Congressional Cosponsorship: A Network Analysis

Holden Huntzinger, UMSI

Introduction

Sponsoring bills is one way in which the 535 members of the US Congress communicate with each other and other political operatives, with their constituents and the broader public, and with donors. It is a way to signal ideology and intent.

Research Questions

In a context of increasing political polarity, are bill cosponsorship networks becoming more partisan? How do political and demographic factors at the member level relate to bipartisanship?

Data

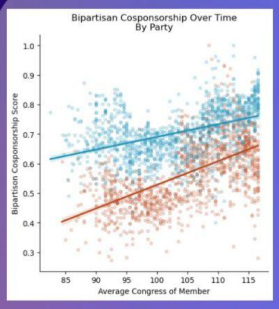
Congressional records are public records made available by the Library of Congress. Information was collected regarding:

- ★ ~325,000 bills sponsored or cosponsored between 1973 and 2021 by
- ★ ~2,500 members of Congress together representing
- ★ ~1.7 million cosponsorship relationships

Methods

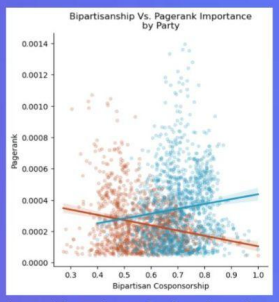
Connections between members were cast as a network in which nodes represent members and the edges between them are weighted according to the frequency of cosponsorship. This network was then analyzed by standard quantitative network measures of each node, e.g., variations of connectedness and centrality, as well as by bipartisanship score, a measure of how connected the member is outside of their own political party. This network data, aggregated at the member level, is then connected to other political and demographic information about each member; thus, these network and cosponsorship measures can be compared across political and demographic groups of members.

Results

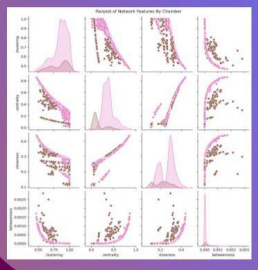


★ Despite growing political division, bipartisan cosponsorship is on the rise

★ Democrats engage in more bipartisan cosponsorship than Republicans



★ Bipartisan cosponsorship is negatively correlated with pagerank centrality for Republicans over the last 5 Congresses (2009 - 2021); in other words, more legislatively important Republicans engage in less bipartisan behavior; the opposite is true for Democrats.



★ Each pair of network measures (centrality, clustering, closeness, and betweenness) is clearly and tightly correlated for Representatives, but the Senate is much more complicated than the House in terms of the interrelationship of members.

Congressional Cosponsorship: A Network Analysis

Holden Huntzinger, 12/13/22

Introduction

Bill cosponsorship is one way in which members of Congress communicate with each other and various political operatives, with their constituents and the public more broadly, and with donors or prospective donors to their campaigns. Proposing or cosponsoring a proposed bill is a way of taking responsibility for the ideas therein; hence, cosponsorship signals political or ideological alignment, or at least collaboration, on the relevant issues in a bill. This report casts these connections between members as a network in which nodes represent members and the edges between them are weighted according to the frequency of cosponsorship. This network is then analyzed by standard quantitative network measures of each node, e.g., variations of connectedness and centrality, as well as by bipartisanship score, a new measure of how connected the member is outside of their own political party. This network data, aggregated at the member level, is then connected to other political and demographic information about each member; thus, these network and cosponsorship measures can be compared across political and demographic groups of members. In addition to this written report analyzing findings, relevant code and data, as well as instructions to reproduce this analysis, are available at github.com/hhuntz/congressional_networks.

Data Sources

Congressional records, including the text and relevant metadata of proposed and revised bills from both chambers, are a matter of public record and the Library of Congress helpfully provides well-curated access to this information. The files used in this analysis were downloaded in a conveniently collated format from ProPublica, a government-focused investigative journalism nonprofit.[1] This collection includes a total of just under 325,000 bills that were introduced in 20 Congressional sessions from 1973 through 2021. Collecting and processing this data resulted in an error rate of approximately 3 bills in 10,000; errors occurred when bills had no known sponsors/cosponsors and so were safely ignored.

