Advice-Seeking Network of a Pre-Professional Business Club at an Elite Private University in the U.S.

Julanna Zhang, Christine Zhou, Cathy Zhuang

## Abstract

In the business industry, there is a particularly strong emphasis on building professional connections. There are many business schools that vie for high rankings to attract the best connection networks. Therefore, the topic of this paper is the advice-seeking network of a pre-professional club in an elite private university in the United States. We conducted an exponential random graph model and coerced conversion with maximum pseudo-likelihood estimator, discovering that year level and position in the club are indicative of high in-degree centrality. Future areas of study include investigation of other factors that may impact in-degree centrality such as race, and job status.

## Introduction

Social and professional networks are a major aspect of the college experience. Many network studies center around social relationships built in universities (Brewer) and workplace connections (Bowler et. al.). In the business industry, there is a particularly strong emphasis on building professional connections. This paper explores which factors impact the advice-seeking network of a pre-professional business club at an elite private university in the U.S. The student organization mapped in this study is a selective, pre-professional business club. Members are accepted through a rigorous application process involving a written application and an hour-long interview with four current members. While this process acts as a form of social control on the network, all business clubs at the elite private university chosen require some combination of written and spoken application processes. The structure and rigor of such clubs imitate work organizations.

Previous studies on student networks have focused on student network impacts on major selection (Baker). Historically, students in community colleges have found completion of higher education to be difficult due to a lack of class planning guidance and low commitment to majors. However, deeper research that analyzes both major considerations and major choice found that groups of students may be more likely to be impacted by a network intervention program when they exhibit higher homophily.

Additional studies find that student networks grow closer through collaborative work (Kapucu et. al.). Through comparing before and after degree and closeness centrality measures, there is evidence to support that friendship and work-related ties increased as a result of collaborative learning. Since pre-professional student organizations often involve collaborative project work, we may see a closely tied network in our study.

We also decided to look into previous studies on workplace networks, as the similarities between pre-professional business clubs and companies could imply network-structural similarities. A study on leaders, followers and coworkers in a workplace environment focused on trust transference in the network as impacted by power dynamics (Bowler et. al.). The discovery that actors tend to trust those who focal members trust may be observed in most organizations involving level attributes.

These papers inform our chosen methods and attributes in our study of a business school club network in an elite university.

**We first hypothesize that one’s position in the club has a statistically significant impact on actor nominations. Our second hypothesis is that graduation year is a statistically significant attribute, namely that an increase in graduation year decreases the nominations one actor sends out while increasing the nominations that actor receives.**

## Data and Methods

Responses were collected from a survey created via Google Forms and sent in the business club’s group chat on Slack by an executive member. Members were given the option to participate and showed consent by completing the survey. For confidentiality reasons, each student was assigned a number. Participants had five days to respond to the survey, with no incentive to participate. The survey asked for students’ names, graduation year, whether they are a BBA major, the highest position they have held in the club, and whether they live in on-campus or off-campus housing. We did not ask for age, as it is heavily correlated with year. Race was also considered, as multiple articles have shown racial homophily to be present (Wimmer); however, given that this club is not diverse and mainly consists of one race, we omitted measuring race. Members were then asked, “Who do you actively seek advice from (it can be any advice, from life advice to school advice)? Choose up to 10 people.” They were given a list of every member in the club. Additionally, some members were directly requested to answer by executive members. We limited responses to up to 10 people to control for strength of ties and to prevent members from nominating someone if they had rarely sought advice from them.

The responses to the survey were then transformed into matrices and vertex attributes. The vertex attributes were year, BBA, level, and housing. Year was constructed as an ordinal variable, with 1 representing freshmen, 2 representing sophomores, 3 representing juniors, and 4 representing seniors. BBA was constructed as a binary variable, with 0 representing a non-BBA major and 1 representing a BBA major. Level was constructed as an ordinal variable, with 1 representing an analyst, 2 representing a team lead, 3 representing a director, and 4 representing an executive member. Housing was constructed as a binary variable, with 0 representing on-campus housing and 1 representing off-campus housing. Nominations from individuals were put into a matrix, where 1 represented a nomination and 0 represented no nomination. Members that did not respond to the survey were not included in the matrix or network; we decided to omit members that did not respond because they would not have had the chance to send out ties by nominating people.

All data analysis was conducted using RStudio version 2022.07.1+554.

**Actor Centrality and Attributes**

First, a network was created using the statnet package. Each actor’s in-degree, out-degree, betweenness, closeness, and eigenvector centralities were calculated. Correlations among these centrality measures were also calculated. Both individual and group indegree were of interest as well and were calculated, and the total number of unique incoming connections of the top four members was found. We used the k-means method as opposed to the edge-removal method for calculating group in-degree because our network is directed, and edge-removal method assumes that the network is undirected, so edge-removal would not fit the model as well. We also ran a linear regression to predict the in-degree centrality using the members’ attributes: year, BBA, level, and housing. We also looked for cliques, with groups of three, four, and five.

**Network Structure**

We calculated the mutuality, density, centralization, and transitivity of the network. We then compared these statistics in the actual network with those in randomized networks. We also used “brokerage” and “level” to calculate a brokerage index (i.e., the total number of times an actor serves as a broker). We found which three actors had the highest brokerage power, and also found how much the brokerage index correlated with the betweenness centrality. We created a structural equivalence matrix based on euclidean distance, and created a hierarchical clustering of actors. We then divided the actors into three blocks based on structural equivalence and created the density matrix of the three blocks.

