Appendix for Mapping Influence

Contents

A Appendix Outline

In this online Appendix I provide details for the manuscript XXXX. I first provide a complete explanation of the simulation used to show the impact of networks among donors. Next I provide a step-by-step discussion of Neal's (2014) backboning method and demonstrate it with several examples. Finally I describe the ERGM used in the validation. This includes a new ERGM term developed to capture an important part of party networks.

B Simulation Details

In brief, I simulate six policy demanders that are trying to decide which of three candidates to nominate. Each group has an ideal point and each candidate has an ideal point. At the initial stage they select the candidate that is closest to their ideal point. Then one group is selected at random, they look to see if their neighbors (those they have a relationship with) have a majority preference that is counter to their own preference. If the majority of their neighbors are selecting another candidate then they consider updating their preference to that candidate. In the model of Klimek, Lambiotte and Thurner (2008) they would pick the majority candidate at this point. In my model they compare the relative loss of utility based on their ideological preferences to the number of groups that support the alternative candidate times a preference for conforming (which I vary across simulations). They only alternate candidates if they gain more by conforming than they lose by changing their ideal candidate.

The full details of the simulation are below:

Algorithm 1: Voter Simulation Process

1. Let there be G groups and C candidates. The groups are part of an undirected network that is defined by the adjacency matrix $\mathbb{N}$ which is $G \times G$. θ_g and θ_c are the positions of each group g or candidate c in a two dimensional ideological space. Set ω, the threshold, to a value greater than or equal to 0.

2. Draw θ_g and θ_c from $\mathrm{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}0.5 & 0\\ 0 & 0.5\end{bmatrix}\right)$

3. Compute a $G \times C$ matrix $\mathbb{D}$ with the euclidean distance between each candidate and group, so that $\mathbb{D}_{g,c} = \sqrt{(\theta_{g1} - \theta_{c1})^2 + (\theta_{g2} - \theta_{c2})^2}$

4. Compute a G length column vector $\mathbb{P}$ with the preference of each group where

$$\mathbb{P}_g = c \text{ such that } \mathbb{D}_{g,c} = \min_{c \subset C}(\mathbb{D}_{g,c})$$

   this is equivalent to the index of the candidate that minimizes the distance between the group and the candidate.

5. Compute a $G \times C$ matrix $\mathbb{D}^r$, the relative distance between the preferred candidate and alternatives by setting

$$\mathbb{D}^r_{g,c} = \frac{\mathbb{D}_{g,c}}{\min_{c \subset C}(\mathbb{D}_{g,c})}$$

6. Set counter = 1, while counter < 100

   (a) Let $\tilde{g}$ be a randomly selected group with neighbors $\tilde{n}$.

   (b) Let $\tilde{p}$ be the candidate most preferred by those in $\tilde{n}$ and let $\tilde{p}'$ be the number of neighbors that prefer that candidate. If there is a tie, then randomly assign $\tilde{p}$ from those candidates with the most support. If $\tilde{p}$ is already preferred by $\tilde{g}$ then set counter = counter + 1 and return to (a).

   (c) If $\tilde{p}' > \mathbb{D}^r_{\tilde{p},\tilde{g}}\omega$ then update $\tilde{g}$'s preference to $\tilde{p}$. Then set counter = counter + 1 and return to (a).

C Backboning Details

Here I describe in more detail the process of backboning used in this manuscript. I start first by explaining the technical procedure, before moving on to showing several randomly selected examples.

The general backboning algorithm employed here is a five step process and can be found detailed in Algorithm 1. In brief, I start with the data on group donations to candidates, model the amount donated from each group to each candidate, and then use that model to make a large number of simulated datasets. The modeling process, Step 2, requires additional detail. The data used here, donation data, is strictly 0 or greater requiring a GLM that fits count data. In addition the large number of zeros (where no donation took place) necessitates either a truncated or zero-inflated model. After testing multiple functional forms and data transformation I selected a hurdle model with the count process modeled using a Poisson distribution. In addition, the donations were transformed into \$100 increments (rounded to the nearest hundred) and the covariates were logged (after being shifted by 1/2 the smallest non-zero value). This was found to be robust across all datasets and also fit the data well.

