The Senate's First Year of the 115th and 117th Congresses

Matt Allen

12/5/2021

Description of Data

The data for this report comes from Voteview.com. Voteview is the continuation of more than 40 years of work by Keith Poole. Over the years he has curated data related to the United States Congress. This data has been made freely available (Boche et al. 2018).

This paper will focus on the first year of the 115th and 117th Congresses (2017 and 2021 respectively). Specifically it will use the first year of roll call voting in the Senate. There are a range of votes that can be cast. For this analysis, it will be limited to Yea and Nay votes.

The data is further filtered to Senators who have voted more than a threshold. The threshold is the average number of votes cast by all senators.

Methods

Analysis of vote data with non-metric multidimensional scaling is described in Everitt and Hothorn (2011). In that analysis they use the R package MASS (Venables and Ripley 2002). The data set that is used in the text as input to isoMDS function from the MASS package is from Romesburg (1984). It is a symmetric square matrix of counts between pairs of Congress members. The counts represent the number of times a pair of representatives voted differently. The main diagonal of the matrix contains zeros, because a member will never vote differently from themselves.

For the current analysis, there were two major challenges. The first was to transform the Voteview data into a form that could be used with the isoMDS function in the MASS package. The Voteview data was joined to have member and party information connected to roll call votes. At the time of this analysis, the 117th Congress was in the first year of its session. All data from beginning to current was used. For the 115th Congress, data from the first vote to the same number of votes as were currently available in the 117th were used. The data was further filtered based on Senators who voted more than the average number of times of all members of the Senate. Also, only Yea or Nay votes were used. An R function was created that compared each Senator's votes to all other Senators' votes across all the votes. For each pair of Senators if they voted differently two cells of the symmetric matrix were updated by 1.

Once a method to create the matrix was completed, the second challenge was off diagonal zeros. An off diagonal zero represents a pair of senators who whenever they were in the same roll call vote, voted the same way. This causes an error in the isoMDS function. There are two ways that this can be fixed. One is to add a small number like 1 in place of the zero. A second way is to remove one of the pairs of senators all together. The second way was

chosen, because the end goal of this analysis is to show differences between Senators, so having two senators who have the same voting record does not add any information related to differences. Off diagonal zeros were found in the 117th Congress data. For each of the Senators with the same voting record, one of the pair of the cloned Senators was removed and the dissimilarity matrix was rebuilt.

Finally, once the dissimilarity matrix was created, the matrix was used as input to the isoMDS function. The resulting distances obtained from non-metric multidimensional scaling were plotted to reveal patterns in voting records of Senators in the 115th and 117th Congresses.

Results

Non-metric multidimensional scaling was performed separately on data from the 115th and 117th Congresses, then Shepard diagrams were used to informally assess the quality of the multidimensional scaling (Everitt and Hothorn 2011). Figure 1 and 2 show Shepard diagrams from the two analyses. The Shepard diagram is a plot of the original dissimilarities against the distances obtained from multidimensional scaling. The points should lie along the bisecting line. These plots have some scatter outside of the line, but overall are good.
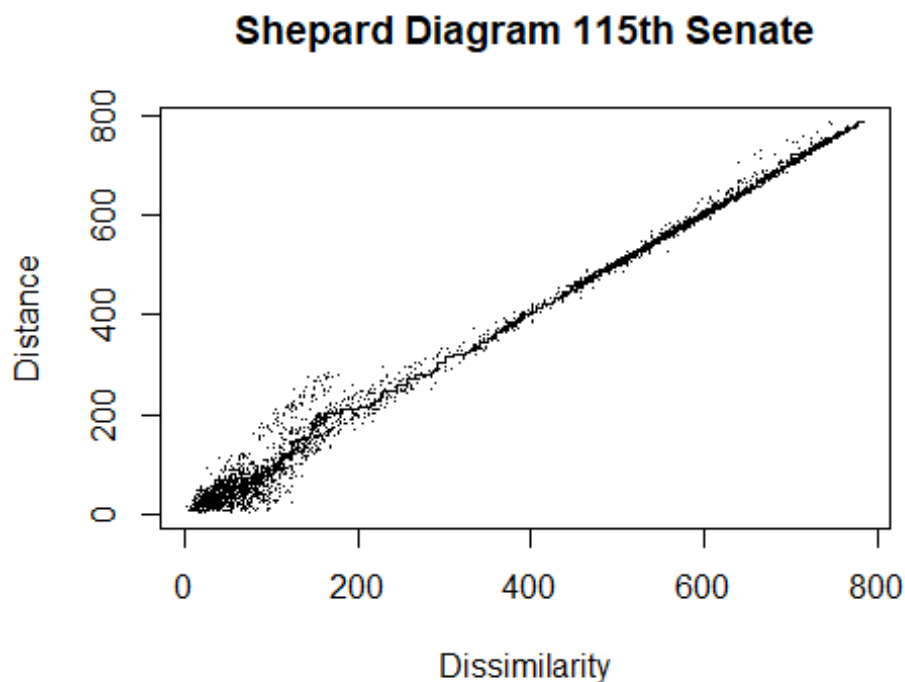
Figure 1. 115th Senate Distance vs Dissimilarity

