) and find that replacing affinity measures with ideal point estimates makes the original findings robust to modeling the dynamics appropriately.

Literature, Data, and Method

The Use of UN Votes in the Literature

Scholars have used UN votes to measure foreign policy preferences virtually since the institution was established (e.g., Ball 1951; Lijphart 1963; Moon 1985; Vengroff 1976; Russett 1966). This interest shows little sign of abating. We found seventy-five articles published between 1998 and 2012 that used UN votes to construct measures of national preferences.

Many papers use UN-based preference measures as independent variables. These include analyses of whether foreign policy interests affect the likelihood of interstate disputes (e.g., Gartzke 1998; Oneal and Russett 1999; Reed et al. 2008), the severity of those disputes (Sweeney 2003), the occurrence of acts of terrorism (Dreher and Gassebner 2008), the distribution of foreign aid (e.g., Alesina and Dollar 2000), the lending behavior of the World Bank and International Monetary Fund (IMF, e.g., Thacker 1999; Dreher and Jensen 2007), the probability of signing treaties (Koremenos 2005), and the reception of diplomatic missions (Neumayer 2008).

Others have used UN-based measures as dependent variables to ascertain whether socialization through intergovernmental organizations leads to convergence in member state interests (Bearce and Bondanella 2007), whether leadership and regime changes alter foreign policy orientations (Dreher and Jensen 2013), whether the European Union has started to form a cohesive foreign policy (Drieskens 2010), whether the United States is becoming increasingly isolated on foreign policy issues (Voeten 2004), or whether US or Chinese foreign aid or trade successfully buys foreign policy adherence (Flores-Macías and Kreps 2013).

Voting in the UN differs from some other voting bodies in that three explicit vote options are widely and deliberately used in the UN: yea, nay, or abstain. The informative abstentions are an important feature of voting and virtually all studies treat a nay vote as a stronger signal of disapproval than an abstention.[1] Abstentions differ from absences. Absences typically do not reflect a country's views but instead are due to causes such as a temporary lack of government. Absences are strongly correlated with the occurrence of civil wars and coups (Voeten 2013). In 68 percent of votes where a state is absent, it will also be absent on the next roll call. This suggests strongly that absences are not usually indications of discontent with a resolution and should thus not be equated with abstentions, as some scholars do (Rosas, Shomer, and Haptonstahl 2015).

Most existing UN-based preference measures are based on dyadic similarity of vote choices. The most widely used is Signorino and Ritter's (1999) S score.[2] S scores treat UN votes as interval-level measures of preference expression with abstentions halfway between a yea and a nay vote (they exclude absences).[3] The

S score is a Euclidean distance measure between every dyad in the UN. It is calculated as:

$$S_{ab} = 1 - \frac{\sum |Y_{a\upsilon} - Y_{b\upsilon}|}{V},$$

where $\upsilon = 1, \ldots V$ indexes votes, a and b refer to two countries, and Y refers to votes, taking on one of three alternatives: $\{$yea $(Y = 1)$, abstain $(Y = 2)$, and nay $(Y = 3)\}$. The S score equals 1 if two countries agree on all votes and -1 if two countries maximally disagree on all resolutions.

The core weakness of S scores and related similarity indices is that they assume a straightforward relationship between how often two states vote together and preference similarity. Yet, voting coincidence is also affected by what resolutions states vote on. Suppose there are ten votes and that country A and country B vote identically on nine of them. In the next session, suppose that the countries preferences do not change, but there happen to be ten additional votes on the single issue that divided the two countries. The similarity index of the two countries would drop from 90 percent on the first ten votes to 50 percent on all twenty votes even as their preferences did not change.

The S score approach differs sharply from contemporary measurement models used for other voting bodies, where scholars invariably use spatial models and item response theory (IRT) empirical methods (e.g., Poole 2005).

Spatial models offer an explicit theory of how states translate preferences into votes. Two features of these models allow them to measure preferences in a way that does not ebb and flow according to the composition of votes in a voting body. First, spatial models estimate vote cut points that identify points along a preference spectrum where countries vote in a particular way. This makes it possible to attribute changes in similarity due to changes in agenda rather than simply assuming all changes in voting similarity reflect preference changes. Second, spatial models estimate "discrimination parameters" that tell us how well individual roll calls distinguish between states with different ideal points along a dimension. These parameters effectively allow us to weight votes by how much they reflect the main preference dimension. If a series of votes appear that have little to do with the main dimension of preferences, they will not exert much influence on ideal point estimates. This contrasts with similarity indices which weight all votes identically and will ebb and flow as these votes influenced by idiosyncratic factors come and go from the agenda.

Figure 1 illustrates how spatial models account for agenda change. There are seven countries with ideal points θ_1 to θ_7. On vote 1, the cut points γ are such that two countries vote nay (solid triangles), two abstain (open circles), and three vote yea (solid squares). On vote 2, only two countries vote nay, four abstain, and two vote yea. On vote 3, the polarity of the vote is reversed so that the high θ countries are those more inclined to vote yea. The coalition is similar to that on vote 1, only for this vote most yeas from vote 1 have become nays and vice versa.

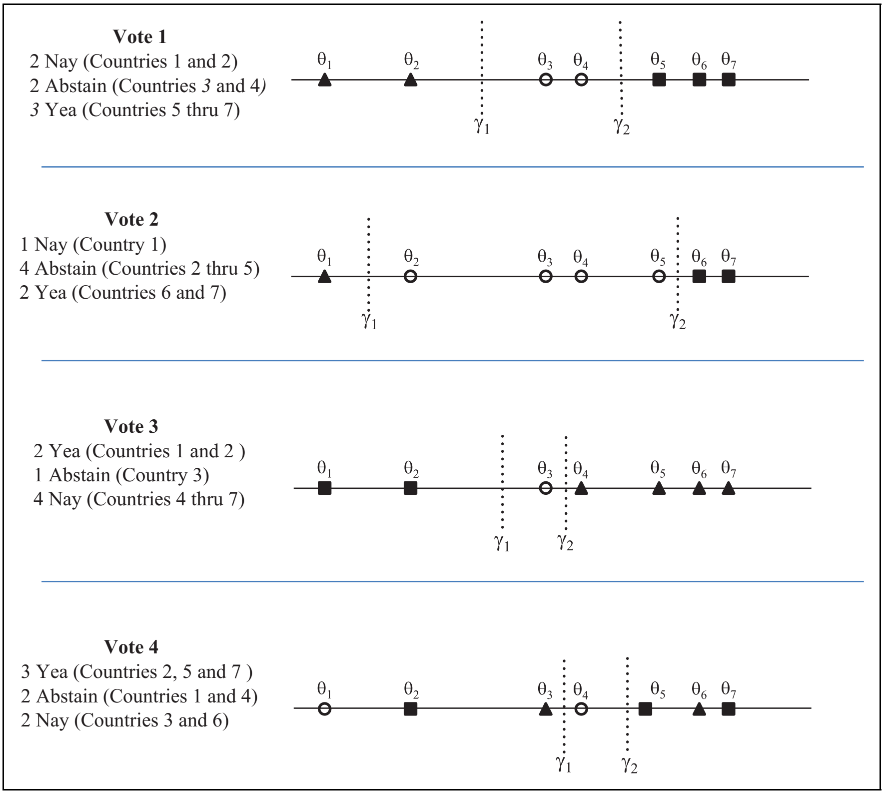

Figure 1. Hypothetical votes for seven countries.

If the agenda changed from having many votes like vote 1 to many votes like vote 2, a similarity index would indicate a major split in countries 5 and 6 and a major warming of relations of countries 4 and 5. But in fact, these countries could have completely unchanged preferences and their voting would only change due to the changes in the types of votes on which they expressed their views.

Spatial models also give us tools to lessen the influence of idiosyncratic votes on preference estimates. An idiosyncratic vote is one that divides countries based on factors unrelated to the main preference dimension. For example, in some years, North–South divisions over UN supranationalism pitted the powerful states (including the USSR and the United States) against developing nations (Voeten 2000). These votes provide no meaningful information on the position of states on a pro-anti-West dimension. Vote 4 in Figure 1 depicts such a case as votes are poorly explained by the countries' ideal points: countries 1 and 4 abstain, while countries 3 and 6 vote nay and the rest vote yea. A similarity index such as the S score would take this vote as evidence of changes in preferences among countries, while a spatial model will place less weight on such votes because they are poorly explained by ideal points. We discuss this parameter in more detail when we present our model subsequently.