**ERGM**

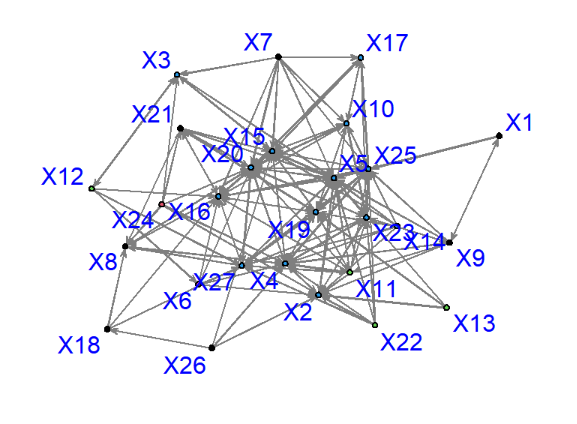
We ran two models using ERGM. The first model included covariate effects and homophily. We accounted for sender and receiver effects of class year and level. We also included effects for the BBA major and housing situation. Our second model consisted of everything in the first model with the addition of more endogenous network formation properties, such as mutuality, preferential attachment, differential sociality, transitivity, and two-path. MPLE (maximum pseudo-likelihood estimator) was forced on this model since the model did not converge otherwise. Finally, goodness of fit was calculated for the model.

## Results

Our network is a directed network, and according to Figure 1, our network shows that the club is very well-connected, and has a few central nodes. Since we omitted people that did not respond, all our nodes have at least one tie. Essentially, actors have at least one tie; either actors nominated or were nominated by someone.

**Figure 1. Plot of Advice-Seeking Network**

Each node represents an actor in the network. Nodes are colored by level (the highest position the actor has held in the club). Ties are represented by the gray arrows.

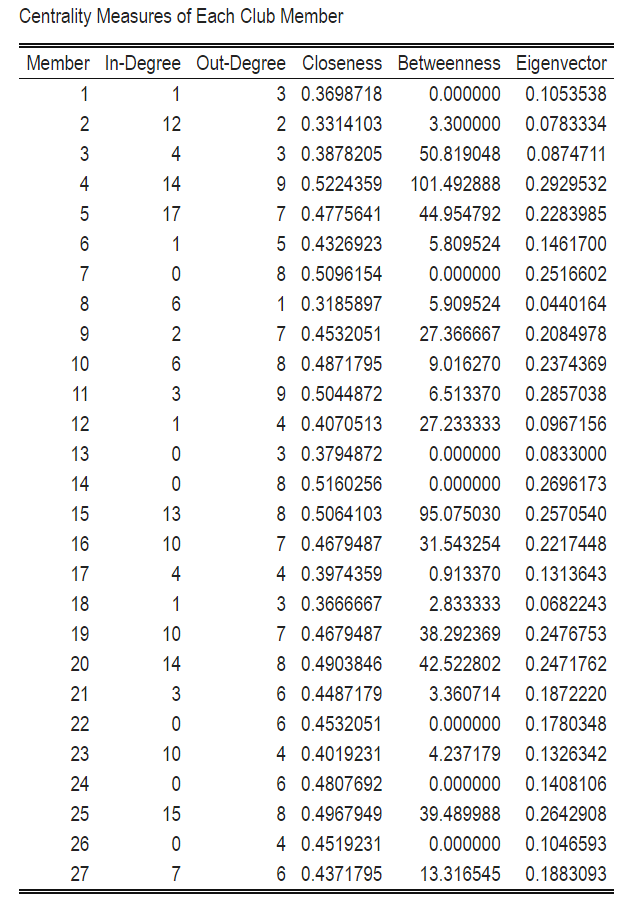
i

According to Figure 2, members 4, 5, 15, 20, and 25 have 13 or more in-degree connections, of which the highest is 17 connections for member 5. In-degree centrality measures the number of times an actor was nominated, so member 5 had 17 nominations. The actors with the most out-degree centrality are 4 and 11, each nominating 9 people which was the maximum that one could nominate. The members with the highest closeness centrality of more than 0.5 are 4, 7, 11, 14, 15, meaning that these members can reach others with the least effort and are best placed to influence the network. In practice, this means these members should be informed first if distributing by word of mouth. The individuals with the highest betweenness of more than 50 are 3, 4, 15. These individuals have the highest sum of shortest paths between two other people passing through. Individual 4 has the highest betweenness measure of 101. Eigenvector centrality is the highest for actors 4, 7, 11, 14, 15, and 25, and they have centralities of over 0.25. These individuals are connected to the most important alters within the network.

The top four members with the highest individual indegree connections were members 5, 25, 4, and 20, with 21 total unique incoming connections. In contrast, the top four members with the highest group indegree using the k-means method were 4, 5, 16, and 20, with 23 total unique incoming connections. This suggests that group indegree may be a better measure here, as the top four members had more unique incoming connections.

**Figure 2. Centrality Measures of Each Club Member**

Centrality measures of in-degree, out-degree, closeness, betweenness, and eigenvector are calculated for each actor.



