

Regions of Opportunity and Willingness Data Codebook

v 3

February 4, 2019

J. Patrick Rhamey Jr. PhD

Virginia Military Institute

The Regions of Opportunity and Willingness Data

The Regions of Opportunity and Willingness (ROW) Data provides a means for international relations scholars to incorporate analytically derived geographic spaces comprised of states with unique patterns of interaction into their empirical research. This definition of regions differs from those which use common historical, cultural, or ethnic definitions by focusing upon the actual observed capabilities and political interactions of states. Using an operationalization based upon political behaviors may be more appropriate for much international politics research, while traditional approaches may be more appropriate to others. Regardless, selection of regional definition and identification should be substantively justified at the outset of any analysis. What follows is a brief discussion of the theoretical underpinnings of the database's coding, the logistics of the database's construction, and finally a brief guide to the available data, including a key to regional codes. Please forward any questions to the author.

Please cite the following:

Rhamey, J. Patrick Jr. 2019. Regions of Opportunity and Willingness Data Codebook v3.
<http://patrickrhamey.com/row>.

&

Volgy, Thomas J., Paul Bezerra, Jacob Cramer, and J. Patrick Rhamey Jr. 2017. "The Case for Comparative Regional Analysis in International Politics." *International Studies Review* 19:452-480.

Theoretical Underpinnings

Conceptually, a region is a stable geographic space consisting of three or more contiguous states capable and willing to interact whose interactions with the international system are unique apart from others. Regions thusly are relatively fixed in that they are bounded by geographic proximity, but also are constructed by the actions of regional members and therefore capable of experiencing changes in composition over time. While this definition focuses on observable interactions to determine regions analytically, it also mirrors common definitions in the literature, where regions are, in the broadest sense, "a cluster of states that are proximate to each other and are interconnected in spatial, cultural and ideational terms in a significant and distinguishable manner" (Paul 2012, 4). States' purposeful interactions determine existing relevant regions, whether cooperative or conflictual in nature. This conceptualization of regions is centered on the observed political behaviors of a sub-group of states in the international system. As a result, while culture or history likely impacts a state's behaviors, they are not directly included as an element in the identification, making this conceptualization of regions less useful for some research questions. For example, depending upon the research question, comparativists may seek to define regions by culture or history to generate similar cases for comparison of internal political dynamics. For researchers interested in identifying regions and understanding the types of outcomes interactions between states within regions produce and doing so apart from the dependent variable they wish to study, this attempt at defining the region may prove particularly useful. However, IR research consistently either arbitrarily defines

regions with no explanation or tautologically defines a region using the dependent variable (see Volgy et al. 2007 for an inventory of recent research and further discussion).

This empirically driven identification of regions is intellectually indebted to the work of Lemke (2002) seeking to identify local hierarchies within regions and their implications for regional security. His regional hierarchies are also constrained by geography, using Boulding's (1967) loss of strength gradient, amended from Bueno de Mesquita's (1981) operationalization, to determine the point at which states are able to reach one another militarily given their capabilities, and the degree to which the resulting subsystems demonstrate relative parity or hierarchy. However, preliminary to his analysis of each state's reach is the assumption of meta-regions. For example, Lemke determines the point at which Venezuela and Colombia are capable of reaching one another, but does not determine when Colombia might be able to reach Panama, as the neighboring state exists outside the South American meta-region. Understanding regions as geographic spaces consisting of states with similar patterns of behavior, nothing is necessarily fixed or primordial in the identification contained here. While culture and common historical memory are important components to regions (Katzenstein 2005), in this conceptualization these slow changing processes likely inform, but do not determine, the state behaviors observed to identify regional spaces. This allows dramatic systemic events, such as the collapse of the Soviet Union and the end of the Cold War, to not go unnoticed in regional composition.

Coding Regions

To identify stable geographic spaces comprised of a subset of states with a unique pattern of interactions, I employ an opportunity and willingness framework (Most and Starr 1984). The following summarizes the steps taken to determine regions present annually.