As the focus here is on connections between members rather than between bills, this bill-level data was supplemented by information about members of Congress[2]

[1] ProPublica. “U.S. Congress: Bulk Data on Bills.” ProPublica Data Store. December 2022. https://www.propublica.org/datastore/dataset/congressional-data-bulk-legislation-bills

[2] The /Unitedstates Project. “congress-legislators.” November 19, 2022. https://github.com/unitedstates/congress-legislators

collated from government sources by the open source The /Unitedstates Project.[3] Data transformation and processing, including extraction of cosponsorship relationships from each bill and reconfiguration toward members as the record-level objects of interest, are discussed in the Methodology section below.

Related Work

Though this is the first research centered on an analysis of bipartisan cosponsorship, it is not a novel idea to consider members of Congress as a network. Political scientist James H. Fowler, currently at UC San Diego, has written or co-written several papers on the topic. Two of these focus on modularity as an extra-party measure of political polarization;[4,5] a third explores the application of small-world network theory to Congress.[6] A 2012 paper from Justin H Kirkland and Justin H Gross applies modern social network analysis to the US Congress and focuses on the social and communicative aspects of Congressional networks.[7]

Methodology

In order to represent and analyze congressional cosponsorships as network data, the data sources were significantly transformed in the following ways. First, sponsorship and cosponsorship information was extracted from the metadata for each individual bill. Relationships between sponsors were then reorganized as a list of individual edges between members of Congress; these original connections would be interpreted during network analysis such that the number of shared sponsorships between two members is the weight of the edge between those members, each represented as nodes. Approximately 1.7 million individual cosponsorship relationships were derived from about 325,000 bills. Finally, information on members of Congress was aggregated into a more convenient JSON structure. This processing was performed via a custom Python program labeled “compile_data.py” in the related Github repository.

---

[3] Mill, Eric. “A Modern Approach to Open Data.” Sunlight Foundation, August 20, 2013. https://sunlightfoundation.com/2013/08/20/a-modern-approach-to-open-data/

[4] Waugh, Andrew Scott, Liuyi, Pei, James H., Fowler, Peter J., Mucha, and Mason A., Porter. "Party Polarization in Congress: A Network Science Approach." (2009).

[5] Zhang, Yan, A. J. Friend, Amanda L. Traud, Mason A. Porter, James H. Fowler, and Peter J. Mucha. ‘Community Structure in Congressional Cosponsorship Networks’. Physica A: Statistical Mechanics and Its Applications 387, no. 7 (2008): 1705–12.

[6] Tam Cho, Wendy K., and James H. Fowler. “Legislative Success in a Small World: Social Network Analysis and the Dynamics of Congressional Legislation.” The Journal of Politics. University of Chicago Press, January 2010.

[7] Kirkland, Justin H., and Justin H. Gross. ‘Measurement and Theory in Legislative Networks: The Evolving Topology of Congressional Collaboration’. Social Networks 36 (2014): 97–109.

The resulting data were then organized and analyzed as a network using NetworkX software.[8] Network-specific attributes were collected for each node/member, including:

- Clustering coefficient, or the extent to which the node is part of a complete graph, i.e., all of its neighbors are interconnected; this is a measure of how tightly knit are the other members to which the member is connected.
- Degree centrality, or the fraction of total nodes to which the node is connected; this represents how broadly connected the member is via cosponsorship.
- Closeness centrality, or the reciprocal of the sum of the shortest path distance from the node to all other nodes normalized by the sum of minimum possible distances; this is one way to measure how directly connected the member is to all other members via both direct and higher degree (or 'friend of a friend…') connections.
- Betweenness centrality, or the sum of the fraction of shortest paths between other nodes that pass through the node; this represents how often the member serves as an intermediate connection between two other members that are not directly connected.
- Pagerank, a variation of eigenvector centrality originally used by Google to rank search results by counting an incoming link as a vote of support and weighting incoming links based on the score of the linking page; applied to the network of members of Congress, this is one rough proxy for measuring the 'legislative importance' of a member – members with more connections and whose connections have more connections are very central, and so important.