While spatial models provide a structure for distinguishing agenda change from preference change, they do not automatically distinguish wholesale shifts in preferences from wholesale shifts in vote parameters. For example, if all country ideal points (θ) moved one step to the left, this is mathematically equivalent to all cut points (γ) moving one step to the right. Therefore, Bailey (2007, 2013) and Voeten (2004) pin down the location of vote cut points by using information about the location of vote cut points over time. By utilizing information that vote cut points on a subset of votes were the same from one year to the next, their models could appropriately interpret any change in votes on these votes would have to be due to preference change.

With some exceptions (Voeten 2000, 2004; Reed et al. 2008; Mattes, Leeds, and Carroll 2015; Andersen, Harra, and Tarp 2006), spatial theory and IRT models have not been applied to the UN. One reason is that standard software packages for ideal point estimation do not easily incorporate votes with three choices. Moreover, existing ideal point approaches are not dynamic and do not separate agenda from preference change.[4]

Statistical Model

Our approach is to use IRT statistical models to estimate one-dimensional preferences that are comparable over time based on votes in the UN. The model is based on the multiple rater ordinal data model of Johnson and Albert (1999, 166). Let $i = 1, \ldots N$ index states and $\upsilon = 1, \ldots V$ index votes. States have unidimensional "ideal points" in each year (θ_{it}).

The vote by a country on a given resolution is a function of its ideal point, characteristics of the vote, and random error. Specifically, the spatial preference of each country on each vote is $Z_{it\upsilon} = \beta_\upsilon \theta_{it} + \in_{i\upsilon}$, with $\in_{i\upsilon} \sim N(0, 1)0$, which is a latent variable. β_υ is analogous to the discrimination parameter in standard IRT models. The sign of β_υ indicates the "polarity" of vote υ: on some resolutions (such as vote 1 in Figure 1), high θ_{it} countries are inclined to vote yea, and β_υ will be positive for these votes. β_υ will be negative when high θ_{it} countries are inclined to vote nay (such as vote 3 in Figure 1). The magnitude of β_υ indicates how well vote υ separates countries with high and low ideal points. A β_υ near zero indicates vote υ is poorly explained by ideal points and is associated with a muddle of yea, abstain, and nay voting across the ideological spectrum, as in vote 4 of Figure 1.

Each vote has three alternatives: $\{$yea ($Y = 1$), abstain ($Y = 2$), and nay ($Y = 3$)$\}$.The observed choice, $Y_{it\upsilon}$, depends on $Z_{it\upsilon}$, the latent vote-specific preference of country$i$, and cut points $\gamma_{1\upsilon}$ and $\gamma_{2\upsilon}$. Formally, the conditions that determine which alternative a country chooses on vote υ are:

$$
\begin{array}{lll}
Y_{it\upsilon} = 1 & \text{if} & Z_{it\upsilon} < \gamma_{1\upsilon} \\
Y_{it\upsilon} = 2 & \text{if} & \gamma_{1\upsilon} < Z_{it\upsilon} < \gamma_{2\upsilon} \\
Y_{it\upsilon} = 3 & \text{if} & Z_{it\upsilon} > \gamma_{2\upsilon}.
\end{array}
$$

Assuming normally distributed errors, the probability that nation i chooses option k is:

$$\Pr(Y_{it\upsilon} = k) = \Phi(\gamma_{k\upsilon} - \beta_{\upsilon}\theta_{it\upsilon}) - \Phi(\gamma_{k-1,\upsilon} - \beta_{\upsilon}\theta_{it\upsilon}),$$

where Φ is the normal cumulative distribution function and $\gamma_{0\upsilon} = -\infty$ and $\gamma_{3\upsilon} = \infty$ for all votes. We assume that absences are distributed as missing at random as in Clinton, Jackman, and Rivers (2004). In a possible future extension, one might examine more nuanced models of absences along the lines of those proposed by Rosas and Shomer (2008) and Rosas, Shomer, and Haptonstahl (2015).

All ideal point models need to be identified with a normalization. We fix ideal points to be centered on zero with a standard deviation of 1. We implement our across-time bridging by imposing a restriction that resolutions with the same content have the same outpoints, γ_1 and γ_2. The assumption is that a resolution at time t has the same resolution parameters as an identically phrased resolution at time $t + 1$. This is a weaker assumption than that made in studies that jointly scale legislators with constituents or judges, where it may not necessarily be true that an identically phrased proposal means the same thing to a legislator and a constituent (Lewis and Tausanovich 2013). Context could change too with time, which is why we limit to five the number of consecutive years in which the resolution parameters are fixed.

We implement a Bayesian prior on the estimate for θ_{it} based on $\theta_{i,t-1}$. The variance of this prior determines how much smoothing occurs. If variance of the prior is set at a very large value, then almost no smoothing occurs and we essentially estimate preferences separately for each year for each country with the preference in the previous period providing no information about preferences in the next period. If the variance of the prior is set at a very small value, then preferences change very little from one period to the next, meaning we essentially estimate a single ideal point for each country for the entire time period. As discussed in Martin and Quinn (2002, 147), specific estimates can be sensitive to the setting of such a parameter and there is no consensus way to determine its value. As with Martin and Quinn, we set this value at a point at which the estimates do indeed move from period to period but not too dramatically. The advantage to this approach compared to estimating ideal point change as a polynomial function over time (as in Voeten 2004) is that it allows for discrete shifts in ideal points, for example, responding to regime change. The prior will soften these shifts but will not make them conform to a specific functional form over time.

We use a hybrid Metropolis–Hasting/Gibbs sampler to estimate the parameters of the model following the process described in Johnson and Albert (1999, 135, 166) and Cowles (1996). The Appendix contains more details on the estimation process.

Like all efforts to extract preference estimates from observed votes, we can only measure revealed preferences rather than underlying "true" preferences. The vote of a state on any given resolution may be a function not only of their unrestricted preferences but also of other motives. Given that UNGA votes are nonbinding, we suspect that strategic voting is less prevalent than in other arenas of world politics, such as the UN Security Council. The influence of some factors such as foreign aid could be modeled empirically based on the ideal points we estimate (Dreher, Nunnenkamp, and Thiele 2008; Dreher and Sturm 2012; Carter and Stone 2015).

Data

The data come from Strezhnev and Voeten (2013) and include 4,335 divisive roll calls over sixty-seven sessions from January 1946 to December 2012. As in all ideal point estimations, we limit the analysis to resolutions adopted with a vote. About a quarter of all UNGA votes are adopted without a representative requesting a roll call (Hage and Hug 2013). Such votes provide no information for ideal point estimation.

We analyze preferences for each country for each UN session. Sessions usually start in September and occasionally run into the new year. A more detailed discussion is in Voeten (2013).

In order to identify preference change over time, we coded whether resolutions are identical. Resolutions are nonbinding. Even adopted resolutions do not necessarily alter the status quo. It is therefore common for consecutive sessions to vote on duplicate resolutions. We identified duplicates by first downloading pdf files of all General Assembly Resolutions from the UN website and converting the files to machine-readable text. We then used plagiarism detection software WCopyFind 3.0.2 (Bloomfield 2011) to estimate the similarity between resolution texts.[5] Small changes in words can meaningfully alter political support for resolutions. We therefore first selected resolutions that were plausibly identical because they returned high similarity scores in WCopyFind and then manually inspected them to confirm that no substantial differences existed.[6] If resolutions differ only in inconsequential ways (such as dates or references to past resolutions), we treated them as identical. This procedure led to 799 pairs of identical resolutions.

There are some differences between the matched and unmatched resolutions. Based on a model with session fixed effects, the percentage of yes votes on matched resolutions is on average about 3 percent higher than on nonmatched votes (from an average of 57 to 60 percent). Matched resolutions are also somewhat more likely to be resolutions on colonialism and the Middle East and somewhat less likely to be about human rights and disarmament. Yet, these differences are modest. For example, human rights resolutions, the most underrepresented category, constitute 16.5 percent of unmatched resolutions and 14.6 percent of matched resolutions.

