Measuring International Dissatisfaction

Daniel Kent*

March 13, 2020

Abstract

A state's international dissatisfaction is central to theories of conflict and cooperation. However, due to dissatisfaction's unobservability, relevant empirical treatments either avoid measuring the concept altogether or rely on conceptually constrained proxies. In response to these methodological challenges, I propose a novel measure of international dissatisfaction spanning from 1816 to 2012 which explicitly operationalizes Gilpin's framework: the difference between a state's expected and actual benefits from the international status quo. I estimate a state's expected international benefits by building upon recent efforts to train machine learning ensembles on war outcomes, which I then use to weight a state's observable material capabilities. I estimate actual international benefits by averaging across a state's centrality in valued international networks. The measure provides multiple advantages over alternative estimates both conceptually and statistically. Indeed, beyond its conceptual value, the measure's association with militarized conflict is robust to model specifications, unlike the current go-to measure when modeling country-level sentiments: ideal point estimates from United Nations voting records.

*kent.249@osu.edu; Ph.D. Candidate, Department of Political Science, The Ohio State University

1 Introduction

A state's level of satisfaction or dissatisfaction with the international status-quo plays a pivotal role in theories of international conflict and cooperation. However, and unfortunately for scholars of international politics, a state's general sentiment toward the international system is directly unobservable, meaning empirical inferences must either be drawn from often-messy observable quantities or avoided altogether. Indeed, insofar as one's focus is a state's dissatisfaction with the international status-quo, the closest available proxy can be found in ideal point estimates derived from United Nations voting records.[1] The popularity of these estimates is unsurprising, given the range of issues that are brought before the United Nations and the related methodological tradition in American Politics around congressional voting records. Moreover, in a dyadic context the estimates are easily interpretable as the ideological distance between two states in a given year.

That being said, while ideal point estimates provide a concise comparison of the ideological difference between pairs of states, they are less clearly applicable if one's focus is on any one state's attitude toward the makeup of the international system as a whole. In response to this methodological challenge, I develop a novel measure of each state's international dissatisfaction from 1816-2012. Drawing on the conceptual framework originally put forward by Gilpin (1983) and more recently applied by Renshon (2016, 2017) to international status, I treat international dissatisfaction as the distance between a state's international *expectations* and the *reality* it faces. While the exercise is an unsupervised one – we lack consistently matchable outcomes that can statistically validate the final measure – I am able to draw upon popular techniques in network science and machine learning when measuring each relevant component.

Beyond the measure's conceptual distinction, it also provides multiple quantitatively appealing features. First, the United Nations-based ideal point estimates are temporally bounded, spanning only from 1946 to 2012. The dissatisfaction estimates instead span the entirety of the Correlates of War data set, allowing for a greater range of analytical applications. Second, the international dissatisfaction measure is a considerably more robust predictor of whether or not a state initiates international conflict. Fringe ideal points do predict conflict onset, but the magnitude and direction of coefficient estimates is sensitive to model specification with ideal points – which is not the case with international dissatisfaction.

[1] The most prominent dataset is found in Bailey et al. (2017). Another set of ideal points are included in Braumoeller (2013), which look at the great powers up to the end of the Cold War and are based on responses to historian surveys.

Lastly, I find that the relationship between dissatisfaction and conflict initiation is consistent in both statistical significance and effect direction across all years, demonstrating that the relationship is not just driven by a handful of choice outlier cases, as is often a concern around research on international conflict.

Previewing the measure's composition, the final values are produced by estimating and then differencing two component quantities: a state's expected and actual benefits from the international status-quo. Starting with a state's actual benefits, I employ techniques for social network analysis and consider networks that are both social and material. Conceptually drawing on the relation school in Sociology (e.g., Burt et al., 2005; Emirbayer and Goodwin, 1994; Emirbayer, 1997; Erikson, 2013) and International Relations Theory (e.g., Jackson and Nexon, 1999; MacDonald, 2018; Qin, 2016), I estimate international benefits by calculating each state's position in networks of various valued international goods. These include the: military alliance, interstate trade, shared diplomatic tie, and arms trade networks. Each state's relative centrality in each network is calculated and averaged across all networks in a given year, providing an estimate of access to the distribution of valued international goods – its actual international benefits.

Turning to a state's expected benefits, in the Gilpinian framework states become dissatisfied because they believe their share of international goods falls short of what they should be receiving based solely on their power-position. In Gilpin's formulation, expectations are composed of two parts: the possession of and reputation for using material capabilities.[2] The former is relatively straightforward empirically. A state's relative material capabilities can be proxied for by its CINC score, which measures its percent of the globe's total: population, military capacity, resource consumption, and iron and steel production. Measuring the latter – reputation for using these capabilities – is more difficult. To proxy for a state's material reputation, I replicate and extend Carroll and Kenkel's (2016) recent work, which uses CINC components to predict militarized interstate dispute (MID) outcomes. After building a machine learning algorithm that accurately predicts which side wins in a MID (ignoring stalemates), I use the model to make predictions about outcomes if all pairs of states in a given year were to engage in a MID. These predicted probabilities are then aggregated and provide general estimates of a state's reputation for using its capabilities.

The final estimate is the difference of these two components – a state's international expectations

[2]Gilpin labels this reputation-based component of power 'prestige'.

and reality. States whose expectations outstrip actual benefits are dissatisfied and, inversely, those whose expectations rest equal to or below their actual benefits are satisfied. The rest of the paper walks through the process of building and validating the measure in the following steps. First, I review the relevant literature. Second, I elaborate upon the my formalization of Gilpin's theory of international dissatisfaction and change. Third, I produce the dissatisfaction measure. Fourth, I test the measure's validity as a predictor of conflict onset and compare its predictive capacity to the aforementioned ideal point estimates.

2 Literature Review

International dissatisfaction is fundamental to understanding both why change does or doesn't occur and, if change does occur, why it sometimes is peacefully managed versus marred by great power conflict. Whether it be understanding why outcomes vary greatly around rising powers (Allison, 2017; Edelstein, 2017; Fearon, 1995; Gilpin, 1983; Goddard, 2009, 2018a; Goh, 2005, 2013; Kennedy, 2010; MacDonald and Parent, 2018; Organski and Kugler, 1981; Schake, 2017; Shifrinson, 2018; Trachtenberg, 2012), how states seek desired levels of status (Chan, 2004; Duque, 2018; Larson and Shevchenko, 2010; Paul et al., 2014; Renshon, 2016, 2017; Ward, 2017), or debates around the fundamental origins of revisionist states, (Davidson, 2006; Goddard, 2018b; Johnston, 2003; Lyall, 2005; Schweller, 1994, 1999, 2015), actual or potential dissatisfaction with the status-quo is conceptually fundamental. After all, why would an actor change any situation if they are reasonably satisfied with it? Or in the international context, why would a state attempt to alter an environment that appears to be reasonably beneficial? If a state seeks change, then there likely is some aspect of the status-quo they are displeased by, meaning its dissatisfaction is a critical quantity and merits investigation.

However, despite international dissatisfaction's conceptual importance, it rarely, if ever, receives thorough statistical treatment. This is a particularly strange omission when considered alongside the rich tradition of quantitative analysis in International Relations scholarship. Most likely, the relative lack of such work stems from the simple fact that dissatisfaction is a latent, unmeasurable variable. And, as noted by Lyall (2005), when considering the relevant sample of cases, there is a tendency to only look ex-post at the handful of clear examples of intensely dissatisfied revisionists (e.g., Pre-WWI and WWII Germany, Imperial Japan, Communist Russia, and Revolutionary Iran), which then risks inducing

inferential biases through sampling on the dependent variable or ignoring the control group. In tandem, these methodological challenges suggest that for studies of international dissatisfaction, revision, and change, not only are dissatisfaction or satisfaction unobservable, but gathering an analytically useful sample can be particularly difficult.

Notably, estimates for a state's ideal point are one potential indicator that has grown in recent popularity and can be applied to all cases where yearly data is available. These estimates are most prominently used when based upon voting pattern at the United Nations, where it is assumed that voting records generally represent a state's true preferences (Bailey et al., 2017).[3] Although these ideal points are a valuable contribution to many areas of international politics, they face important methodological and conceptual limitations if one's goal is to study a state's dissatisfaction with the international status-quo as a whole. Methodologically, the data start in 1946, which is a relatively limited time-span historically and does not include some of the most important cases of both revisionist and status-quo states. Conceptually, the estimates place states on points along an abstract ideological spectrum. Quantifying what one location on this ideal point spectrum actually means on its own is a difficult, if not impossible, task. Rather, the estimates primarily provide value if one is interested in a state's general views, *relative to others* based upon the distance between their positions.

While the ideal point estimates allow valuable inferences about how ideologically different pairs or groups of states are from each other, they do not necessarily speak to any single state's attitude toward the international system as a whole. Comparing values certainly sheds light on the similarity or dissimilarity of visions for a best-case international scenario. But comparing values does not necessarily tell us anything about a single state's actual satisfaction or dissatisfaction with their international environment.[4] Put differently, a state's leadership and population may desire some abstract form of the world, but that does not mean anyone actually believes such a world is possible. Only the most powerful states even consider such a world to be in the realm of possibility.[5] Yet, despite the implausibility of actually achieving a best-case scenario, states have consistently taken the side of an international status-quo which they presumably were satisfied by enough to fight for. In this sense, an accurate understanding of any state's attitude toward the status-quo is rooted in something more conservative than ideal visions of some best-case scenario.

[3] See Braumoeller (2013) for ideal point estimates for the great powers before 1946 based on historian surveys.
[4] I later find that there is almost zero correlation between a state's ideal point and ,my dissatisfaction estimates.
[5] Though even for the greatest of empires, moments of achieving one's ideal are rarely lasting.(Kennedy, 2010)

One alternative understanding of international preferences (and this paper's focus) can be found in Gilpin's (1983) approach, where dissatisfaction is treated as the difference between a state's international expectations and reality. Gilpin's theory of war and change argues that when states grow more powerful, their desires similarly expand. Should the international environment fail to recalibrate and accommodate these newfound desires, then states tend to turn to war as their preferred mechanism of producing the desired changes. At the core of this conception is the assumption that a state's expected international benefits – labelled as the sum of a its "territorial, political, and economic arrangements."[6] – are a function of relative power.[7] Gilpin breaks power into two categories, the first being a state's material – primarily military and economic – capabilities. The second, which he labels 'prestige', is a state's reputation for the capacity and willingness to use their material capabilities. Change is sought when a state's power substantially outgrows its share of the aforementioned international goods. Or, as Gilpin argues: "Thus, a precondition for political change lies in a disjuncture between the existing social system and the redistribution of power toward those actors who would benefit most from a change in the system."[8] Generally understood as a occurring through rising powers (though in the next chapter I empirically investigate whether this phenomena applies to all states), Gilpin's precondition for dissatisfaction-driven change is a gap between power and benefits.