A domain-specific measure was calculated in addition to these more standard network analysis measures: the bipartisan cosponsorship score measures how often a member of Congress has sponsored bills alongside members outside of their own political party relative to how many total cosponsorship relationships they have. In the analysis below, this is generally a dependent variable as we measure how time, demographic attributes, and the network attributes described above relate to members' bipartisan cosponsorship scores. The one exception is a look at whether this measure of bipartisan collaboration correlates with pagerank scores and how political and demographic variables influence that relationship. A bipartisanship cosponsorship score is calculated for all historical members for which bill data is available, but the visualizations below contain information only from members of the Democratic and Republican political parties because there were far too few members from any other party to derive meaningful analysis for them.

---

[8] Aric A. Hagberg, Daniel A. Schult and Pieter J. Swart, "Exploring network structure, dynamics, and function using NetworkX", in Proceedings of the 7th Python in Science Conference (SciPy2008), Gäel Varoquaux, Travis Vaught, and Jarrod Millman (Eds), (Pasadena, CA USA), pp. 11–15, Aug 2008.

This collected data was collated into a single set of records listing each member alongside their particular attributes; this information is *MOC_features.csv* in the Github repository linked from the introduction of this paper. Relationships between these attributes were inspected using ordinary least squares linear regression. Limitations of this very simple regression technique are noted where appropriate, and only more significant trends are discussed in detail. The figures presented here were all generated using the Seaborn library[9] for Python.

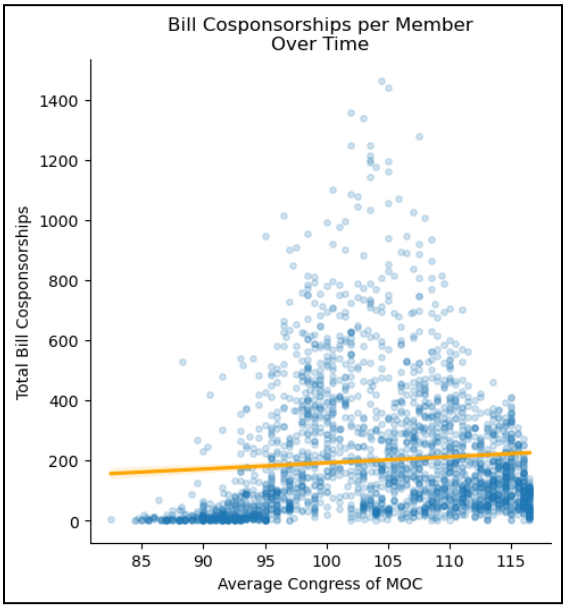
It should be noted here that, because this data is organized at the level of members of Congress, there is not a specific variable or attribute that represents time. Not only do many members serve multiple terms, but these terms are intentionally staggered so that only a portion of each chamber's membership turns over between each Congress. Additionally, members are sometimes elected to serve non-consecutive terms. Because of this complexity, two proxy variables are used to represent time in the analysis below: the average congressional session number of a member and the member's birthdate. These findings and visualizations were all measured based on both of these proxies, and in no case did using one or the other significantly impact the trends discussed below. The average congress of each member is used below for the purposes of visualizing trends over time.

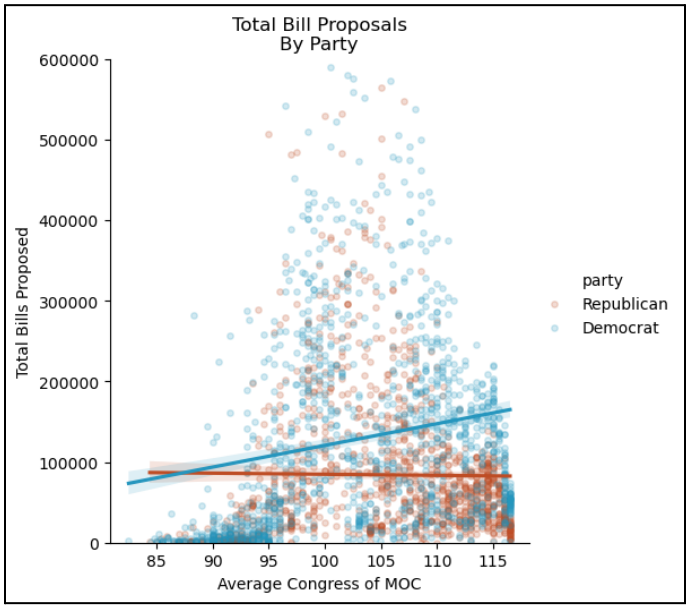
It is also important to mention that this research does not account for potential confounding variables that may influence the trends and correlations identified below. It is unclear from this work whether the age of a particular member affects any of these variables, and it is possible that a generalized age difference between members of political parties, for example, might affect these results. Similarly, about two-thirds of the historical female members of Congress belong to the Democratic party; this and similar underlying truths of this dataset are noted in the following sections where they are relevant.