Most of the time, states do not alter their positions on matched resolutions: 97.2 percent of nonmissing vote choices were the same from one year to the next. The United States changed somewhat more frequently than average (4.2 percent of the time). Yet these are very informative data points if we wish to identify changes in foreign policy preferences.

We also estimated the model on the subsample of 336 votes identified as important by the State Department since 1983. Similarity indicators based on this subset of votes are often used in applied work.

Validity

This section examines the validity of our ideal point estimates in four ways. First, we assess whether the changes in ideal points make sense, given what we know about post–World War II history. Second, we examine whether ideal point estimates are affected by changes in the agenda. Third, we explore which substantive areas most reflect underlying preference differences in UN voting. Fourth, we ascertain whether the ideal points are associated with democratization, financial liberalization, and government ideology in the ways scholars would expect.

To compare ideal points to S scores, we transform ideal points into dyadic measures by taking the absolute distance between the ideal points of two countries (multiplied by -1 to give both measures the same sign). The bivariate Pearson correlation between the distances in ideal points and the S scores is .82. This is fairly high but, as we shall see, the differences are consequential especially in intertemporal comparisons.

Face Validity

Figure 2 plots the S scores and the absolute distance between the ideal points for the United States and the USSR/Russia. Both measures capture a dramatic change immediately when the Cold War ends. Yet there are important differences. First, S scores imply implausible intertemporal comparisons. If we believe S scores, the United States and the USSR had much more similar interests during several periods of the Cold War (mid-1950s and mid-1970s) than during most of the post–Cold War period with the exception of a few years in the early 1990s. Indeed, 1989, 2008, and 2009 are the darkest years in the relationship between these two countries according to S scores, with an equally implausible warming of ties since. This defies common sense.

Second, S scores fluctuate tremendously during the Cold War plausibly due to changes in the relative importance of the second dimension. In the early 1970s, the group of seventy-seven gained control of the UNGA's agenda, introducing many new resolutions on UN supranationalism and other North–South issues on which the USSR and the United States took similar positions (Voeten 2000). This led to large increases in the number of times the two states voted together (and consequentially S scores) but not in their positions on the main underlying dimension of contestation, as correctly captured by relatively stable ideal point estimates. Similarly, in the mid-1950s, the Suez crisis was a particular event (with many associated votes) that pitted the United States against its Western allies. It did not, however, reflect a fundamental easing of relations between the superpowers as the S scores would suggest. By contrast, the ideal point estimates are fairly consistent during the Cold War and accurately keep the two countries much further apart during the Cold War than at any point after.

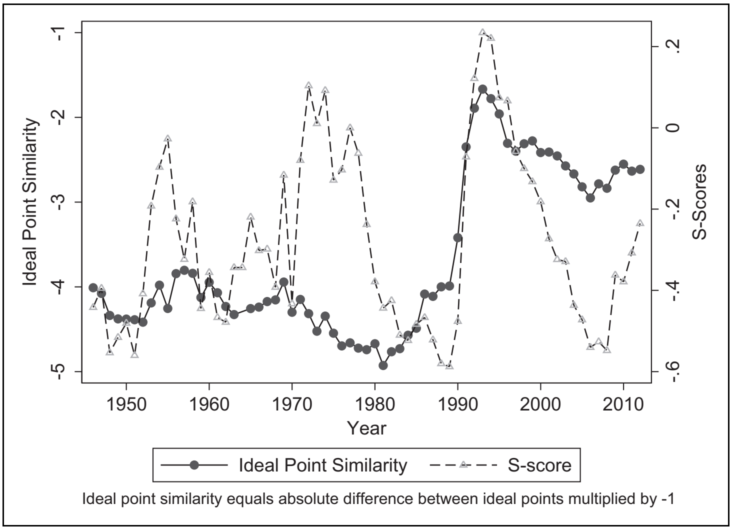

Figure 2. United Nations voting similarity between Russia/Union of Soviet Socialist Republics (USSR) and the United States.

Third, despite greater volatility, S scores sometimes fail to detect real shifts in preferences. Most notably, the ideal point measure picks up on Gorbachev's efforts to reconcile with the West in the mid-1980s, while S scores pick up on a change only in 1991.

This latter point is made more clearly in Figure 3, which plots ideal point estimates for the five permanent members of the UN Security Council (P-5). It is the Soviet Union rather than the United States, which moves in the 1980s.

Figure 3 also paints a broader picture of great power ideology in the post–World War II period. In the 1950s and early 1960s, France and the United Kingdom took a more extreme "Western" position than the United States, as they desperately clung to their empires. The gap between the United States and its Western allies has increased since the Reagan Administration, which notably departed from the Carter Administration. The People's Republic of China, which did not take a UN seat until 1971, initially canvassed with the nonaligned states near the center of the ideological space. Tensions with the West increased after Tiananmen square. Moreover, the collapse of the Soviet Union left China without a foe on the opposite side of the ideological spectrum. Since the late 1990s, China has taken a much more pragmatic and active approach in the UN. China's ideal point never moves as far in opposition to the Western liberal order as the USSR's did during the Cold War. By contrast, S scores between the United States and China during the 1990s (not shown) were lower than those between the United States and the USSR at *any* time during the Cold War.

One concern is that spatial models (like S scores) may not be able to distinguish sincere preferences from strategic concerns, such as a desire to get aid or other

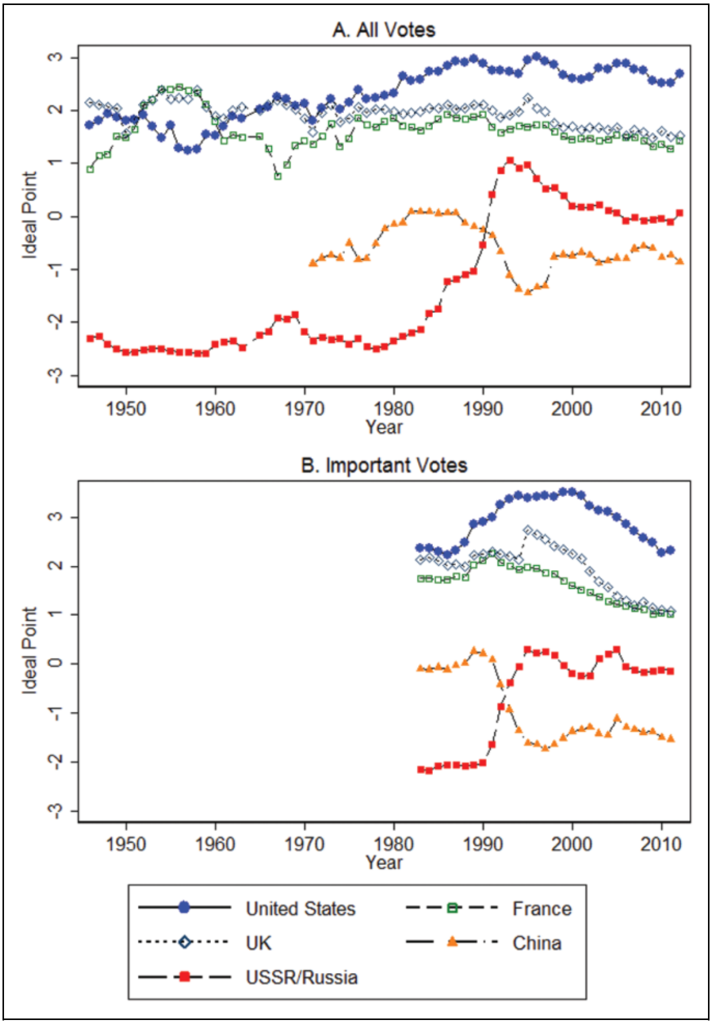

Figure 3. Ideal points of the five permanent members of the UN Security Council (P-5) in the United Nations General Assembly (UNGA).

concessions from powerful states in exchange for votes. Although UNGA votes are nonbinding and less strategic than those in venues such as the UN Security Council, there is evidence of strategic voting (Carter and Stone 2015). This may be especially pertinent on votes important to the United States, which are resolutions on which the US lobbies and where the correlation between US foreign aid and voting patterns is especially high (Wang 1999). Therefore, we also estimated ideal points based only on a subsample of votes on which the US state department declared it lobbied other countries in *Voting Practices in the United Nations*, an annual report the state department has issued to Congress since 1983. The bottom panel of Figure 3 shows ideal point estimates for P-5 countries based on these votes. The broad patterns are the same, but minor differences do exist and these may be the more appropriate measures for some analyses. The bivariate correlation between ideal points estimated based on important votes and all votes is .92.

