

THE AFFINITY OF NATIONS INDEX, 1946-2002

Version 4.0
10 March 2006

Erik Gartzke
Columbia University

Data Description:

This is a dataset that attempts to measure state preferences, or more precisely, the interest similarity among pairs of states (dyads). Having stated the purpose of the dataset, truth in advertising mandates that I point out that this objective is problematic. Any attempt to measure state preferences is at best a defensible compromise among intractabilities, implausibilities, and unknowns. Use these data with circumspection!

Background

Students of international relations have long debated the fundamental motivations of leaders and states. To simplify, one can characterize state motives in terms of power, preferences, or perceptions. Surprisingly, the dominant position for much of the post-World War II period (foremost in the form of neo-realism, but also as neo-liberal institutionalism) was that preferences are extraneous. States were said to possess uniform interests dictated by a common desire for security or wealth. Growing impetus behind the discovery of the democratic peace and the collapse of the bipolar system in the early 1990's lead realist (Schweller) and liberal (Moravcsik) interpreters to lay claim to interest variability, arguing that preferences had been lying dormant as a causal variable within each respective paradigm all along. While such claims may be accurate, the period of dormancy was substantial. In the mean time, Bueno de Mesquita's *War Trap* offered an expected utility theory that both precedes and eclipses recent institutionalist accounts involving preferences.

Expected utility theory is an explanation of the role of preferences in human and social behavior, where desire is weighted by the relative ability of actors to realize certain outcomes. In *The War Trap*, and later in Bueno de Mesquita and David Lalman's *War and Reason*, preferences are measured using matrices of alliance ties (portfolios). States that share all the same alliance partners are said to have the highest "utility" for each other's policies (This use of the term "utility" is inaccurate, since the index relies on comparisons in and across dyads. The most that indices of this type can achieve is to measure variability in the similarity of interests between pairs of states). States with no allies in common are thought to have dissimilar interests. The domain bounded by these two extremes constitutes a continuum measuring the similarity of state interests. Bueno de Mesquita uses the τ_B statistic (values range -1 to 1) to construct his utility index.

Advantages and Disadvantages of Various Approaches

Alliance data are that they are available over a long period beginning in 1815. However, the length of these time-series must be balanced with the "utility" of the index. Preferences are treated in rationalist theory as a rank ordering over outcomes. As such, preferences constitute a representation of a supposed psychological condition. Researchers typically cannot observe what actually occurs within the confines of a person's mind. Yet, most students of human behavior will acknowledge that something akin to preferences exist; individuals behave "as if" they prefer apples to oranges, or vice versa, or are indifferent between the two. A key reason social science is so challenging is that this key motivation is not directly observable. Ignoring preferences seems likely to lead to inferential errors, since what actors want appears to matter to behavior.

Yet, how one treats preferences is open to some debate, as researchers can only reliably chart the effects of preferences, not the preferences themselves. Uncertainty invites caution but the inability to perfectly represent preferences should not discourage researchers from taking reasonable steps to approximate their measurement, given the importance of this conceptual element in explaining international and social behavior.

What should an indicator of preferences look like, given the compromises inherent in such an effort? The ideal “instrument” mirrors what one seeks to measure, but which is not directly observable. To measure preferences, one should seek to build an index based on behaviors that minimize the distortion inherent in converting preferences into behavior. Since behavior involves cost, all behavior will look slightly different from actual preferences. A favorite example involves the consumption of automobiles. Graduate students and assistant professors might agree that Ferrari’s are preferred as modes of transportation to VW’s. Yet, one is unlikely to observe those on the lower rungs of academia (or many on the upper rungs, for that matter) cruising around campus in the latest confection from Maranello. Similarly, alliances are pricey acts. States are not free to ally with whatever state they might like in part because alliances involve commitments and costs which preclude the pursuit of other objectives. Thus, a distorting feature of alliances as indicators of interest is that they are costly and so unlikely to be pursued by states that lack strong motives or threats.

One advantage of the Affinity index is that it relies on an information source that is less distorted than are alliance portfolios. Votes in the United Nations General Assembly are often thought to be largely only of symbolic value. If so, then they serve the purposes of an index of interest similarity well. Even issues that are of significant importance to members of the Assembly seldom have the impact of forming an alliance (or, for that matter, of decisions in the Security Council). There are bound to be distortions in any index of state preferences, but these distortions will be least intense where the value in making choices is most modest.

A second advantage of the Affinity index is that it contains more information than data on alliance portfolios. Changes in alliance structure are relatively infrequent, while dozens or hundreds of resolutions appear in each session of the General Assembly. The difference in variance between indices using alliance portfolios and UN voting is particularly acute during the post-World War II period, when alliance patterns became ossified. Arguably, compromises between length of time-series and transparency are best achieved using the Affinity data. Yet, one need not choose between the alternative indices. Use of both alliance-based and UN voting-based indices facilitates flexibility in conducting research and allows for intercomparability of results.