[9] Waskom, Michael L. "Seaborn: Statistical Data Visualization." Journal of Open Source Software 6, no.60 (2021): 3021.

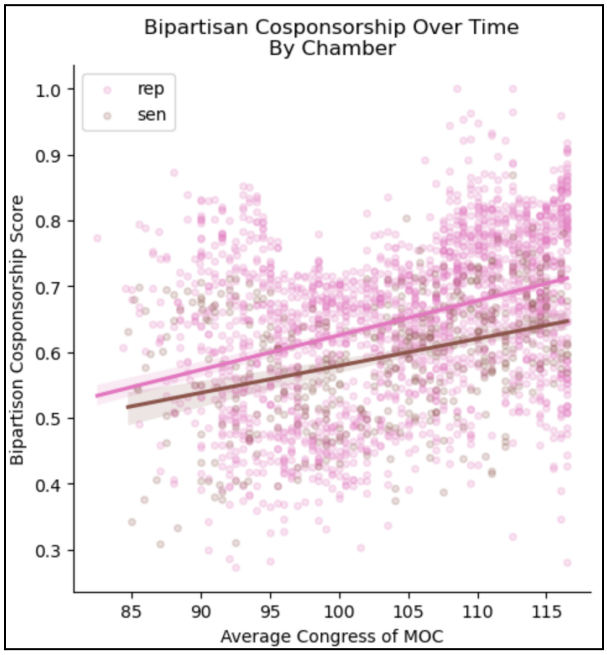
Findings and Interpretation

In order to provide some context for this analysis, the figures below are provided to demonstrate larger trends in bill sponsorship over time. Bill cosponsorship per member varies significantly over the last seventy years, but there are significantly fewer high outliers in the last five Congresses. The overall trend line is slightly positive, meaning cosponsorship generally is becoming more common. However, as the second figure demonstrates, this trend is not equal between the two primary US political parties: Democrats are cosponsoring more bills over time while Republicans are not.





For a slightly deeper look, the figure to the right shows changes in bipartisan cosponsorship scores over time – note that this score is relative to the number of bills each member has produced. It is a way to compare how 'bipartisan-ness,' in general, of the proposed bills to which a member has added their name. Though the linear regression plotted here does not reveal more nuanced trends, such as the extent to which bipartisan cosponsorship scores of the major parties diverge and converge over time, it does demonstrate an obvious

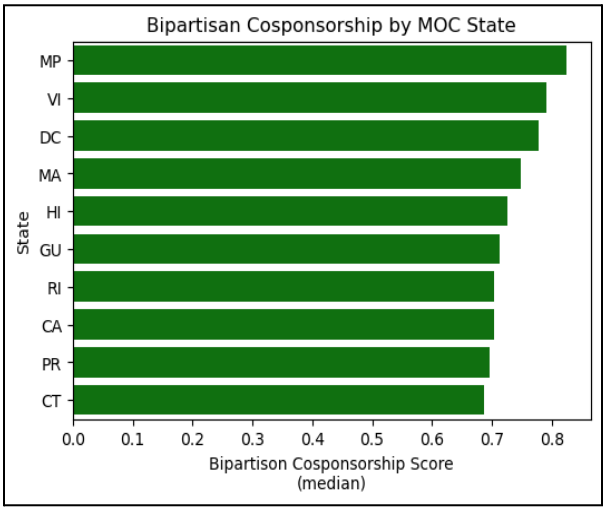


general trend: bipartisan cosponsorship is on the rise.