Algorithm 1: Backboning Process

1. Let G be the number of donors, C the number of candidates, and δ_{cg} the donation from donor d to candidate c. Then $\sum_{c=1}^{C} \delta_{cg}$ is the total number of donations made by donor g and $\sum_{g=1}^{G} \delta_{cg}$ the total received by candidate c. Finally Δ is a $C \times G$ matrix made up of all D_{cg}.

2. Estimate a model of each candidate-donor relationship:

$$\delta_{cg} = \beta_1 \sum_{c=1}^{C} \delta_{cg} + \beta_2 \sum_{g=1}^{G} \delta_{cg} + \beta_3 \Big(\sum_{c=1}^{C} \delta_{cg} \Big) \Big(\sum_{g=1}^{G} \delta_{cg} \Big) + \epsilon$$

3. For each candidate-donor pair take a 1,000 draws from the estimate model to generate 1,000 matrices of estimate donations—$\hat{A}_i$

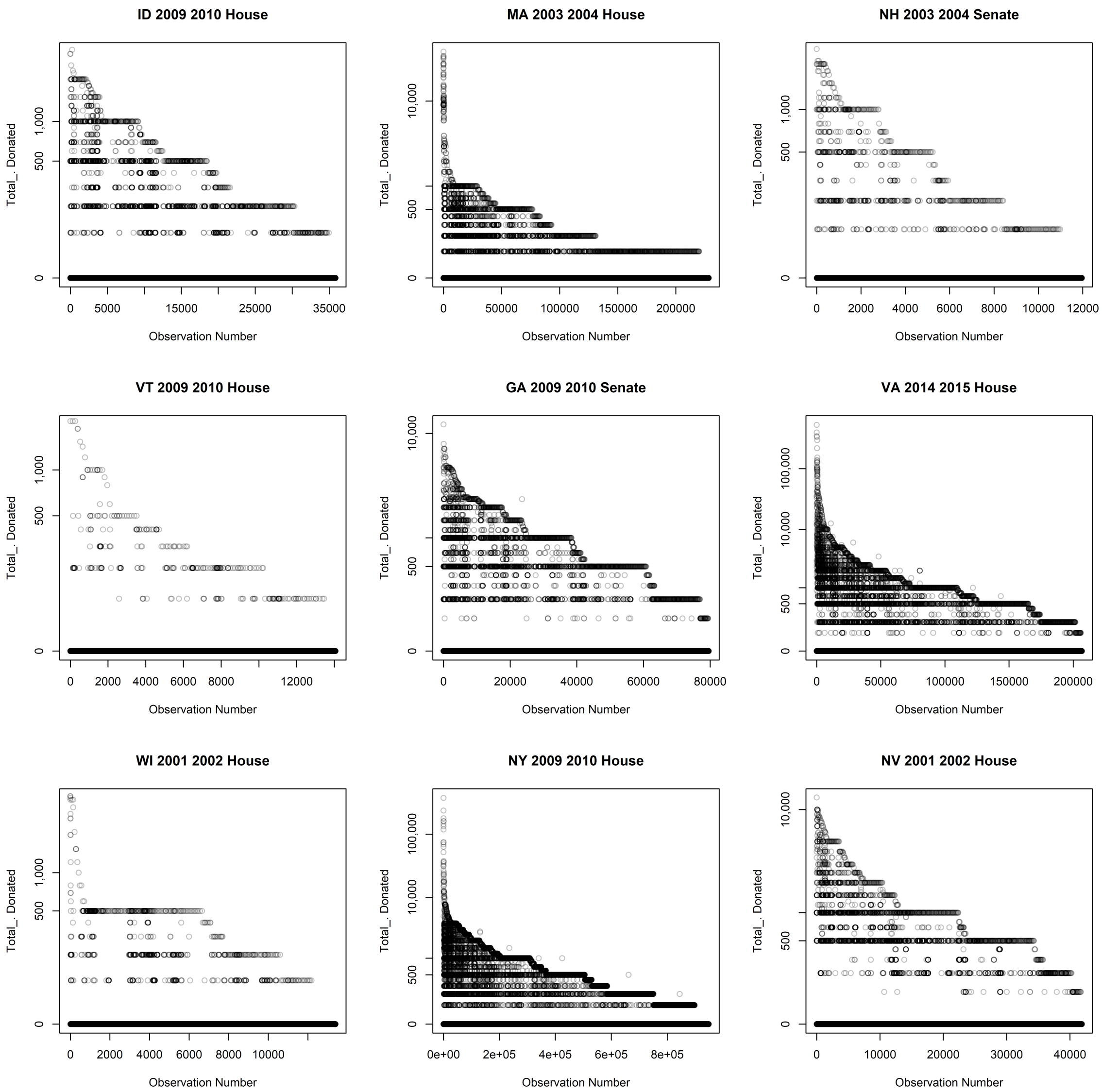
4. Take the crossproduct of all simulated matrices and of the real matrix. So $\hat{\Delta}_i \cdot \hat{\Delta}_i^t = \hat{A}_i$ and $\Delta_i \cdot \Delta_i^t = A_i$