Other comparisons also reveal substantial differences between ideal points and S scores. Figure 4a plots the ideal point distance (multiplied by -1 for comparability) between several Latin American states and the United States. Regime changes widely known to have affected foreign policy orientations also affected ideal point shifts. Cuba shifted rapidly following the 1959 revolution. Chile quickly and temporarily shifted away from the United States during the Allende regime but moved back after the 1973 coup. The fall of the Pinochet regime led to a small shift away from the Western position but not to the extent of the left-wing governments. Venezuela shifted away from the United States following the election of Hugo Chávez. Argentina became much more pro-West following the election of Carlos Menem in 1989. Nicaragua's changes correspond closely to the rise and fall of Sandinista governments.

Figure 4b shows the same data but now calculated as S scores with the United States. The lines also capture the main regime shifts, but they are much harder to distinguish from other bumps. For example, the Allende government is barely distinguishable from other small shifts in Chile's S scores. Moreover, the Latin American states are very difficult to differentiate since 1970. For the most part, Figure 4b shows a uniform shift away from the United States. For example, S scores portray Cuba and Chile as similar to each other in foreign policy preference with a massive gap between Chile and the United States. According to S scores, Chile was as far from the United States in the 2000s as Cuba was in the late 1960s. In contrast, ideal points properly put Chile and Argentina quite a bit closer to the United States than Venezuela and Cuba.

Agenda Change

What explains the large differences between the ideal points and the S scores? The major difference in the approaches is that the ideal points are based on a model that explicitly addresses agenda change. Therefore, we turn to examining how much the two preference measures are affected by substantive content of the votes.

Table 1 shows results from a lagged-dependent variable model in which the dependent variables are country ideal points (column 1) and S scores with the United States (column 2). The independent variables are annual issue proportions. We coded whether resolutions concerned colonialism, the Middle East, nuclear issues, disarmament, human rights, and/or economic issues (see Voeten 2013). Resolutions could fit into multiple categories or none.

The results indicate that the two preference measures respond very differently to changes in the agenda. None of the issue proportions correlate significantly with

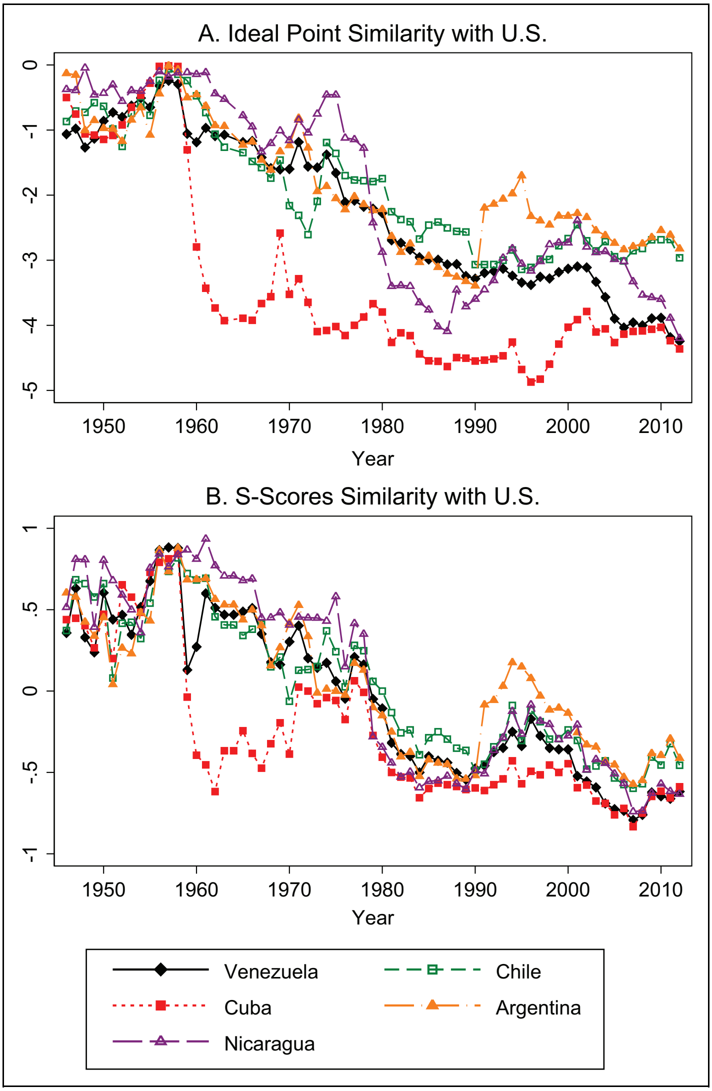

Figure 4. Latin American ideal points and S scores with the United States.

ideal points (model 1). By contrast, most of the issue proportions are associated with statistically significant changes in S scores with the United States (model 2). Variation in the proportion of nuclear issues is especially strongly correlated with S scores. To illustrate just how strongly issue proportions correlate with S scores, we also estimated a static model with fixed country effects and annual issue proportions. This model accounts for 36 percent of the variation within country S scores

Table 1. Lagged-Dependent Variable Regressions of Ideal Points and S Scores on Annual Issue Proportions.

	(1) Ideal Points	(2) S Score with United States	(3) S Scores, all dyads	(4) Ideal Point Similarity, All Dyads
Lagged-dependent variable	.982***	.870***	.880***	.965***
	(.002)	(.013)	(.002)	(.000)
Middle East	.003	−.053	−.127***	−.035***
	(.041)	(.036)	(.004)	(.006)
Nuclear	.072	−.764***	.020***	.058***
	(.080)	(.062)	(.007)	(.012)
Disarmament	−.055	−.071*	−.028***	.027***
	(.049)	(.037)	(.005)	(.008)
Human rights	.016	.189***	.051***	−.031***
	(.039)	(.029)	(.004)	(.006)
Colonialism	.005	−.182***	−.270***	−.031***
	(.031)	(.031)	(.006)	(.006)
Economic	−.020	.054	−.043***	.095***
	(.049)	(.052)	(.006)	(.009)
Constant	−.014	.100***	.150***	−.035***
	(.011)	(.014)	(.002)	(.002)
Observations	8,579	8,580	645,843	632,242
R^2	.969	.849	.782	.933
Number of countries/dyads	196	195	18,874	18,874

Note: Robust standard errors clustered on countries within parentheses.
***$p < .01$. **$p < .05$. *$p < .1$.

with the United States but only 3 percent of the variation in country ideal points (results available from authors). Qualitatively identical results hold with an error-correction model (ECM) and if we add fixed country effects. This confirms that a large proportion of variation in dyadic S scores has less to do with country-specific preference change than with changes in the UNGA's agenda.

In columns 3 and 4, we show results from the same model estimated on all dyads. Again, issue proportions correlate strongly and significantly with S scores (model 3). Here, we also find significant effects when we use ideal point differences as the dependent variable (model 4). Yet, the substantive effects are small compared to those on S scores. When we standardize the coefficients from the table, we find that none of the standardized coefficients on ideal point similarity exceeds 0.004 (the standardized coefficients on economic and Middle East). All but one (nuclear disarmament) of the standardized coefficients on S scores exceed 0.004.

Other analyses lead to similar conclusions. In a country fixed effect model, annual issue proportions explain 17 percent of the within country variation in dyadic S scores and just 3 percent in the variation in ideal point similarity. In an ECM with fixed country effects for ideal points, many of the issue proportions are no longer significant, whereas all of the issue proportions remain significant in such a specification for S scores.