According to Figure 3, the correlations among centrality measures are positive and high, with closeness centrality being slightly lower than the rest, but still positive. This means that a member with a high centrality in one measure of centrality will most likely have high centrality in other measures of centrality. In other words, a member with more ties will most likely have other ways that they are more connected to the network, such as being more connected to important alters.

**Figure 3. Correlations Among Centrality Measures**

How each centrality measure is correlated with other centrality measures.

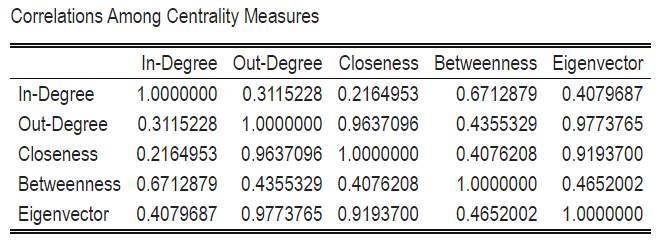
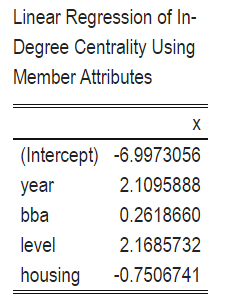


Figure 4 reveals that the coefficient for BBA is slightly higher than 0, meaning BBA majors are more likely to be asked questions. The positive coefficient for year means that the higher the individual’s year level, the more likely they are to be asked for advice and be central to the network. The positive coefficient for level means that the higher one’s status in the club, the more likely they are to be central. The negative coefficient for housing means that people who live off-campus (1) are less likely to be central to the network.

**Figure 4. Linear Regression of In-Degree Centrality Using Member Attributes**

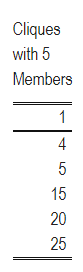
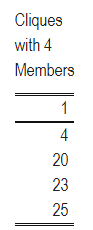
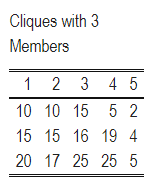
Coefficients of the equation based on college year level, BBA status, level in club, on or off-campus housing.



It is notable that actors 4, 20, and 25 are in at least three different cliques (Figure 5).

**Figure 5. Cliques**

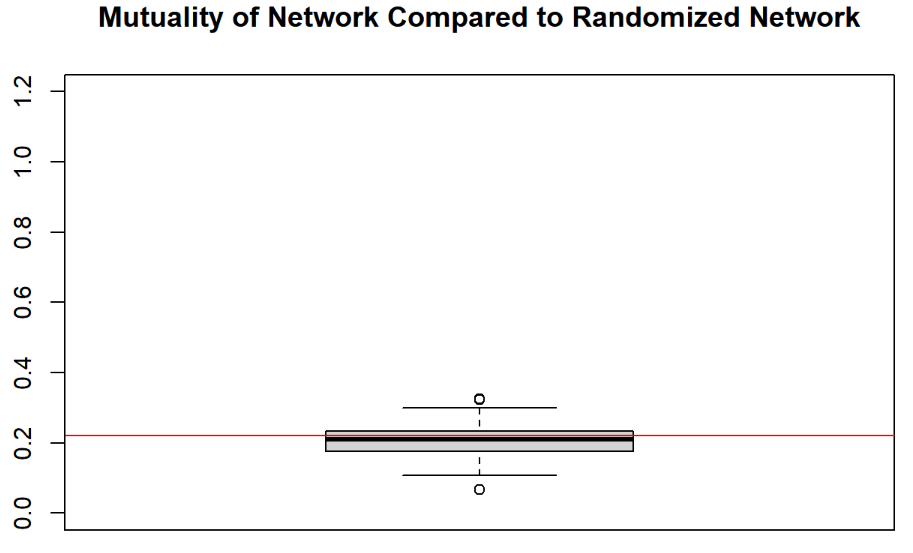
Members present in cliques of three or more members.