5. The backbone B is 1 for each candidate-donor where the observed donation is greater than the α-centile of the simulated edge weights. Formally $\forall b_{cg} \in B$, $b_{g_1 g_2} = \mathcal{I}\Big(a_{g_1 g_2} > \text{quantile}_\alpha(\hat{a}_{g_1 g_2})\Big)$ where $a_{g_1 g_2} \in A$ and $\hat{a}_{g_1 g_2 i} \in \hat{A}_i$

To demonstrate this process I have selected 9 State-House-Full-Cycle at random. Figure 1 plots the real donations from a sample of states. Each observation is a group-candidate pair.

As discussed above, in the majority of cases this is a 0 as there was no donation made from the group to the candidate. To model the data, I used the data transforms discussed above along with a hurdle model with a Poisson count process.

Figure 1: Real Donations from a Sample of States

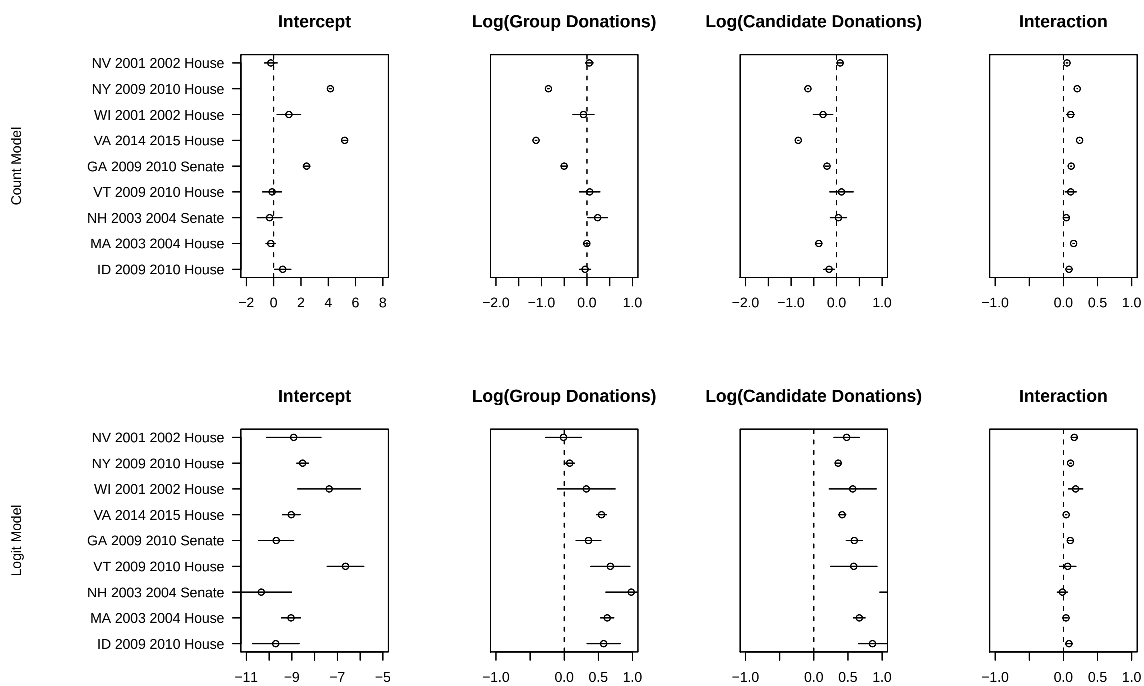


Note: Donation data from 9 sample states. Each observation is group-candidate pair, showing the amount donated from a group to a candidate. This shows how the majority of donations are zero and that there is a large amount of variation in the amount donated.

