A Latent Factor Approach to Measuring State Preferences

Abstract

State preferences play an important, yet under discussed role in international politics. This is in large part because actually observing and measuring these preferences is impossible. In general, scholars have tried to infer preferences using either UN voting or alliance behavior. The two most notable measures of state preferences that have flowed from this research area are ideal point (Bailey et al., 2015) and S Scores (Signorino & Ritter, 1999). The basis of both these models is a spatial weighting scheme that has proven useful but discounts higher-order effects that might be present in relational data structures such as UN voting and alliances. We begin by arguing that both alliances and UN voting are simply examples of the multiple layers upon which states interact with one another. To estimate a measure of state preferences from this multilayer structure, we introduce a latent factor model that provides a reduced-rank approximation of the main patterns across the layers. Our new measure of preferences plausibly describes important state relations, and yields important insights on the relationship between preferences, democracy, and international conflict. Most importantly, a model of conflict that uses this measure of state preferences decisively outperforms models using extant measures when it comes to predicting conflict in an out of sample context.

Why we care about preferences

Compared to other concepts in the study of international conflict and cooperation, the foreign policy preferences of states have been understudied. Perhaps this is a consequence of black box theories of international relations, where it is assumed that all state preferences can be reduced to an axiomatic desire for more power. Alternatively, it may be because it is more difficult to measure concepts such as preferences compared to tangible material or institutional factors.

This dearth of attention belies the importance of preferences in our theories of international processes. A number of formal theories of international relations require measures of preferences to be tested: the expected utility theory proffered by (Bueno de Mesquita, 1983) has similarity of preferences as an important input. Further attempts to expand studies of crisis bargaining to include mediation (Kydd, 2003), coalitional dynamics (Wolford, 2014), or the possibility of additional disputants (Gallop, 2017) require a measure of state preferences in order to predict whether war will be the result of bargaining failure. Preferences have been used in empirical studies predicting bilateral trade, foreign aid, stability of international institutions and the incidence of conflict (DeRouen and Heo, 2004; Stone, 2004; Gartzke, 2007; Kastner, 2007; Braumoeller, 2008).

A substantive theoretical reason for why we need a good measure of preferences is to correctly understand the democratic peace. It is difficult to entangle whether democracies avoid war with other democracies because of the intrinsic nature of democracy, or simply because they appear to share similar ends. Farber and Gowa (1995) argue that democracies were only peaceful during the Cold War period because they had similar preferences and alliance structures. Similarly, Gartzke (1998) argues that dissimilar preferences are a necessary condition for conflict. Oneal and Russett (1999) respond by arguing that democracy has both a direct inhibiting effect on conflict, and an indirect one through influencing state preferences.[1] While there has been some impressive development with our measures of preferences in recent years, a more accurate measure is essential to disentangle the extent to which peace is the product of shared preferences, and the extent to which institutions and norms are driving peace.

Much of the extant literature has focused on estimating state preferences by utilizing spatial weighting models on either alliance behavior or United Nations (UN) voting scores. These approaches have proven to be useful but there are two reasons to desire a different approach. First, alliances are rare and voting together in the UN is very common, so, by only focusing on the direct dyadic behavior, we risk mischaracterizing important relationships. Second, we would expect a better understanding of state preferences to help us predict state behavior, but as we show in figure **??** adding measures of state preferences to a traditional model of interstate disputes yields relatively scant increases in our predictive ability.

We argue that we can improve on these measures of preferences using the same raw material by acknowledging that both alliance membership and UN voting are examples of relational data that takes place in a network. A bevy of research has shown that accounting for network structure necessitates an approach that can account for higher-order dependence patterns such as homophily and stochastic equivalence. As such, we make two contributions to the existing literature on state preferences. We introduce a latent factor model that accounts for higher-order dependence patterns across multiple layers. We show that our revised approach of measuring state preferences both better characterizes relationships that have counterintuitive results using

[1] Gartzke (2000) argued that even though democracies might have similar preferences, the residual of preferences from democracy explains conflict much better than the residual of democracy from preferences.

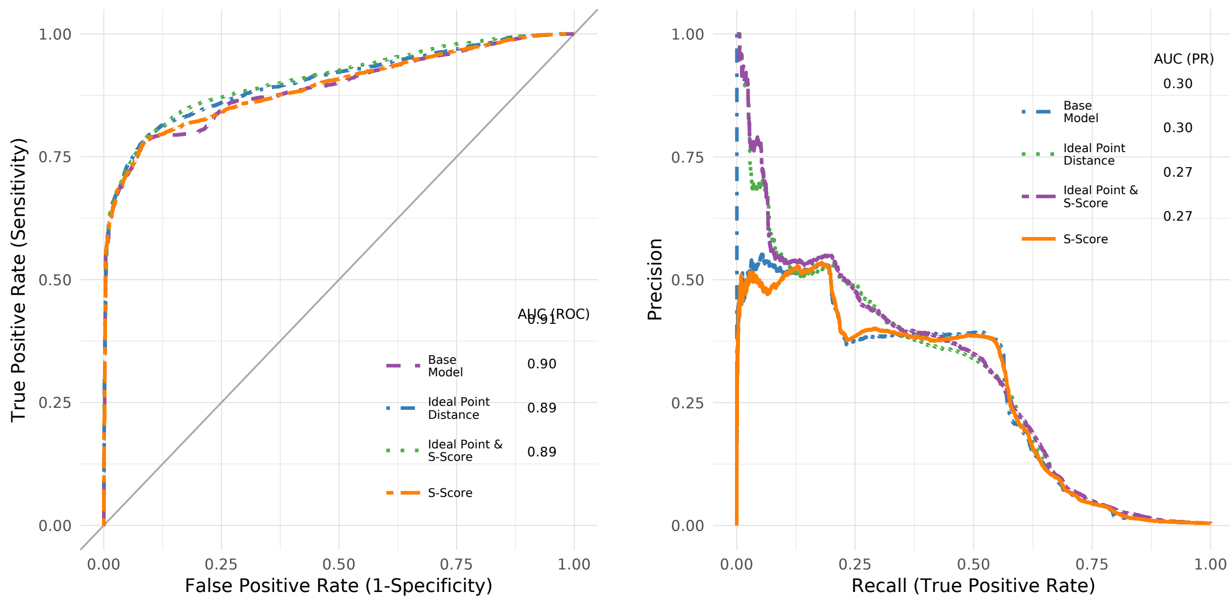

Figure 1: Assessments of out-of-sample predictive performance of Militarized Interstate Disputes using ROC curves and PR curves. AUC statistics are provided as well for both curves.

existing measures of preferences, and this measure greatly enhances our ability to predict instances of conflict between countries when compared to existing measures.