An alternative test is to examine how susceptible ideal point estimates are to annual shocks. We estimated a regression with country fixed effects and indicator variables for each year. The year dummies explain only 4 percent of the variation in ideal points.[7] Thus, there is little evidence that ideal points move up and down due to common shocks rather than individual countries changing their positions.

Year dummies explain 50 percent of the variation in S scores between the United States and other countries. This compares to 27 percent of the variation in ideal point distances with the United States (remember S scores are dyadic so we must compare with ideal point distances). If we examine S scores and ideal point similarities in all dyads (with dyad fixed effects), year dummies explain 11% of the variation in S scores and 2 percent of the variation in ideal point distances. In sixty out of sixty-six years, all dyadic S scores significantly (at the 5 percent level) move in the same direction. This is again consistent with the notion that a large portion of variation in S scores is a consequence of common shocks to the agenda rather than individual states changing their positions.

Simply regressing S scores on year dummies and using the residuals as a preference measure would be one way to improve on existing practices. Nevertheless, this is an ad hoc solution, which does not address the other advantages of ideal point methods, and it would incorrectly adjust for instances where there are real common shocks to preferences, such as with the end of the Cold War.

Substantive Content of Preferences

We can also use variation in issue proportions over time to gain some insight into the substantive interpretation of the preference dimension that dominates UN voting. To do this, we regressed the absolute values of the estimated discrimination parameter β_υ on the issue codes. To recall, the discrimination parameter is an estimate of the weight a resolution has, which is based on the extent to which conflict on the resolution reflects conflict along our single dimension. On average, human rights resolutions have a discrimination parameter that is higher by 0.35 (half a standard deviation) than the reference category (resolutions that do not fit any of the issue categories). Colonialism issues (0.31), economic issues (0.12), and disarmament issues (0.09) also have significantly higher than average discrimination parameters, but Middle East issues and nuclear issues do not.[8]

This fits the substantive interpretation of the dimension as being about a liberal world order. Human rights resolutions weigh heavily, especially after the end of the Cold War, as they signify conflict between liberal and nonliberal states. By contrast, resolutions on nuclear weapons often (though not always) separate the nuclear haves from the have-nots. This means that all permanent members of the UN Security Council tend to vote on the same side. The proportion of resolutions that deal with this issue varies considerably, from zero in some years to about 30 percent in others (see Voeten 2013). This may lead us to wrongly conclude that the P-5 are much closer one year than they are the next simply because there are more resolutions on nuclear issues. This may be especially problematic for applications that use UN votes to construct measures of great power agreement (e.g., Copelovitch 2010).

We do not argue that conflicts over nuclear issues or the Middle East are unimportant. Yet they are not always conflicts over the liberal world order. To be clear, we do not a priori exclude any issue areas or resolutions from our analysis. Instead, resolution weights are estimated based on the extent to which votes discriminate countries along the primary dimension of conflict. Analysts primarily interested in nuclear or Middle East issues should use our data and method to estimate a separate set of ideal points. We did so for Middle East issues and found that the resulting estimates are reasonably correlated with our first dimension estimates (Pearson correlation is .83). Yet, these estimates reflect some distinct position taking on the Middle East and are therefore probably more suitable for predicting votes on the admission of Palestine as a state or other predictive tasks with regard to the Middle East.

Correlation with Democracy, Financial Liberalization, and Government Ideology

We can also examine the validity of our ideal point measures by analyzing temporal variation in ideal points. Changes within a country should reflect changes in state characteristics rather than common shocks to the agenda. In this section, we compare how ideal point similarity with the United States (distance in ideal points multiplied by -1 for comparability) and S scores respond to domestic changes that should theoretically correlate with a Western liberal foreign policy outlook. We focus on three factors: democracy, financial openness, and left-wing governments.

We estimate ECMs with fixed country effects, as we are interested in explaining shifts within countries. The models estimate both short- and long-term effects of domestic attributes (see De Boef and Keele 2008). Coefficients on lagged variables are short-run effects, and coefficients on changes in variables are long-run effects.

We report separate results for each independent variable due to concerns that the various domestic liberalism measures are endogenous to each other and due to large numbers of missing values for two of our measures.

Table 2 presents results from multiple specifications. Ideal points are the dependent variables in the three columns on the left; S scores are the dependent variables in the three columns on the right.

We can compare the results in columns 1 and 4 to see the association of democracy and ideal points and S scores. These columns include Polity 4 scores (rescaled to a 0–1 scale), a widely used measure of democracy. Column 1 indicates that both lagged levels and annual changes in levels of democracy are correlated with shifts toward the United States. That is, there is both an immediate and a long-run link between countries democratizing and moving closer to the US ideal point. Column

Table 2. ECMs with Country Characteristics on Ideal Point and S score Similarity with the United States.

	Ideal Point Change toward the United States			S score Change toward the United States		
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged dependent variable (lagged level)	−.069***	−.039***	−.115***	−.135***	−.109***	−.192***
	(.004)	(.006)	(.00506)	(.006)	(.009)	(.00757)
Lag polity2 (lagged level)	.056***			−.027***		
	(.012)			(.009)		
Change polity2 (first difference)	.067**			−.065***		
	(.031)			(.021)		
Lag open economy (lagged level)		.055**			−.003	
		(.025)			(.016)	
Change open economy (first difference)		.257***			.108**	
		(.068)			(.043)	
Lag left (lagged level)			−.013			.002
			(.009)			(.006)
Change left (first difference)			−.045***			−.016*
			(.013)			(.009)
Constant	−.231***	−.140***	−.346***	−.019***	−.0109	−.071***
	(.011)	(.0186)	(.0143)	(.005)	(.00814)	(.003)
Observations	6,391	2,852	5,037	6,588	2,969	5,034
R^2	.056	.023	.097	.080	.056	.118
Countries	165	82	173	166	82	174

Note: Standard errors within parentheses. ECM = error-correction model.
***$p < .01$. **$p < .05$. *$p < .1$.

4 reports a strong and significant *negative* correlation with democratization and S scores with the United States, exactly opposite of natural expectations that democratization is related to affinity with the Western-led world order.

We can compare the results in columns 2 and 5 to see the association of financial openness and ideal points and S scores. These specifications include a single scale that combines Quinn and Toyoda's (2008) measures of current and capital account openness (also rescaled to the 0–1 interval). This measure is used in the literature on the capitalist peace (Gartzke 2007), although it is only available for a limited number of countries. The effects are strong. A country that shifts from a fully closed to a fully open economy is expected to move about 0.2 toward the US side of the ideal point dimension, with additional long-run effects implied by the positive and significant coefficient on the lagged level. We only find a weak correlation with S scores.

Columns 3 and 6 assess the relationship between government ideology and closeness to the United States. Column 3 shows that shifts from a non-left- to a left-wing government are associated with a small but significant shift away from the Western pole of the first dimension. Column 6 shows that this effect is only significant with S scores at the 10 percent level.

We also estimated the last model on the subsample of Latin American countries, a group viewed to have shifted away from the United States during left-wing regimes and toward the United States during right-wing regimes. The difference between the left and right coefficient is significant at $p < .001$ when the dependent variable is either the ideal point of a Latin American country or the absolute distance with the United States. By contrast, the difference between left and right governments does not reach conventional levels of significance for S scores. Here, again the ideal point estimates reflect conventional wisdom about foreign policy change while the S scores do not.

None of the models in Table 2 include year fixed effects. Including year fixed effects could change results if either (a) all countries move the same amount toward or away from the United States in many years or (b) the preference measures themselves are moving systematically in each year due to agenda change. The ideal point measures are designed not to be affected by agenda change while the S scores could be affected by both types of year fixed effects. When we estimate the abovementioned models and include year fixed effects, the results for the ideal point models are virtually unchanged. The results for the S scores change substantially, however. With fixed year effects, the coefficients on democracy and financial openness become positive and significant.[9]

Using both S scores and ideal point estimates, we would conclude that countries become more similar to the US foreign policy position if they democratize, move toward an open market economy, and move away from left-wing governments. Yet, with S scores, these effects only become visible after factoring out common annual shocks due to agenda changes. Without these, sometimes the correlations point in the opposite direction than expected. This is a serious problem for applications in which S scores are independent variables that supposedly capture within-country variation in foreign policy positions.

