**Note: While this is online only, we would like a NWS-specific header followed by a standardized reference to the work that it accompanies. This will become our standard for online appendices.**

**Supplementary On-Line Material.**

*Purpose*

Why write a 1-page research note and what do you expect readers to do with it?

*Disciplinarily*, we have two motivations. First, social science articles are getting longer, and many papers are often buried with technical detail and dense literature reviews. Such papers seek to wrap up problems, rather than open them. A short, empirical note or, since the real substance of this project is a figure, a “portrait” is designed to *open* problems for investigation. The advantage of a new journal and new format such as that afforded by *Network Science* is to reshape the basic form of an empirical finding: the goal here is to present a simple data-based puzzle for the scientific community to ponder. Here the massive shift in polarization, and the ways in which this intersects with political party politics, majority membership, presidential party and individual careers provides a deeply complex puzzle to think about. This sort of puzzle, we think, is just the sort of problem that *Network Science* is poised to work on.

Second, data visualization is a hallmark for network science, but exemplary data-deep models are rare. Often network visualizations are wedded to program defaults and generate images that provide very little substantive information. But visual approaches to data can allow viewers to make connections between parts – to literally see a micro-macro relation (Tufte, 2001) – which we seek to take advantage of here. As such, we have built an image that informs immediately (the increasing trend in political polarization), but also rewards detailed scrutiny (the coalition patterns within Senate sessions, or the career trajectories of particular Senators over time, or the intersection of executive and legislative control, to name just a few). In this, we are similar to work on the visualization of agreement groups in the Senate (Friggeri, 2012), which also readily demonstrates the political trajectories of individual Senators through time in a compact visualization.

*Substantively*, the goal of this piece is to document an interesting political phenomenon in a way that might lead to new insights. The U.S. Senate is experiencing a level of polarization, as measured by the structural pattern of co-voting ties, unseen for over a century. The potential causes of this pattern may be many, but it seems plausible that endogenous coalition behavior captured through networks might contribute, and this figure is designed to help us explore such possibilities. Current work in political sociology suggests that such extreme professional polarization is at odds with American popular opinion. This suggests that a simple voter-representation explanation cannot hold, and that internal dynamics could be driving the divisions (Baldassari & Bearman, 2007; Baldassari & Diani, 2007).

*Methodologically,* this image reflects advances in both network analysis and visualization. First, network studies still have a strong cross-sectional bias and much work tends to emphasize connectivity features (centrality, cliques) rather than positional features (roles & systems patterns; see Burt, 1987). We seek here to expand on both of these dimensions. First, we fold time into the network *as* *a type of structure* with identity arcs (see Moody, 2010 for visualization and diffusion implications, Mucha et. al. (2010) for community detection tools on such graphs). This allows one to capture network dynamics in a straightforward way and ask how types of positions emerge over time. Second, we add role information to the connectionist subgroups by combining new community-detection tools (Blondel et al., 2008; Newman, 2006) with structurally-equivalent blockmodels (White et al., 1976). The connectionist point of view captures the extent of political polarization, while role positions help us understand how such polarization evolves.

The most direct methodological advance of this piece is in the form of a careful new visualization template. Visualizations of these sorts of data – the one-mode projections of two-mode networks (also known as “persons-through-groups” data – Breiger, 1976) -- typically suffer from high density, since there is a weighted line between every pair in the network, and standard space-based layouts typically result in large, uninformative clusters. One solution is to plot only the strongest ties, but this can still leave a large cluster of nodes in a fully-connected clique. This problem is amplified in co-voting networks, where (nearly) every member votes on every item. Here, we simplify the network by aggregating over the structurally equivalent positions. Since every Senate session includes at least two positions consisting (almost) entirely of party-line voters, we chose to anchor the layout with these “party loyalist” positions.

*Data*

The data are roll-call votes in each completed Congressional session as constructed by Poole (2009; Poole & Rosenthal, 1997). We cleaned the data for name consistency (in the raw data, party-changing Senators are sometimes assigned different ID numbers; we instead continue to identify them by individual). For the purposes of investigating differences in voting, we include only non-unanimous votes (less than 97% agreement; as in Waugh et al., 2009). Data on the 112th congress runs through June 7th, 2012 and was taken from: <http://adric.sscnet.ucla.edu/rollcall/>. To construct identity arcs in this figure, all data were resolved to “name.state” level, with corrections for any instances with two people having the same name (such as Jr. or when spouses are appointed for a deceased member). In rare instances where the same senator had different ids within the same year, we treated the data as given.