One potential explanation for this increase in bipartisan cosponsorship is signaling: sponsoring a bill, because it is relatively easy for members to do and because there is no necessary negative consequence to sponsoring an unsuccessful bill, is a useful way for members to communicate their ideology and intentions. This is one reason for which this dataset is not limited to bills that reached some level of success (i.e., only those on which members voted or only those signed by the President into law); even a proposal that is doomed to fail is done for a purpose. Though polls consistently find that Americans say they would prefer more bipartisan cooperation, there is also evidence that "preferences for bipartisan legislating do not outweigh party desires in the evaluation of public policy."[10] In other words, stated preferences for the ideal of bipartisanship can fall away when there is actual work to be done. Though communication may be the primary goal of some bill proposals, bills are also the fundamental way in which legislation happens – the (at least ostensible) goal of a bill proposal is that it becomes law – signaling is an important but insufficient explanation for variations in bipartisan cosponsorship.

Perhaps more convincing, then, is the argument that bipartisanship is a political necessity. Recent political science research from Curry and Lee finds that, despite major parties being "more cohesive than at any point in the modern era," coalitions that successfully enact important legislation are as bipartisan now as they were fifty years ago.[11] In light of this explanation, a rise in bipartisan cosponsorship seems to be an adaptation to pressures – successful bills tend to involve at least some amount of bipartisan support.

This suggests that bipartisanship is more common only because (and when) it is necessary. This idea is further supported b y the finding that members of Congress from places with the least political power tend to engage in more bipartisan cosponsorship. In the chart to the right, which shows the bipartisan cosponsorship scores for the highest-scoring ten geographical areas with representation in Congress. Though some larger states (MA and CA) appear, the list is dominated by the smallest states and by territories and the District of Columbia.