Preferences and Conflict

To what extent does it matter whether scholars use ideal point measures or an affinity measure such as S scores? If ideal point measures contain less noise and more signal, then they should strengthen empirical findings if the theory is correct. To evaluate this, we replicate Gartzke (1998), a seminal study that popularized the use of UN votes as measures of state preferences. This article argues that the relative absence of conflict between democratic states is an artifact of the common preferences of democratic dyads. Including a measure of UNGA voting similarity in the original democratic peace model from Oneal and Russett (1997), the article finds that affinity strongly correlates with the incidence of militarized interstate disputes (MIDs) and possibly explains away the effect of joint democracy. Gartzke (2000) updated this model in response to a critique by Oneal and Russett (1999).

We reestimate the logistic regression reported in table 1, model 2 of Gartzke (2000), attempting to match the original data set as closely as possible. The sample consists of all politically relevant dyads from 1950 to 1985. In addition to the preference similarity measure, the model includes seven covariates.[10] We construct all of the variables exactly as in Gartzke (2000), except for the original affinity scores, which we update to reflect corrections in the UNGA roll call data set.[11] Because of data revisions, the affinity measure from the original article differs from the most current S score implementation. It is important to make this minor adjustment to the original model in order to ensure that differences between the ideal point measure and S scores are not simply due to changes in the data being used. The dependent variable takes on a value of 1 if an MID occurred or was ongoing for a given dyad year and 0 otherwise.

We first estimate the original reported regression model presented without any corrections to account for temporal dependence that have become standard in binary time series cross-sectional regressions. As reported in the original article, the estimated coefficient on the affinity score variable is negative and statistically significant at the $\alpha = .05$ level. However, when working with time series data, failure to account for temporal confounding can lead to misleading and overconfident inferences (Beck, Katz, and Tucker 1998). To adjust for this, we follow the recommendations of Carter and Signorino (2010) and include a cubic polynomial of the number of years since the last dyadic dispute as a set of additional controls.

We find that after controlling for temporal dependence, affinity scores are not statistically significantly associated with the incidence of dyadic MIDs at the .05 level, contrary to the findings reported in Oneal and Russett (1999) and Gartzke (2000). However, when we replace affinity scores with the negative absolute difference in the ideal points of both sides of each dyad, preference similarity becomes a statistically significant ($p < .05$) predictor of reduced dispute propensity. Figure 5 shows the predicted changes in MID probability that correspond to a comparable increase in preference similarity for the affinity and ideal point models. The estimated effects are on the same order of magnitude, but using ideal points increases the precision of the estimated effect. Substituting ideal points for S scores also appreciably improves model fit.[12] These results suggest that our proposed measure may be a more accurate indicator of interstate preference similarity as well as one that is more robust to modeling dynamics.

Conclusion

Votes in the UNGA contain valuable information about the preferences of states. Most political science applications that take advantage of this fact distill preference information via dyadic vote-agreement scores such as S scores. Unfortunately, such scores can move dramatically due to changes in agenda even when the underlying preferences do not change. They produce estimates that in some cases beggar belief, implying, for example, that the United States and Russia today are more at odds than the United States and Soviet Union ever were.

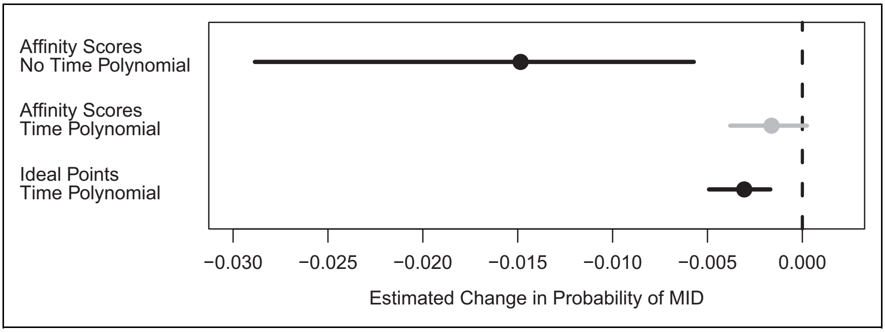

Figure 5. Estimated changes in probability of militarized interstate disputes for a given change in dyadic preferences. Changes in probability correspond to an increase in preference similarity from the twenty-fifth percentile to the seventy-fifth percentile. All other covariates held at their medians. Distributions of first differences computed using Monte Carlo simulation following King, Tomz, and Wittenberg (2000). Points denote the means of the simulated distributions. Lines denote 95 percentile intervals.

We present a spatial theory IRT model that estimates national foreign policy preferences based on UN voting from 1946 to 2012. We identify preference change over time by taking advantage of the fact that the UN voted on some resolutions in multiple years. Our estimates systematically identify widely understood foreign policy consequences of regime changes within countries, such as the switch toward (and away) from the United States when military (and left-wing) regimes take control in Latin American countries. Moreover, the ideal point estimates correlate in expected ways with democratization, financial liberalization, and government ideology.

We end with two recommendations. First, scholars should be specific about what coefficients on similarity indicators derived from UN votes mean. Scholars should not refer to measures based on UN votes as assessments of "common interests" writ large. These measures have nothing to say about whether two countries agree on how to resolve a border conflict or regional issues. The UN deals with issues of global importance. Thus, measures based on UN votes are only useful if we believe, theoretically, that a state's position on global issues matters for the outcome under consideration. If measures are based on all UN votes since World War II, there is a relatively consistent underlying substantive dimension that dominates contestation. We could estimate ideal points for subsets of issues, such as the Middle East, colonialism, or votes, on which the United States has lobbied. Those could yield different substantive results.

Second, scholars should approach dyadic measures such as S scores with caution. We have shown that these yield implausible results, such as the United States and Russia now more divided than the United States and the Soviet Union ever were during the Cold War. Moreover, S scores are very sensitive to annual agenda changes. We see little reason why the analysis of roll call votes in the UN should not follow the example set by roll call analysts anywhere else and utilize an explicit model of how actors translate their preferences into vote choices: the empirical spatial model. Our article has removed the barrier imposed by ordinal vote choices and has shown how consistent dynamic estimates can be derived using the spatial model.

Appendix

The likelihood is:

$$L(\gamma, \theta, \beta) = \prod_i \prod_\upsilon \quad \Phi(\gamma_{y_{i\upsilon}} - \beta_\upsilon \theta_{it\upsilon}) - \Phi(\gamma_{y_{i\upsilon}-1} - \beta_\upsilon \theta_{it\upsilon}),$$

where $\gamma_{y_{i\upsilon}}$ is the γ associated with the choice of country i on vote v. Johnson and Albert (1999, 163) show that the likelihood may be reexpressed in terms of the latent variable Z:

$$L(\gamma, \theta, \sigma) = \prod_i \prod_\upsilon \quad \varphi(Z_{it\upsilon} - \beta_\upsilon \theta_{it\upsilon}) I(\gamma_{y_{i\upsilon}-1} \leq Z_{it\upsilon} < \gamma_{y_{i\upsilon}}),$$

where $I(\gamma_{y_{i\upsilon}-1} \leq Z_{it\upsilon} < \gamma_{y_{i\upsilon}})$ is an indicator function indicating that $Z_{it\upsilon}$ is between the cut points bounding country i's choice (e.g., if the country abstained, then $Z_{it\upsilon}$ is between $\gamma_{\upsilon 1}$ and $\gamma_{\upsilon 2}$).

The estimation process proceeds in the following steps.