The most careful quantitative application of this framework can be found in Renshon (2016, 2017), who applies the expectations-reality approach to a state's desire for international status. When a state's power outstrips its actual status, measured by shared diplomatic ties, then Renshon argues it suffers from a 'status deficit'. Estimates for status deficits are then demonstrated to predict the onset of interstate conflict. Renshon models a state's expected status as its CINC score (Singer et al., 1972) and actual status as its PageRank centrality in the diplomatic exchange network. This methodological approach is valuable, as it is the first to explicitly take Gilpin's popular theory and demonstrate a way to operationalize the central parts. But it is limited to a single sphere of international benefits (status) and avoids the reputational portion of power. In the next section I take Renshon's general approach and expand upon his methodology to improve in these areas. Put differently, Renshon provides a careful

[6] p11

[7] Gilpin's explicit statement about what constitutes the international status-quo is in itself a valuable contribution. As Johnston (2001) points out, most studies of international dissatisfaction and revision are vague about what revisionists actually seek to change and what it means for them to be dissatisfied. But Gilpin actually lays his cards on the table about what international quantities dissatisfied revisionists are concerned about.

[8] p9

and useful starting point for measuring international dissatisfaction, but only provides a first piece of the broader puzzle.

Lastly, we should note that this exercise is more than just semantics and inside-baseball. A Gilpinian understanding of international dissatisfaction is at the core of current analyses of great power politics. Friedman Lissner and Rapp-Hooper (2018) represent this view well, pointing out that, to the extent they are antagonistic, Russian and Chinese challenges to the existing international order are rooted in such pressures: "As American hegemony has eroded, so too has the willingness of the United States' near-peer competitors to tolerate a liberal international order which reflects a distribution of benefits that decreasingly resembles the global distribution of economic and military power."[9] In current great power politics, justifications for spoiling and revising the status-quo are often understood or assumed to to be based upon an expectations-reality calculation. But, at the end of the day, these claims are made based on theoretical expectations. Whether or not they are accurate descriptions of current and past moments, however, is an empirical question and the focus of this paper's remainder.

3 Model and Theory

Gilpin theorized that the international system is in equilibrium for a member state if the expected costs of international change outweigh the expected benefits.[10] In other words, a state is satisfied with its international environment if they face no incentives to seek significant alternatives; any potential benefits of change, whether they be territorial, institutional, social, or another, fall short of the expected costs. This does not mean the status quo is *ideal* for any country, rather no better alternative reasonably exists.[11] Herein, international disequilibrium occurs for any state when its current capabilities, relative to other states, outstrip its relative benefits. Put in less formal terms, a country is unhappy with the current world if their share of the material pie falls short of what they think they should be receiving according to power-based expectations. In a situation where a state's relative size is greater than its relative share of international goods there should be opportunities for a state to remedy the gap in expected benefits by coercive force.[12]

[9] p13

[10] pp10-11

[11] This is the core difference between United Nations-based ideal point measures and this estimate.

[12] Admittedly, this runs into Fearon's (1995) puzzle of why countries would ever need to use force, when war is costly and a preferable war-avoiding negotiated settlement should exist. However, in this context if a disequilibrium persists, then I

The general equation form of this argument can be understood as:

$$\text{Dissatisfaction}_{it} = \text{Capabilities}_{it} - \text{Benefits}_{it} \tag{1}$$

where i is an individual state and t a given year. This formulation brings forward at least one immediate question: What does it mean for a country to receive *international benefits*?[13] A state's well-being is linked to prosperity and security, which of course are at least partly product of domestic resources such as institutional design, leadership, and more. But the international system also provides opportunities for improving a state's strategic position and general prosperity. Countries benefit from trade, diplomatic exchanges, alliances, and more. In other words, the relations that constitute international relations are a source of international benefit. While functioning domestic institutions and markets are necessary for a country to exert itself internationally, so too are robust relations with other states. International politics are not just conducted under the shadow of security concerns; states are also concerned with making the most out of opportunities to prosper through relations with each other. [14]

I operationalize a state's international benefits as the aggregation of its strategic relations with other states. However, rather than simply summing the number of a state's relations across a network in a given year, I weigh relations with more important states more heavily. I do so by measuring each state's PageRank centrality in its international network. (Brin and Page, 1998, 2012) PageRank centrality aggregates the number of ties to a node in a network, but it weighs ties to more central nodes more heavily than ties to nodes that are less central.[15] The underlying principle in its estimation process is that a node is more central the more connections it has, and connections are more important when they are made to other highly connected nodes. Formally, PageRank centrality is expressed as:

$$x_i = \alpha_A \sum_j a_{ij} \frac{x_j}{g_j} + (1 - \alpha_A) \frac{1}{N} \tag{2}$$

assume that for some reason the relevant states have been unable to reach a war-avoiding negotiation. If this were possible, the incentives to avoid costly war should have kicked in and avoided the disequilibrium altogether. That being said, the mechanics of why states might fail to reach these negotiations in specific cases are not the focus for this paper. Why states might fail to accommodate these *systemic* imbalances is an open and interesting question for future work.

[13] The determinants of a country's international capabilities are certainly far from obvious, but are a less abstract concept than international benefits.

[14] This class of argument draws heavily upon the relational-network approach to social systems. For more on this type of theory see Burt et al. (2005); Erikson (2013); Emirbayer and Goodwin (1994); Jackson and Nexon (1999); Hafner-Burton et al. (2009); Wasserman and Faust (1994)

[15] The original idea was that the internet can be viewed as a network where nodes are sites and ties are links to one site that are included on another site. Google's search engine then ranks results based upon a version of PageRank centrality, where a site's ranking on the search algorithm is based upon its PageRank centrality.

where $g_j = \sum_i a_{ji}$ (node j's total number of ties). x_i is the PageRank centrality for node i and x_j is the centrality for node j.[16] a_{ij} is equal to 1 if a tie exists between nodes i and j, but it is equal to 0 otherwise. α_A is a constant damping factor that weighs how much a node's ties matter, as opposed to treating each node as equally central (if $\alpha_A = 1$); in the context of Google's search engine the damping factor approximates the probability that a user will stop clicking links at any time. A visual example of PageRank centrality applied to an example from international politics can be found in Figure 1.

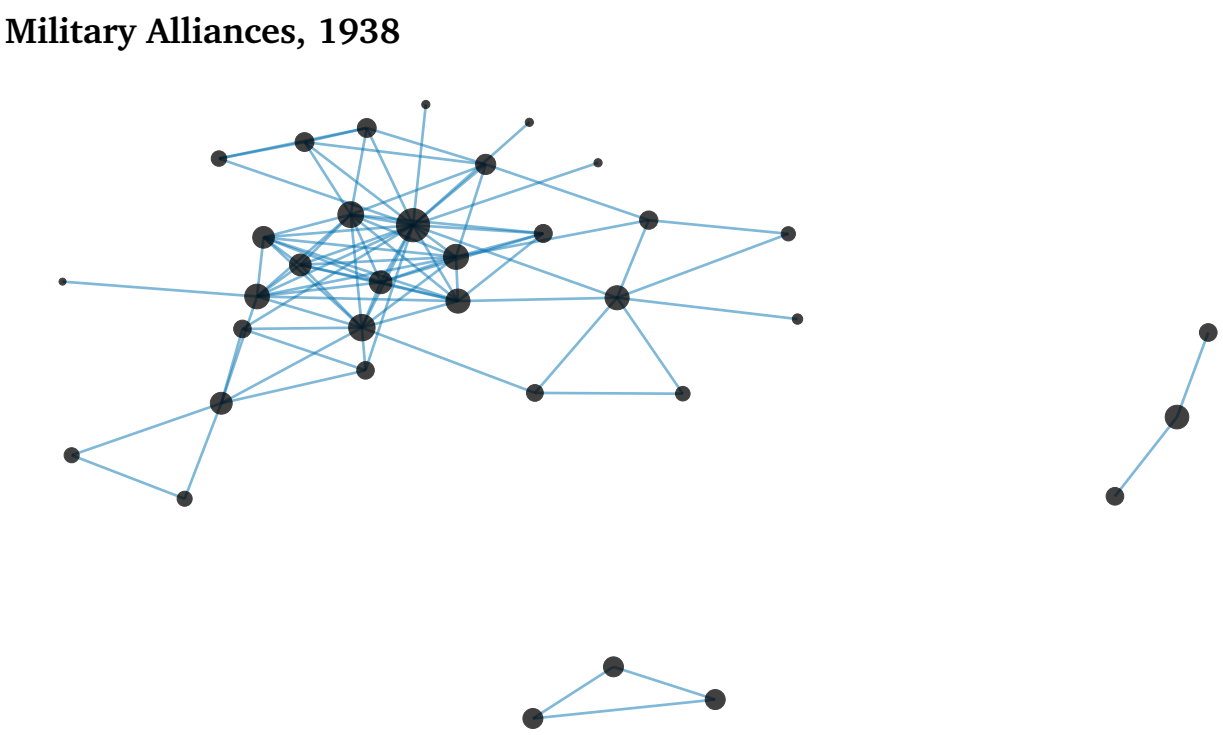

Figure 1: The interstate military alliance network in 1938, before the outbreak of World War II. Node size is a state's PageRank centrality, with larger nodes representing more central nodes. Each tie represents the presence of a military alliance between two states (either a defensive alliance, offensive alliance, or neutrality pact). Country labels are not intentionally left out to not distract from varying node sizes. Isolates (nodes with no ties) are intentionally left out as well.