*Methods*

For each two-year Senate, we construct a network with Senators as nodes and their agreement on votes cast as edges. The value of the edge is the proportion of all votes made by the two Senators where they agreed on the vote (positive or negative). These network “slices” are then joined over time by *identity arcs* to generate a fully time-connected network. An identity arc captures the number of Senators who move from position *i* at time *t* to position *j* at time *t*+1. If the positions are individuals, then the arc simply links a person from their past to future selves. If the positions are aggregates, then the value of the arc shows the number of people moving between such positions over time.

Measuring polarization

We operationalize polarization as the tendency for votes to fall within clearly defined groups (called variously “clusters” or “communities” or “cohesive groups” in the literature; we stick with “communities” here for simplicity). Communities were identified endogenously using an algorithm that partitions the network into mutually exclusive and exhaustive sets such that the weighted volume of ties within groups is maximized relative to a null model (see Waugh et. al,. 2009).[[1]](#footnote-1) The code for this program can be found at at <http://netwiki.amath.unc.edu/GenLouvain> at the file named klnb.m (Jutla, Jeub, & Mucha 2011-2012). Specifically, communities were identified here using the Kernighan-Lin (see Newman, 2006) iterative improvement of a Louvain method greedy partitioning algorithm (Blondel, et al. 2008) for maximizing modularity. Modularity measures the relative weight of ties within versus between communities, calculated as eq. 1 below.:



(1)

Where:

s indexes communities in the network

*l*s is the total weight of lines in community s

ds is the sum of the weighted degrees of s

L is the total weight of lines in the network

With the exception of the 103rd Senate (1993-‘94) which included a small 3rd group, all networks plotted in figure 1 were maximized with a 2-cluster solution corresponding (largely) to political party. The resulting modularity score is plotted as the y-axis in the figure and the inset, and traces the extent of polarization over time (Waugh et al., 2010). Since the division is typically split between two groups, we visually divide the distance from pure integration (modularity=0) evenly between the two parties resulting in a symmetric trend around the zero line (though the contribution to modularity made individually by the two clusters are not equal).

Identifying Roles

To identify role positions, we modified the standard CONCOR strategy for identifying structurally equivalent positions (White et al., 1976). The identification of positions is done *independently* of the cohesive group detection algorithm, and the code for the CONCOR split, construction of edge-identity links and initial graph output can be found at <http://www.soc.duke.edu/~jmoody77/congress/simpleread3_redone_wc111_c112.sas>. CONCOR is a divisive clustering algorithm that splits a whole into two parts, and then continues to split each resulting group in two again until a user-defined threshold is met. This split is defined by the “convergence of iterated correlations;” where an initial correlation between pairs based on the pattern of ties to all others is repeatedly correlated until all cells converge on +1 or -1. A critique of CONCOR is that the bifurcation strategy strongly constrains the resulting partition (Wasserman & Faust, 1994). To solve this, instead of specifying a global number of splits, we specify a minimum correlation threshold within each group, and only split *along* *that branch* if the internal similarity of the group is below our correlation threshold (initially 0.85), ensuring a consistent level of similarity amongst positions. Individuals will occupy unique positions if the internal correlation of the branch they are part of never reaches the threshold.

Finally, the extent of polarization in later periods is so great, that a single threshold over the entire time period set high enough to distinguish unique positions in later periods resulted in little aggregation in the earlier period (often returning 50-60 positions). Thus, we adjusted the threshold correlation dynamically to ensure between 5 and 20 positions; if the process yielded more (less) than 20 (5) positions at a threshold of 0.85, we lowered (raised) it by 0.005 until the size bound was met. Substantively this means simply that the global pattern (voting diversity early, polarized party-line later) would be magnified by any constant threshold. Seen another way, cross-party vote becomes more salient when most votes fall within party, and thus looking distinctively “middle ground” requires fewer cross-party votes.