1. **Opportunity.** Each dyad is coded for whether the two states have the opportunity to reach one another given their existing capabilities, tempered by distance. For capabilities, I use a state's GDP as a proportion of global GDP.¹ These proportions of global capabilities are then filtered through a loss of strength gradient (Bueno de Mesquita 1981) to calculate each state's projected power to every other state in the system by the following formula:

$$\text{Adjusted Power} = \text{Power}^{\log\left(\frac{\text{miles}}{\text{miles per day}} + (10-e)\right)}$$

Calculating the distance between two states' capitals (miles), with "power" ranging from 0 to 1, the formula functionally provides a "bubble" of capabilities emanating outward from a state's capital with distance scaled by the number of miles a state is capable of traveling daily (miles per day). Given the time frame of the data is limited to post-World War II, I fix miles per day to five hundred consistent with Bueno de Mesquita (1981).²

¹ For those states missing GDP figures in the World Bank data (Taiwan, North Korea, Somalia, and a number of small microstates depending on the year), the Correlates of war CINC score is used (Singer, Bremer, and Stuckey 1973; Singer 1987).

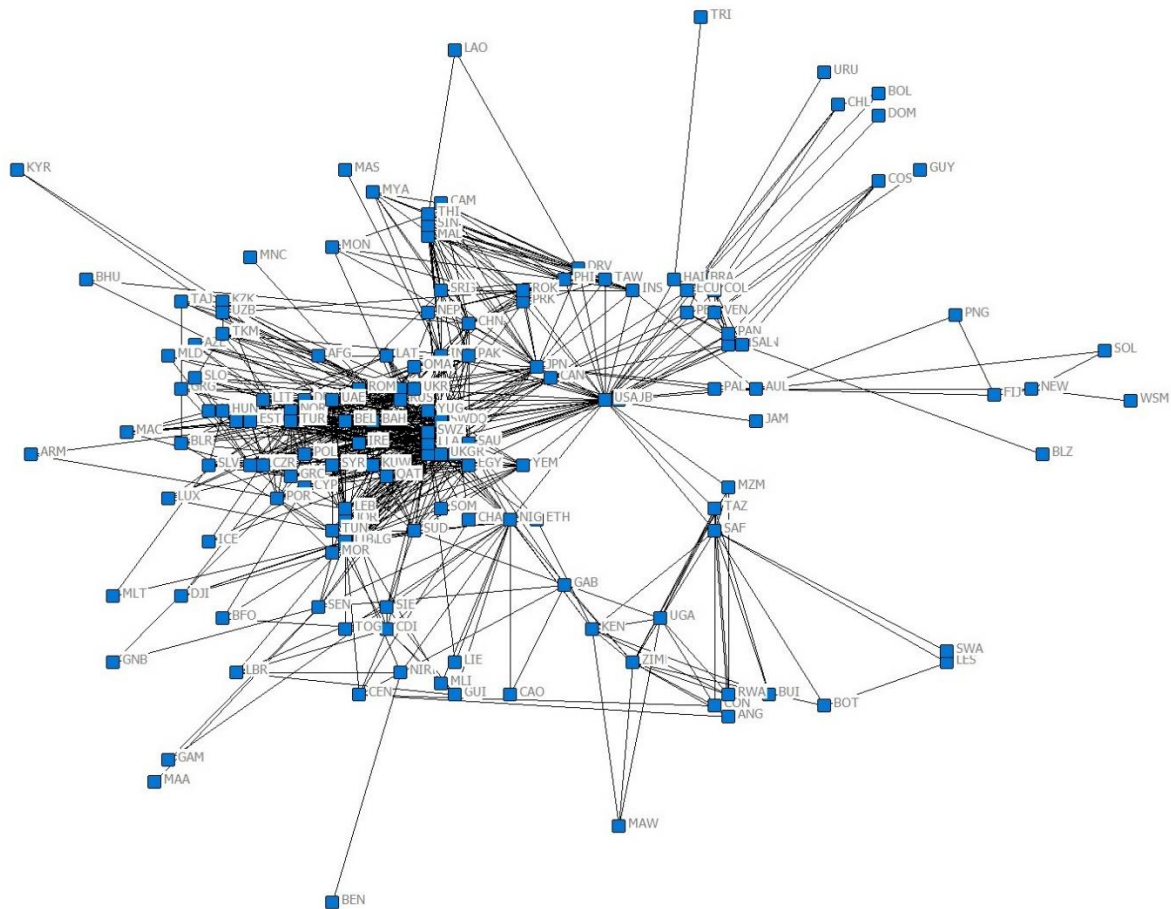
² For further discussion see Lemke (2002). Lemke scales the miles per day component further by the type of terrain being traversed to reach the capital of the target state from the state of origin. Doing so for every state reaching