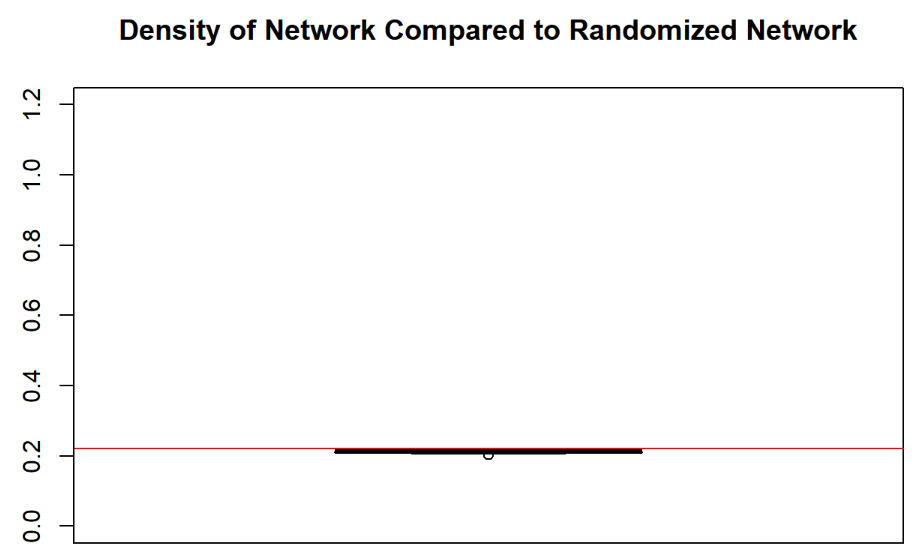


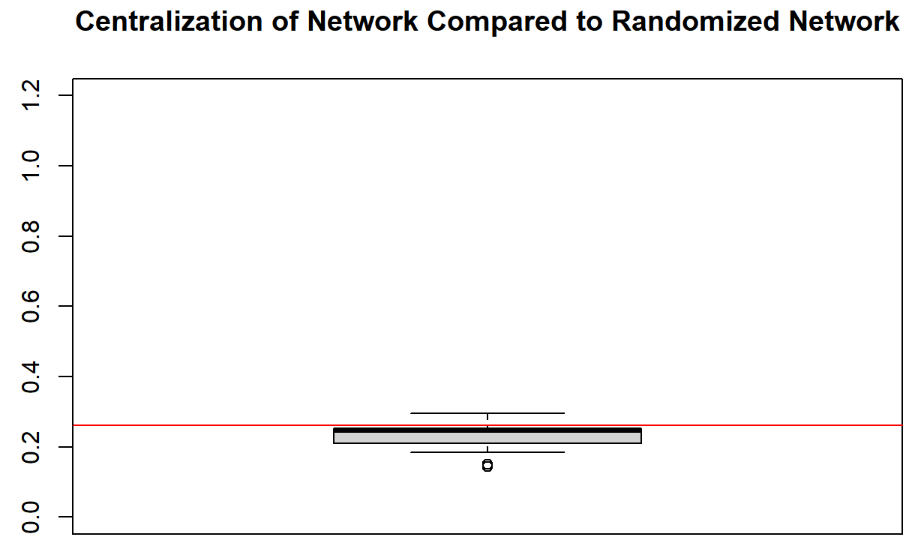
The mutuality of the network is 0.44, the density is 0.22, the centralization is 0.26, and the transitivity is 0.49. We see higher transitivity for the network than randomized networks (Figure 6). This means that in our network of members, the proportion of transitive triangles out of possible triangles is significantly higher than in randomized networks. The transitivity is therefore significantly different from what we would get purely by chance. The density, the proportion of all member ties out of all possible member ties, of our network appears to be about the same as that of randomized networks (Figure 6). The centralization is overall higher compared to randomized networks but is not higher than all the results in randomized networks (Figure 6). Overall, the extent to which ties are concentrated is higher but not statistically significant enough to rule out chance. The mutuality, the proportion of ties that are reciprocated, of our network also appears to be about the same as that of a randomized network (Figure 6).

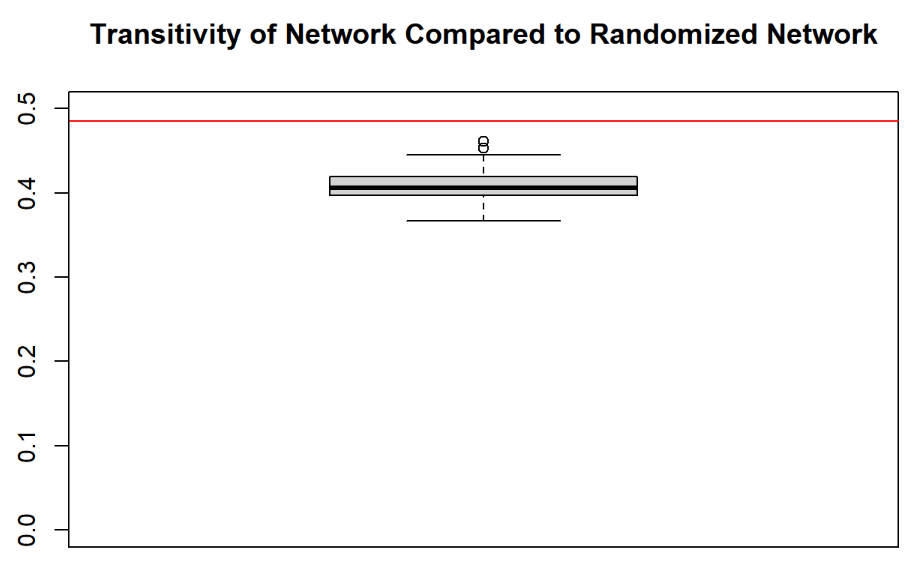
**Figure 6. Mutuality, Density, Centralization, and Transitivity of Network Compared to Randomized Network**

Red line indicates the measure of observed network, compared to a boxplot of randomized network.





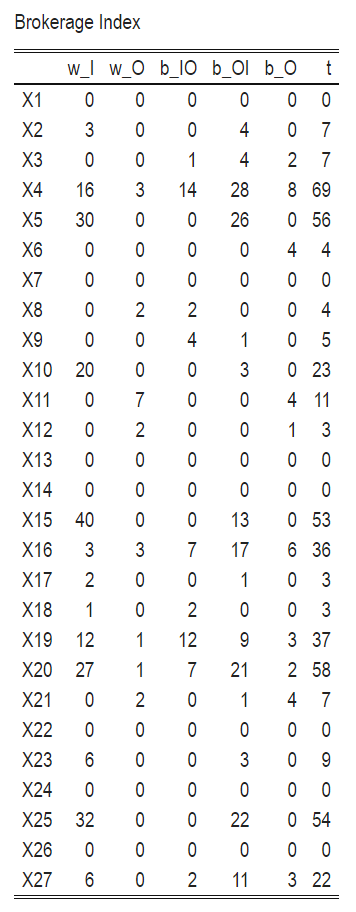




The actors that have the highest brokerage power are members 4, 20, 5, 25, and 15 (Figure 7). They served the most times as a broker total. Our brokerage index is positively correlated with betweenness at 0.82 overall (Figure 8). This means that the members that are more important to the flow of the network tend to also have more brokerage power.