Sources of Preference Measures: Alliance Portfolios and UN Voting

Given that we cannot directly observe state preferences, scholars have attempted to estimate preferences using two main behavioral indicators: who states choose to ally with and how states vote at the United Nations (UN). The idea behind alliance portfolio measures is that we can infer a state's foreign policy by looking at the states they choose to align with. In the extreme case, if two states have all of the same allies, it is likely that their foreign policy goals are quite similar. Conversely, if all allies of one state are not allied to another, and vice versa, our best guess is that these states would have different aims and desires in foreign policy. Bueno De Mesquita and Lalman (2008) encapsulate the logic when they note that "alliance commitments reflect a nation's position on major international issues". Measures of alliance behavior do, however, suffer from the fact that these measures are largely static and sparsely occurring. Formal alliances are relatively constant over time, whereas in many cases state preferences will be more fluid, and therefore these scores will be at best a lagging indicator of preferences. Furthermore, as Häge (2011) points out, the fact that links are so rare creates an artificial similarity of alliance portfolios.

We also have a relatively large corpus of behavioral information in UN Voting Records. The cost of voting in the UN is low, and so, scholars have argued that measures of affinity based on UN voting are relatively representative of the underlying distribution of preferences (Gartzke, 1998). This is especially fortuitous because the methodology of inferring preferences from voting in a legislature is relatively advanced. A few issues with these measures are that the potential benefit of winning UN votes is low, and so states might have incentives to vote against their preference as they are not costly signals, and the distribution of UN voting is prone to large supermajorities of the type rarely seen in “ordinary" legislatures.

Current measures of preferences: S-Scores

The initial measure used to measure preference similarity based on alliance portfolios was Kendall’s τ_B (Bueno De Mesquita and Lalman, 2008). This measure is:

$$\tau_B = \frac{n_c - n_d}{\sqrt{(n_0 - n_1)(n_0 - n_2)}} \tag{1}$$

where n_c is the number of pairs where both actor i and j have the same rank ordering (for example both the UK and the US are more closely allied to Israel than to Iran), n_d is the number of pairs where they have discordant rankings (the US is more closely allied to Saudi Arabia than to Russia, Syria is more closely allied to Russia than to Saudi Arabia). The denominator attempts to adjust the total number of pairs with the number of ties: n_0 is the total number of pairs ($n(n-1)/2$), n_1, n_2 are measures for ties in both i and j's rankings respectively.

Signorino and Ritter (1999) convincingly pointed to flaws in this measure, notably its focus on rank-ordering as applied to a context where we instead care mostly about the presence or absence of an alliance. In addition, if we add additional strategically irrelevant states, we will create artificially high τ_B statistics. Thus, Signorino and Ritter introduce the S score, which has since been the most widely used alliance similarity measure.[2]

The equation for the S score is:

$$S(P^i, P^j, W, L) = 1 - 2w_k \frac{d(P^i, P^j, W, L)}{d^{\text{max}}(W, L)} \tag{2}$$

Where:

$$d(P^i, P^j, W, L) = \sum_{k=1}^{N} \frac{w_k}{\Delta_k^{\text{max}}} |p_k^i - p_k^j| \tag{3}$$

Where w_k is the k'th element of a weight matrix, $d^{\text{max}}(W, L)$ is the maximal distance on a given dimension, and Δ is a normalizing constant. For the weight matrix, generally analysis has used S scores calculated with a weight matrix of ones–giving each potential ally equal weight–though the other plausible choice would be to weight states by import, for example using their share of world military capability, as calculated by Singer and Small (1995). Gartzke (1998) attempted to apply a similar S-score methodology to UN voting data and created the "Affinity of Nations" index.

An advantage of utilizing UN General Assembly Voting, is that it allowed the field to take advantage of methodological advances that have been made in the study of legislatures. Bailey and Voeten (2015) do so by using an Item Response Theory model on

[2]Bennett and Rupert (2003) also find a stronger relationship between theoretical predictions and results when using S-scores than when using τ_B.

UNGA voting. This model seeks to place states on a unidimensional latent preference space using their voting behavior. The assumption of this model is that states' votes on a resolution are a function of states' ideal points, characteristics of the vote, and random error. In particular, for each bill v, a state's vote will be based on the latent variable Z_{itv}

$$Z_{itv} = \beta_{iv}\theta_{it} + \epsilon_{iv} \tag{4}$$

such that the state will vote yes if $Z_{itv} < \gamma_{1v}$, no if $Z_{itv} > \gamma_{2v}$ and otherwise abstain. Here, θ_{it} is state i's ideal point at time t, and β_{iv} is the discrimination parameter of a particular bill v. When β_v is positive, states with high ideal points will be more likely to vote no. When it is negative, they will be more likely to vote yes.

The authors specifically fix the parameters γ_{1v} and γ_{2v} such that the same bill will have the same value in different years, and they standardize and normalize θ. They also use θ_{it-1} as a prior on θit. With these constraints, they solve for the ideal points using a Metropolis Hastings Markov Chain Monte Carlo (MCMC).

Both methods relying on UN data, and those relying on alliances have difficulties distinguishing within '0's and '1's. For example, if we know two states are allies, we have reason to believe they have similar preferences, but if we know they are not allies, it is not clear whether they are enemies or they are indifferent – the United States is "not-allies" with both Bhutan and North Korea for example. As of 2012, using the Correlates of War projects alliance data, only about 1/8th of all dyads were between countries with any sort of alliance. Similarly, with UN voting, so many UN votes contain super-majorities and states vote together a huge proportion of the time. If we only look at yes and no votes in the UN general assembly, the median pair of states has voted together about 96% of the time. If we include abstentions, they have voted together 86%

of the time. So when two states vote together it is hard to distinguish between states voting together because of similar preferences, or just preferences that are not radically dissimilar. Both S-scores and Item Response theory succeed in adding granularity and nuance to these rough measures, but they are both limited by focusing only on a relationship between two states.