Figure 2 displays the estimated coefficients from the 9 models for each state along with their 95% confidence intervals. Although the coefficients vary across the units, there are broad similarities. The top row shows the coefficient from the count proportion of the model, the bottom row shows the estimated coefficient from the logit model predicting if a donation takes place.

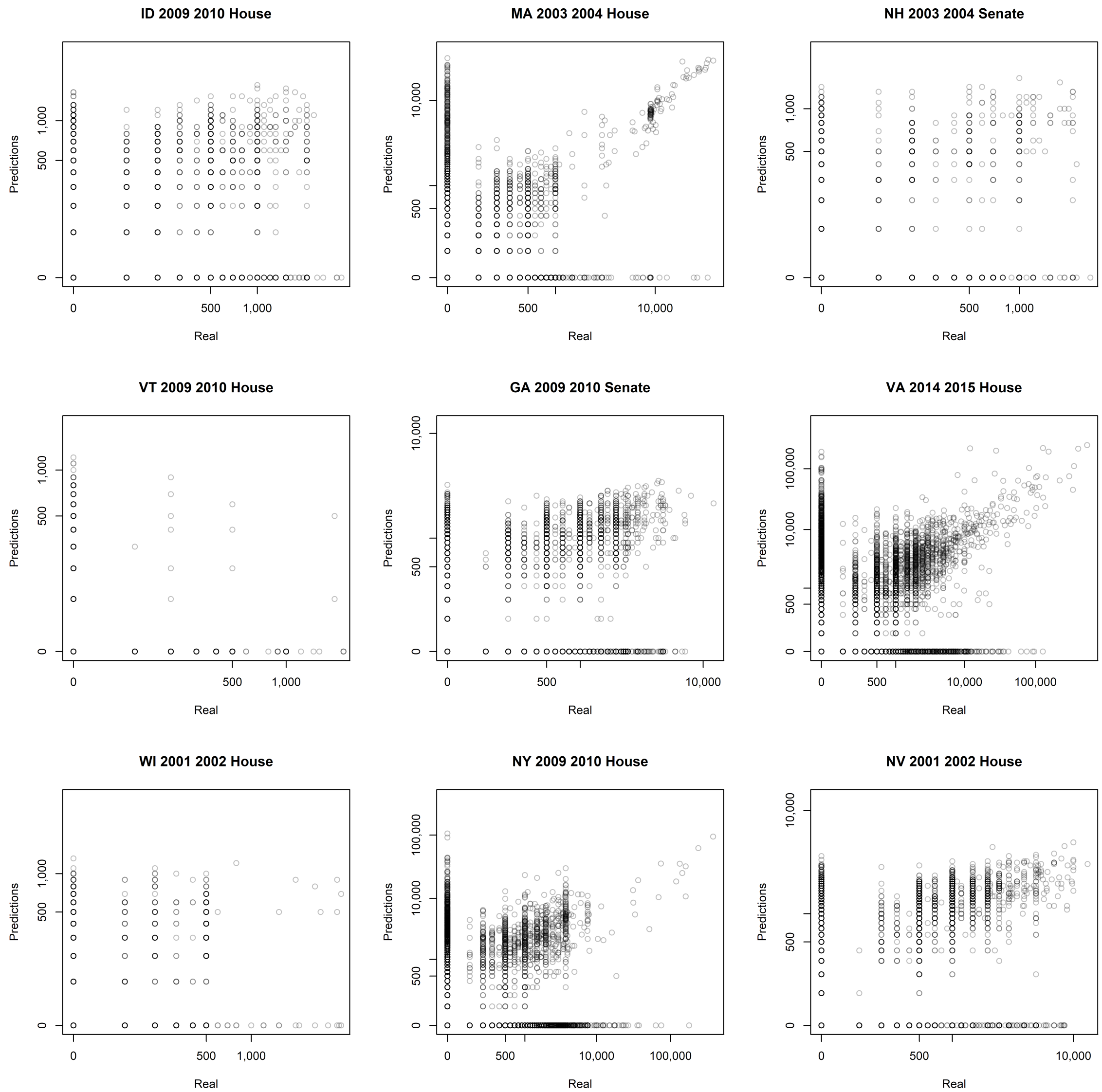
The intercepts in the logit model are all negative, reflecting the fact that any given group donating to any given candidate is relatively unlikely. The interaction terms in the count model are also all significant, indicating that the coefficients for the propensity of a group to donate, and the propensity of a candidate to receive, depend on each other. Figure 3 plots a set of predictions for each State-House-Full-Cycle against the real donations. Although many of the dots fall on the 45° line, there are also a substantial proportion where either no donation was predicted and one was made or where no donation was made and one was predicted.