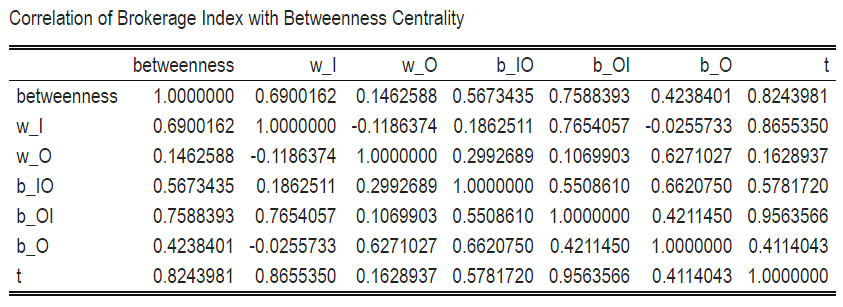
**Figure 7. Brokerage Index**

How each actor is measured by brokerage roles. “w\_I” is a coordinator role (all nodes belong to the same group). “w\_O” is a consultant role (broker belongs to a different group). “b\_IO” is a representative role (recipient belongs to a different group). “b\_OI” is a gatekeeper role (source belongs to a different group). “b\_O” is a liaison role (all nodes belong to different groups). “t” is how many times an actor served as a broker in all roles.



**Figure 8. Correlation of Brokerage Index with Betweenness Centrality**

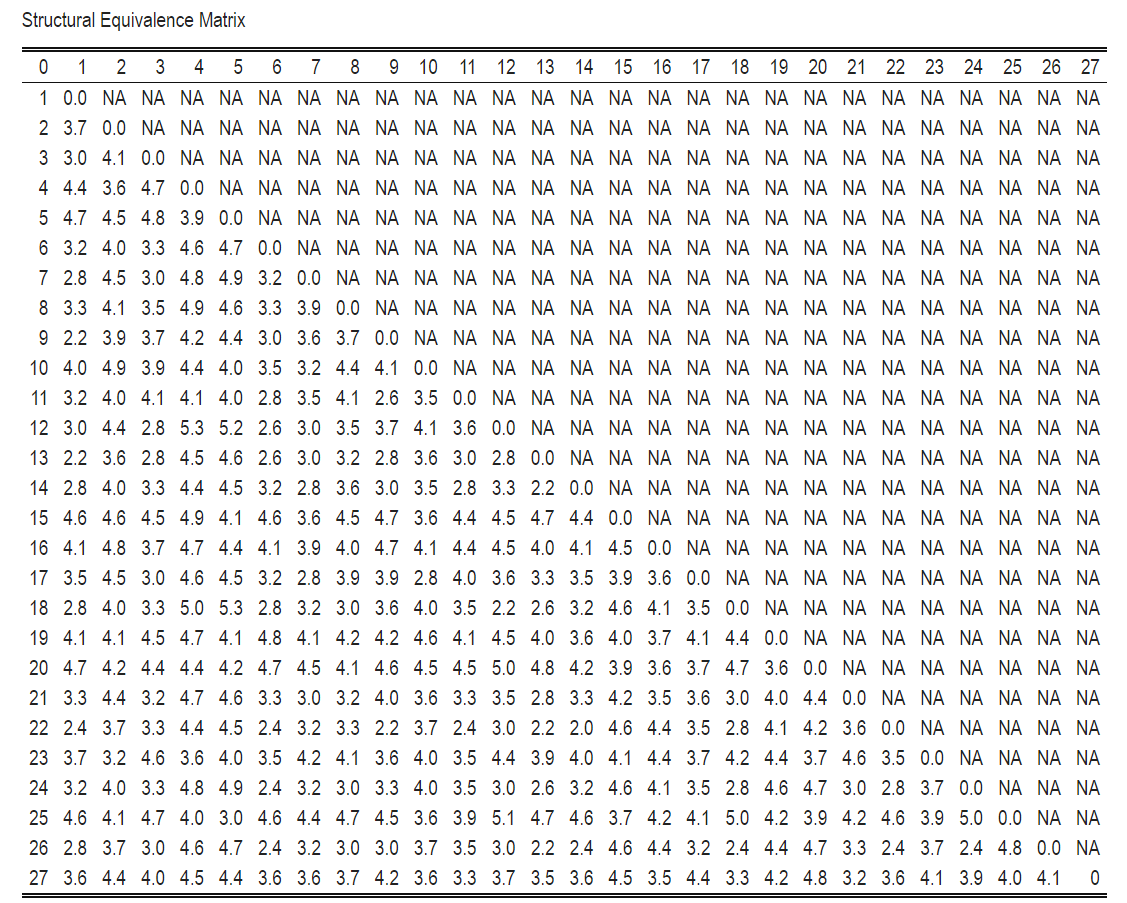
How each brokerage index correlates to other brokerage indices.