We can see the danger of focusing only on direct relations when we view how these two measures of preference treat relations over the Korean peninsula. If we take China, North Korea, South Korea, and the United States, we would expect the US and South Korea to have preferences that are similar to each-other and dissimilar from China and North Korea (and vice versa). Yet, if we look at extant measuresof preferences (as of 2012), as depicted in SOMETHING **??**, they do not seem to effectively characterize this relationship. S-scores based on alliance portfolios posit that China and the two Koreas are closely clustered, with the US distant from all three, while ideal point distances put South Korea as equidistant between the US and the North. Now it could be that these measures are producing a novel, counterintuitive, result, but given the failures of extant preference measures to add much to our predictive models of conflict, one might be skeptical.

	US/ROK	DPRK/ROK	US/DPRK	China/ROK	US/China	China/DPRK
idPtDist	1.95	2.69	4.65	1.58	3.53	1.11
sScore	0.22	0.95	0.17	0.94	0.16	0.99

The sources of these surprising results become more clear when looking at the raw data on which they are based. In 2012, according to the correlates of war, North Korea only had 3 alliances – a non-aggression pact with the South, and alliances with Russia and China. Similarly, South Korea's only alliances were that non-aggression pact, and an alliance with the United States. Given the US's many other allies, there was much more divergence between their alliance portfolio and South Korea's, than there was between the two Koreas' portfolios. Similarly, when looking at voting patterns at the UN, South Korea voted 50% of the time with the US, 63% of the time with North Korea, and 70% of the tie with China.

Even using this data, we can do better at measuring state preferences when we treat alliances and UN voting as relational data. When it comes to alliance behavior, the fact that North Korea was allies with China and Russia, and South Korea with the United States could give us additional information, because the alliance behaviors of the US on the one hand, and Russia and China on the other hand, are so divergent. Similarly, while South Korea only voted with the US 50% of the time at the UN, this was in the top 15% of all countries in terms of voting with the US, whereas South Korea was in the bottom 20% in terms of the proportion of time voting with China.

We thus argue that two changes can substantially improve our measures of state preferences. First, both alliances and UN voting contain information about state foreign policy preferences, and given the limits of this information, we should find a way to use both. Second, by using network techniques and treating this data as relational data, we can wring more information from the stone, and get both a more nuanced, and more accurate view of state affinity and state preferences.

Synthesizing Measures of State Preference

The approach that we introduce here to measure state preferences starts by assuming that both UN voting and alliance relationships are sources of information on how states relate to one another in the international stage. By accounting for the multiple layers upon which states interact with one another we can synthesize a better measure of state preferences than if we relied on any one measure alone. The idea of using multiple metrics to get a better handle on preferences is not new, in fact, Signorino and Ritter suggested it in introducing S scores, which were designed to allow for aggregation of similarity on multiple dimensions (such as alliances and UN voting). The downside of this extant approach, however, is that it does not account for structural patterns that we often see in relational data.

Relational data is composed of observations between pairs of actors, or dyads. For both alliance relationships and UN voting, we are able to observe how the actors in the international system interacted with one another across time. This system of interactions taken in its totality defines a network, and within these types of structures a bevy of research has shown that we need methods that go beyond assuming that interactions are taking place between just two actors in a vacuum (Wasserman and Faust, 1994; Snijders and Nowicki, 1997). As such we reformulate the problem of determining state preferences in terms of network analysis. The goal of our approach is summarized in Figure 2. In the top row, we represent UN voting and alliance patterns at time t in a pair of adjacency matrices that form a multiplex.[3] This multiplex represents the relations between states across two dimensions, and our goal is to extract a lower dimensional representation that accounts for higher order network patterns such as stochastic equivalence. The end result then will be a single $n \times n$ matrix, where n represents the number of actors, in which the cross-sections denote our predictions for the similarity of preferences between two countries.

We generate these predictions through a matrix decomposition technique that estimates a latent Gaussian score for each country pair. We will show that by combining different measures of state preferences, and better accounting for network dependencies through our approach, we are able to generate a measure for preference that: maintains the insights of both UN voting scores and S-scores; and which can yield new insights, in particular, when it comes to predicting and explaining interstate conflict.

[3]The approach that we describe here can be generalized to a multiplex with more than two dimensions.

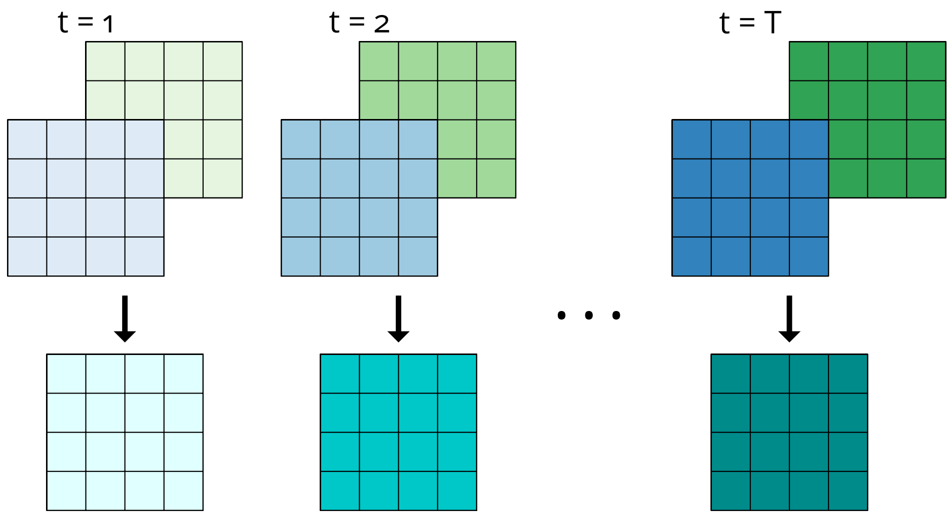

Figure 2: The green and blue colors represent different relational measures and darker shading indicates later time periods. Our goal is to reduce the patterns found across those layers of relationships into a single measure.

