

# A Latent Factor Approach to Measuring State Preferences

---

## Abstract

State preferences play an important, yet under discussed roll 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, 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 automatic 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.

Yet, this relative 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.<sup>1</sup> 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.

To provide this more accurate measure we take a multilayer network based approaches to estimating state preferences. Much of the extant literature has employed spatial weighting models (Signorino and Ritter, 1999; Bailey and Voeten, 2015) on relational data such as alliance behavior and UN voting scores. We make

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 they ignore the fact that the interactions occurring between states within the alliance or UN voting system are examples of networks. 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. First, we introduce a latent factor model that accounts for higher-order dependence patterns across multiple networks. Second, we show that our revised approach of measuring state preferences enhances our ability to predict instances of conflict between countries when compared to existing measures.

---

<sup>1</sup>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.

**Sources of Preference Measures: Alliance Portfolios and UN Voting**

Given that we cannot directly observe state preferences, scholars have attempted to ascertain it 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 also 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  $\tau_B$  Bueno De Mesquita and Lalman (2008). This measure is:

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

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  $\tau_B$  statistics. Thus, Signorino and Ritter introduce the S score, which has since been the most widely used alliance similarity measure.<sup>2</sup>

The equation for the S score is:

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

---

<sup>2</sup>(Bennett and Rupert, 2003) also find a stronger relationship between theoretical predictions and results when using S-scores than when using  $\tau_B$

Where:

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

Where  $w_k$  is the  $k$ 'th element of a weight matrix,  $d^{\max}(W, L)$  is the maximal distance on a given dimension, and  $\Delta$  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).

One important distinction for these scores is that they are purely dyadic. One can look at the S-score between two states, but one cannot look at a state's preferences in comparison to a larger cluster, or note the movement in a states preferences over time. In monadic analysis, these score measures are not even available, and once we are dealing with situations involving more than two states, the number of S-scores necessary to fully characterize the preferences balloons quickly (it is the number of actors choose two). 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 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  $Z_{itv}$

$$Z_{itv} = \beta_{iv}\theta_{it} + \epsilon_{iv} \quad (4)$$

such that the state will vote yes if  $Z_{itv} < \gamma_{1v}$ , no if  $Z_{itv} > \gamma_{2v}$  and otherwise abstain. Here,  $\theta_{it}$  is state  $i$ 's ideal point at time  $t$ , and  $\beta_{iv}$  is the discrimination parameter of a particular bill  $v$ . When  $\beta_v$  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  $\gamma_{1v}$  and  $\gamma_{2v}$  such that the same bill will have the same value in different years, and they standardize and normalize  $\theta$ . They also use  $\theta_{it-1}$  as a prior on  $\theta_{it}$ . With these constraints, they solve for the ideal points using a Metropolis Hastings Markov Chain Monte Carlo (MCMC).

The issues here are that the voting behavior, especially in the European Union (EU) general assembly, is not well behaved in the way that voting in the US Congress is. We actually can get some sense of it by the existence of multiple identical resolutions: UN resolutions have no legal force, and so most votes are symbolic. Thus UN voting is rife with nearly unanimous voting and other super-majorities, which means that the requirements to distinguish between state preferences are more onerous. Another issue here, not necessarily with use of voting in general, but with this application, is the limitation to one dimension: it could be that two states which have very similar preferences on issues of trade – say the United States and Saudi Arabia – might differ mightily on questions related to the Middle East, and in particular Israel.<sup>3</sup> However, the dearth of contentious UN votes makes it difficult to add additional dimensions.

---

<sup>3</sup>You can see issues like this on legislative positions in the United States: during the post WW2 era, two democrats who might agree on the need to expand the social safety net might be diametrically opposed on issues of civil rights.

## Synthesizing Measures of State Preference

The approach that we introduce here to measuring 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.