Once estimated, we can take each state's centrality estimate across networks and estimate total benefits by averaging:

[16] If one carefully reads the formula, then they will see that calculating a node's PageRank centrality requires having already estimated every other node's PageRank centrality. This produces a chicken-egg problem where there is no obvious answer about how to estimate initial values. For a mathematically detailed explanation of how this problem is overcome see: http://www.ams.org/publicoutreach/feature-column/fcarc-pagerank.

$$\text{Benefits}_{it} = \frac{1}{N}\sum_{i=1}^{N}\text{PageRank}_{it} \tag{3}$$

where N is the number of networks measured in a given year, i is an individual state, and t is a single year. The aggregated benefit estimates across all states in a given year end up summing to 1, meaning each state's benefits are a percentage. I subsequently outline the data used in more detail, but the networks included here are: alliances, trade, shared international organizational membership, shared diplomats, arms-trading, and defense cooperation agreements.

Turning to a state's capabilities, Gilpin theorizes that a state's power in any given year is a function of *observed capabilities* and *reputation for resolve and capabilities*.[17] Gilpin labels the latter "prestige", elaborating: "Prestige is the reputation for power, and military power in particular. Whereas power refers to the economic, military, and related capabilities of a state, prestige refers primarily to the perceptions of other states with respect to a state's capacities and its ability and willingness to exercise its power."[18] In other words, a state's power is a function of both its aggregated oversable material capabilities *and* other states' perception of a state's ability and willingness to use those capabilities. Moreover, while the former – a state's observable capabilities – has received a good deal of attention in the literature and reasonable measures are readily available, the latter – a state's reputation for its material capabilities – requires more careful treatment in the subsequent pages.[19]

I separately produce estimates for observed capabilities and perceptions of state capabilities and then combine the two into a single measure:

$$\text{Capabilities}_{it} = \text{CINC}_{it} \times \frac{\sum_{j=1}^{J}\Pr(\text{Win}_j)}{J-1} \tag{4}$$

Where CINC_{it} refers to the composite index of national capabilities, a measure of a state's percent of all material goods for a given year. CINC scores capture the observed portion of Gilpin's conceptualization. The subsequent perceptions-based portion of the equation $\frac{\sum_{j=1}^{J}\Pr(\text{Win}_j)}{J-1}$ is more complicated. Extending Carroll and Kenkel (2016), I build a machine learning ensemble to predict the outcome of militarized interstate disputes (MIDs) that have taken place between countries. Then I take the predictive ensemble

[17] Gilpin labels these as capabilities and prestige.

[18] p.31

[19] One option is to understand reputation for use of material capabilities as a state's *resolve*. For a micro-oriented treatment of resolve in international politics, see Kertzer (2016).

and use it to predict the probability of each state winning a MID *if every dyad from 1816-2012 were to engage in a MID*. These predicted probabilities can then be aggregated to provide us with a relative sense of each state's expectations about its capabilities and those of others – a proxy for Gilpin's perception-based quantity. Turning to the equation, for every state in every year, I average each states' probability of winning across all possible MIDs in a given year. The aggregated probabilities is $\sum_{j=1}^{J} \Pr(\text{Win}_j)$, where J is the number of states in a given year (I subtract 1 because a country does not fight itself in interstate conflict) and $\text{Win}_j = 1$ if state i expects a victory against state j; summing to the number of expected victories. I then divide by $J-1$ to standardize between 0 and 1.

Before combining the two components into a final dissatisfaction measure, I implement a smoothing function on both to minimize the presence of inaccurate spikes due to measurement issues. For each state, the smoothing function takes all input values from the past (prespecified number) of years and computes the average. The formulation is as follows:

$$\mathcal{S}(x_{it}) = \frac{\sum_{t-n}^{t}(x_{it})}{n} \tag{5}$$

where x_{it} is the input value for state i at time t, n is the number of time periods considered for the smoothing function[20], and $\mathcal{S}$ represents the smoothing function itself. Past years only are considered for the smoothing function because conflict onset – which I theorize is related to dissatisfaction – can influence a country's subsequent dissatisfaction, either furthering social exclusion or successfully remedying dissatisfaction-inducing grievances.

In sum, combining the two components gives us our final measure:

$$\text{Dissatisfaction}_{it} = \log\left[\mathcal{S}\left(\text{CINC}_{it} \times \frac{\sum_{j=1}^{J} \Pr(\text{Win}_j)}{J-1}\right)\right] - \log\left[\mathcal{S}\left(\frac{1}{N}\sum_{i=1}^{N} \text{PageRank}_{it}\right)\right] \tag{6}$$

where $\text{Dissatisfaction}_{it}$ is state i's international dissatisfaction with the status quo in a given year t. $\mathcal{S}\left(\text{CINC}_{it} \times \frac{\sum_{j=1}^{J} \Pr(\text{Win}_j)}{J-1}\right)$ is a state's smoothed expected international benefits in year t and $\mathcal{S}\left(\frac{1}{N}\sum_{i=1}^{N} \text{PageRank}_{it}\right)$ is a state's smoothed actual international benefits in year t. I log both components because they are heavily right skewed, with most states being low in capabilities and low in centrality. Much like the common practice in Economics of logging right-skewed variables, such as an in-

[20] I settle on 4 years for each smoothing function.

dividual's income, to produce a relatively normal distribution, this helps us parse out more fine-grained variation across states for each component. On the final scale positive values correspond to dissatisfaction and negative values to satisfaction. The theory then predicts that the more dissatisfied a state is (increasingly positive values), then the more likely that state is to start an interstate conflict. Dissatisfied states believe they are being materially and socially shorted by the international system. These states then are expected to turn to militarized coercion in order to remedy their dissatisfaction.

4 Data

Expected Benefits

Each measure – a state's relative capabilities and its access to international goods – is a composite of multiple data sources. Breaking down each portion of the relative capabilities measure, recall that a state's relative capabilities are estimated as:

$$\text{CINC}_{it} \times \frac{\sum_{j=1}^{J} \Pr(\text{Win}_j)}{J-1}$$

and serve as a combination of observable capabilities and a state's reputation for using those capabilities. The observed portion of capabilities are a state's composite index of national capabilities (CINC) score (cites) for a given year. Started by Singer et al. (1972), a CINC score represents a state's percent of the globe's material capabilities in a given year. The variables used relate to: industrial capacity, population, wealth, and military size.

While the CINC score captures readily observable quantities, the equation's second portion: $\frac{\sum_{j=1}^{J} \Pr(\text{Win}_j)}{J-1}$ is meant to approximate a state's reputation for its use of capabilities. This is the perceptions-based portion of Gilpin's formulation, where, building on work by Carroll and Kenkel (2016), I train a machine learning ensemble on the components of each state's CINC score to predict the outcome of past militarized interstate disputes. Once a reliable model is trained on the conflicts that have occurred – where the model is evaluated using a test set of data not included in training – the model is used to make predictions for the outcome of all possible MIDs from 1816-2012. This provides predictions about how many MIDs each state could win, given all possible MIDs, in a given year. Although certainly not the exact same as a state's reputation for using force, a model which predicts conflict outcomes well (both

in and out-of sample) serves as a reliable proxy and best guess for what would occur if any two pairs of states were to fight. Importantly, the two quantities appear to capture separate phenomena, with a correlation of 0.469 between the two.

Notably, I have to address one substantial issue when adapting the DOE scores to my purposes. Most MIDs end in a stalemeate. Indeed, in the DOE dataset 84% of MIDs end in a stalemate. This likely occurs because states only select into conflicts if they think they have a chance of winning. However, in practice, this means that when predictions are made about hypothetical MIDs, such as the United States vs. Fiji, the prediction is that the hypothetical MID will end in a stalemate. Although hypothetically a selection process could occur where a MID would only break out between the two if Fiji believed it had a chance of standing up against the U.S. military, a prediction which treats the two as equally capable is an extreme stretch. Ultimately the original DOE estimates predict that essentially all possible dyadic conflicts from the end of World War II-forward will end in stalemates. If the goal is estimate *relative* capabilities, then a model that considers all states to be *equally* powerful is not necessarily useful.

To avoid this issue, I shrink the training set substantially and only consider MIDs where the outcome is not a stalemate. The final predictions then consider: *if each dyad were to fight and one side were to win, then how likely is each side to be the victor?* Although this lowers us to a training set of n = [270?], the machine learning algorithm is able to converge upon a reasonable fit both in training and test set. Indeed, in the the test set, where prediction is inherently more difficult, 72.3% of cases are predicted correctly. Lastly, turning back to the formula of interest – $\frac{\sum_{j=1}^{J} \Pr(\text{Win}_j)}{J-1}$ – we can sum a state's predicted probability of winning a MID for each possible dyad in a given year and then divide by the number of pairs to get an average probability of victory. I summarize the predictions in the next section.

Actual Benefits

The actual benefits portion of international dissatisfaction is estimated as:

$$\text{Benefits}_{it} = \frac{1}{N} \sum_{i=1}^{N} \text{PageRank}_{it}$$

Where PageRank_{it} refers to a state's PageRank centrality in a network in a given year, which is then aggregated across every available network and averaged. This gives us a measure of how central a country is across networks, on average. Here I opt for parsimony and only included variables which

are available across most, if not all, states over time and space. These are a state's: military alliances, interstate trade, shared diplomatic relations, and arms transfers. The most notable exclusion from this list is shared membership in international organizations. However, and unfortunately for inference purposes, most states are members of most organizations. So the network of shared organizational memberships is too dense to reliably parse out which states are more central than others. While the trade network is also dense, the weight of ties (the amount of trade between two countries) differs enough across dyads to consistently extract meaningful variation.

Estimates for a state's centrality within the international alliance network are produced using the Alliance Treaty Obligations and Provisions (ATOP) dataset (Leeds et al., 2002). Trade data comes from the correlates of war interstate trade dataset (Barbieri et al., 2009; Barbieri and Keshk, 2016), shared diplomatic ties are a combination of data from the correlates of war diplomatic exchange dataset (Bayer, 2006) and data used in Duque (2018), and arms transfers were downloaded from the SIPRI conventional arms transfers dataset (Stockholm International Peace Research Institute, 2019). The datasets vary in their temporal span, with the alliance dataset being the most expansive, covering the entirety of the estimates (1815-2012). See Figure 2 for a correlation matrix with each component centrality variable.