In our case, two members are considered structurally equivalent if they are connected to other members in a similar way–they are nominated in a similar manner. Many of the individuals are structurally similar as seen by the values on this structural equivalence matrix ranging from 2.2 to 5.0 (Figure 9).

**Figure 9. Structural Equivalence Matrix**

How club members are structurally equivalent to others.



The dendrogram (Figure 10) tells us that members 11, 1, 9, 13, 14, and 22 are more structurally similar than they are to other members not in the same prong.

**Figure 10. Hierarchical Clustering of Network Positions**

Dendrogram grouping similar actors.

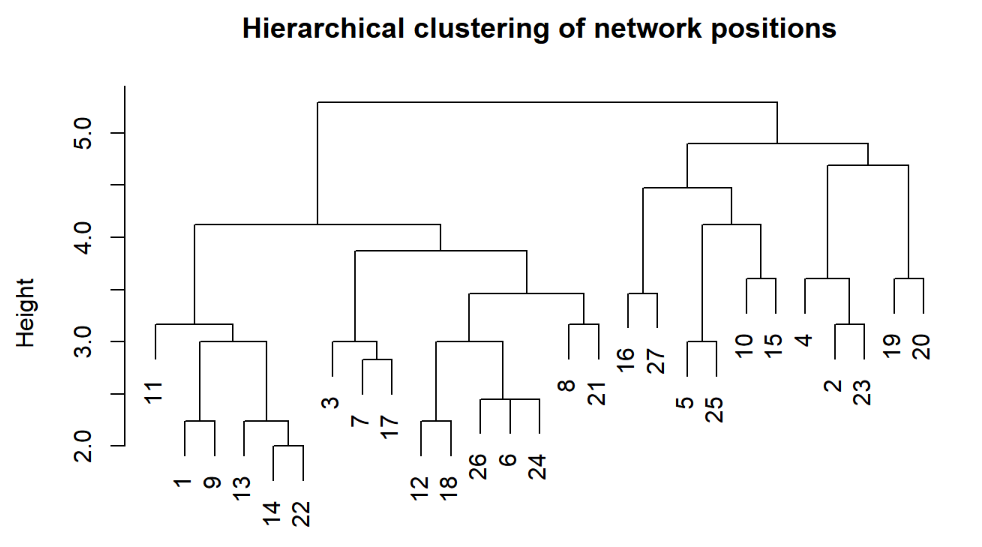
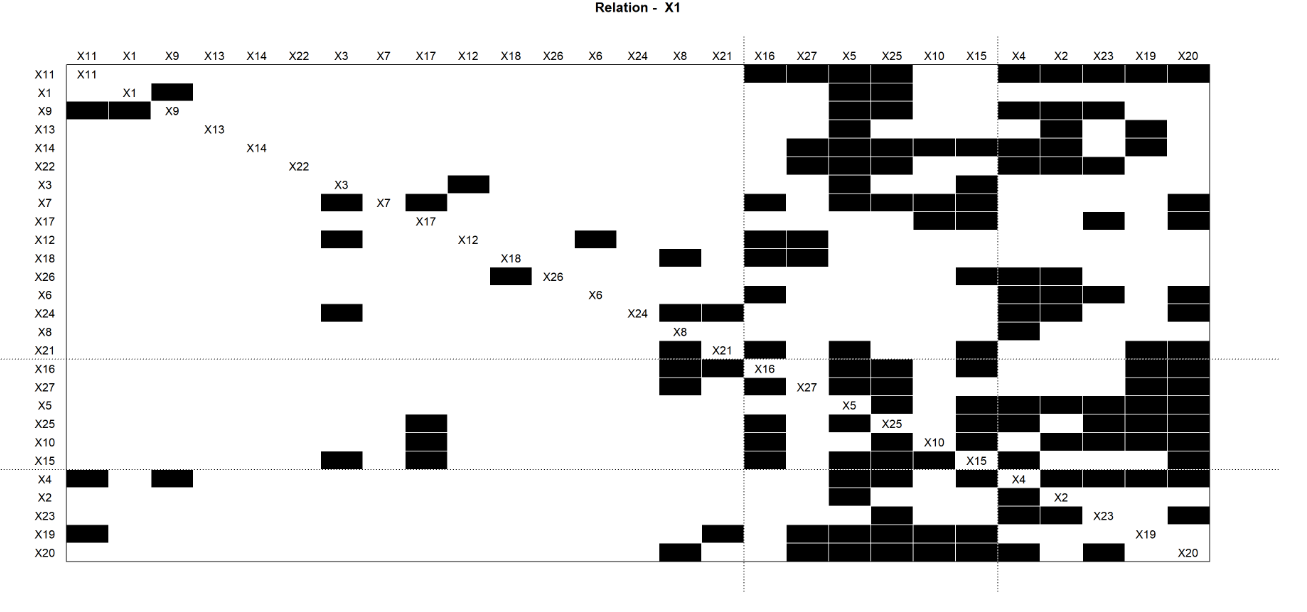


Figure 11 is the block model of the advice-seeking network. This was divided into three blocks based on the hierarchical clustering of network positions dendrogram. The groups were split by one level into the hierarchy.