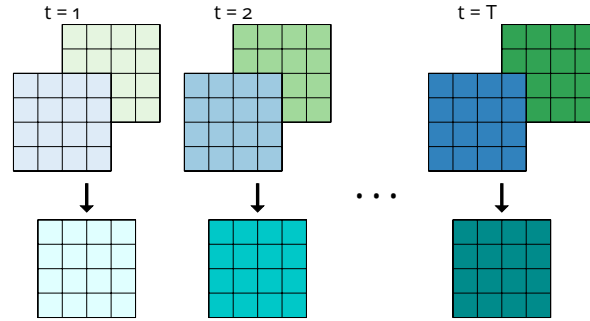
Relational data is composed of observations between pairs of actors, or dyads. For both alliance relationships and UN voting, we are able to observe how the actors in the international system interacted with one another across time. This system of interactions taken in its totality defines a network, 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 1. In the top row, we represent UN voting and alliance patterns at time  $t$  in a pair of adjacency matrices that form a multiplex.<sup>4</sup> 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

---

<sup>4</sup>The approach that we describe here can be generalized to a multiplex with more than two dimensions.



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 strength of relations between two countries.



**Figure 1:** Tensor representation of longitudinal dyadic, representational measures. The green and blue colors represent different relational measures and darker shading indicates later time periods. Specifically, we show a tensor with dimensions of  $4 \times 4 \times 2 \times 3$ , where 4 represents the number of actors, 2 the number of relational measures, and 3 the number of time points.

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.

### *Latent Factor Model*

A number of techniques have been developed in recent years to estimate low-dimensional representations of network structures.<sup>5</sup> 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 com-

<sup>5</sup>For a literature review of these approaches see Goldenberg et al. (2010).

munities. 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  $i, j$  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 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. Ignoring these patterns when generating a measure of state preferences 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  $T \times n_t \times n_t \times p$  arrays ( $Y_t$ ), where  $T$  represents the number of periods,  $n_t$  represents the number of actors in period  $t$ , 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  $t$ .<sup>6</sup> Obtaining a lower-dimension relational measure of state preferences that captures network dependence patterns, we use the LFM to decompose each array:

$$Y_t = f(\theta_t)$$

$$\theta_t = \beta^\top X_t + Z_t$$

$$Z_t = M + E$$

$$M = U\Lambda U^\top, \text{ where}$$

$$u_i \in \mathbb{R}^k \text{ and}$$

$$\Lambda \text{ is a } k \times k \text{ diagonal matrix}$$

where  $f(\cdot)$  is a general link function corresponding to the distribution of  $Y_t$  and  $\beta^\top \mathbf{X}_t$  is the standard regression term for dyadic and nodal fixed effects, and  $Z_t$  represents any additional patterns in data unrelated to the specified 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 factored out the effect of distance, for example, than this could be accomplished within the context of this framework.

The key part of this model lies in the decomposition of  $Z_t$ . Hoff (2009) notes that we can write  $Z_t = M + E$  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  $u_i \in \mathbb{R}^k$  is a latent vector associated to node  $i$  and  $\Lambda$  is a  $k \times k$

---

<sup>6</sup>The approach we describe below has already been generalized to the case where  $y_{ij} \neq y_{ji}$ .

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  $\Lambda$ ,  $\lambda_k > 0$  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 both positive or negative homophily in varying degrees and stochastically 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 Markov Chain Monte Carlo (MCMC) procedure that enables us to construct approximate samples from the posterior distributions of the latent variables.

A Bayesian procedure to determine the eigenvalue decomposition is available in the **amen** R package (?).

### Data Sources, Modeling choices

We use the AME model on the two aforementioned measures of state amity to generate a combined measure of state preference similarity which accounts for network effects. We use the distance between states’ ideal points (as calculated by Bailey and Voeten (2015) using UN data) and S-score for two states alliance portfolios. However, to facilitate comparison between the metrics, we first transform the S-score into a measure of distance between alliance portfolios.<sup>7</sup> We then standardize and normalize these two measures. This gives us an N by N by Y by 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 XXX and YYY at the first year of our data (YYYY), simi-