Returning to the earlier discussion of Renshon's (2016; 2017) arguments about status dissatisfaction, while the diplomacy network (which is the sole focus of Renshon's analyses) displays a correlation with the other networks (0.622 and 0.665, respectively), there is a substantial swath of variation which it does not capture. Indeed, a closer look at the bottom-left portion of the matrix suggests that much of the correlation across all variables stems from most countries being low in centrality within all networks. Like many social processes the diagonal elements – which contain density plots of the distribution for each variable – are heavily right skewed, with fat-tails. (Barabási and Albert, 1999; Clauset et al., 2009) But as countries reach farther from zero in centrality for each network, then their centrality in other networks starts to diverge. In this sense, particularly among the great powers that tend to have the capability to carry out genuine systemic revision, one's position within the diplomacy network is only a partial indicator of one's position within the trade, alliance, and arms-transfer networks.

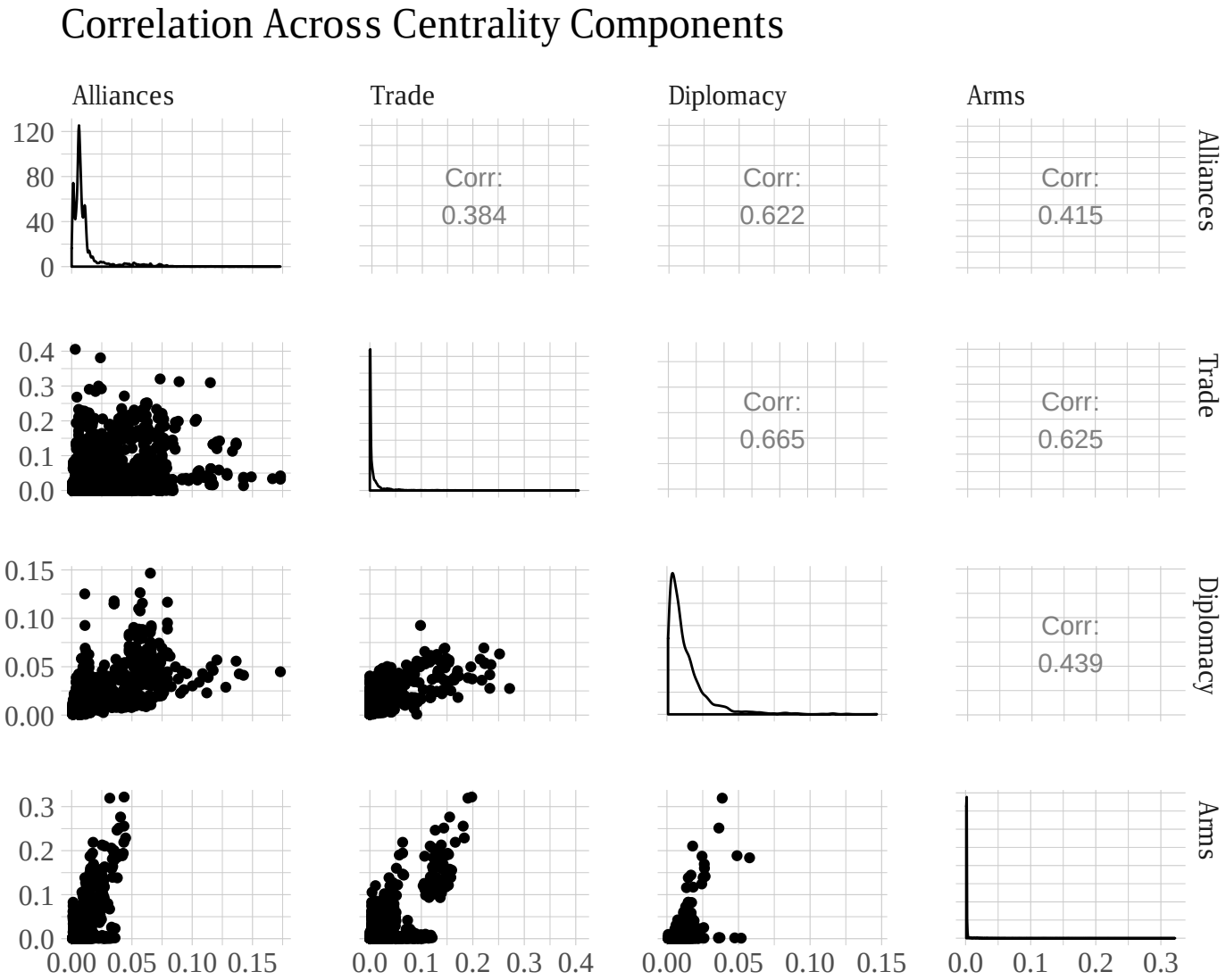

Figure 2: Correlation matrix of each component included in a state's average centrality score. In the bottom-left plots, each point is a state's centrality measure for the variable of interest in a given year. Diagonal elements are a density plot of the distribution of each variable for all country-year combinations. The top-right plots list the correlation between each variable combination. While there is some correlation across variables – it would strange if there were not – the highest correlation between variables is at 0.665 between diplomatic ties and trade. Each variable likely is picking up substantial independent variation.

5 Results

In this section I, in the following order, introduce component estimates, the final estimates, and compare statistical models of interstates conflict onset based upon ideal point estimates and the dissatisfaction measure. Models are fit to both estimate whether a state initiated any militarized interstate disputes (MIDs) and the number of initiated MIDs in a given year. Across model specifications, the dissatisfaction measure maintains a positive and significant relationship with a state's propensity to initiate MIDS. Extreme ideal points are also associated conflict onset, but the relationship is far less consistent – flipping directions and varying in significance depending upon model specification.

Expected Benefits

Starting with the expectations-based component of a state's international dissatisfaction, the main challenge is producing the expectations-based component of material capabilities. A countries observable material capabilities are simply treated as its CINC score in a given year, but the expectations-based component are an extension of Carroll and Kenkel's (2016) 'Dispute Outcome Expectations' scores, which are the predicted probability in all possible dyads that State A wins, State B wins, or as stalemate occurs. Unfortunately, due to patterns in MID outcomes, almost all dyads from around World War II-forward are predicted as almost-guaranteed stalemates. This limits the utility of the original DOE estimates, so I update them to only be trained on MIDs where there is no stalemate. While this limits the size of the training set, which always risks removing useful variation, in this case it ultimately provides more useful estimates. Indeed, these estimates can be treated as the predicted probability that one side will get a better deal in a war-ending bargain than the other side.

I produce these estimates through the automated machine learning (AutoML) algorithm, a process developed by H2O[21]. In the AutoML algorithm, the following models are fit to a specified training set across potential hyperparameter combinations: a random forest, an extremely-randomized random forest, a grid of gradient boosting machines, a grid of deep neural nets, a grid of GLMs, and two stacked ensembles.[22] (The H2O.ai team, 2015) Models are fit with k-fold cross-validation and then their training performance (here with PR-AUC and AUC-ROC since the outcome is a binary catgory) is stored so the user can decide which of the available models to use for predictions, with the best-performing model

[21] https://www.h2o.ai/

[22] A simple overview can be found http://docs.h2o.ai/h2o/latest-stable/h2o-docs/automl.html.

(or leader) being the default choice.

H2O's algorithm is admittedly not the only AutoML option (other popular options include, but are not limited to, AutoSklearn, AutoKeras, and Darwin). This raises the question of why I choose H2O's software over other options. Ultimately, while no single AutoML software is universally agreed to outperform all others, recent research suggests that H2O's algorithm is equally as good as, if not marginally better than, any other AutoML algorithms at binary classification tasks such as this. (Truong et al., 2019) In this case, after fitting various models to the training set, the AutoML algorithm settles upon a gradient boosting machine.[23] (Friedman, 2001, 2002) It is not necessarily surprising that the AutoML algorithm settles upon a tree-based model, rather than a deep neural network (which are often considered the state-of-the-art for predictive modeling), because neural networks are generally understood to require much more data than is available in a dataset like the population of MIDs. Turning to the test set, the algorithm predicts approximately 72% of all out-of-sample MID outcomes correctly. Test set predictions and actual outcomes are compared below in Table 1.

		Actual		
		Victory A	Victory B	Accuracy
Prediction	Victory A	20	7	0.74
	Victory B	4	10	0.71
	Total	24	17	0.73

Table 1: Confusion matrix of predicted and actual outcomes in the *test set*. Training set performance is nearly perfect, though is this is not an indicator of model accuracy because of the tendency of machine learning models to overfit. (Neunhoeffer and Sternberg, 2019) Because the model is trained on dyadic conflicts where one side is the clear winner, Victor A and B represent which of two states is predicted to be or actually is the victor. Elements in the same row and column ((Victory A, Victory A) and (Victory B, Victory B)) represent accurate test set predictions.

For a more fine-grained breakdown of the CINC scores, machine learning estimates, and composite measure over time, see Table 2 in the appendix.[24] In Figure 3, I present the estimated expected international benefits for various great powers over time. The estimates generally follow conventional understandings of power rankings in the literature, but with some important nuance. The United Kingdom initially starts with the greatest expectations, with a shift around the United States power transition in the 1890s. Germany spikes in expectations before both World Wars. During the Cold War there is a

[23] Technically, the AutoML algorithm settles upon an XGBoost model, which is a type of gradient boosting machines. (Chen and Guestrin, 2016)

[24] Tables include the average CINC scores, average probability of wining a hypothetical MID, and the average estimated expected international benefits for the top-10 highest expecting states in the following time periods: Pre-WWI, Interwar, Cold War, and Post-Cold War.