**Figure 11. Block Model**

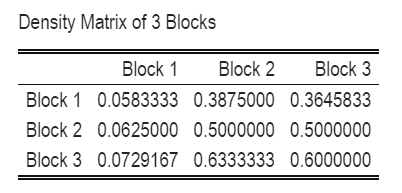
Actors are divided into groups based on structural equivalence.



The density matrix of the three blocks in Figure 12 is as follows: block 1 seems to be the least dense compared to blocks 2 and 3. Blocks 2 and 3 are the most dense, which suggests there is more advice-seeking among the members in these blocks.

**Figure 12. Density Matrix of the Blocks**

The density of each block is based on the brokerage of actors.

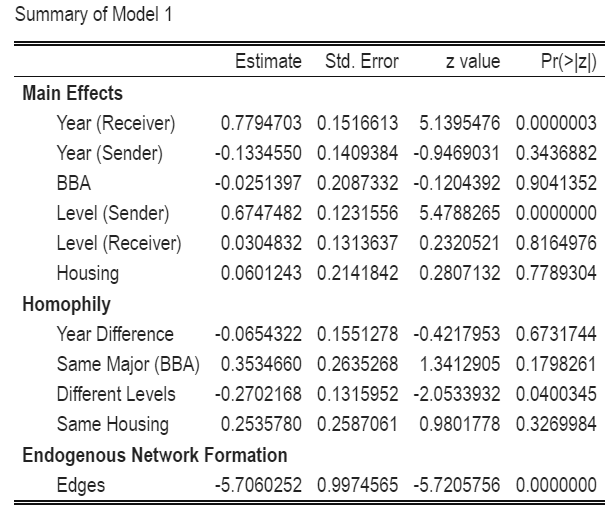


After comparing ERGM for two models, we chose model 2 because it has a smaller AIC (550 compared to 628). Of the statistically significant results, older members are more likely to receive and less likely to send out nominations (Figure 14). Increasing the age by one year is associated with an increase in the odds of receiving a tie by e0.4685 - 1 = around 60% and a decrease in the odds of sending a tie by e-0.3735 - 1 = around 31%.

Members of higher levels are more likely to send out ties (Figure 14). Increasing the level by 1 is associated with an increase in the odds of sending out a tie by e.4847 - 1 = around 62%. These confirm our first and second hypotheses since level and age are statistically significant. None of the homophily measures were statistically significant, meaning we cannot conclude that members of the same year, major, level, or housing situation were more likely to seek advice from one another. The negative coefficient on edges indicates that the network is sparser than expected by chance, and this result is statistically significant. There is also significant mutuality in tie formations. The odds for a tie to form a mutual relation is e2.077 = around 7 times the odds for a tie to form a non-mutual relation. The negative coefficient for preferential attachment indicates ties are more likely to concentrate on a few nodes. The negative coefficient on two-path indicates ties are less likely to form open triangles.

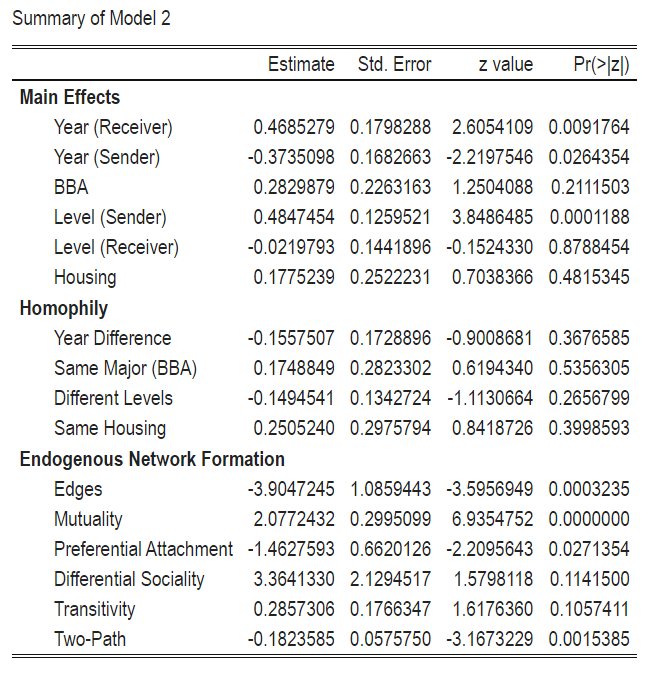
**Figure 13. Summary Statistics of Model 1**

Main effects, homophily, and endogenous network formation from ERGM model 1.