Following Lemke (2002), a state is said to have the opportunity to interact with another if the loss of strength gradient at the target's capital is at least 50% of the projecting state's unadjusted capabilities. Thus, for two states to surpass the opportunity threshold, both state's projected capabilities to the other state's capital must be greater than or equal to half of their GDP as a proportion of global GDP in that year.

2. **Willingness.** The willingness threshold is the presence of above average amounts of political interaction, cooperative or conflictual, with the target state, as measured by the IDEA dataset (Bond et al. 2003) from 1990-2012 and COPDAB from 1950-1978 (Azar 1980). A state surpasses the willingness threshold if the sum of all scaled events with a target state in a year is greater than the originating state's total foreign policy events with all countries divided by the number of states in the international system. In other words, given a state's foreign policy portfolio, does it direct at the target state an above average amount of its total foreign policy activity?
3. **Network Ties.** For a dyad to have a tie, coded as "1", the states must mutually surpass both opportunity and willingness thresholds. All other dyads are coded as "0". All ties thereby represent a pair of states with both the opportunity and willingness to interact. This set of relationships is then analyzed using clique analysis in UCI network software (Borgatti et al. 2002), which identifies "a collection of actors all of whom 'chose' each other" (Wasserman and Faust 1994, 254). Figure 1 illustrates the network for the year 2000. Cliques must be at minimum 3 or more states.
4. **Analysis of Cliques.** Rather than set an arbitrary cut point at which a node's (state's) ties in the network correlate sufficiently with another to be considered members of the same region, it is recognized that the degree of intra-regional correlation likely varies. Some, like Europe (the densest collection of nodes in Figure 1), have a very strong correlation of ties relative to the system at large, whereas others (i.e. Southern Africa) are less correlated. Indeed, some regions may have more in common with others (e.g. the Middle East and Europe) than some states within a region. Therefore, identified regions are groups of contiguous states over land or up to 500 miles of water that cluster consistently with one another uniquely apart other groups of states in the system. That distinction in uniqueness can occur at any degree of correlation, so long as the group of states is relatively more correlated with one another than they are other groups. In the case of states that are further than 500 miles across water from any other state, they are considered "contiguous" for these purposes with the state that is the most proximate. Identifying these clusters is done through analysis of the endogram output. Starting with the most highly correlated clique (typically Europe), groups of states are in separate regions if they cluster in a "branch" of the endogram apart from others and are contiguous. If a state's ties are correlated similarly between two or more contiguous, more highly correlated groups, it is designated as a border state. If a state lacks sufficient ties in the year to register in the clique analysis, it is omitted.

every other state globally for all years would have been logistically infeasible. Given that the data is limited to the post-World War II era, doing so is unlikely to have a significant effect on the regions identified.

Figure 1: Network of Ties, 2000, ROW Data



5. **Annual Consistency.** While the definition of regions included here acknowledges that regions are prone to change reflecting the “observable power and purpose” of states (Katzenstein 2005, 2), it is likewise recognized that regions conceptually should entail some degree of stability, not prone to dramatic annual change sparked by any single event. A single event, whether an attack or cooperative arrangement, should not alone possess the ability to form a region, which then dissipates in the following year. To solve the problem of single events having an inordinate impact on a year’s network, and to conceptually reflect regional stability, states are coded as members with other states who they cluster with most often in a five-year period. Therefore, for the year 2000, France clusters with Britain from 1996-2000 every year, and is therefore in the same region with Britain in the final 2000 coding. States that do not appear in the clique analysis due to an absence of ties in a majority (3 of 5) years are omitted. States that do appear in the clique analysis for 3 of 5 years, but are identified as a border states in a sufficient number of years so that they do not cluster with a specific group for a majority, are placed with the group of states with which they most frequently cluster. If they cluster an equal number of years across the five-year period with two or more groups, they are coded as a border state.