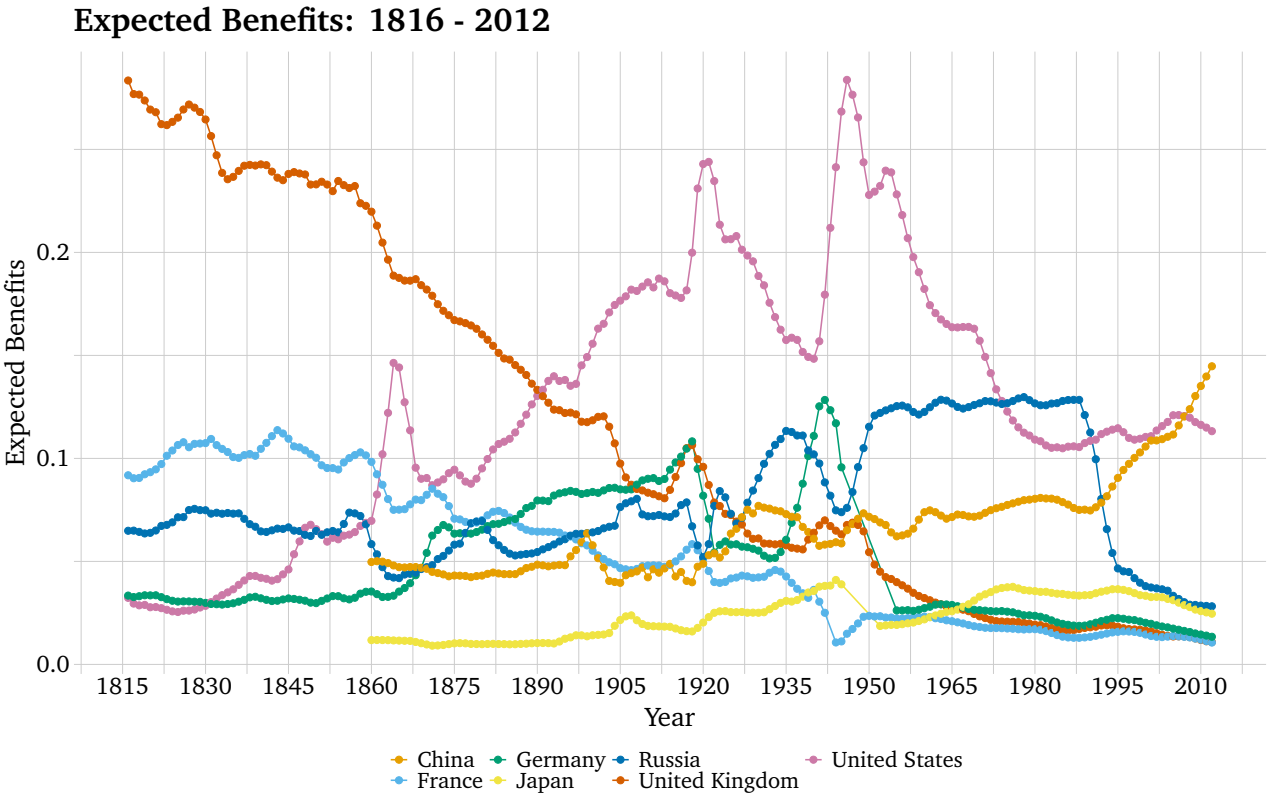

Figure 3: Estimates for a state's *expected* international benefits from 1816-2012. The y-axis is the estimate for each state and the x-axis is the estimate's year. Values are only included for a handful of great power, so that the plot is easily interpretable.

moment of Russian surpassing the United States.[25] The post-Cold War era is initially characterized by United States dominance, but we see China overtaking the US in its desires near the end of the time series. Russia's modern expectations fall considerably short of both the United States and China.

Actual Benefits

Second, let's consider the measure's actual benefits component. To reiterate, I estimate a state's access to valued goods within international networks and then take the average of each centrality score. While the expected benefits component was estimated based upon a iterative process of tuning a machine learning algorithm until test set predictions were sufficiently accurate, the centrality estimates were more straightforward to estimates because I use an already-established centrality formula.

The estimates are included below in Figure 4. Like Figure 3, each point represents a country's

[25]Indeed, this captures the general sense that during the Cold War it was far from obvious which side was more powerful.

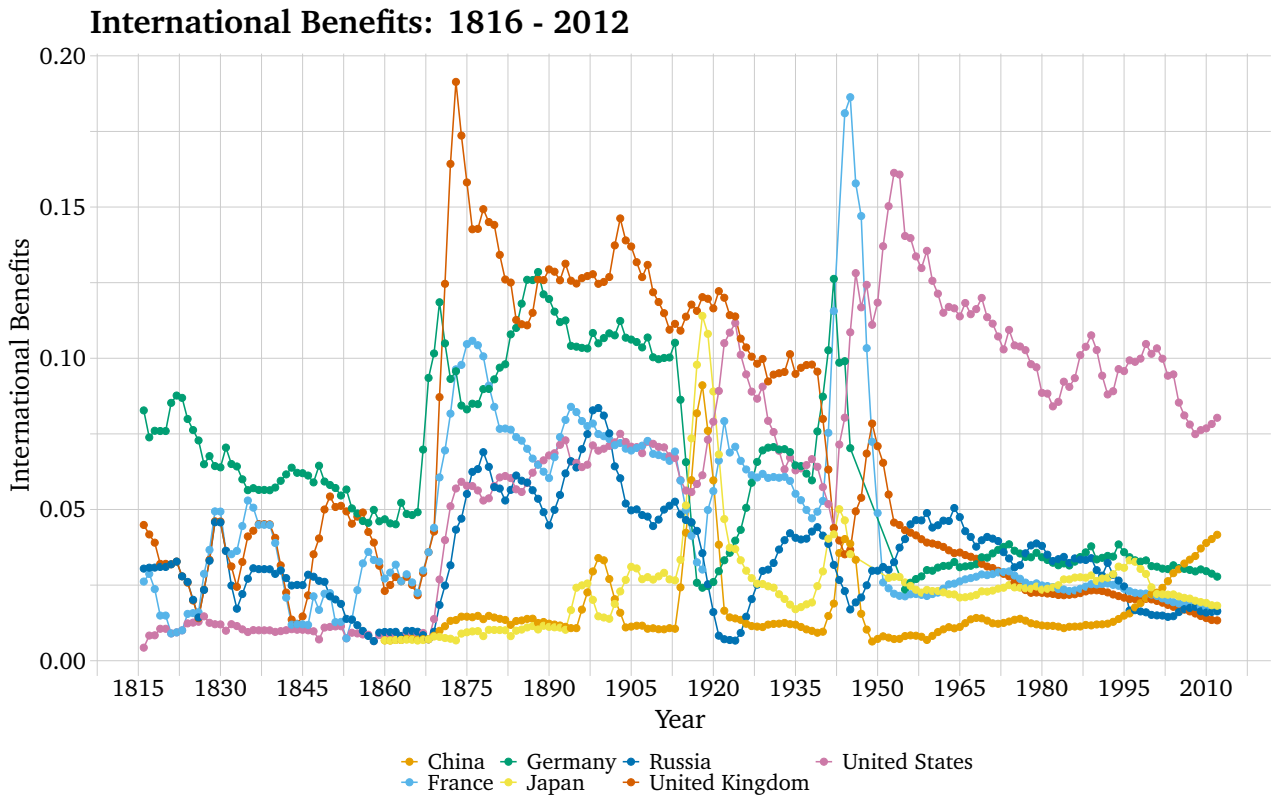

Figure 4: Estimates for a state's *actual* international benefits from 1816-2012. The y-axis is the estimate for each state and the x-axis is the estimate's year. Values are only included for a handful of great power, so that the plot is easily interpretable.

actual estimate for a given year. The y-axis includes decimal values because estimates represent the percent of global benefits that any single state accesses in a given year. Much of the data follows general understandings of international history, with the United Kingdom and United States receiving the majority of international benefits for much of the data. Interestingly, we see China rising in terms of benefits from the early-1990's forward. But China's portion of international benefits in 2012 falls substantially short of the United States. While China is a greater beneficiary than the other included great powers, the gap between China and the United States illustrates an important point. During the last available estimate (2012), The United States nearly doubles China's estimates.

On this note, the gap between the United States and China in estimated international benefits sheds light on one important question regarding China's rise. When it comes to the debate about whether China's growth poses a revisionist threat, a popular question is to ask why a country would alter a system which benefits it? While China has reason to consider itself as powerful, if not more powerful

than the United States, it lags behind in actual benefits from the international system. In this sense, while the international system may be growing increasingly beneficial for China, that does not mean the international system is sufficiently beneficial, nor is guaranteed to be so, in China's eyes. Turning to the final estimates, we see that this has been the case for some time.

Dissatisfaction

Figure 5 includes the final dissatisfaction estimates for select great powers. The figure's format is the same as the preceding figures, where the y-axis includes the measure, the x-axis is the year, and each state is represented by a different color. Positive values on the y-axis correspond to increasingly *dissatisfied* countries. Negative values on the y-axis correspond to increasingly *satisfied* countries.

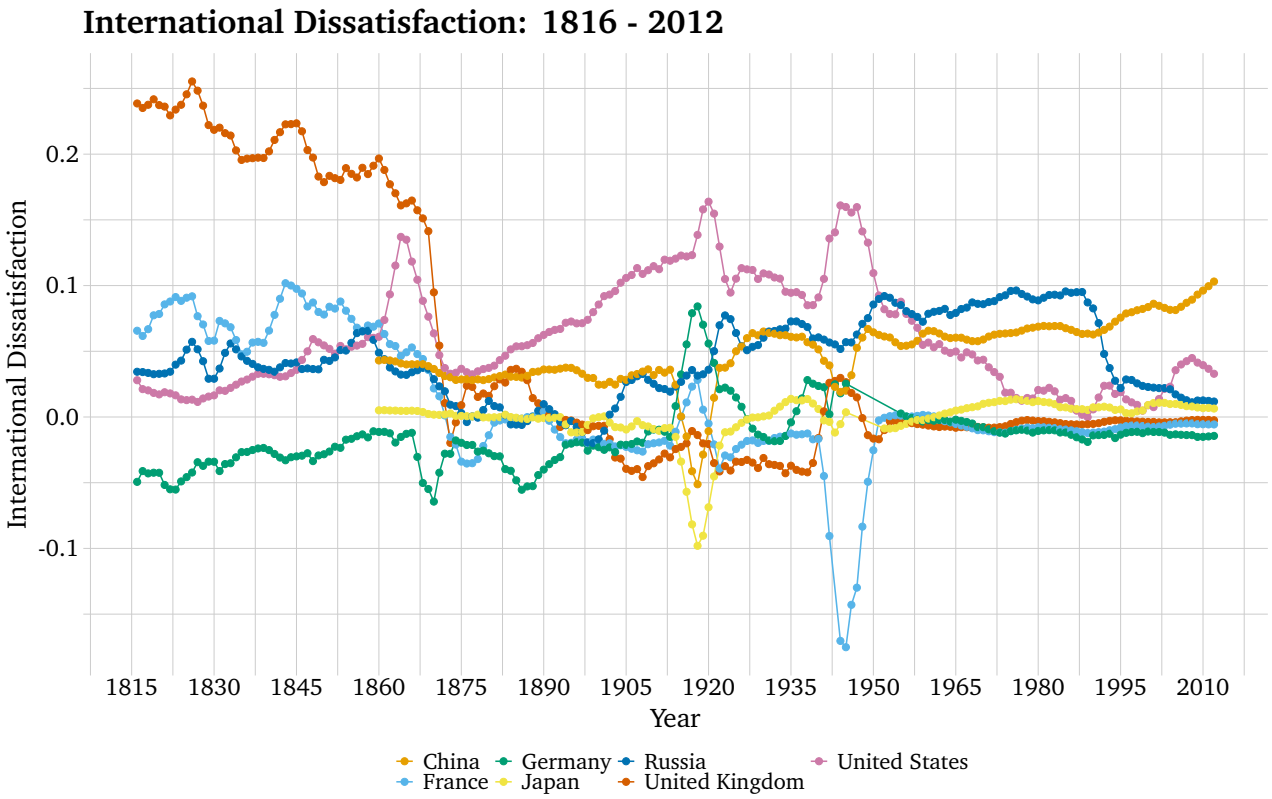

Figure 5: Estimates for a state's international dissatisfaction from 1816-2012. The y-axis is the estimate for each state and the x-axis is the estimate's year. Values are only included for a handful of great power, so that the plot is easily interpretable. Colors correspond to the country. Positive values on the y-axis correspond to increasingly dissatisfied states. Negative values on the y-axis correspond to increasingly *satisfied* states. Note that before each World War Germany spikes in dissatisfaction.