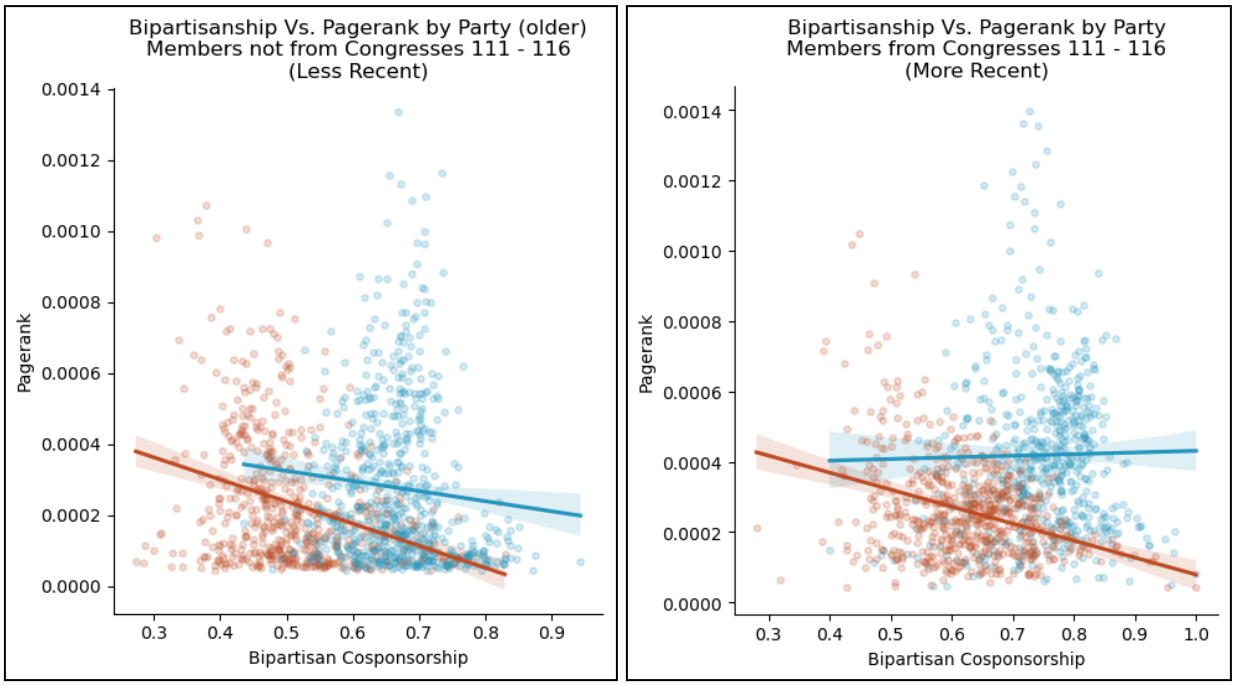


[10] Harbridge, Laurel, Neil Malhotra, and Brian F. Harrison. "Public Preferences for Bipartisanship in the Policymaking Process." *Legislative Studies Quarterly* 39, no. 3 (2014): 327–55.

[11] Curry, James M., and Frances E. Lee. "Non-Party Government: Bipartisan Lawmaking and Party Power in Congress." Perspectives on Politics 17, no. 1 (2019): 47–65.

Regardless of the reasons for which bipartisan cosponsorship is increasing over time, organizing these relationships as a network of interconnections can reveal more about which members are embracing this method of political expediency and ideological communication as well as how they are doing so.

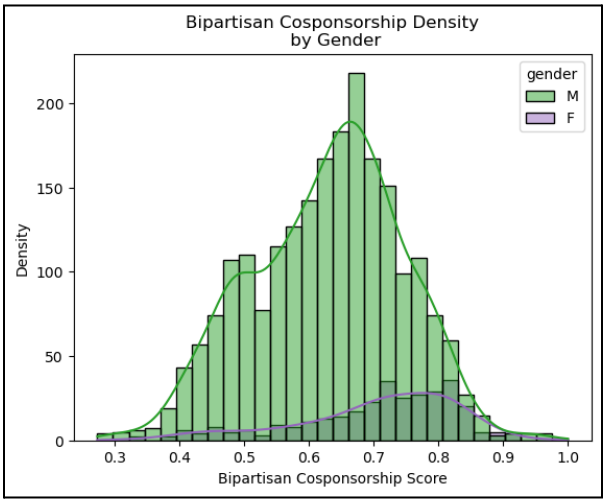
There is a large difference, for example, in the rate of bipartisan cosponsorship between members of the two major political parties. Again it is worth noting that a simple linear regression (represented by solid lines in the figure directly above) does not show nuances but only a general trend. That said, there is an obvious discrepancy here: Democrats have been much more likely to engage in bipartisan cosponsorship than have Republicans. The difference is consistent and clear all through the last 70 years, and it is particularly interesting in light of another significant correlation in this data, seen in the figures directly below.



Pagerank scores, a variant of eigenvector centrality, are used here as a proxy variable for the 'legislative importance' of a member. Historically, the Democrats with higher pagerank scores have tended to engage in slightly less bipartisan cosponsorship; this trend holds true for both parties, but is relatively weak, particularly for Democrats. What's more interesting here is the change in this relationship over time. A look at the last five Congresses (111 to 116, which occurred between 2009 and 2021) reveals an important shift. Where there was previously a slight negative correlation between legislative importance and bipartisan cosponsorship, there is now no such relationship – but only for Democrats. In more recent Congressional sessions, more central Democratic members have, in general, engaged in no less bipartisan bill

cosponsorship than have their less well connected comrades. Perhaps this is a byproduct of the increase in party cohesion discussed above. If Democrats are generally voting the party line, both more protected by and more beholden to the party, lower-ranking party members and those on the ideological fringes of the party have less need to reach across the aisle to find support. If this theory is correct, then it may be that the Republican party is significantly more fractured internally.

In a context of increased polarity of political opinion in the general public[12] to the extent that Americans now rank political polarization as one of the top issues facing the country.[13] A similar partisan difference occurs in these polls, as well: 33% of Americans who identified as Democrats said “Political extremism or polarization” is “among the most important [issues] facing the country,” but 10% fewer Republicans agreed.