Shepard Diagram 117th Senate
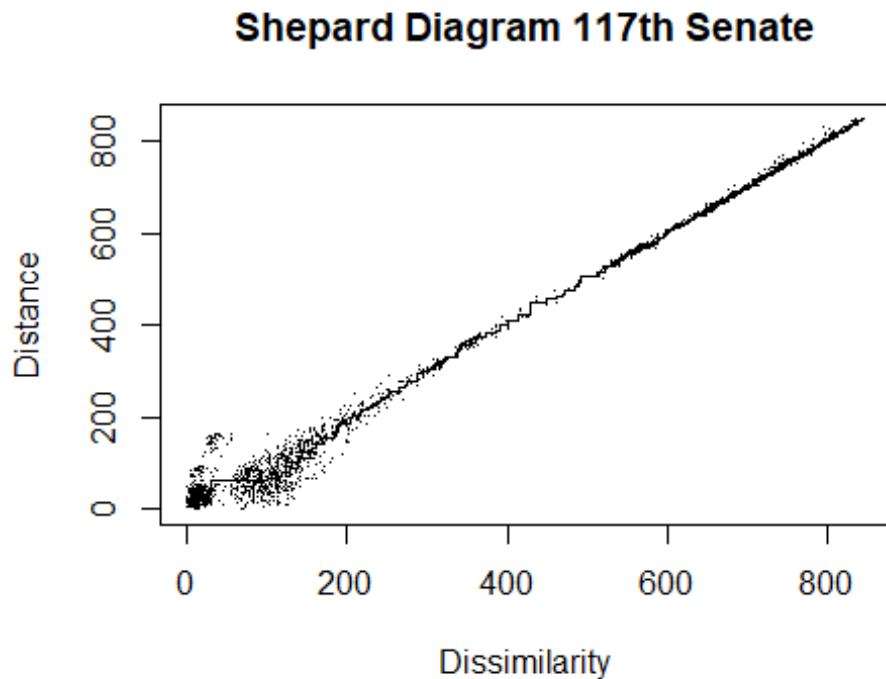
Figure 2. 117th Senate Distance vs Dissimilarity