---

<sup>7</sup>D = 1 - S

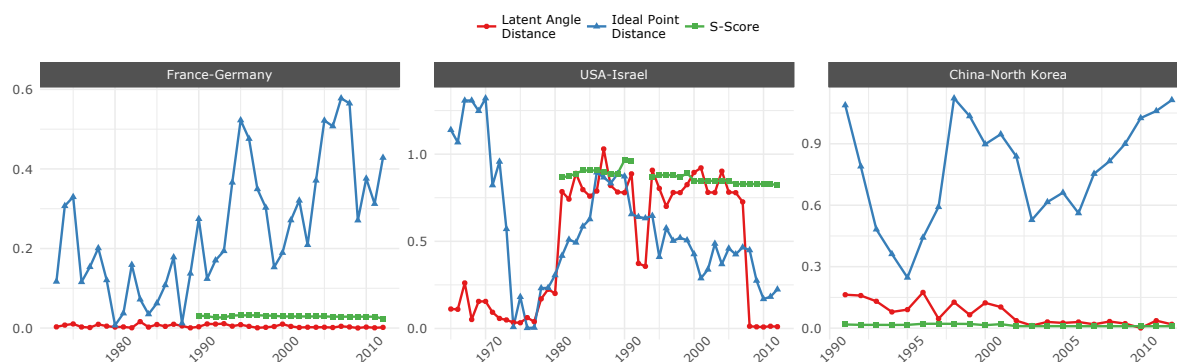
larly (1,2,1,2) would be the UN ideal point distance.

Another important question is the amount of temporal aggregation used. In our baseline model, we treat each year as separate and gain a unique observation of each states ideal point in each year. However, this raises a real risk of temporal inconsistency in the values. An alternative approach would be to have a rolling average for the measures of similarity over a number of years. This would allow us to infer a country's relative position not just by their behavior in a given year, but also their behavior in the past few years. The risk if we use too much temporal aggregation is that we are including data which is no longer relevant to a country's relative preferences. For instance, Turkey and Russia's relationship looks a lot more positive when we look at 2013 and 2014 then when we look at 2015. To that end, in addition to our baseline model where years are seen as independent, we also evaluate models where ...

With this data, we run an AME model with a Gaussian link, and in particular we use the uDv term to estimate each states position in a two-dimensional latent space. We then evaluate whether there is additional utility gained from using this latent position, as compared to the component measures of similarity of alliance portfolio and UN ideal point distance.

### *Face Plausibility*

An important question for these different measures of preferences, are whether they give results that "make sense" for certain prominent pairs of countries. In particular we would hope that the measures both give sensible levels to relationships – sorting states into friends and foes effectively – but also that when these measures change, they do so in sensible ways, and in ways that correspond to changes in the world. We thus present all three measures' accounts of eight dyadic relationships over



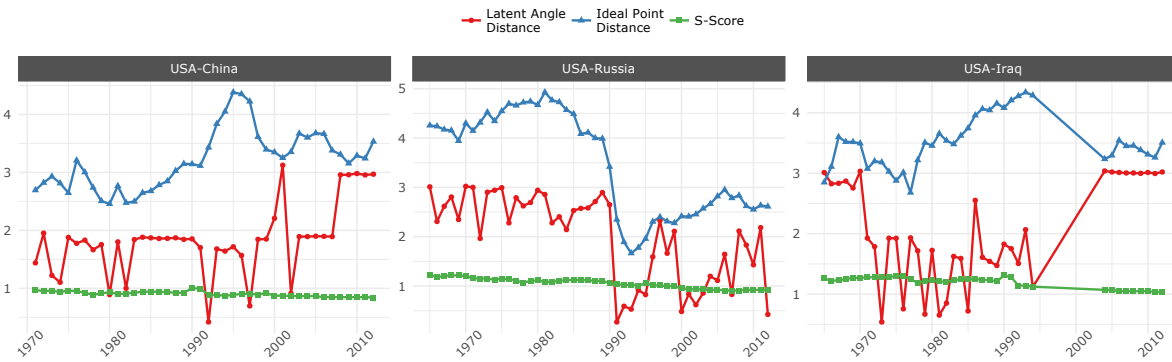
**Figure 2:** Dyadic relationships between 3 pairs of countries according to the different preference measures. Yearly latent angle distances between dyads are denoted by red circles and ideal point distance based on UN voting in blue triangles. S-score values have been rescaled such that they work in the same direction as the other measures (1-S-score, so that the best score is 0), and are depicted by green squares.

time.<sup>8</sup> A note is 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'd expect to see similar preferences: as depicted in figure 2 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 German, and China and North Korea, and low but less stable values for the US/Israel relationship. Alliance S-Scores correctly classify 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 3 we look at the United States' relationship with three countries that are