Figure 2: Estimated Coefficients for Sample Models



Note: This shows the estimated coefficients across the different datasets. The top row are the coefficients in the count model, the bottom the coefficients in the logit model. The lines are 95% confidence intervals around the coefficient.

Figure 3: Predicted versus Real Donations from a Sample of States

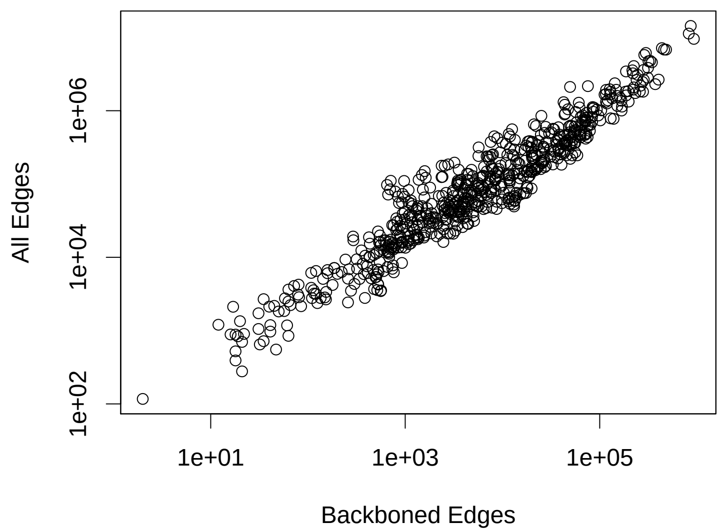


Note: Plots of the predicted versus real donation amounts.

The next step in the process is to draw 1,000 simulated values from each of these models. For each of these simulated sets of donations the crossproduct is taken. This creates an empirical distribution of the expected weight placed on a tie between two groups. Using this empirical distribution I include only ties between groups where the real weight is at or above the 97.5% percentile of the empirical distribution.

This significantly reduces the total number of edges in any given network. Figure 4 plots the total number of edges in each network versus the number of edges identified in the backbone methodology. Note that the scale of the Y axis, the number of initial edges, is orders of magnitude larger than the scale of the X axis, the number of edges identified by the backbone.

Figure 4: Number of Edges Pre and Post Backboning



Note: The number of edges without bacboning compared to the number of edges backboned. Each observation is a dataset. This shows the reduction of edges caused by the backboning process and that the reduction is proportional across the datasets.

D Exponential Random Graph Models

The main difficulty in modeling networks is that we cannot assume independence among units within the network. Although in many social science cases the assumption of independence among units is tentative, the entire purpose of network analysis is to explore the complex dependencies between units and so there cannot be broad assumptions of independences. Exponential Random Graph Models (ERGMs) accommodate complex dependencies by modeling the network as a single realization from a multidimensional distribution. It is possible to include explanatory variables that reflect both individual (nodal) characteristics as well as characteristics about the network as a whole (which will be the focus of the idea of cohesion). The ability to model network level characteristics is important and unique to ERGM methods (Cranmer et al. 2017).

ERGMs have a similar parametric framework as traditional regression models, with a set of terms that are hypothesized to explain the network formation or structure. If Y_m is the network, then

$$P(Y_m) = \frac{\exp(-\sum_{j=1}^{k} \Gamma_{jm}\theta_j)}{\sum_{m=1}^{M} \exp(-\sum_{j=1}^{k} \Gamma_{jm}\theta_j)}$$

Γ_{mj} is equivalent to an independent variable that measures some characteristic of the network. For example this can be an attribute of nodes that might give rise to edge formation, like if a node is a labor union or not. It can also be a network level characteristic that reflects the structure of the network as a whole, such as if nodes often form triangles (where three nodes are all connected to each other). θ_j is akin to a coefficient showing how important a term is. The denominator acts as a normalizing constant across the range of potential networks. The estimation of ERGMs is relatively complicated given the possible number of networks. For more detail on ERGMs and their estimation see Cranmer and Desmarais (2010).

