

# Temp

Runjing Liu

December 5, 2016

Dimensionality reduction on roll call vote data in the House of Representatives in the 110th Congress. (a) Nonnegative matrix factorization on the  $448 \times 1707$  matrix (448 representatives, 1707 bills) of roll call votes into two matrices of dimensions  $448 \times 2$  and  $2 \times 1707$ . The rows of the  $448 \times 2$  matrices were plotted to visualize the distribution of representatives in a 2D space, and we clearly see division along party lines. (b) Principle component analysis on the roll call vote data. The eigenvalues and eigenvectors of the  $448 \times 448$  covariance matrix of representative voting data were computed, and each representative's voting profile was projected onto the space of the two eigenvectors with the two largest eigenvalues.

We chose to model caucus memberships because initial exploratory data analysis suggested that caucus memberships are related to a legislator's voting behavior. Figure 1 plots the number of shared caucuses between two representatives against the proportion of bills on which they voted the same way, and we see that the more caucuses two members share, the more likely they are to vote the same way.

Figure 2 shows the relationship between representatives within several caucuses in an undirected graphical model. We first used roll call vote data to infer the graph structure among the representatives in the entire House; we assumed pairwise interactions described via an Ising model in which each node denotes a binary variable of a representative voting either yes or no. The edges were inferred using neighborhood selection [?], and the graphs shown in figure 2 are subsets of this full graph corresponding to members of a caucus. The connectivity (measured by the fraction of total edges present) of the full graph with 448 representatives is 0.064, while the connectivity within the caucus subgraphs was much higher. This suggests that a representative is more likely to be influenced by a member of his or her caucus than another random representative in the House.

Therefore, this strongly motivates taking into account interactions among the representatives in Congress. In particular, this analysis suggests that caucus memberships may at least partly explain whether two representatives will vote in a similar fashion. Therefore, we proceed in this project by utilizing caucus membership data and connecting them to ideal points using a stochastic block model; specifically, we hope that caucus memberships will inform a latent community structure among the representatives, and exploiting these interactions, we obtain better predictions of ideal points and hence better predictions of roll call votes.

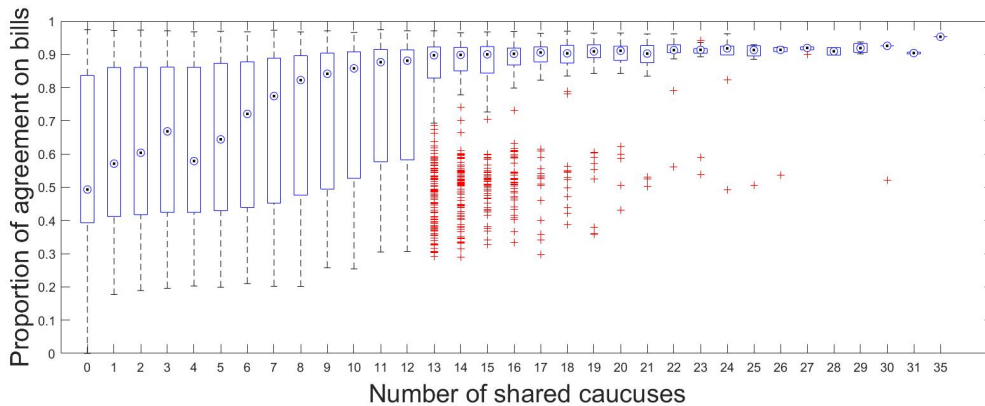


Figure 1: The distribution of agreement on bills as a function of the number of caucuses two representatives share. We see that the more caucuses people share, the more likely they are to agree on a bill.

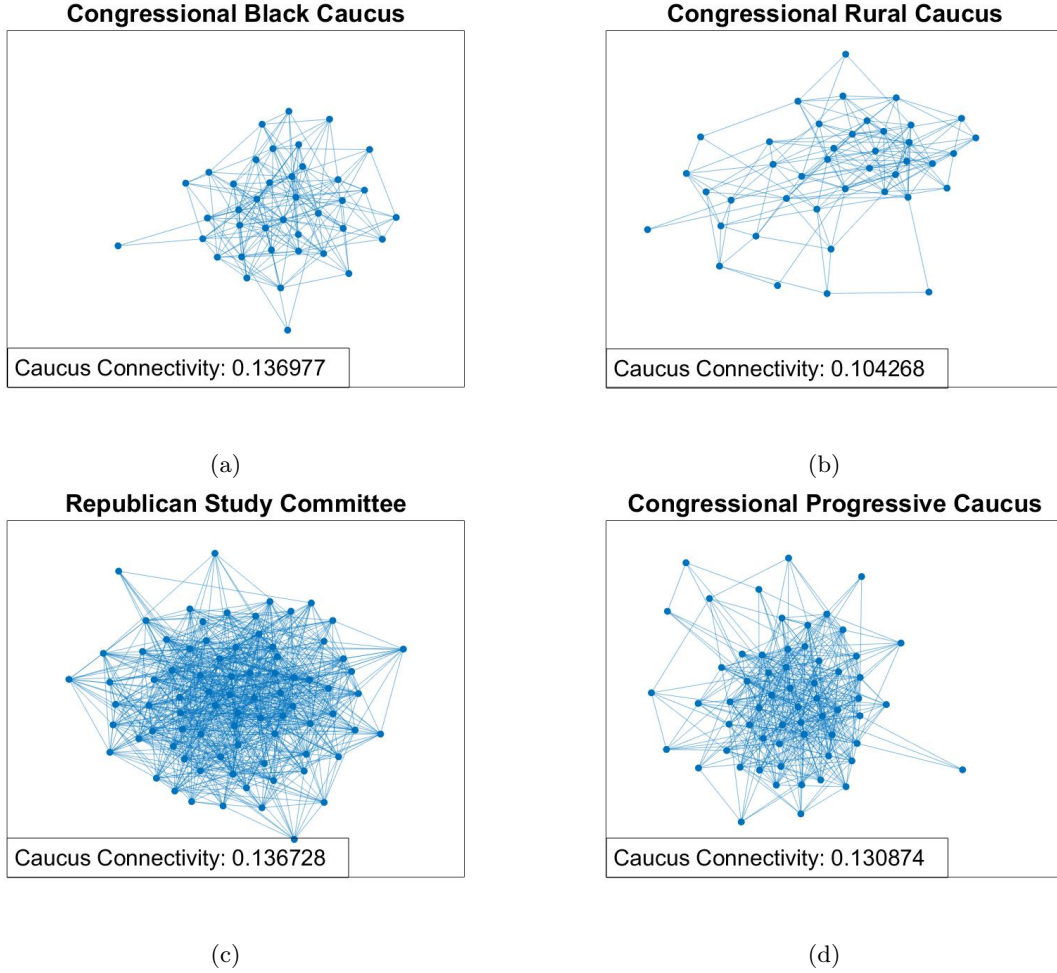


Figure 2: Neighborhood regression on roll call vote data was used to infer an undirected graphical model capturing relationships among members of the House of Representatives. Each node represents a random variable corresponding to a legislator voting “yea” or “nay” on a bill, and we assumed pairwise interactions using an Ising model. Shown here are subgraphs with representatives taken from a given caucus. The caucuses and their connectivities shown here are (a) the Congressional Black Caucus, connectivity 0.137; (b) the Congressional Rural Caucus, connectivity 0.104; (c) the Republican Study Committee, connectivity 0.136; and (d) the Congressional Progressive Caucus, connectivity 0.131. In each case, the connectivity within the caucuses was higher than the connectivity of the whole graph of the House (0.064).