1. Set starting values
   (a) For each vote, set $\beta_\upsilon = 1$ if the United States voted yea and set $\beta_\upsilon = -1$ if the United States voted nay. If the United States did not vote, use the United Kingdom as a reference point; if neither the United States nor the United Kingdom voted, set start value of $\beta_\upsilon = 0$.
   (b) For each vote, estimate an ordered probit model with no covariates to generate γ_υ. This is equivalent to assuming all countries have ideal points of zero and estimating an ordered probit.
   (c) Draw $Z_{it\upsilon}$ from distribution centered on 0 and truncated to be between the outpoints bounding country i's choice (in the interval $(\gamma_{\upsilon, y_{i\upsilon}-1}, \gamma_{\upsilon, y_{i\upsilon}})$).[13]
2. Sample θ_{it} from normal distribution with mean $(\beta'_\upsilon \beta_\upsilon)^{-1} \beta'_\upsilon Z_{it\upsilon}$ and variance $(\beta'_\upsilon \beta_\upsilon)^{-1}$.
3. Update the vote parameters (Johnson and Albert 1999, 136):
   (a) Sample candidate values of outpoints g_j from a $N\left(\gamma_j^{(k-1)}, \sigma^2_{MH}\right)$ distribution truncated to interval $\left(g_{j-1}, \gamma_{j+1}^{(k-1)}\right)$, where $g_0 = -\infty, g_3 = \infty$ and σ^2_{MH} is set to $\frac{0.05}{C}$ as per Johnson and Albert's (1999, 135) rule of thumb for this parameter.

(b) Compute the acceptance ratio R:

$$R = \prod_i \frac{\Phi(g_{y_{i\upsilon}} - \beta_\upsilon \theta_{it\upsilon}) - \Phi(g_{y_{i\upsilon}-1} - \beta_\upsilon \theta_{it\upsilon})}{\Phi(\gamma_{y_{i\upsilon}} - \beta_\upsilon \theta_{it\upsilon}) - \Phi(\gamma_{y_{i\upsilon}-1} - \beta_\upsilon \theta_{it\upsilon})} \times \frac{1 - \Phi\left(\frac{g_1 - \gamma_2}{\sigma_{MH}}\right)}{1 - \Phi\left(\frac{\gamma_1 - g_2}{\sigma_{MH}}\right)}.$$

(c) Accept the candidate values with probability R; otherwise use current value of γ_υ.[14]

4. Sample the latent variable Z for each individual's vote from conditional density, given γ and $\theta_{it\upsilon}$, which is a normal density with mean $\theta_{it\upsilon}$ with variance one truncated to interval $(\gamma_{y_i-1}, \gamma_{y_i})$.
5. Sample from the distribution of β_υ with normal distribution with mean $(\theta_i'\theta_i)^{-1}\theta_i' Z_{it\upsilon}$ and variance $(\theta_i'\theta_i)^{-1}$, where θ_i is a vector of ideal points of countries voting on a vote v. Include a prior so that no single vote has such a large β that exclusively defines ideological dimension. This is equivalent to adding a prior on discrimination parameter in dichotomous item response theory models [see Gelman et al. (1995, 254-260) on implementation of the prior].
6. Repeat steps 2 through 5 for a burn in period and then run for a sufficient number of values to obtain a sample of estimates from the posterior density of the parameters. We ran a model with 20,000-iteration burn-in period and took every twentieth sample from the next 20,000 iterations.

Table A1. Government Ideology and UN Voting in Latin American Countries (Ideal Points and S Scores).

Variables	(1) S Scores with the United States	(2) Ideal Points	(3) Ideal Point Similarity with the United States
Lagged dependent	.831***	.894***	.877***
Variable	(.015)	(.020)	(.015)
Left	−.003	−.024	−.023
	(.007)	(.017)	(.017)
Right	.009	.023*	.033***
	(.009)	(.012)	(.011)
Constant	−.079***	−.045***	−.401***
	(.007)	(.008)	(.041)
Observations	968	968	968
R^2	.763	.897	.908

Note: Robust standard errors clustered on countries within parentheses.
***$p < .01$. **$p < .05$. *$p < .1$.

Table A2. Replication of Gartzke (2000)—Logistic Regression of Militarized Disputes on Affinity Scores/Ideal Point Distances.

Variables	Model 1	Model 2	Model 3
Affinity score	−1.114***	−0.414*	
	(0.309)	(0.238)	
Negative ideal			−0.373***
Point difference			(0.100)
Lower democracy	−0.033**	−0.039***	−0.018
	(0.015)	(0.013)	(0.013)
Higher democracy	0.001	0.010	−0.002
	(0.015)	(0.012)	(0.013)
Lower growth rate	−0.046*	−0.021	−0.029
	(0.026)	(0.023)	(0.022)
Allies	−0.346	−0.274	0.001
	(0.262)	(0.207)	(0.257)
Capability ratio	−0.002**	−0.002***	−0.002**
	(0.00075)	(0.00071)	(0.00073)
Contiguity	1.864***	1.028***	1.159***
	(0.343)	(0.287)	(0.297)
Lower dependence	−43.207**	−15.327	−15.229
	(21.730)	(14.086)	(14.265)
Constant	−3.469***	−1.404***	−2.216***
	(0.362)	(0.373)	(0.428)
Time polynomial	No	Yes	Yes
Observations	18,303	18,303	18,303
AIC	5,431.824	4,302.512	4,239.798

Note: Robust standard errors clustered on countries within parentheses. Coefficients denote expected changes in log odds of militarized dispute in year *t*. AIC = Akaike information criterion.
***$p < .01$. **$p < .05$. *$p < .1$.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Notes

1. An exception is Voeten (2000) who treats both abstentions and no-votes as signals of disapproval.

2. Other approaches include Lijphart's (1963) index of agreement which is a proportion but otherwise identical to the S score, Bueno de Mesquita's (1975) ordinal measure using Kendall's τ_b and Gartzke's (1998, 2000) Spearman rank-order correlation coefficient.
3. Signorino and Ritter offer a generalized version, but scholars use the one discussed here.
4. There are two studies that do this for limited numbers of votes or time frame: Voeten (2004) and Mattes, Leeds, and Carroll (2015).
5. Grimmer (2010) uses a similar approach to identify direct quotation of US Senators' press releases in newspaper coverage.
6. We adjusted the threshold at which we manually inspected resolutions based on year. Due to the poor printing of some of the earlier resolutions, optical character recognition software imperfectly translates the documents into text, leading to more errors. We also eliminated budget and other annual resolutions that could clearly not have an identical interpretation across years.
7. This is the R^2 from STATA's xtreg procedure.
8. Note that these are averages. Some Middle East resolutions separate countries along East–West lines very well.
9. Results available from the authors.
10. These covariates are the lower and higher levels of dyadic democracy (measured using the Polity scale), a dyadic alliance indicator, geographic contiguity, per capita economic growth, trade dependence, and the ratio of the stronger state's military capability to the weaker state's military capability. The exact construction of these variables is detailed in Oneal and Russett (1997).
11. Due to this modification, we have 18,303 dyad-year observations in our data set compared to 17,910 reported in the study by Gartzke.
12. A lower Akaike information criterion (AIC) is preferred. The AIC for the model with the ideal point measure is 4,239.8 compared to the affinity model's AIC of 4,302.5.
13. A draw from a $N(c, d^2)$ distribution truncated to (a, b) is $Y = c + d\Phi^{-1}\left[\Phi\left(\frac{a-c}{d}\right) + U\left(\Phi\left(\frac{b-c}{d}\right) - \Phi\left(\frac{a-c}{d}\right)\right)\right]$, where U is a random uniform variable Gelfand et al. (1990, 977).
14. When there are no yeas, $\gamma_{2\upsilon} = \infty$, when there are no nays, $\gamma_{1\upsilon} = -\infty$, and when there are no abstentions, $\gamma_{1\upsilon} = \gamma_{2\upsilon}$.

References

Alesina, Alberto, and David Dollar. 2000. "Who Gives Foreign Aid to Whom and Why?" *Journal of Economic Growth* 5 (1): 33-63.

Andersen, Thomas Barnebeck, Thomas Harra, and Finn Tarp. 2006. "On US Politics and IMF Lending." *European Economic Review* 50 (7): 1843-62.

Bailey, Michael A. 2007. "Comparable Preference Estimates across Time and Institutions for the Court, Congress, and Presidency." *American Journal of Political Science* 51 (3): 433-48.

Bailey, Michael A. 2013. "Is Todays Court the Most Conservative in Sixty Years? Challenges and Opportunities in Measuring Judicial Preferences." *Journal of Politics* 75 (3): 821-34.

Ball, M. Margaret. 1951. "Bloc Voting in the General Assembly." *International Organization* 5 (1): 3-31.