I now briefly describe the nodal terms that I include to control for other reasons groups might be connected. After discussing the control variables, I discuss Maximally Weighted Edgewise Shared Partners (MWESP) and how I will use it to measure cohesion within a party network.

D.1 Nodal Terms

- **Group Type:** This is a factor variable using a modification of the group type from the NIMSP data. I recode this data so that there are at most 10 categories: Businesses, Candidates, Democratic Party, Republican Party, Third Party (rare) Ideology/Single Issue, Labor Unions, Lawyers and Lobbyists, and Other. The Business category includes multiple categories coded by NIMSP. The group data is included as a factor variables to account for some group types being more or less likely to have connections than other groups.

- **Group Homophily:** Because groups of the same type are likely to have shared interests they are also potentially more likely to connect with each other than with groups outside of their type. This phenomenon is known as homophily (McPherson, Smith-Lovin and Cook 2001). To account for this I included a differential homophily term that allows for the degree of homophily to be unique for each type. I do this as there are also differences across types. Although labor unions might be likely to be tied together, lawyers and lobbyists are a much more diverse set of groups and so are not necessarily likely to be tied together.

- **Logged Total Donations (lagged):** This is a measure of how much a group has donated in the previous electoral cycle. It accounts for the fact that groups that were very active in previous election might be more likely to have connections with groups in the current election.

- **Logged Absolute Difference in Total Donations (lagged):** Edge formation might not only be responsive to the amount of donations a group makes but also to how similar two groups donation patterns are. In particular groups that donate at similar levels might be more likely to be connected. This is accounted for by this term which is the log of the absolute difference in donations between two nodes.

- **Percentage of Donations to Democratic Candidates (lagged):** Calculated:

$$\frac{\sum \text{Dem Donations}}{\sum \text{Dem Donations} + \sum \text{Rep Donations} + \sum \text{Other Donations}}$$

  This term controls for difference in Democratic versus Republican leaning groups. In some states where one party dominates, nodes that participate in that party might be more likely to have ties. A positive value for this indicates that as a group donates more to Democratic candidates the more likely they are to have edges, negative values indicate the opposite.

- **Absolute Difference between Democratic Donation Percentage (lagged):** This is similar to the logged absolute difference in total donations but instead accounts for the similarity in partisan leaning of donations. A significant negative coefficient indicate that groups that donate at the same level tend to be tied together, so Democratic leaning groups are more liked to be tied to other Democratic leaning groups

D.2 Geometrically Edgewise Shared Partners and Extensions

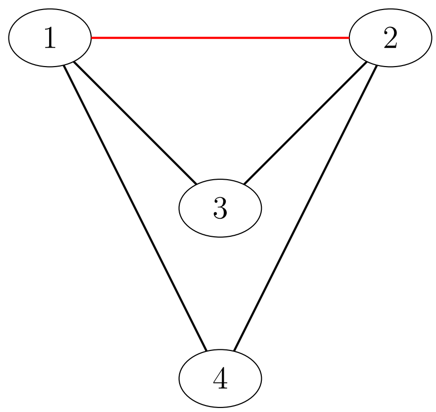
ERGM analysis has developed several set of terms that account for network dependencies. One particular term, GWESP (Geometrically Weighted Edgewise Shared Partners) account for the tendency of nodes to form if they have shared partners. When there is a single shared partner they create a triangle. Triangles are often expected in networks if friends of friends (edgewise shared partners) are likely to also be friends. It is possible though for two friends to have multiple shared partners (forming multiple triangles). Figure 5 shows an example of what happens when two edgewise shared partners exist. Here 1 and 2 are both connected to 3 and 4 and so we might expect this to increase the probability of an edge between 1 and 2 forming. Theoretically we might expect that the additional edgewise shared partner matters, but not as much as the first edgewise shared partner (that the adding more edgewise shared partner has a decreasing marginal effect). The extent that additional edgewise shared partners matter is modeled, which is where GWESP receives the name *geometrically weighed.*