6. **Peripheral States.** Some states, particularly very small ones, may not register due to not possessing ties at all or not possessing ties with any single unique contiguous group. Following steps 1-5, remaining unplaced states are identified as peripheral members of a region if proximate to only one, or, if between regions given contiguity, identified as border states. For example, if Argentina and Brazil both are in the South American region, and Uruguay does not actively cluster a majority of the five-year period, Uruguay is coded as a South American peripheral state. Likewise, San Marino always is a peripheral state to whatever region includes Italy. In the data, a peripheral state is coded as “1” for the “periph” category. Among border states (coded “999” in the region column), those with a “1” for periph are those that do not cluster a majority of the five-year period and are geographically between identified regions. Those with a “0” are those that do cluster a majority of years, but not with any one of the identified regions more than another.
7. **Region Codes.** Each region is then given a three-digit code, representing a commonly referenced regional space and noting when a region dramatically changes composition into something else (e.g. “Europe” is most of the continent, usually including Russia alongside the Western major powers, whereas “Western Europe” is coded when the region does not extend eastward past Germany/West Germany). A list of region numerical codes and their names are provided below.
8. **Regions by Decade.** For the purposes of past research on regional powers, decade aggregations to analyze patterns of behavior of powerful states within consistently present spaces is provided in the “decade” variable. The variable “dperiph”, likewise, is coded 1 if the state is peripheral to the region as described in 6 above for a majority of the years in the decade. Aggregations of this sort can be done by following the above method for any relevant time period (e.g. Cold War and post-Cold War).

Version History

Version 1 (Rhamey 2012 & Cline et al. 2010). Regions only post 1990 in five-year increments. States received a tie in the network if they surpassed an opportunity threshold of either Military Expenditures (SIPRI) or Market Capitalization (World Bank) and a willingness threshold of either events (IDEA only) or trade (Barbieri et al. 2009) in any amount.

Version 2 (Volgy et al. 2017). Added years 1950-1978 using COPDAB to extend the willingness measure. To provide consistency and maintain data availability across the increased time period, the opportunity threshold was limited to projected GDP (proportional to global GDP, World Bank) and the willingness threshold limited to the two measures of events (COPDAB and IDEA). Unlike version 1, measurement distinguishes between “core” and “periphery” states, where core states actively clustered with contiguous neighbors a majority of the preceding five year period, and peripheral states are those that, while not clustering a majority of the time with a group of neighbors, clustered some years and/or are geographically unable to be placed in any other proximate region (e.g. San Marino is always in the same region as Italy). Regions were then identified by decade (50s, 60s, etc.) for use in analysis of regional powers.

Version 3 (Rhamey 2019). Revised post-2000 data using an updated version of the IDEA dataset and made minor corrections to other years. Data is provided annually, with states coded in groups they appear with in a majority of the preceding five-year period. Decade tables replicating those used in Volgy et al. (2017) are also reproduced. Each state is given a regional code, distinguishing between core and peripheral members.

Country changes in decade tables from Version 2 to Version 3:

- 1950s: Cuba from Border State to N. Central America.
- 2000s:
 - From Europe to Middle East: Israel
 - From Border State to Middle East: Oman, Yemen
 - From Maghreb to Europe: Algeria, Libya, Morocco, Tunisia
 - From Border State to European Periphery: Malta, Monaco
 - From Border State to West Africa: Senegal
 - From Border State to West African Periphery: Cabo Verde, The Gambia, Guinea-Bissau
 - From Horn of Africa to Central Africa: Ethiopia, Somalia
 - From East Asia to South Asia: Nepal
 - From Southern Africa to Border State: Angola, Botswana, Namibia
 - From Middle East to Border State: Azerbaijan
 - From Horn of Africa to Border State: Eritrea
 - From West Africa to Border State: Ghana
 - From South Asia to Border State: Tajikistan

Please contact rhameyjp [at] vmi.edu with errors, comments, and suggestions.