Considering the two World Wars, which are given understandable focus in study of international

dissatisfaction and revision, Germany spikes in dissatisfaction right before both the start of both war s. Turning to the modern era, China and Russia are the two most dissatisfied of the great powers. Russia's dissatisfaction is estimated as being in decline, but we can understand this in the context of Russia's general economic decline. While still a formidable world force, its post-Cold War status is not what it once was. Therefore, we should expect its international expectations to be in decline. Moreover, the European great powers – France, Germany, and the United Kingdom – are all satisfied.

Two surprising outcomes are Japan and the United States both being somewhat dissatisfied in the modern era. Though surprising, these are also the types of estimates that make the modeling exercise worth doing. If the estimates perfectly meshed with general historical narratives, then there would be little added value. Indeed, while not necessarily the narrative one would immediately prescribe to both countries; the United States sits atop the international hierarchy and Japan's economic and cultural standing are well-documented. However, the United States is considerably the most powerful country in world history. Though that has translated to international hegemony, it does not mean its post-Cold War benefits are commensurate to its power-based expectations. Similarly, Japan is a world economic power, but a stagnant level of international benefits may be a legacy of its role in World War II and earlier wars.

Lastly, this measure was presented as substantially different from a ideal point measures. Is this borne out? If, after all of these estimation routines, we just end up with similar values reached through different methods and data, then the exercise has little value. Fortunately, that is not the case. Figure 6 plots ideal point estimates against the final dissatisfaction values. The correlation between the two is almost zero (actually at -0.094), meaning there is almost no association between the two variables when they are plotted against each other. The lack of association between the two variables gives us confidence that the modeling exercise is substantively useful and captures a distinct quantity.

Statistical Performance

How can we validate the dissatisfaction measure as a valid statistical construct? According to my theory of international dissatisfaction, the more dissatisfied a country is, then the more we expect it to turn to coercive military force as a means of remedying that gap. The theory itself is agnostic about which country(ies) will be targeted. But some form of militarized force is expected to follow. Moreover, beyond

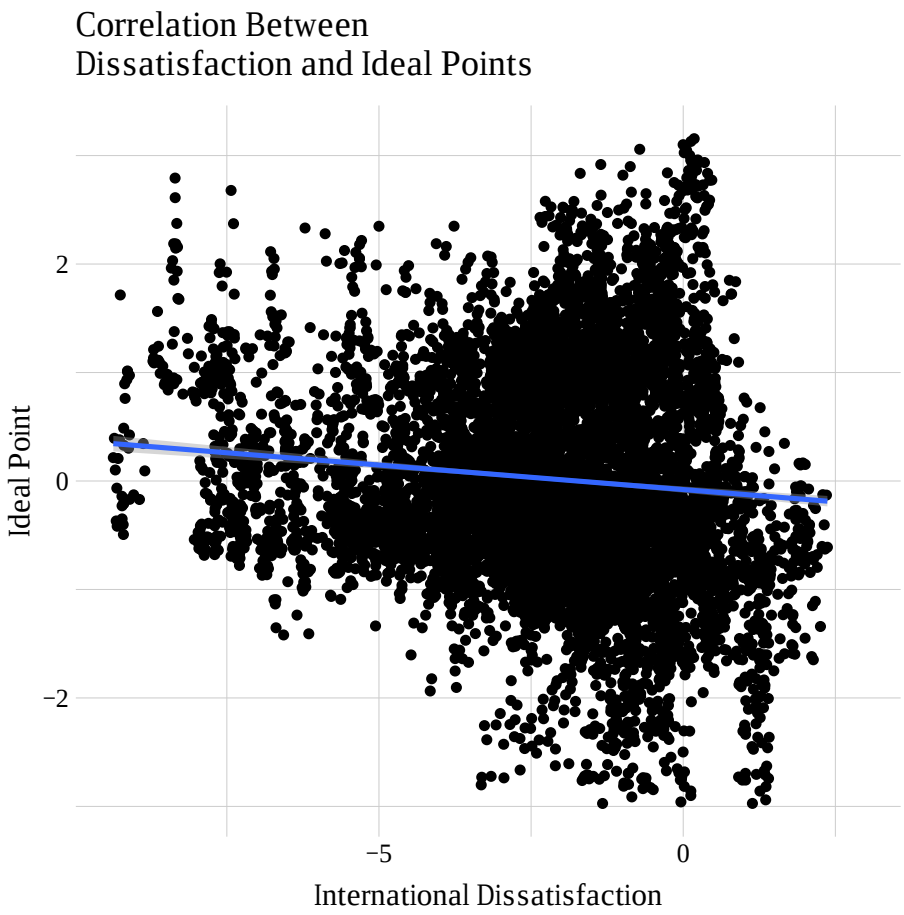

Figure 6: Correlation between United Nations-based ideal point estimates and international dissatisfaction. Each point is a country-year. The correlation between the two is -0.0946, which is demonstrated by the regression line along the plot. The regression line is a gaussian generalized linear model of ideal point values regressed on international dissatisfaction, with standard errors reported.

whether or not the measure has some predictive capacity, we also care about how well it does relative to a popular alternative. At the outset of the paper, I compared international dissatisfaction to a state's ideal point estimate, based on its United Nations voting records. In the following statistical models I estimate a logistic regression of whether or not a country initiated a militarized interstate dispute in a given year and then a poisson regression of the number of MIDS initiated by a country in a given year.

The general formulation is as follows:

$$\text{MID}_{it} = \alpha_i + \gamma_t + \beta \cdot \text{Dissatisfaction}_{it} + \eta \cdot \text{Ideal Point}_{it} + \delta \cdot \text{Control Variables}_{it} + \epsilon_{it} \tag{7}$$

where MID_{it} is whether (or the count) a state i initiates a MID in year t, α_i are individual-level fixed effects, and γ_t are year effects. For control variables, while a standard set of controls are considered in dyadic studies of conflict (joint democracy, peace years, distance, etc.), a standard set of controls for a monadic study such as this one are less apparent. Following Braumoeller (2019), I only consider reciprocated MIDs. Combined with the new MID coding by Gibler et al. (2016), this gives us a population of MIDs that does not involve completely inconsequential MIDs but includes the low-level MIDs that could very well have escalated further than they did (reciprocated). consider a state's polity score and its population, as one can plausibly see a situation in which more autocratic states are both more dissatisfied and have a greater tendency for initiating conflict. Population is also potentially prior to both. I use the logarithm of a country's population as a proxy for size instead of CINC scores, because of the prominent role that CINC scores play in the international dissatisfaction measure. Lastly, fixed effects can reasonably be seen as soaking up other time-invariant unobservable confounders.

Figure 7 includes coefficient estimates. While effect magnitude depends upon whether fixed effects and control variables are included, the primary takeaway point is that international dissatisfaction is, regardless of model specifications, significantly associated with a positive increase in the probability of conflict initiation. Insofar as predicted probabilities are concerned, comparing predictive performance and case-specific predictions is the focus of my next chapter. However, in this regression-specific context on modeling linear associations, when fixed effects are not included, then countries that find themselves further on the fringe of ideal point estimates are more likely to also initiate MIDs. *Yet*, when fixed effects are added to the model, then the relationship between ideal points and MID onset completely reverses and becomes negative. This model-dependence suggests that ideal point estimates are useful in a dyadic context, speaking to how different any two states are. But if one's goal is to understand any one state's general dissatisfaction with the status quo, then the measure provided here has more desirable statistical properties.

Moreover, the relationship between international dissatisfaction and conflict initiation is not just resilient to control variables and GLM family. Indeed, the relationship's direction and significance is also time-invariant. Figure 8, includes the results of a vary-coefficient model (Hastie and Tibshirani, 1993), which are then tested for coefficient changepoints. While changepoints are detected, meaning the relationship's magnitude does vary (Kent et al., 2020), the general point about international dissatisfaction likely driving conflict onset across time and space holds. That said, that the relationship does