The formula for GWESP is:

$$v(\mathbf{y}, \theta_t) = e^{\theta_t} \sum_{i=1}^{n-2} \left(1 - \left(1 - e^{-\theta_t}\right)^i\right) \mathrm{EP}_i(\mathbf{y})$$

θ_t controls the weighting and $\text{EP}_i(\mathbf{y})$ is a count of the number of instances of i edgewise shared partners. For the network in Figure 5 $\text{EP}_1(\mathbf{y}) = 4$ [1] and $\text{EP}_2(\mathbf{y}) = 1$. If $\theta_t = 0$ then only the first edgewise partner is counted and the formula becomes simply $\sum_{i=1}^{n-2} \text{EP}_i(\mathbf{y})$. In contrast at $\theta_t \to \infty$ then each additional edgewise partner counts as 1 and this is equivalent of adding up all the edgewise shared partners within the network.

Figure 5: Examples of Edgewise Shared Partners



Note: Shows an example of edgewise shared partners between 1 and 2. 1 and 2 are connected and also both connected to 3 and 4 so they have 2 edgewise shared partners. If 1 and 2 were not connected then they would have 2 dyad shared partners.

GWESP accounts for a common feature in social networks, that friends of friends are also likely to be friends. This might also apply to donor networks, where a donor that works with another donor is also likely to work with the donors associated with that donor. For this to be the case though it is necessary for the all the donors in that set to have a shared idea of what they are working for. This is exactly that idea of cohesion that I am interested in here, where groups within the network generally cohere in how they see the team they are trying to get elected. A large coefficient on a GWESP term indicates that this sort of friend-of-friend relationship is very important in explaining how the network formed, and so that cohesion likely exists within the network. This is why I use an extension of GWESP to capture cohesion within the party network.

[1] Note that 1-4, 1-3, 2-3, and 2-4 all have a single edgewise shared partner

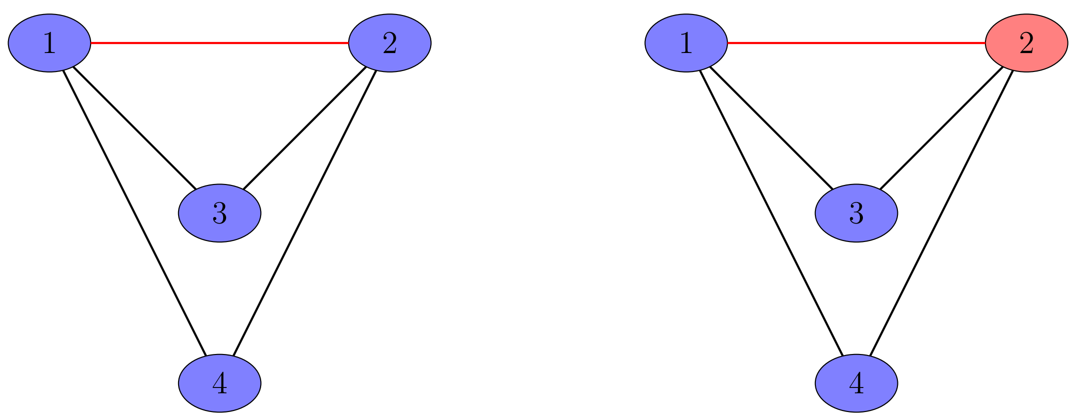
D.3 Maximally Weighted Edgewise Shared Partner

The GWESP term is limited in that it assumes that edgewise shared partners across all nodes within a network are equivalent, and all that matters is the number of edgewise shared partners. The networks examined here though contain nodes that are not necessarily similar. Take for example four groups, start by assuming that all four donate only to Democratic candidates, this is captured in Figure 6a. In this case we would expect that the shared partners between 1 and 2 would increase the likelihood of a connection between them. Now assume that node 2 is replaced with a group that donates only to Democrats (Figure 6b). The probability of an edge forming between 1 and 2 given a similar configuration of edgewise shared partners seems unlikely to be the same as in the previous example. Put differently it is unlikely that a friend of a friend is as meaningful in this context.

Figure 6: MWESP Examples

(a) All Democratic Donors

(b) One Republican Donor



Note: Figures show two examples of nodes with different MWESP value. The set of nodes on the left all donate to Democratic candidates and so the MWESP value is the same as the GWESP value. The set of nodes on the right has one that donates solely to Republicans and so has a MWESP value of 0.