Version 3 Region Codes³

100 – North America
 110 – N. Central America
 120 – S. Central America
 200 – South America
 210 – Andes
 220 – Southern Caribbean
 300 – Europe
 310 – Western Europe
 320 – Northern Europe
 321 – Benelux
 322 – Scandinavia
 330 – Eastern Europe
 332 – Baltics
 333 – Balkans
 340 – Central Europe
 400 – West Africa
 410 – Gold Coast
 420 – Western Coast
 430 – Central Savannah
 500 – Southern Africa
 510 – Central Africa
 520 – Horn of Africa
 530 – Upper Zambezi
 540 – Nyanza
 600 – Middle East
 610 – Maghreb
 620 – Gulf States
 630 – Near East
 700 – South Asia
 710 – Central Asia
 800 – East Asia
 810 – Northeast Asia
 900 – Asia Pacific
 910 – Southeast Asia
 999 – Border States*

*Note: In the five-year annual data only, a state that is coded both 999 for region and 1 for periphery means that it does not cluster for a majority of the five-year period preceding and including that year. A state that is coded 999 for region and 0 for periphery does cluster a majority of the five-year period preceding and including the year but does so equally between 2 or more groups and cannot be placed.

³ Special appreciation to the work of three undergraduate research assistants, Kyle Reavis, Timothy Butt, and Trevor Smith, for inventorying patterns of behavior and identifying existing popular regional names and their composition.

References

- Azar, E. E. 1980. "The Conflict and Peace Data Bank (COPDAB) Project." *Journal of Conflict Resolution* 24 (1): 143-152.
- Barbieri, Katherine, Omar M. G. Keshk, and Brian M. Pollins. 2009. "Trading Data: Evaluating Our Assumptions and Coding Rules." *Conflict Management and Peace Science* 26(5):471-491.
- Bond, Doug, Joe Bond, Churl Oh, J. Craig Jenkins, and Charles Lewis Taylor. 2003. "Integrated Data for Events Analysis (IDEA): An Event Typology for Automated Events Data Development." *Journal of Peace Research* 4 (6): 733-45.
- Borgatti, S. P., M. G. Everett, and L. C. Freeman. 2002. *Ucinet for Windows: Software for Social Network Analysis*. Harvard, MA: Analytic Technologies.
- Boulding, Kenneth E. 1962. *Conflict and Defense: A General Theory*. San Francisco: Harper.
- Bueno de Mesquita, Bruce. 1981. *The War Trap*. New Haven, CT: Yale University Press.
- Cline, Kirssa, J. Patrick Rhamey, Alexis Henshaw, Alesia Sedziaka, Aakriti Tandon, and Thomas J. Volgy. 2010. "Identifying Regional Powers and Their Status," in Volgy, Corbetta, Grant, and Baird, *Major Powers and the Quest for Status in International Politics: Global and Regional Perspectives*.
- Katzenstein, Peter J. 2005. *A World of Regions*. Ithaca, NY: Cornell University Press.
- Lemke, Douglas. 2002. *Regions of War and Peace*. Cambridge: Cambridge University Press.
- Most, Benjamin A., and Harvey Starr. 1984. "International Relations Theory, Foreign Policy Substitutability, and "Nice" Laws." *World Politics* 36(3): 383-406.
- Paul, T. V., ed. 2012. *International Relations Theory and Regional Transformation*. Cambridge: Cambridge University Press.
- Rhamey, J. Patrick Jr. 2012. "Constrained to Cooperate: Domestic Political Capacity and Regional Order." PhD diss. University of Arizona.
- Singer, J. David, Stuart Bremer, and John Stuckey. 1972. "Capability Distribution, Uncertainty, and Major Power War, 1820-1965." In Bruce Russett (ed) *Peace, War, and Numbers*, Beverly Hills: Sage, 19-48.
- Singer, J. David. 1987. "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816-1985." *International Interactions* 14: 115-32.
- SIPRI (Stockholm International Peace Research Institute). "SIPRI Military Expenditure Database." <https://www.sipri.org/databases/milex>.
- Volgy, Thomas J., Paul Bezerra, Jacob Cramer, and J. Patrick Rhamey. 2017. "The Case for Comparative Regional Analysis in International Politics." *International Studies Review*. 19 (3): 452-480.
- Wasserman, Stanley, and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.
- World Bank. World Bank Open Data. <https://data.worldbank.org/>