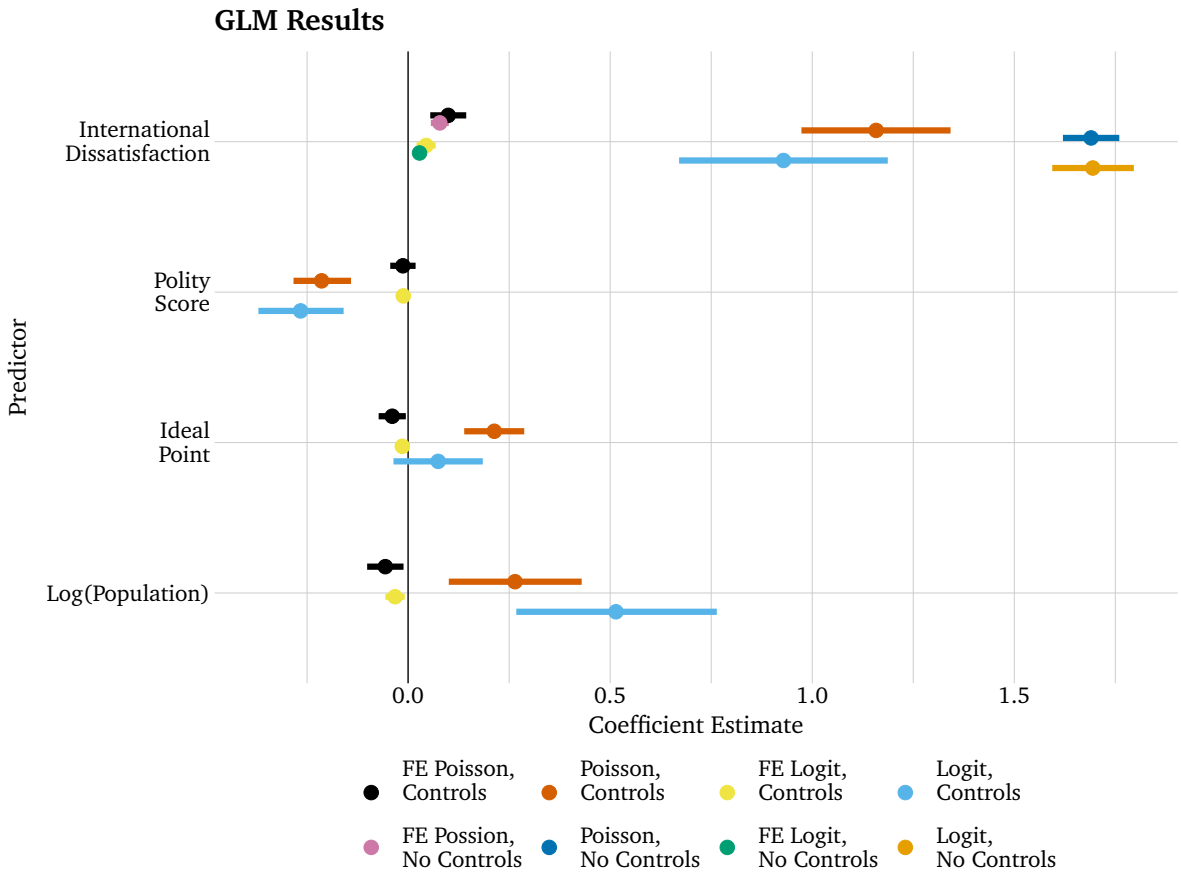

Figure 7: Coefficient estimates and standard errors for GLMs where the outcome is whether a state initiated a reciprocated MID in a country-year. Models with no control variables are a regression of conflict onset on a country's dissatisfaction. Models with control variables include a country's polity score, their ideal point estimate from UN voting records, and their population in a given year. Beyond wanting to compare the dissatisfaction measure to a country's ideal point, the other two control variables are included for two reasons. First, studies of conflict tend to consider regime type and size as control variables. Second, it is possible that both controls are causally prior to dissatisfaction and conflict. All models are fit with and without fixed effects. Unsurprisingly, coefficients are pushed closer to zero with fixed effects models. Count models of the number of MIDs initiated in a country-year are also estimated.

vary over time presents substantively important and interesting variation, which itself is worth investigating. Subsequent work should ask: why does the relationship between dissatisfaction and conflict onset vary? And why the consistent increase from the end of World War II forward?

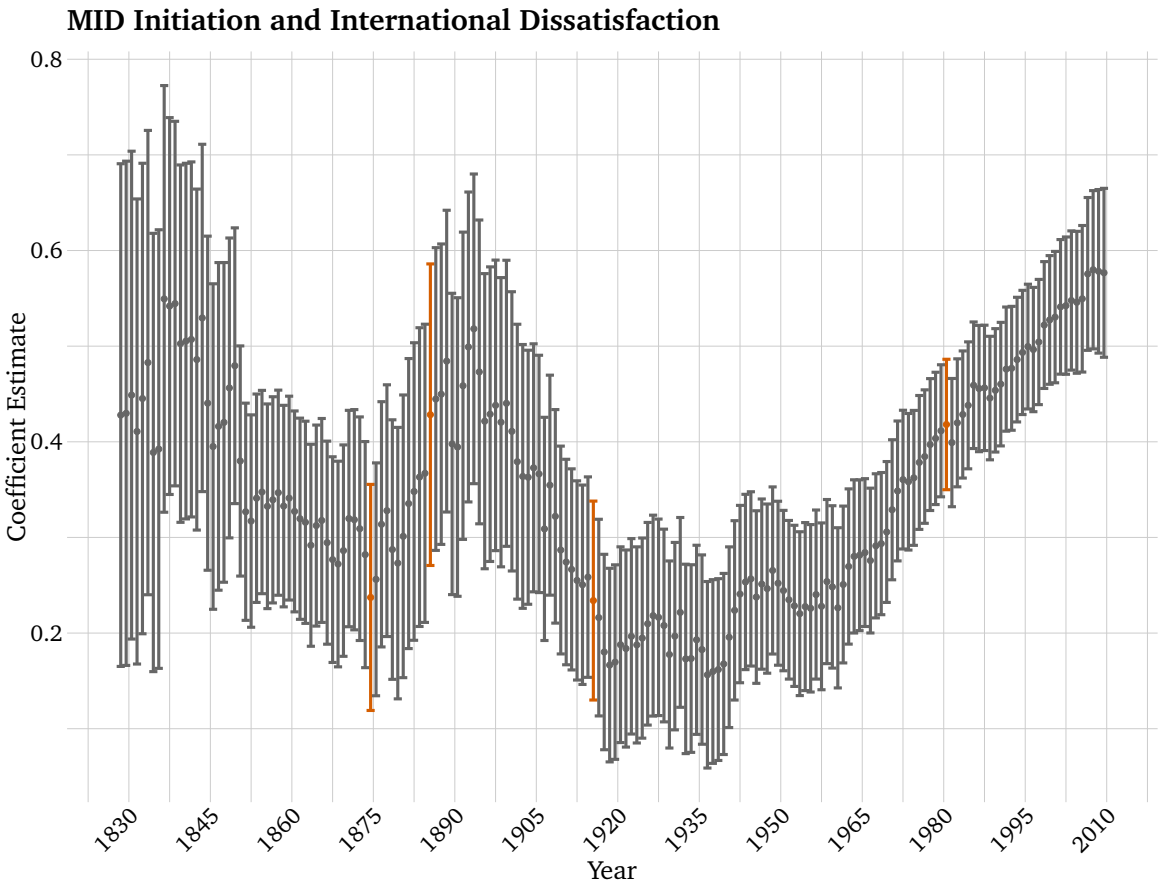

Figure 8: Regression coefficients for a logistic regression of MID initiation in a country-year on international dissatisfaction, controlling for ideal points, polity scores, and the logarithm of a country's total population. Notably, all coefficients are greater than zero and statistically significant. That being said, coefficient magnitude does vary in size. Coefficients which are estimated to be changepoints, based on effect magnitude, are highlighted in red and larger. Whether the relationship between international dissatisfaction and conflict is time-variant and why is an interesting question for future work.

6 Implications

Continuing with the questions raised by coefficient changepoints, the measure's statistical performance not only raises multiple substantive implications, but also novel research questions. Substantively, the measure captures whether or not a country has reason to think that its access to valued international goods falls short of what it should be able to gain in a coercive bargaining scenario. Importantly, the measure represents a form of social exclusion, where a country considers the sum of its relations relative to a its overall capabilities. In this sense, small states may generally be low in international benefits, but also low in expectations about what they can realistically achieve. Furthermore, because expecta-

tions are low, achieving equilibrium is also relatively likely. Joining on to a larger power's hierarchy – exchanging autonomy for welfare (Lake, 2009) – provides ample access to international goods and services.

Rather, by the measure's logic, the greatest threat comes from countries that are large enough to have substantial expectations, but find themselves relatively excluded from the necessary communities and networks that will satisfy unmet expectations. In many ways, this is a fundamental characteristics of revisionist states throughout history. Whether it be Pre-WWI and WWII Germany, Imperial Japan, Communist Russia, or others, a hunger for unmet systemic change drove campaigns of international belligerence. While the story of revisionist states is unlikely to be monocausal – with robust literatures on the role of differential growth rates and various domestic factors demonstrating their importance – the measure's predictive performance does highlight general dissatisfaction's centrality to belligerent states.

Expanding on the point of international revision likely resulting from multiple processes – not just dissatisfaction via social exclusion – a comparison to other explanations for revision is merited. The results in this paper suggest that measuring international dissatisfaction in a Gilpinian manner captures conflictual tendencies more than fringe ideal points, which is a useful validation of the measure. But does the measure itself better predict conflictual tendencies than other popular classes of explanation for international revision? More specifically, does international dissatisfaction better predict a state's propensity to start conflicts than differential growth rates or drastic domestic changes? While a different question than why states become dissatisfied in the first place, knowing which class of variables best predicts conflictual tendencies is valuable for policy and scholarship alike, as it can provide analytical focus. In the next chapter I take on this question, comparing the predictive capacity of each variable in identical machine learning formats with identical training and test sets.

Before concluding, the evidence found in this paper carries provocative and novel implications for debates around whether China is or will be a revisionist state. (Johnston, 2003) *If dissatisfaction stems from a state's ability to access valued international goods, then whether or not China pursues a strategy of revision may depend more on the rest of the international system than China itself.* Put differently, if potential partners are willing to deepen relations with China as its capabilities grow, then there will likely be no incentive for China to alter a system which continues to match its rise. *But if China's growth continues to outpace other state's willingness to expand relations with it, then we should not be surprised*

by a substantial uptick in the coercive use of Chinese military force.

7 Conclusion

International dissatisfaction is widely recognized as central to theories of international conflict and cooperation. However, despites its theoretical significance, its unobservability presents a barrier to empirical operationalization. In this paper I take Gilpin's framework for international dissatisfaction and produce statistical approximations. The result is a dataset with estimates for every state's international dissatisfaction from 1816-2012. When compared to a potential alternative – estimates for a state's ideal point based on its voting record at the United Nations – the measure is a far more robust predictor of conflict onset. Moreover, when estimated with a varying-coefficient model the relationship between international dissatisfaction and conflict onset remains positive and statistically significant across all available years. While dissatisfaction – measured via relative social exclusion – likely does not alone explain revisionist tendencies, the robust relationship between dissatisfaction and conflict onset does support dissatisfaction's centrality to international revision. The measure also provides research opportunities for any area where dissatisfaction plays a conceptual.