Latent Factor Model

A number of techniques have been developed in recent years to estimate low-dimensional representations of network structures.[4] For many of these approaches the goal is to account for higher order dependence pattern such as stochastic equivalence. Stochastic equivalence refers to the idea that there are communities of nodes in a network, and actors within a community act similarly towards those in other communities. Thus the community membership of an actor provides us with information on how that actor will act towards others in the network. Put more concretely, a pair of actors ij are stochastically equivalent if the probability of i relating to, and being related to, by every other actor is the same as the probability for j (Anderson et al., 1992). An additional dependence pattern that often manifests in networks is homophily – the tendency of actors to form transitive links. The presence of homophily in a network implies that actors may cluster together because they share some latent attribute. In

[4] For a literature review of these approaches see Goldenberg et al. (2010).

the context of clustering in alliance relationships, we are likely to find that states like the United States, United Kingdom, and Germany may cluster together because they share some latent state level attribute. We would ignore salient information if we did not use, for example, the UK's behavior towards third parties, when trying to understand the United States' preference similarity with those parties. Doing so is likely to paint an incomplete picture of the preferences that states share with one another.

We account for higher order dependence patterns using a latent factor model (LFM) that allows us to capture both concepts discussed above: the tendency of actors to assort themselves into groups and to form transitive links (Hoff, 2007; Minhas et al., 2016). Using the LFM ensures that similarity in preferences are likely to be transitive, for example, if the US has similar preferences to the UK, and the UK to France, the US's preferences should be relatively close to France's. The most useful feature of the LFM for our purpose is that it provides a visual interpretation of those interdependencies by inferring actor positions in a k-dimensional latent vector space. Actors that have vectors pointing in similar directions are more likely to have similar state preferences based on their alliances and UN voting records. The angle between the vectors for actors i and j provides an estimate of how similar the state preferences of i are to j.

To generate this measure we begin by constructing T different $n \times n \times p$ arrays, where T represents the number of periods, n represents the number of actors,[5] and p the number of observed variables used to synthesize a measure of state preference. In this case, both alliance relationships and UN voting scores are undirected measures, meaning that $y_{ij} = y_{ji} \ \forall \ p$ and t.[6] In order to obtain a lower-dimension relational

[5]The number of actors can vary by period.
[6]The approach we describe below has already been generalized to the case where $y_{ij} \neq y_{ji}$.

measure of state preferences, we use the LFM separately for each time point:

$$
\begin{aligned}
Y &= f(\theta) \\
\theta &= \beta^\top X + Z \\
Z &= M + E \\
M &= U \Lambda U^\top, \text{ where} \\
&\quad u_i \in \mathbb{R}^{\mathrm{k}} \text{ and} \\
&\quad \Lambda \text{ is a } k \times k \text{ diagonal matrix}
\end{aligned}
$$

where $f(.)$ is a general link function corresponding to the distribution of Y and $\beta^\top \mathbf{X}$ is the standard regression term for dyadic and nodal fixed effects. In this application, for the sake of parsimony we abstain from using fixed effects. However, if one was interested in estimating a measure of preference that parsed out the effect of geographic distance between i and j, for example, than this could be accomplished within the context of this framework by simply including that as a covariate in the X design array. Z represents any additional patterns in data unrelated to the specified dyadic and nodal fixed effects. To incorporate multiple measures of similarity into a single ideal point estimation, we treat each different slice of data as arising from a common distribution. In this way, each additional observed relationship between actors that we add serves to provide additional information to the model that can be used to estimate a latent measure of state preference.

The key part of this model lies in the decomposition of Z. Specifically, we can write $Z = M + E$ such that the matrix E represents noise, and M is systematic effects. By matrix theory, we can factorize M into the product of two simpler matrices: $M = U \Lambda U^\top$, where $u_i \in \mathbb{R}^{\mathrm{k}}$ is a latent vector associated to node i and Λ is a $k \times k$ diagonal matrix. Thus under this framework a vector of latent characteristics are estimated for each actor, $u_i = \{u_{i,1}, \dots, u_{i,k}\}$. Similarity in the latent factors between two actors, $u_i \approx u_j$, corresponds to how stochastically equivalent they are and the diagonal entries in Λ, $\lambda_k > 0$ or $\lambda_k < 0$, determine the level of homophily (or antihomophily) in the network (Minhas et al., 2016). Within this framework, the LFM can represent either positive or negative homophily in varying degrees and stochastially equivalent actors may or may not share strong relationships with others in their "community".

Inference of the latent vectors for each actor takes place within the context of a MCMC procedure that enables us to construct approximate samples from the posterior distributions of the latent variables. For the latent factor model, a diffuse normal prior is placed on Λ and the prior distribution on U is taken to be a uniform distribution. The MCMC proceeds by sampling the parameters from their full conditional distributions for each k: sample $\{u_i, \dots, u_n\}$ from a multivariate normal distribution and then sample Λ from its multivariate normal distribution.[7]

The key output from the LFM for our purpose here is $U\Lambda U^{\top}$. Actor positions in the latent factor space are characterized by the concepts of homophily and stochastic equivalence discussed above. Interpretation of this space needs to be done with care. Distances between actors cannot be interpreted using Euclidean metrics, as actor's latent positions are actually embedded within a k-dimensional hypersphere. This means that an actor's vector direction within the k-dimensional hypersphere indicates which latent preferences an actor i has and does not have. Comparing the similarity of preferences between two states, $\{i, j\}$, can be accomplished by comparing the direction to which their respective factor vectors point. A commonly used metric for this sort of problem in the recommender system literature from computer science is the cosine

[7] A Bayesian procedure to determine the eigenvalue decomposition is available in the **amen** R package (Hoff et al., 2017).

of the angle formed by the latent vectors of both actors.[8] We refer to this distance metric as latent angle distance. Thus, if the estimated latent vectors of two states are in the same direction, they are apt to have both alliances and UN co-voting to similar partners. The way we measure this similarity in dimension is by looking at the absolute distance of the angles created by each states position and the center of the latent space.