Visualization layout & display

The layout is achieved in four steps that combine fixed and space-based layout strategies. The community detection was done in MATLAB and CONCOR was conducted in SAS. Visualization was done using the PAJEK network analysis software for the base layout and adjusting color and label details in Adobe Illustrator. First, for party loyalist positions, the y-axis is set by the polarization measure and party (-1\*modularity for Republicans, + modularity for Democrats) and the x-axis is set by date. Second, for non-loyalist positions, we apply a Kamada-Kawai spring-embedder algorithm that adjusts the position of non-loyalists between the loyalist anchors. For this stage of the layout process (only), we remove all ties except those linking the party-block anchors to middle-level positions. Since each non-loyalist position is effectively “suspended” between the anchors with a weighted link to each, the resulting y-position of non-loyalists reflects their balance of agreement with the anchors. For presentation clarity, we display the co-voting similarity edge if agreement is > 0.60 (green edge in the display). The x-position is adjusted manually within the date-range to minimize line crossings. Third, we size edges and nodes proportional to volume, generate labels and export to Illustrator. Fourth, after the images are exported to Illustrator, we place labels and keys and adjust shades and colors. Individual Senators are labeled so that one can trace the careers of all Senators who appear more than once in non-loyalist positions (though you may have to zoom in some to do it!). Aggregate positions are then shaded relative to the average vote agreement (resulting in light to dark red/blue). For example, the lowest “democratic loyalist” agreement level was in the 96th Congress, with members agreeing 72% of the time, which increases to 89% of the time in 108th Congress.

Presidential term and Senate party balance is given in the timeline at the bottom of the figure. We use a purple shade reflecting the Red-Blue balance of the senate, interpolating the proportion Red/Blue in the RGB color balance to set purple as a 50/50 split. For convenience, we label the dates of each Congress ignoring any final few January dates of that Congress (e.g., we label the 110th Congress as ’07-’08 even though it was officially in session from January 4, 2007 to January 3, 2009).

*References*

Baldassarri, D. & Bearman, P. S. (2007). Dynamics of Political Polarization, *American Sociological Review, 72, 784-811*.

Baldassarri, D. & Diani, M.. (2007). The Integrative Power of Civic Networks. *American Journal of Sociology, 113, 3, 735-780*

Blondel, V. D., Guillaume, J. L., Lambiotte, R. & Lefebvre, R. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *10, P10008.*

Breiger, R. L. (1974). The Duality of Persons and Groups. *Social Forces, 53, 181-90.*

Burt, R. S. (1978). Cohesion Versus Structural Equivalence As a Basis for Network Sub-Groups. *Sociological Methods and Research, 7, 189-212.*

Friggeri, A. (2012). Agreement Groups in the United States Senate. Retrieved from: <http://friggeri.net/research/senate/> (retrieved 2012).

Jutla, I. S., Jeub, L. G. S., & Mucha, P.J. A generalized Louvain method for community detection implemented in MATLAB, <http://netwiki.amath.unc.edu/GenLouvain> (2011-2012).

Moody, J. (2010). Static Representations of Dynamic Networks Manuscript (R&R at *Social Networks*)

Mucha, P. J., Richardson, T., Macon, K., Porter M. A., & Onnela. J.-P. (2010). Community Structure in Time-Dependent, Multiscale, and Multiplex Networks. *Science, 328, 876-878*.

Newman, M. E. J. (2006). Finding community structure in networks using the eigenvectors of matrices. *Physical Review E, 74, 3, 036104-19.*

Poole, K. T. & Rosenthal, H. (1997). *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.

Poole, K. T. (2012). Voteview. http://voteview.com (retrieved 2012).

Tufte, E. R. (2001). *The Visual Display of Quantitative Information, 2nd edition*. Cheshire, CT: Graphics Press

Wasserman, S. & Faust, K. (1994). *Social Network Analysis*. Cambridge: Cambridge University Press.

Waugh, A. S., Pei, L., Fowler, J. H., Mucha, P. J., & Porter, M. A. (2009). Party Polarization in Congress: A Network Science Approach *SSRN* (<http://ssrn.com/abstract=1437055>).

White, H. C., Boorman, S. A., & Breiger R. L.. (1976). Social Structure From Multiple Networks I. *American Journal of Sociology, 81, 730-780.*

1. Why cluster instead of simply using the party affiliation labels? Because some nominal party members frequently vote with the other party. For example, western Senators from states split between liberal urbanites and conservative rural constituencies, (Oregon, Montana, or Main, for example) or socially conservative Southern Democrats are commonly in such positions. [↑](#footnote-ref-1)