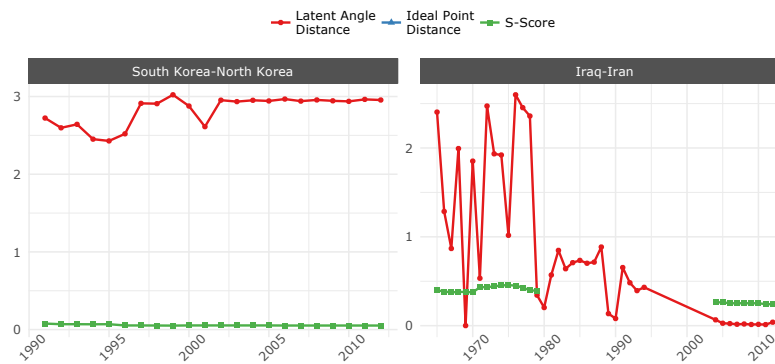
<sup>8</sup>For the sake of having each measure operate in the same direction, we transform the S-score such that the value shown is 1 - S. Thus for all three variables, 0 implies perfect symmetry of preferences.



**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 work in the same direction as the other measures (1-S-score, so that the best score is 0), and are depicted by green squares.

characterized by change and major events. The biggest difference between the measures is that alliance 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 preferences following the US occupation, though only Latent Angle distance finds an increase in enmity in the run-up to the US's invasion. Finally, for the US and China, the S-score has a consistent neutral relationship, while both UN ideal points and Latent Angles have more both more negative relationships, and more variance, with Latent Angles finding particularly large spikes at the beginning of the Bush and Obama administrations.

One big difference between these measures is that Latent Angle distance has a more full time series than the other two measures, this is particularly relevant when looking at the preferences of certain rogue states, as we do in figure 4. In both the case of Iran/Iraq and North Korea/South Korea there is no data for UN voting (and for



**Figure 4:** 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 work in the same direction as the other measures (1-S-score, so that the best score is 0), and are depicted by green squares.

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 Latent Angles show more instability, and notable 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 than the other measures at conforming to our expectations of the relationships. Of course, this could just be a case of us picking particularly propitious cases. These 8 dyads show the plausibility of our measure of state preferences, but the real test is in the large-N analysis of conflict in the next section.



*Model Competition*

To evaluate different measures of state preferences, we compare them in a model 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 (as in ??, one using the ideal points determined by UN voting (as in ??), 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.

*Data, Controls*

In each of these models, we look at a logistic regression of Militarized Interstate Dispute participation on measures of state preferences and a vector of control variables. These control variables are most of the standard ones used most famously in O’Neal and Russett’s work on the democratic peace (Oneal and Russett, 1997).<sup>9</sup> 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 Capabilities (CINC). We also account for temporal interdependence using a peace year spline, as in ??.

*In sample explanation*

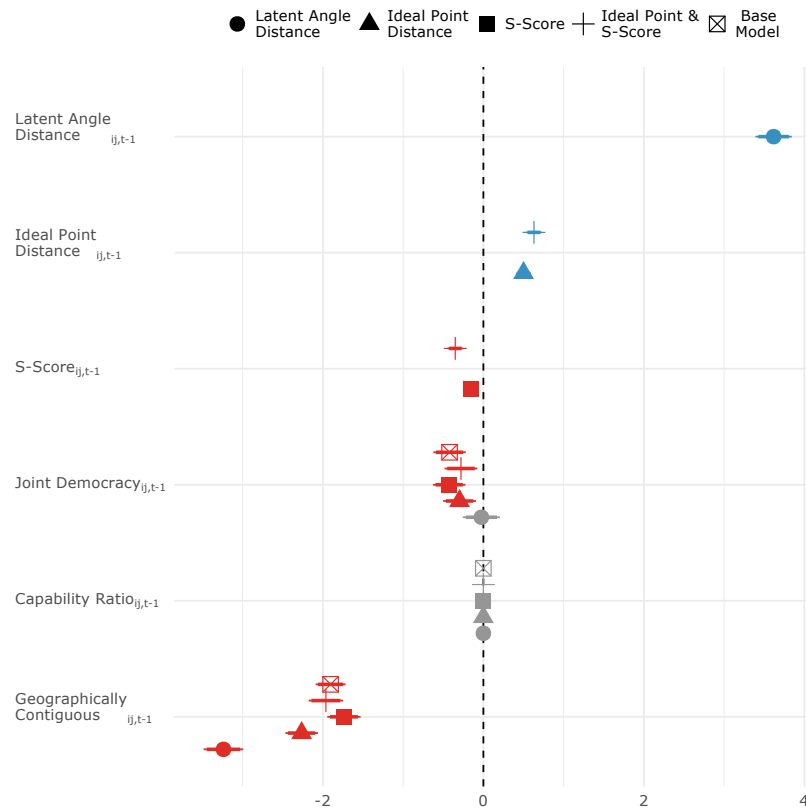
As detailed in figure 5, each measure of state preferences performs as we would expect them to. Our measure of state preference using latent angle difference is highly

---

<sup>9</sup>The exception is that our models ignore trade interdependency, as including that data drastically decreases the number of observations.