Conclusions

The 115th and 117th Senate Sessions both had the party in control of the Senate also in control of the White House. In the 115th Congress, Republicans were in control and in the 117th Congress Democrats have control. Figure 3. and 4. show similar clustering, but on opposite sides of the political spectrum. When a party has control of the Senate, they tend to all vote together. This can be seen by the tight clustering of the party in power. The party not in power tends to be more dispersed. Interestingly, there are a few Senators from the party not in power that are willing to reach across the aisle. In the 115th Congress, Democratic Senators Heitkamp of North Dakota and Manchin of West Virginia tended to reach across the aisle. In the 117th Congress, Republican Senators Collins of Maine and Murkowski of Alaska tend to reach across the aisle. Three of the four Senators just mentioned also happen to be women.
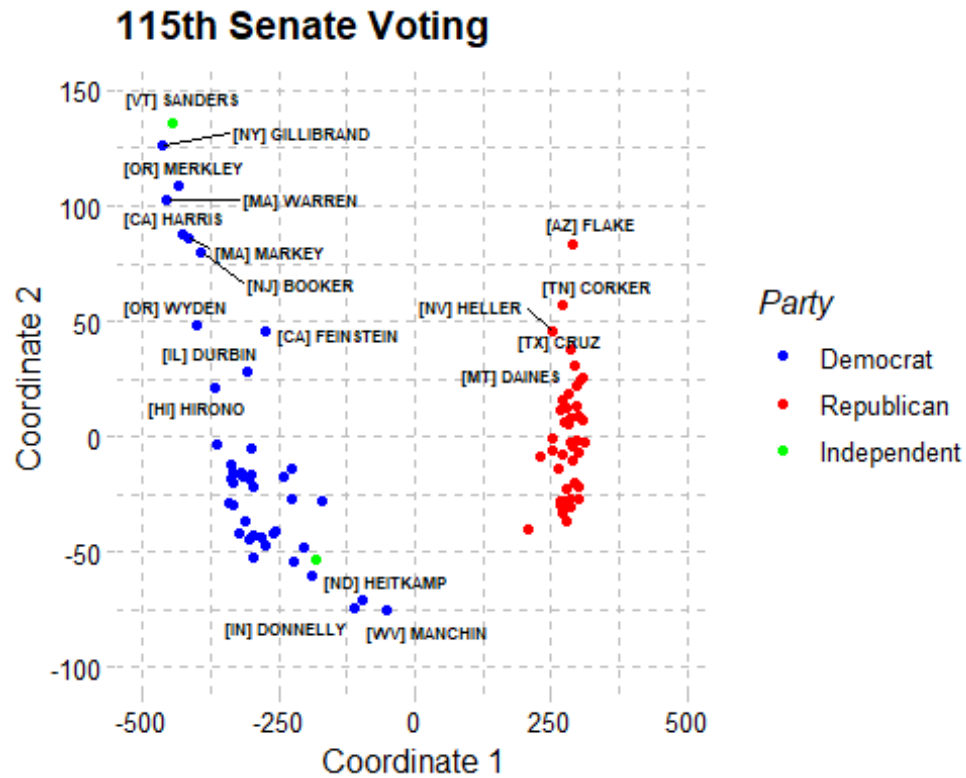
Figure 3. 115th Congress, Republicans Control Senate and White House
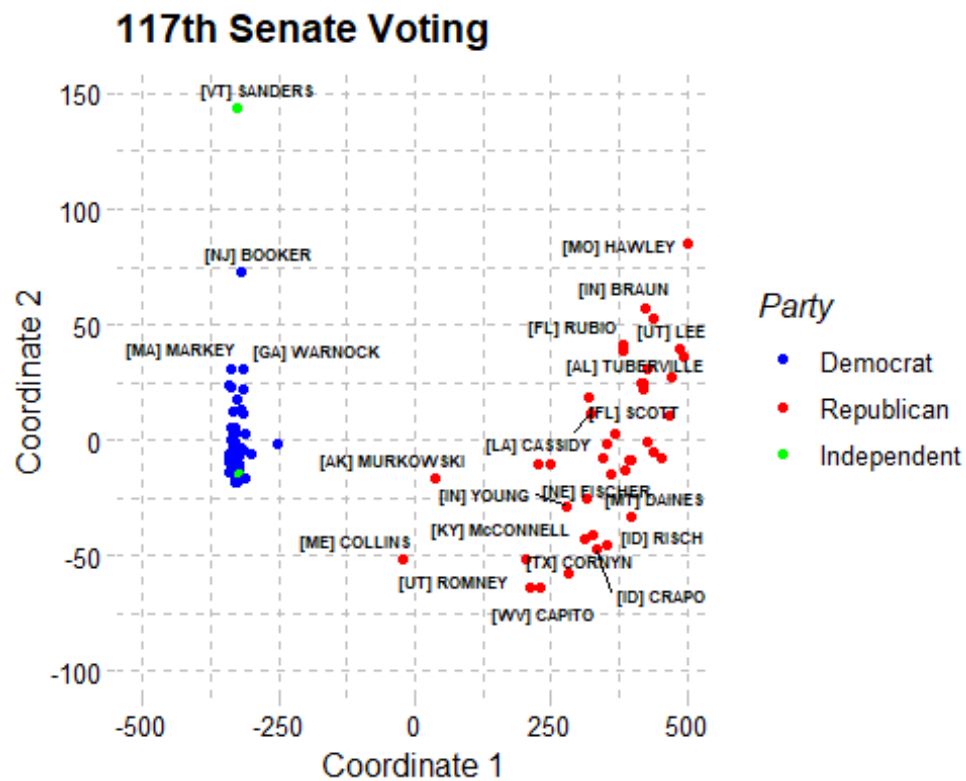

Figure 4. 117th Congress, Democrats Control Senate and White House

References

Boche, Adam, Jeffrey B. Lewis, Aaron Rudkin and Luke Sonnet. 2018. The new Voteview.com: preserving and continuing Keith Poole's infrastructure for scholars, students and observers of Congress. Public Choice, Springer, vol. 176(1), pages 17-32, July.

Everitt, B. and T. Hothorn. 2011. An introduction to applied multivariate analysis with R. Springer, New York.

Venables, W. N. and B. D. Ripley. 2002. Modern Applied Statistics with S. Fourth Edition. Springer, New York.

Romesburg, H.C. 1984. Cluster Analysis for Researchers. Lifetime Learning Publications, Belmont, CA.