Bearce, David H., and Stacy Bondanella. 2007. "Intergovernmental Organizations, Socialization, and Member-state Interest Convergence." *International Organization* 61:703-33.

Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. "Taking Time Seriously: Time-series-cross-section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42 (4): 1260-88.

Bloomfield, Louis. 2011. "WCopyFind 3.0.2." Software. Accessed September 1, 2011. http://plagiarism.bloomfieldmedia.com/z-wordpress/.

Bueno de Mesquita, Bruce. 1975. "Measuring Systemic Polarity." *Journal of Conflict Resolution* 19(2): 187-216.

Carter, David B., and Curtis S. Signorino. 2010. "Back to the Future: Modeling Time Dependence in Binary Data." *Political Analysis* 18 (3): 271-92.

Carter, David B., and Randall Stone. 2015. "Democracy and Multilateralism: The Case of Vote Buying in the UN General Assembly." *International Organization* 69 (1): 1-33.

Clinton, Joshua, Simon Jackman, and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98 (2): 355-70.

Copelovitch, Mark S. 2010. "Master or Servant? Common Agency and the Political Economy of IMF Lending." *International Studies Quarterly* 54 (1): 49-77.

Cowles, Mary Kathryn. 1996. "Accelerating Monte Carlo Markov Chain Convergence for Cumulative-link Generalized Linear Models." *Statistics and Computing* 6 (2): 101-11.

De Boef, Suzanna, and Luke Keele. 2008. "Taking Time Seriously." *American Journal of Political Science* 52 (1): 184-200.

Dreher, Axel, and Martin Gassebner. 2008. "Does Political Proximity to the U.S. Cause Terror?" *Economic Letters* 99 (1): 27-29.

Dreher, Axel, and Nathan Jensen. 2007. "Independent Actor or Agent? An Empirical Analysis of the Impact of U.S. Interests on International Monetary Fund Conditions." *The Journal of Law and Economics* 50:105-24.

Dreher, Axel, and Nathan Jensen. 2013. "Country or Leader? Political Change and UN General Assembly Voting." *European Journal of Political Economy* 29:183-96.

Dreher, Axel, P. Nunnenkamp, and R. Thiele. 2008. "Does US Aid Buy UN General Assembly Votes? A Disaggregated Analysis." *Public Choice* 136 (1): 139-64.

Dreher, Axel, and Jan-Egbert Sturm. 2012. "Do the IMF and the World Bank Influence Voting in the UN General Assembly?" *Public Choice* 151 (1-2): 363-97.

Drieskens, Edith. 2010. "Towards a Systematic Analysis of the EU as an Actor in the UN System." *Journal of International Organizations Studies* 1 (1): 105-8.

Flores-Macías, Gustavo A., and Sarah E. Kreps. 2013. "The Foreign Policy Consequences of Trade: China's Commercial Relations with Africa and Latin America, 1992–2006." *The Journal of Politics* 75 (2): 357-71.

Gartzke, Erik. 1998. "Kant We All Just Get along? Opportunity, Willingness, and the Origins of the Democratic Peace." *American Journal of Political Science* 42 (1): 1-27.

Gartzke, Erik. 2000. "Preferences and the Democratic Peace." *International Studies Quarterly* 44 (2): 191-212.

Gartzke, Erik. 2007. "The Capitalist Peace." *American Journal of Political Science* 51 (1): 166-91.

Gelfand, Alan E, Susan E Hills, Amy Racine-Poon, and Adrian Smith. 1990. "Illustration of Bayesian inference in normal data models using Gibbs sampling." *Journal of the American Statistical Association* 85 (412): 972-985.

Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin. 1995. *Bayesian Data Analysis*. Boca Raton, FL: Chapman and Hall.

Grimmer, Justin. 2010. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases." *Political Analysis* 18 (1): 1-35.

Hage, Frank M., and Simon Hug. 2013. "Consensus Voting and Similarity Measures in IOs." Paper Presented at the 2013 EPSA Annual Meeting, Barcelona, Spain, June 2013.

Johnson, Valen E., and James H. Albert. 1999. *Ordinal Data Modeling*. New York: Springer.

King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44 (2): 347-61.

Koremenos, Barbara. 2005. "Contracting around International Uncertainty." *American Political Science Review* 99 (4): 549-65.

Lewis, Jeffrey B., and Chris Tausanovich. 2013. "Has Joint Scaling Solved the Achen Objection to Miller and Stokes?" In *UCLA Mimeo*. Accessed July 13, 2015. http://www.vanderbilt.edu/csdi/miller-stokes/05_MillerStokes_LewisTausanovitch.pdf.

Lijphart, Arend. 1963. "The Analysis of Bloc Voting in the General Assembly: A Critique and a Proposal." *The American Political Science Review* 57 (4): 902.

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): 134-53.

Mattes, Michaela, Brett Ashley Leeds, and Royce Carroll. 2015. "Leadership Turnover and Foreign Policy Change: Societal Interests, Domestic Institutions, and Voting in the United Nations." *International Studies Quarterly* 59(2): 280-90.

Moon, Bruce E. 1985. "Consensus or Compliance? Foreign-policy Change and External Dependence." *International Organization* 39 (2): 297-329.

Neumayer, Eric. 2008. "Distance, Power and Ideology: Diplomatic Representation in a World of Nation-states." *Area* 40 (2): 228-36.

Oneal, John R., and Bruce M. Russett. 1997. "The Classical Liberals Were Right: Democracy, Interdependence, and Conflict, 1950–1985." *International Studies Quarterly* 41 (2): 267-94.

Oneal, John R., and Bruce M. Russett. 1999. "Is the Liberal Peace Just an Artifact of Cold War Interest? Assessing Recent Critiques." *International Interactions* 25 (3): 213-41.

Poole, Keith T. 2005. *Spatial Models of Parliamentary Voting*. Cambridge, UK: Cambridge University Press.

Quinn, Dennis P., and A. Maria Toyoda. 2008. "Does Capital Account Liberalization Lead to Growth?" *Review of Financial Studies* 21 (3): 1403-49.

Reed, William, David H. Clark, Timothy Nordstrom, and Wonjae Hwang. 2008. "War, Power, and Bargaining." *Journal of Politics* 70 (4): 1203-16.

Rosas, Guillermo, and Yael Shomer. 2008. "Models of Nonresponse in Legislative Politics." *Legislative Studies Quarterly* 33 (4): 573-601.

Rosas, Guillermo, Yael Shomer, and Stephen Haptonstahl. 2015. "No News Is News: Non-ignorable Non-response in Roll-call Data Analysis." *American Journal of Political Science* 59 (2): 511-28.

Russett, Bruce M. 1966. "Discovering Voting Groups in the United Nations." *The American Political Science Review* 60:327-39.

Signorino, Curtis S., and Jeffrey M. Ritter. 1999. "Tau-b or Not Tau-b: Measuring the Similarity of Foreign Policy Positions." *International Studies Quarterly* 43 (1): 115-44.

Strezhnev, Anton, and Erik Voeten. 2013. "United Nations General Assembly Voting Data." Accessed July 13, 2015. http://hdl.handle.net/1902.1/12379.

Sweeney, Kevin J. 2003. "The Severity of Interstate Disputes." *Journal of Conflict Resolution* 47 (6): 728-50.

Thacker, Strom. 1999. "The High Politics of IMF Lending." *World Politics* 52 (1): 38-75.

Vengroff, Richard. 1976. "Instability and Foreign Policy Behavior: Black Africa in the U.N." *American Journal of Political Science* 20 (3): 425.

Voeten, Erik. 2000. "Clashes in the Assembly." *International Organization* 54 (2): 185-215.

Voeten, Erik. 2004. "Resisting the Lonely Superpower: Responses of States in the United Nations to U.S. Dominance." *Journal of Politics* 66 (3): 729-54.

Voeten, Erik. 2013. "Data and Analyses of Voting in the United Nations General Assembly." In *Routledge Handbook of International Organization*, edited by B. Reinalda, 54-66. London, UK: Routledge.

Wang, Te-Yu. 1999. "US Foreign Aid and UN Voting: An Analysis of Important Issues." *International Studies Quarterly* 43 (1): 199-210.