Preference Endogeneity and Other Issues

Where do preferences come from? The truth is that we do not have a compelling answer to this question. There exists no coherent, empirically testable theory of preference origin. While it is not taxing to posit an account of preference-formation, such an effort acquires a substantial mystical element when it passes from cause to effect. Since preferences are not directly observable, one can at most posit empirical claims that occur *through* preferences (a social scientific *ether*). Variance in outcomes could be the result of changes in preferences, or a consequence of changing material conditions. Since in any scientific endeavor some element must be treated as exogenous, it might be preferable to assign this role to the aspects of the explanation that are already difficult to endogenize. Other researchers may approach the issue differently, aware that some factor must be outside the scope of the explanation and ready to grapple with the indirect nature of inference.

A final caveat involves the notion of state preferences. There is obviously fiction in the assertion that states have interests (see Arrow). Like other assumptions, however, this is one that has utility in practice. Should one choose to disaggregate state behavior and model sub-state actors, then some other operationalization of

preferences is in order. While this is appealing in some contexts, there still remain questions in international relations that are most approachable with state-level behavior. I invite researchers to extend indicators of preferences to the sub-state level and also to further refine state-level measures of national interest similarity. For now, this is what I can offer, which I hope is better than the practice of ignoring preferences altogether.

How is the Affinity of Nations Index Created?

The Affinity index constitutes values on a two unit interval scale (–1 to 1) for all countries that are members of the United Nations for the period 1946 to 2002. Initial UN vote data for the period 1946 – 1985 were collected and deposited in the Inter-University Consortium for Political and Social Research (ICPSR dataset # 5512). These data contain a large number of coding errors, some of which are documented in an earlier version of this dataset. Gartzke and Jo (2002) created a dataset of United Nations General Assembly votes. Here, I use UNGA data collected by Erik Voeten (<http://home.gwu.edu/~voeten/UNVoting.htm>).

Code Book:

DYADIDYR: Unique identifier for the dyad year unit-of-analysis. $DYADIDYR = STATEA * 10^6 + STATEB * 10^3 + YEAR - 1000$ (subtract 2000 from years ≥ 2000).

CNTRYERA: Unique ID for STATEA in YEAR. $CNTRYERA = STATEA * 10^4 + YEAR$.

CNTRYERB: Unique ID for STATEB in YEAR. $CNTRYERB = STATEB * 10^4 + YEAR$.

STATEA: Correlates of War Project Country Code for country A.

STATEB: Correlates of War Project Country Code for country B.

YEAR: year (1946 to 2002)

S2un: Values for the Affinity data range from –1 (least similar interests) to 1 (most similar interests). The Affinity data are coded with the “S” indicator (“S” is calculated as $1 - 2 * (d) / d_{max}$, where **d** is the sum of metric distances between votes by dyad members in a given year and **d_{max}** is the largest possible metric distance for those votes, see Signorino and Ritter 1999) from 2 category UNGA vote data (1 = “yes” or approval for an issue; 2 = “no” or disapproval for an issue.), coded as follows:

Code for Votes

1 for “Yes”

2 for “Abstain”

3 for “No”

8 for “Absent (country cast no vote and no evidence of non-participation)”

9 for “Non-member” (South Africa is coded as “55” for the 30th to 47th sessions)”

S3un: Values for s3un (–1 to 1) and are coded as for s2un. These data use the 3 category United Nations Voting data (1 = “yes” or approval for an issue; 2 = abstain, 3 = “no” or disapproval for an issue.).

S2uni: s2un data interpolated for missing values (There are no UNGA votes in 1964).

S3uni: s3un data interpolated for missing values (There are no UNGA votes in 1964).

References:

- Arrow, Kenneth J. 1951. *Social Choice and Individual Values*. John Wiley & Sons.
- Bueno de Mesquita, Bruce. 1981. *The War Trap*. New Haven, CT: Yale University Press.
- Bueno de Mesquita, Bruce, and David Lalman. 1992. *War and Reason: Domestic and International Imperatives*. New Haven, CT: Yale University Press.
- Gartzke, Erik. 1998. "Kant We All Get Along?: Opportunity, Willingness, and the Origins of the Democratic Peace." *American Journal of Political Science* 42(1):1-27.
- Gartzke, Erik. 2000. "Preferences and the Democratic Peace." *International Studies Quarterly* 44(2):191-210.
- Gartzke, Erik, and Dong-Joon Jo. 2002. *United Nations General Assembly Voting, 1946-1996*. Version 3.0. <http://www.columbia.edu/~eg589/datasets>.
- Moravcsik, Andrew. 1997. "Taking Preferences Seriously: A Liberal Theory of International Politics." *International Organization* 51(4):513-553.
- Schweller, Randall. "Neorealism's Status-Quo Bias: What Security Dilemma?" *Security Studies* 5(3):90—121.
- Signorino, Curtis S. and Jeffrey M. Ritter. 1999. "Tau-b or Not Tau-b: Measuring the Similarity of Foreign Policy Positions." *International Studies Quarterly* 43(1):115-144.

Acknowledgments:

I thank Dong-Joon Jo and Richard Tucker for assistance in preparing an earlier version of these data. Erik Voeten has collected the most recent version of United Nations General Assembly roll-call data (<http://home.gwu.edu/~voeten/UNVoting.htm>). Several others have helped at various times.