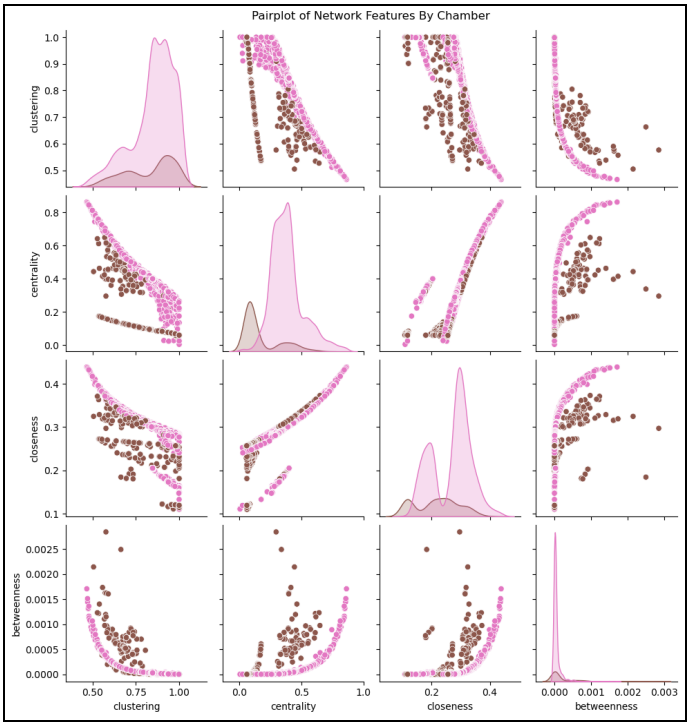
From this analysis, it seems that political factors such as party affiliation relate to cosponsorship network features in interesting new ways. It turns out that demographic information matters, too. Of the 318 female members of Congress included in this dataset, fewer than 30% (89 total) are Republicans. This skew confounds a simple analysis by gender, but it is still noteworthy that, in general, female members have both higher pagerank centrality scores and higher bipartisan cosponsorship scores.

One notable finding without a simple explanation is that the House of Representatives is much more simple than the Senate. As is evident from the chart below, Representatives (in pink here) line up in neat little rows – for each of the paired network measures, there is an obvious and consistent pattern. The Senate, however, is much messier – though similar correlations persist, they are much less straightforward. For Senators, betweenness does not cleanly allow prediction of closeness the way it would for Representatives. The Senate is designed to be the more slow and level-headed chamber – Senators are on average older and have more political experience – but it is not obvious why that would relate to less perfect relationships here.

---

[12] Dimock, Michael, and Richard Wike. “America Is Exceptional in Its Political Divide.” The Pew Charitable Trusts. The Pew Charitable Trusts, March 29, 2021. https://www.pewtrusts.org/en/trust/archive/winter-2021/america-is-exceptional-in-its-political-divide.

[13] Skelley, Geoffrey, and Holly Fuong. “3 In 10 Americans Named Political Polarization as a Top Issue Facing the Country.” FiveThirtyEight. ABC News Internet Ventures, June 14, 2022. https://fivethirtyeight.com/features/3-in-10-americans-named-political-polarization-as-a-top-issue-facing-the-country/.



Recommendations for Further Study

There is significant further research to be done with this data. To start, a closer analysis of bill- or congress-level data would allow more nuance to be seen in this existing data; computational limitations made that impossible for this work. Additionally, more network-like connections might be drawn between members of Congress, such as connections to donors or lobbying groups. Similarly, more demographic data would allow for broader analysis of trends between social groups within Congress.

Conclusion

Casting congressional cosponsorship relationships as network data allows for insights into the connections between individual members, but it also grants access to more general trends when members are aggregated by political and demographic information. Political factors, demographic attributes, and time are all useful in making meaning from the complex web of cosponsorship. Though differences between political parties seem to be particularly descriptive, especially in a context of higher-than-ever partisan polarization, this analysis also reveals ways in which intra-party dynamics may be changing unevenly. There is more information to be mined from this data, but this work can serve as an example of the ways in which fundamental network analysis can be applied to the political domain.