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 the measure of preferences is 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 because of one of the major criticisms of democratic peace theory is that, for one reason or another, democracies have similar preferences, and this is what actually causes peace among democracies. Despite this dispute, most attempts to include preferences in the standard democratic peace regressions still found 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.



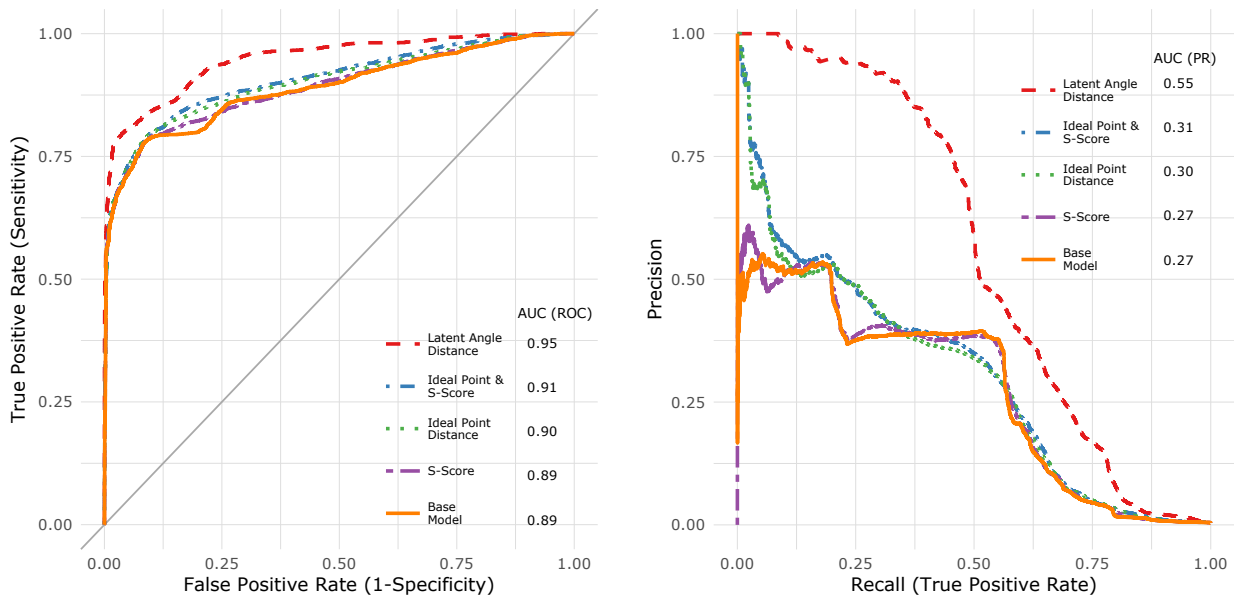
**Figure 5:** Parameter estimates from models with different measures of state preference. Point represents average estimate, line through the point represents the 95% confidence interval.

### *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. The way we do this is by partitioning our data into 30 different folds, and for each fold we generate a predicted value for each case the other 29 folds for each model. This leaves us with a set of predicted values generated entirely out of sample. We then compare the models performance using the area under both the Receiver Operator Characteristic Curve (AUC ROC) which examines the tradeoffs between true positives and false positives, and the area under the Precision Recall Curve (AUC PR)

which looks at the tradeoffs between making only correct predictions, and predicting all the disputes which 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 6, 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. Thus we have real 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 6:** 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. 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.

- 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  $\tau_b$  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.
- 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.
- 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., 2009. Multiplicative latent factor models for description and prediction of social networks. *Computational and Mathematical Organization Theory* 15 (4), 261–272.
- Hoff, P. D., 2007. Extending the Rank Likelihood for Semiparametric Copula Estimation. *Annals of Applied Statistics* 1 (1), 265–283.
- 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.
- 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.