Appendix

Top Pre-WWI Capability Estimates

Country	Average CINC	Average Win Probability	Average Expected Benefits
United Kingdom	0.245	0.749	0.186
United States	0.112	0.823	0.0933
France	0.113	0.714	0.0814
Russia	0.122	0.534	0.0636
Germany	0.0847	0.630	0.0538
China	0.148	0.326	0.0467
Austria-Hungary	0.0621	0.590	0.363
Spain	0.0237	0.617	0.0147
Turkey	0.0351	0.414	0.0143
Japan	0.0264	0.490	0.0131

Top Interwar Period Capability Estimates

Country	Average CINC	Average Win Probability	Average Expected Benefits
United States	0.230	0.749	0.183
Russia	0.137	0.684	0.0939
China	0.128	0.537	0.0678
United Kingdom	0.0873	0.732	0.0639
Germany	0.0889	0.704	0.0627
France	0.0577	0.795	0.0407
Japan	0.0042	0.632	0.0280
Italy	0.0350	0.688	0.0241
Poland	0.0212	0.656	0.0139
Spain	0.0164	0.649	0.0106

Top Cold War Capability Estimates

Country	Average CINC	Average Win Probability	Average Expected Benefits
United States	0.202	0.787	0.159
Russia	0.168	0.732	0.123
China	0.111	0.663	0.0738
India	0.0517	0.648	0.0335
Japan	0.0459	0.654	0.0301
United Kingdom	0.0397	0.732	0.0290
West Germany	0.0339	0.720	0.0245
East Germany	0.0293	0.714	0.0209
France	0.0266	0.710	0.0189
Italy	0.0191	0.689	0.0132

Top Post-Cold War Capability Estimates

Country	Average CINC	Average Win Probability	Average Expected Benefits
United States	0.146	0.783	0.114
China	0.166	0.681	0.113
India	0.0715	0.674	0.0482
Russia	0.0524	0.718	0.0377
Japan	0.0463	0.676	0.0313
Germany	0.0256	0.716	0.0183
Brazil	0.0251	0.722	0.0181
South Korea	0.0231	0.649	0.0150
United Kingdom	0.0210	0.706	0.0149
France	0.0196	0.698	0.0137

Table 2: Each table references a time period in international politics and ranks the top ten average expected benefit scores. Each states's average CINC score and average win probability across all dyads is included also. The average expected benefit score represents how much of the international sphere a state expects it should have control over, based on power calculations alone. While the traditional great powers tend to be included in the tables, rankings and percentages vary substantially over time.

References

G. Allison. *Destined for War: Can America and China Escape Thucydides's Trap?* Houghton Mifflin Harcourt, 2017.

M. A. Bailey, A. Strezhnev, and E. Voeten. Estimating dynamic state preferences from united nations voting data. *Journal of Conflict Resolution*, 61(2):430–456, 2017.

A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

K. Barbieri and O. Keshk. Correlates of war project trade data set codebook, version 4.0. *Online: http://correlatesofwar.org*, 2016.

K. Barbieri, O. M. Keshk, and B. M. Pollins. Trading data: Evaluating our assumptions and coding rules. *Conflict Management and Peace Science*, 26(5):471–491, 2009.

R. Bayer. Diplomatic exchange data set, v2006. 1. *Online: http://correlatesofwar. org*, 2006.

B. F. Braumoeller. *The great powers and the international system: systemic theory in empirical perspective*. Cambridge University Press, 2013.

B. F. Braumoeller. *Only the dead: the persistence of war in the modern age*. Oxford University Press, 2019.

S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems*, 30(1-7):107–117, 1998.

S. Brin and L. Page. Reprint of: The anatomy of a large-scale hypertextual web search engine. *Computer networks*, 56(18):3825–3833, 2012.

R. S. Burt et al. *Brokerage and closure: An introduction to social capital*. Oxford university press, 2005.

R. J. Carroll and B. Kenkel. Prediction, proxies, and power. *American Journal of Political Science*, 2016.

S. Chan. Can't get no satisfaction? the recognition of revisionist states. *International relations of the Asia-pacific*, 4(2):207–238, 2004.

T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794. ACM, 2016.

A. Clauset, C. R. Shalizi, and M. E. Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.

J. Davidson. *The origins of revisionist and status-quo states*. Springer, 2006.

M. G. Duque. Recognizing international status: A relational approach. *International Studies Quarterly*, 2018.

D. M. Edelstein. *Over the Horizon: Time, Uncertainty, and the Rise of Great Powers*. Cornell University Press, 2017.

M. Emirbayer. Manifesto for a relational sociology. *American journal of sociology*, 103(2):281–317, 1997.

M. Emirbayer and J. Goodwin. Network analysis, culture, and the problem of agency. *American journal of sociology*, 99(6):1411–1454, 1994.

E. Erikson. Formalist and relationalist theory in social network analysis. *Sociological Theory*, 31(3): 219–242, 2013.

J. D. Fearon. Rationalist explanations for war. *International organization*, 49(3):379–414, 1995.

J. H. Friedman. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pages 1189–1232, 2001.

J. H. Friedman. Stochastic gradient boosting. *Computational statistics & data analysis*, 38(4):367–378, 2002.

R. Friedman Lissner and M. Rapp-Hooper. The day after trump: American strategy for a new international order. *The Washington Quarterly*, 41(1):7–25, 2018.

D. M. Gibler, S. V. Miller, and E. K. Little. An analysis of the militarized interstate dispute (mid) dataset, 1816–2001. *International Studies Quarterly*, 60(4):719–730, 2016.

R. Gilpin. *War and change in world politics*. Cambridge University Press, 1983.

S. E. Goddard. When right makes might: how prussia overturned the european balance of power. *International Security*, 33(3):110–142, 2009.

S. E. Goddard. *When Right Makes Might: Rising Powers and World Order*. Cornell University Press, 2018a.

S. E. Goddard. Embedded revisionism: Networks, institutions, and challenges to world order. *International Organization*, pages 1–35, 2018b.

E. Goh. Meeting the china challenge: The us in southeast asian regional security strategies. 2005.

E. Goh. *The struggle for order: Hegemony, hierarchy, and transition in post-Cold War East Asia*. Oxford University Press, 2013.

E. M. Hafner-Burton, M. Kahler, and A. H. Montgomery. Network analysis for international relations. *International Organization*, 63(3):559–592, 2009.

T. Hastie and R. Tibshirani. Varying-coefficient models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 55(4):757–779, 1993.

P. T. Jackson and D. H. Nexon. Relations before states: Substance, process and the study of world politics. *European journal of international relations*, 5(3):291–332, 1999.

A. I. Johnston. Treating international institutions as social environments. *International Studies Quarterly*, 45(4):487–515, 2001.

A. I. Johnston. Is china a status quo power? *International security*, 27(4):5–56, 2003.

P. Kennedy. *The rise and fall of the great powers: economic change and military conflict from 1500 to 2000*. Vintage, 2010.

D. Kent, J. D. Wilson, and S. J. Cranmer. A randomization approach to dynamic analysis of panel data. *Invited to Revise and Resubmit at Political Analysis*, 2020.

J. D. Kertzer. *Resolve in international politics*. Princeton University Press, 2016.

D. A. Lake. *Hierarchy in international relations*. Cornell University Press, 2009.

D. W. Larson and A. Shevchenko. Status seekers: Chinese and russian responses to us primacy. *International Security*, 34(4):63–95, 2010.

B. Leeds, J. Ritter, S. Mitchell, and A. Long. Alliance treaty obligations and provisions, 1815-1944. *International Interactions*, 28(3):237–260, 2002.

J. M. Lyall. *Paths of Ruin: Why Revisionist States Arise and Die in World Politics*. Cornell University, 2005.

P. K. MacDonald. Embedded authority: a relational network approach to hierarchy in world politics. *Review of International Studies*, 44(1):128–150, 2018.

P. K. MacDonald and J. M. Parent. *Twilight of the Titans: Great Power Decline and Retrenchment*. Cornell University Press, 2018.

M. Neunhoeffer and S. Sternberg. How cross-validation can go wrong and what to do about it. *Political Analysis*, 27(1):101–106, 2019.

A. F. Organski and J. Kugler. *The war ledger*. University of Chicago Press, 1981.

T. V. Paul, D. W. Larson, and W. C. Wohlforth. *Status in world politics*. Cambridge University Press, 2014.

Y. Qin. A relational theory of world politics. *International Studies Review*, 18(1):33–47, 2016.

J. Renshon. Status deficits and war. *International Organization*, 70(3):513–550, 2016.

J. Renshon. *Fighting for status: hierarchy and conflict in world politics*. Princeton University Press, 2017.

K. Schake. *Safe Passage: The Transition from British to American Hegemony*. Harvard University Press, 2017.

R. L. Schweller. Bandwagoning for profit: Bringing the revisionist state back in. *International Security*, 19(1):72–107, 1994.

R. L. Schweller. Managing the rise of great powers: history and theory. *Engaging China: The management of an emerging power*, pages 1–31, 1999.

R. L. Schweller. Rising powers and revisionism in emerging international orders. *Russia in Global Affairs*, 7, 2015.

J. R. I. Shifrinson. *Rising Titans, Falling Giants: How Great Powers Exploit Power Shifts*. Cornell University Press, 2018.

J. D. Singer, S. Bremer, J. Stuckey, et al. Capability distribution, uncertainty, and major power war, 1820-1965. *Peace, war, and numbers*, 19:48, 1972.

Stockholm International Peace Research Institute. *SIPRI Yearbook 2019: Armaments, Disarmament and International Security*. Oxford University Press, 2019.

The H2O.ai team. *h2o: R Interface for H2O*, 2015. URL http://www.h2o.ai. R package version 3.1.0.99999.

M. Trachtenberg. Audience costs: An historical analysis. *Security Studies*, 21(1):3–42, 2012.

A. Truong, A. Walters, J. Goodsitt, K. Hines, B. Bruss, and R. Farivar. Towards automated machine learning: Evaluation and comparison of automl approaches and tools. *arXiv preprint arXiv:1908.05557*, 2019.

S. Ward. *Status and the Challenge of Rising Powers*. Cambridge University Press, 2017.

S. Wasserman and K. Faust. *Social network analysis: Methods and applications*, volume 8. Cambridge university press, 1994.