Data Sources, Modeling choices

We use the LFM with $k = 2$ on the two aforementioned measures of state amity to generate a combined measure of state preference similarity which accounts for network effects. As inputs, we use the UN voting similarity index (Voeten, 2013)[9] and number of alliance relationships between states by time t (Gibler and Sarkees, 2004). However, to facilitate comparison between the metrics, we first standardize and normalize these two measures. This gives us an $n \times n \times T \times 2$ array, where the first two dimensions represent countries, the third dimension is the year, and the fourth is the particular measure of similarity. So the item at index (1,2,1,1) would be the transformed value of the S-score for countries i and j at the first year of our data, similarly (1,2,1,2) would be the UN ideal point distance measure.

With this data, we run an LFM with a Gaussian link, and in particular we use the $U\Lambda U^\top$ term produced by each yearly model to estimate each states position in a two-dimensional hypersphere. We then evaluate whether their is additional utility gained from using this latent angle position, as compared to the similarity of alliance portfolios and UN ideal point distance measures.

[8] For a review of this literature see (Xavier Amatriain and Pujo, 2015).

[9] Specifically, the "agree3un" variable.

Face Plausibility

An important question for these different measures of preferences, are whether they give results that "make sense" for prominent pairs of countries. In particular we would hope that the measure both provides plausible levels to relationships – sorting states into friends and foes effectively – and that when these measures change, they do so in ways that correspond to changes in the world. We thus present all three measures' accounts of eight dyadic relationships over time.[10] Additionally, note that each of these measures is on a different scale, and so just because one measure has a higher value, does not necessarily mean that it posits more dissimilarity of preferences.

We first look at three close relationships where we would expect to see similar preferences. Figure 3 depicts the relationships between France and Germany, the US and Israel, and North Korea and China. In all three cases our measure using latent angle distance has consistently low values – nearly 0 in the case of France and Germany, and China and North Korea, and low but less stable values for the US/Israel relationship. Alliance S-Scores correctly classifies Franco-German and Sino-North Korean friendship, while the US/Israel relationship is characterized as being relatively neutral. The measure based on UN voting is most out of step in terms of characterizing these relationships, with notable divergence in the preferences of many of these pairs.

In figure 4 we look at the United States' relationship with three countries that are characterized by change and major events. S-scores are the only measure that does not detect a marked improvement in the US/Russian relationship at the end of the Cold War. Both the UN ideal points and latent Angle space find significant improvements followed by a drift toward enmity, whereas S-scores have a constant (though slightly improving) neutral relationship. For the US and Iraq no measure depicts more similar

[10] For the sake of having each measure operate in the same direction, we transform the S-score such that the value shown is 1 - S. Thus for all three variables, 0 implies perfect symmetry of preferences.

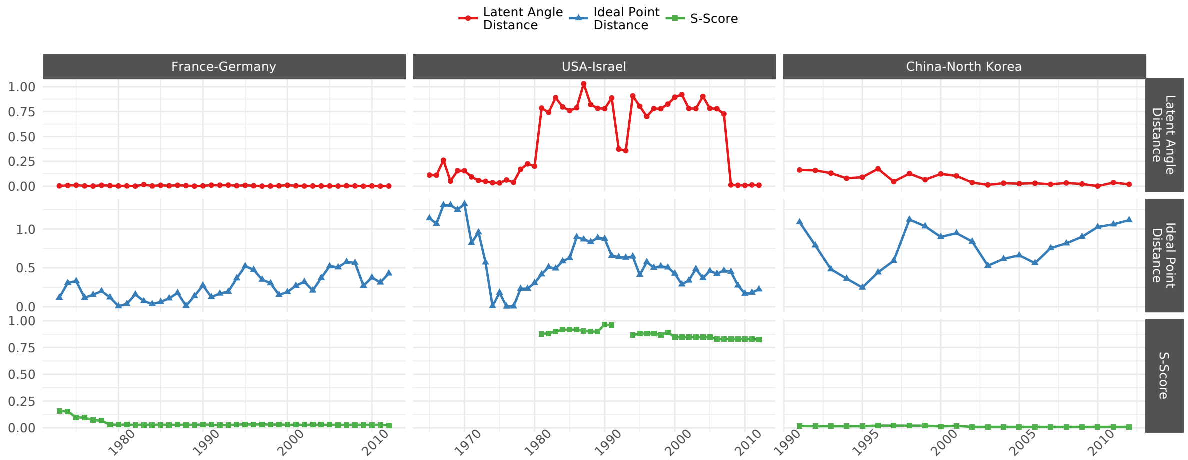

Figure 3: Dyadic relationships between 3 pairs of countries according to the different preference measures. Yearly latent angle distances between dyads are denoted by red circles and ideal point distance based on UN voting in blue triangles. S-score values have been rescaled such that they works in the same direction as the other measures (1-S-score, so that the best score is 0), and are depicted by green squares.

preferences following the US occupation, though only latent angle distance finds an increase in enmity in the run-up to the United States' invasion. Finally, for the US and China, the S-score has a consistent neutral relationship, while both UN ideal points and latent angles both have more negative relationships, and more variance – with the latent angles metric finding particularly large spikes at the beginning of the Bush and Obama administrations.

A big difference between these measures is that latent angle distance has a fuller time series than the other two. This is particularly relevant when looking at the preferences of certain rogue states, as we do in figure 5. In both the case of Iran/Iraq and North Korea/South Korea there is no data for UN voting (and for Iran/Iraq missing data for S-scores). For the Korean relationship, latent angle distance better characterizes the relationship as enmity, while S-scores treat the two as having very similar preferences. Whereas for Iran and Iraq's relationship S-scores give a consistent close preferences (where data exists), while the latent angle metric show more instability, and notable