GWESP terms, as described above, assume uniformity across all nodes within the network and so cannot account for these differences in nodes. To account for the way that groups donate to parties I extend the GWESP term so that the similarity of partisan donation strategies can be included. I do this in the case where the weighting term (θ) is equal to 0 and so only the initial edgewise shared partner counts. It is possible to extend this to allow for other values of θ but for simplicity I start here.

To extend GWESP start by redefining $(sp)_{lj}$ the shared partner statistics between two nodes (l and j). Normally this counts the number of shared partners between two nodes and so is equal to an integer value greater than or equal to 0 (Goodreau 2007). This can be modified to a maximally weighted shared partners where

$$\text{mwsp}_{lj} = \max_{k \in V} \left[((y_{lj})(y_{lk}) \cdot \min(n_j, n_l, n_k)) \right]$$

where n_j, n_i, and n_k are the value of nodes and the maximum is taken across all nodes in the network. In the networks presented here the values of node range from 0 to 1 (no donations to a party and all donations to a party).

The formula for MWESP is:

$$r(\mathbf{y}) = \sum_{w=0}^{\mathbb{W}} w \text{EP}'_i(\mathbf{y})$$

. Where EP'_w is similar to $\text{EP}_i(Y)$ in the GWESP equation but now counts the instances of maximally weighted edgewise shared partners with weight w instead of the instances with i edgewise shared partners. This is summed across the set $\mathbb{W}$ which contains all the *mwsp* values observed.

D.4 Estimation

As discussed above, the estimation of ERGMs can be tricky. Unlike traditional regression models, misspecification can lead to degeneracy where the underlying estimation schema tends towards networks that are either compete or empty (Schweinberger 2011; Hunter, Krivitsky and Schweinberger 2012). Estimating ERGMs across a range of networks requires specification that is robust across these networks. In addition, the large size of some of the networks can make traditional MC-MLE process computationally infeasible.

The latter case is more easily accommodated. In cases where there are more than 1,000 nodes I use MPLE parametric bootstrap as described in Schmid and Desmarais (2017). This is only necessary in 38 of the networks and in these cases I take 100 draws in order to estimate the distribution of the coefficients of interest.

In addition a subset of estimations did not converge if the percentage of Democratic donations were included either as the difference in the percentage of donations or the summation of the percentage of donations. These networks were estimated with those variables dropped. The problem from convergence was likely a result of their being collinearity between those variables and the MWESP variables. After estimation, models with an absolute value of MWESP greater than 15 were checked for additional convergence issues and were often re-estimated after dropping the other party variables.

References

Cranmer, Skyler J and Bruce A Desmarais. 2010. "Inferential network analysis with exponential random graph models." *Political Analysis* 19(1):66–86.

Cranmer, Skyler J, Philip Leifeld, Scott D McClurg and Meredith Rolfe. 2017. "Navigating the range of statistical tools for inferential network analysis." *American Journal of Political Science* 61(1):237–251.

Goodreau, Steven M. 2007. "Advances in exponential random graph (p*) models applied to a large social network." *Social networks* 29(2):231–248.

Hunter, David R, Pavel N Krivitsky and Michael Schweinberger. 2012. "Computational statistical methods for social network models." *Journal of Computational and Graphical Statistics* 21(4):856–882.

Klimek, Peter, Renaud Lambiotte and Stefan Thurner. 2008. "Opinion formation in laggard societies." *EPL (Europhysics Letters)* 82(2):28008.

McPherson, Miller, Lynn Smith-Lovin and James M Cook. 2001. "Birds of a feather: Homophily in social networks." *Annual review of sociology* 27(1):415–444.

Neal, Zachary. 2014. "The Backbone of Bipartite Projections: Inferring Relationships from Co-Authorship, Co-Sponsorship, Co-Attendance and other Co-Behaviors." *Social Networks* 39:84–97.

Schmid, Christian S and Bruce A Desmarais. 2017. "Exponential Random Graph Models with Big Networks: Maximum Pseudolikelihood Estimation and the Parametric Bootstrap." *arXiv preprint arXiv:1708.02598* .

Schweinberger, Michael. 2011. "Instability, sensitivity, and degeneracy of discrete exponential families." *Journal of the American Statistical Association* 106(496):1361–1370.