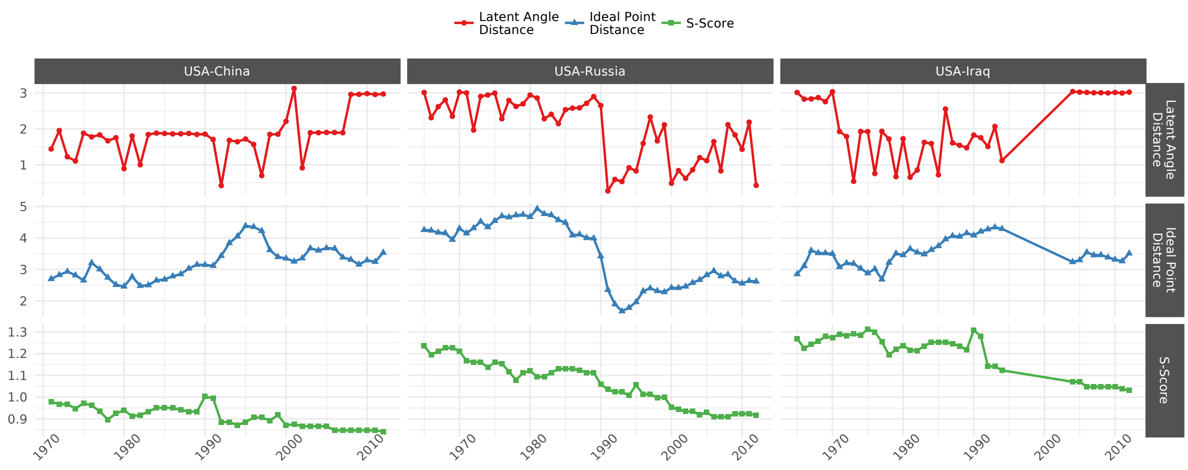

Figure 4: Dyadic relationships between 3 pairs of countries according to the different preference measures. Yearly latent angle distances between dyads are denoted by red circles and ideal point distance based on UN voting in blue triangles. S-score values have been rescaled such that they works in the same direction as the other measures (1-S-score, so that the best score is 0), and are depicted by green squares.

elevation (though not as large as we would expect) during their war.

As can be seen from these relationships, the measure of preferences based on latent angle distance is in many cases as plausible or more than incumbent measures of preferences. While the measure has more temporal instability than S-scores or UN ideal points, in these 8 cases it often does better, and rarely worse at conforming to our expectations of the relationships. Of course, this could just be a case of us picking particularly propitious cases. Next, we conduct a the large-N analysis of conflict and compare the added utility of these various measures of state preferences.

Model Competition

To evaluate the different measures of state preferences, we assess their utility in predicting the occurrence of interstate disputes. Here we look at four non-nested models: a model using no measures of state preferences, one using an S-Score based on similarity in alliance portfolio (Signorino and Ritter, 1999), one using the ideal points

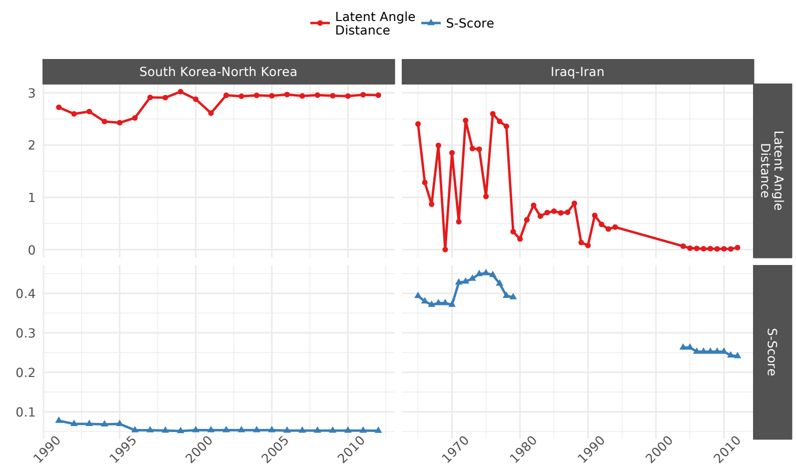

Figure 5: Dyadic relationships between 2 pairs of countries according to the different preference measures. Yearly latent angle distances between dyads are denoted by red circles and ideal point distance based on UN voting in blue triangles. S-score values have been rescaled such that they works in the same direction as the other measures (1-S-score, so that the best score is 0), and are depicted by green squares.

determined by UN voting (Bailey and Voeten, 2015), a model using both UN ideal points and alliance S-scores, and finally, a model using our latent angle approach to combine data from UN voting and alliances. We evaluate the models on two criteria: whether state preferences have a consistent effect in the predicted direction, and how well each model does at predicting disputes on out of sample data.

In each of these models, we utilize a logistic regression of Militarized Interstate Dispute (MID) participation on measures of state preferences and a vector of control variables. These control variables overlap with the standard framework used in O'Neal and Russett's canonical work on the democratic peace (Oneal and Russett, 1997).[11] In particular, we include a binary measure of joint democracy (whether both states have Polity IV scores $geq7$), whether the states are contiguous, and the ratio of state capabilities as measured by the Correlates of War Project's Composite Index of National Capabil-

[11]The exception is that our models ignore trade interdependency, as including that data drastically decreases the number of observations.

ities (CINC). We also account for temporal interdependence using a peace year spline (Carter and Signorino, 2010).

In sample explanation

As detailed in figure 6, each measure of state preferences behaves as we would expect. Our measure of state preference using latent angle difference is highly significant and positive: states with more dissimilar preferences will have greater difference in their latent angles, and this is highly associated with a greater risk of conflict. Similarly, both incumbent measures of preferences pass this test. The measure using UN voting ideal points is positive and clearly distinct from 0, indicating that states with more distant ideal points, and thus more dissimilar preferences, are more likely to find themselves in conflict. Similarly, higher alliance S-scores are consistently associated with lower probabilities of conflict – so states with more similar preferences as measured using alliance portfolios are less likely to quarrel. These results hold when those measures of preference are used in isolation, or in tandem.

The models have one major difference in terms of the controls: in the model using latent angle distance, joint democracy is indistinguishable from 0. This is particularly interesting as one of the major criticisms of democratic peace theory is that democracies have similar preferences. Some argue that it is similarity in preferences that leads to peace among democracies. Despite this dispute, most attempts to include preferences in the standard democratic peace regressions still find a consistent pacifying effect of democracy. (as do those models with UN voting and S-scores presented here). With our measure of preferences, however, democracy's effect is negligible. More research is necessary to disentangle whether the criticism is now credible, but work such as this provides a template for us to parse out the effect of democracy versus preferences on conflict.

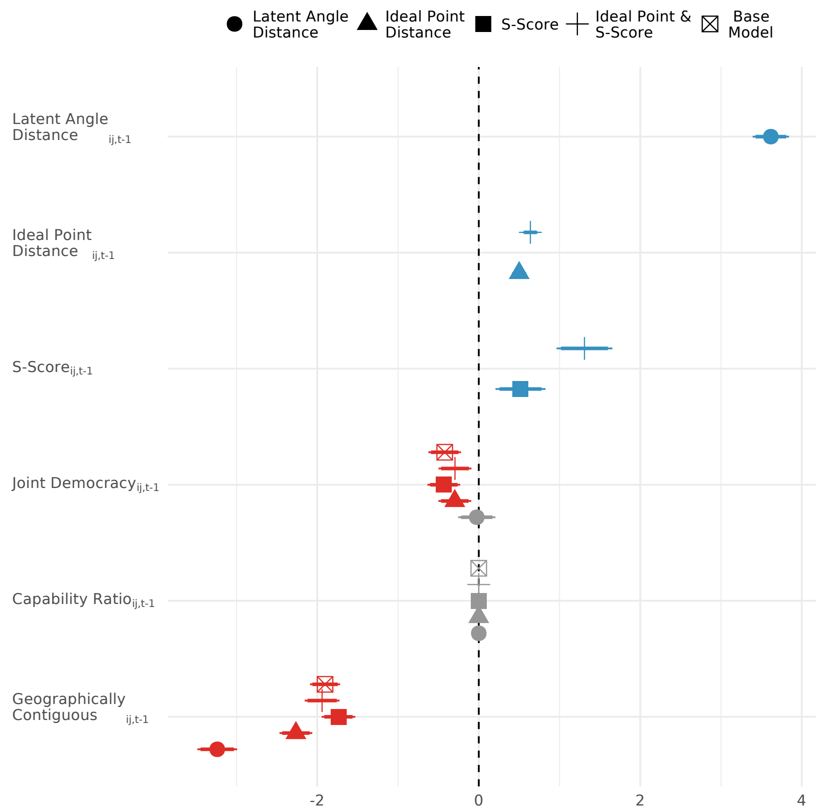

Figure 6: Parameter estimates from models with different measures of state preference. Point represents average estimate, line through the point represents the 95% confidence interval.

Out-of-sample prediction

Given these results, we can say that all of the measures of preferences behave as we would expect. To adjudicate which measures best capture state preferences, and what we should make of the differing effect of democracy, we turn to out of sample prediction. We undertake a cross-validation procedure in which we partition our data into 30 different folds. This process works by taking each fold, k, and running a logistic model excluding data from that fold. Once we have parameter estimates from a model that excluded fold k, we predict the probability of a MID in fold k using only the parameter estimates and the covariates from fold k.

We then compare the performance of these models using the area under both the

Receiver Operator Characteristic Curve (AUC ROC) and under the Precision Recall Curve (AUC PR). The AUC ROC curve examines the tradeoff between true positives and false positives, and the AUC PR examines the tradeoff between making only correct predictions and predicting all the disputes that occurred. In general, the AUC ROC will disproportionately reward those models that predict 0 well – and we can interpret the AUC ROC as the likelihood a prediction is correct. The numeric value for the AUC PR has less of a clear interpretation, but models with a higher AUC PR do a better job of predicting when events actually occur.

As shown in figure 7, the model using latent angle distance decisively outperforms all the other models. While the AUC ROC is somewhat higher with the latent Angle model, the real difference between the measures shines through in the AUC PR, where the model using this measure performs twice as well as the base model. In contrast, models using other measures of state preferences yield only minimal improvements in prediction over the base model. This is especially relevant because AUC PR specifically focuses on the difficult task of predicting conflict, compared to the relatively easier task of predicting non-conflict that is rewarded by the AUC ROC measure. Thus we have reason for confidence in both the usefulness of this measure of state preferences and renewed reason for skepticism in the effect of joint democracy once we control for state preferences.

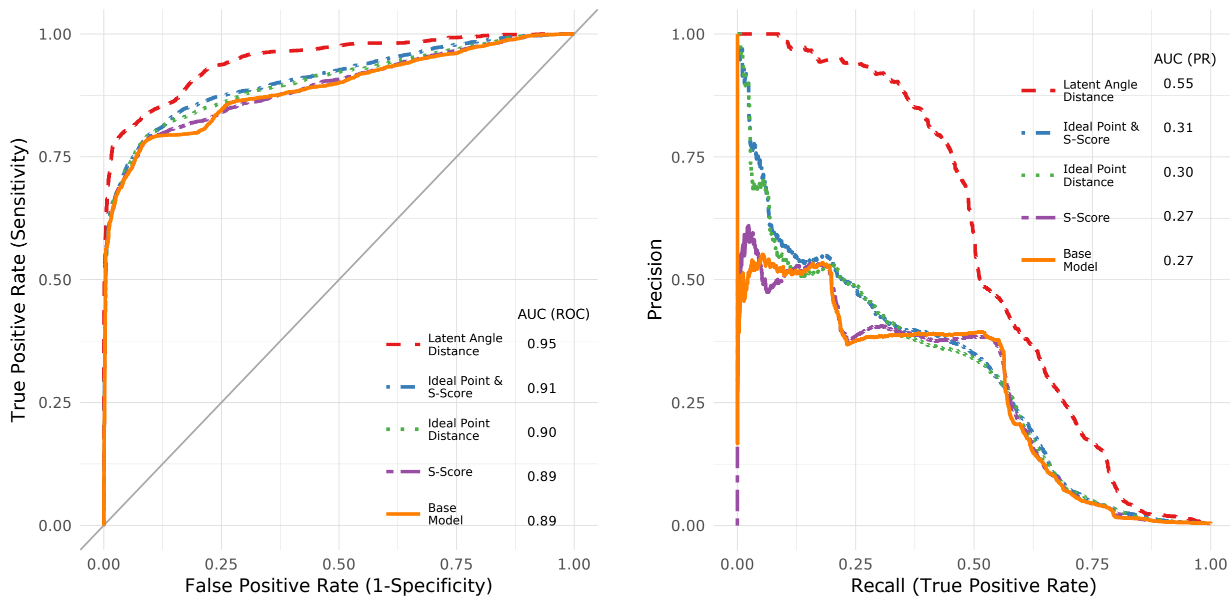

Figure 7: Assessments of out-of-sample predictive performance using ROC curves and PR curves. AUC statistics are provided as well for both curves.

Conclusion

We use a network methodology to create a new measure of state preferences using both UN Voting data and alliance data. This measure of state preferences demonstrates face plausibility comparable or superior to existing measures when it comes to capturing the dynamics of a number of notable dyadic relationships. We then attempt to use this measure of state preferences in a predictive model of interstate disputes – in doing so, we find that the measure of preferences has the expected effect (states with similar preferences are less likely to be involved in disputes) and inclusion of our measure leads to the measure of joint democracy becoming indistinguishable from 0. Most excitingly, a model of interstate conflict that includes our measure of preferences decisively outperforms models that include both of the most prominent existing measures of preferences.

While this measure of state preferences has yielded leverage in predicting conflict, it should also be useful in answering a number of other questions. One possible use would be to look at the measure of state preferences as an outcome variable, rather than a predictor. We could here look at how preferences change in tandem with leadership change – can we find evidence, for example, that the election of Donald Trump moved the United States' foreign policy preferences away from the major Western European states and towards Russia's? We could also see how well this measure of state preferences predicts more collaborative behavior – treaty membership or trade for example. Finally, this particular method allows us to control for confounding fixed effects. A closer examination of the relationship between democracy, preferences, and conflict could create a measure of the component of state preferences not explained by democracy.[12] Similarly, the underlying methodology could be usefully applied to other questions. It could be used to estimate the cliques and alignments of elites in non-democratic states, or as a way of combining a number of related variables on development or democracy, in order to get an underlying latent measure.

[12]We look at the entire measure of state preferences because our primary goal here is not to examine this relationship, but to create the most accurate measure of preferences possible.

Anderson, C. J., Wasserman, S., Faust, K., 1992. Building stochastic blockmodels. Social Networks 14 (1), 137–161.

Bailey, Michael A., A. S., Voeten, E., 2015. Estimating dynamic state preferences from United Nations voting data. Journal of Conflict Resolution 61 (2), 430–456.

Bennett, D. S., Rupert, M. C., June 2003. Comparing measures of political similarity: An empirical comparison of S versus τ_b in the study of international conflict. Journal of Conflict Resolution 47 (3), 367–393.

Braumoeller, B. F., 2008. Systemic politics and the origins of great power conflict. American Political Science Review 102 (1), 77–93.

Bueno de Mesquita, B., 1983. The War Trap. Yale University Press, New Haven, CT.

Bueno De Mesquita, B., Lalman, D., 2008. War and reason: Domestic and international imperatives. Yale University Press.

Carter, D. B., Signorino, C. S., 2010. Back to the future: Modeling time dependence in binary data. Political Analysis 18 (3), 271–292.

DeRouen, K., Heo, U., 2004. Reward, punishment or inducement? US economic and military aid, 1946–1996. Defence and Peace Economics 15 (5), 453–470.

Farber, H., Gowa, J., 1995. Politics and peace. International Security 20 (2), 123–146.

Gallop, M., 2017. More dangerous than dyads: How a third party enables rationalist explanations for war. Journal of Peace Research.

Gartzke, E., 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, E., 2000. Preferences and the democratic peace. International Studies Quarterly 44, 191–212.

Gartzke, E., 2007. The capitalist peace. American Journal of Political Science 51 (1), 166–191.

Gibler, D. M., Sarkees, M., March 2004. Measuring alliances: The correlates of war formal interstate alliance data set, 1816-2000. Journal of Peace Research 41 (2), 211–222.

Goldenberg, A., Zheng, A. X., Fienberg, S. E., Airoldi, E. M., 2010. A survey of statistical network models. Foundations and Trends in Machine Learning 2 (2), 129–233.

Häge, F. M., 2011. Choice or circumstance? Adjusting measures of foreign policy similarity for chance agreement. Political Analysis, 287–305.

Hoff, P. D., 2007. Extending the Rank Likelihood for Semiparametric Copula Estimation. Annals of Applied Statistics 1 (1), 265–283.

Hoff, P. D., Fosdick, B., Volfovsky, A., He, Y., Minhas, S., July 2017. Additive and multiplicative effects models for networks and relational data, version 1.4. Comprehensive R-Archive Network.

Kastner, S. L., 2007. When do conflicting political relations affect international trade? Journal of Conflict Resolution 51 (4), 664–688.

Kydd, A., 2003. Which side are you on? Bias, Credibility, and Mediation. American Journal of Political Science 47 (4), 597–611.

Minhas, S., Hoff, P. D., Ward, M. D., October 2016. Let's say amen for latent factor models, working paper.
URL `https://arxiv.org/abs/1611.00460`

Oneal, J., Russett, B. M., 1999. Is the liberal peace just an artifact of cold war interests? Assessing recent critiques. International Interactions 25 (3), 213–241.

Oneal, J. R., Russett, B. M., 1997. The classical liberals were right: Democracy, interdependence, and conflict, 1950-1985. International Studies Quarterly 41 (2), 267–293.

Signorino, C., Ritter, J., 1999. Tau-b or not tau-b. International Studies Quarterly 43 (1), 115–144.

Singer, J. D., Small, M., 1995. National Military Capabilities Data. Correlates of War Project, Ann Arbor, MI.

Snijders, T. A., Nowicki, K., 1997. Estimation and prediction for stochastic blockmodels for graphs with latent block structure. Journal of Classification 14 (1), 75–100.

Stone, R. W., 2004. The political economy of IMF lending in Africa. American Political Science Review 98 (4), 577–591.

Voeten, E., 2013. Data and analyses of voting in the United Nations. Routledge, Ch. 4.

Wasserman, S., Faust, K., 1994. Social Network Analysis: Methods and Applications. Cambridge University Press, Cambridge.

Wolford, S., 2014. Showing restraint, signaling resolve: Coalitions, cooperation, and crisis bargaining. American Journal of Political Science 58 (1), 144–156.

Xavier Amatriain, Alejandro Jaimes, N. O., Pujo, J. M., 2015. Data mining methods for recommender systems. In: Recommender Systems Handbook. Springer, pp. 227–262.