**Figure 14. Summary Statistics of Model 2**

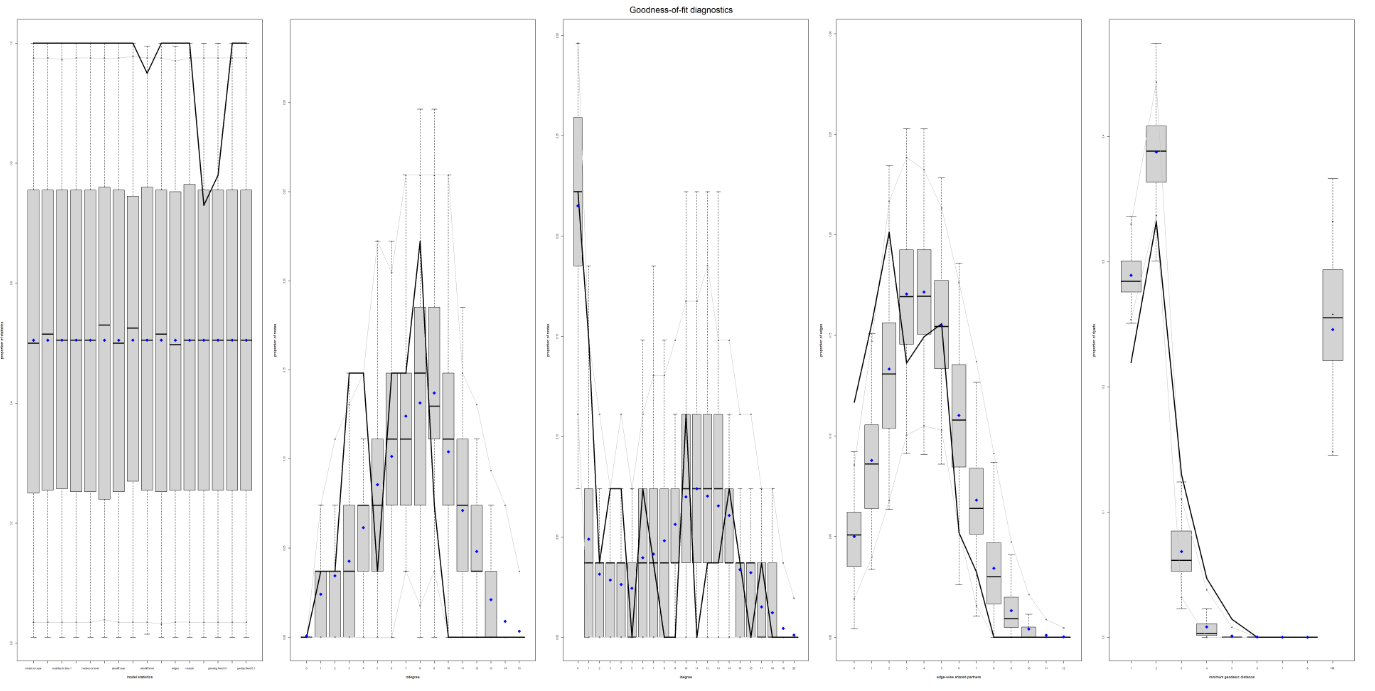
Main effects, homophily, and endogenous network formation from ERGM model 2.



For the goodness of fit comparisons, model statistics are much higher than in a fitted model for the most part, with the exception of one effect (Figure 15). The outdegree centrality and indegree centrality all seem to be slightly left-skewed. This may mean that for lower x values, our model has a higher proportion of nodes in lower values than in a fitted dataset, and lower than average proportion of nodes for higher x-values. The edgewise shared partners is also left-skewed for the proportion of edges compared to the fitted model, comparing across x-values. The minimum geodesic distance has no clear trend between any of the observed measures, however the proportion of dyads for observed, lower, minimum geodesic distances are lower than the fitted values, and higher for the minimum geodesic values.

**Figure 15. Goodness of Fit Diagnostics**

Model statistics, outdegree and indegree centrality, edgewise shared partners, and minimum geodesic distance of our observed model compared to a fitted model.



## Discussion

Our results indicate that more advice is sought from older members in the club, as well as those in higher positions, proving both our hypotheses. However, there were various limitations in building out this network. Starting from data collection, there was selection bias from the voluntary nature of the responses. Members who did not respond may have similarities that change the network map. For instance, non-respondents may not seek advice from anyone in the club, which would cause the network to be less closely tied together. Additionally, the club has a rigorous application process, a form of social controlling similar to network-intervention controlling. Certain attributes that correlate with in-degree centrality may surface as early as the application process. We also did not specify which types of advice each club member seeked from one another, which could differ for social versus career-related advice.

Additionally, the network model does not capture the full detail of certain attributes. Some attributes, such as level and BBA, are constantly changing. In data collection, members were only asked for current BBA status and the highest position they served or currently hold. Not only is the organization’s hierarchy horizontal, but it is also common for members to move positions. Sometimes members move to higher positions, or, in other cases, take an intermediate position after serving in a higher one. This means potentially misleading data related to levels of members in the club.

From the current model, it became clear that further attributes could be analyzed to reveal more nuances behind the network. Factors include strength of ties, gender, having internship experience or a job offer, and aspiring field of work. The job offer may increase indegree centrality since this is a pre-professional business club, which focuses on career progression and finding a job. We did not ask for racial identification because this club does not have a diverse spread of race. However, due to the racial homophily detected in literature reviews, a larger network would benefit from this analysis.

Methodologically, limitations include not conducting edge removal analysis because it assumes an undirected network. Additionally, our second ERGM model forced MPLE on the ERGM which assumes naive logistic regression.

## Works Cited

Baker, Rachel. “Understanding College Students’ Major Choices Using Social Network Analysis.” Research in Higher Education, vol. 59, no. 2, 2018, pp. 198–225. JSTOR, http://www.jstor.org/stable/26451635. Accessed 3 Dec. 2022.

Bowler, Wm. Matthew, et al. “Understanding Trust Transference between Leaders, Followers and Coworkers: Testing Two Three-Way Interactions.” Journal of Managerial Issues, vol. 30, no. 4, 2018, pp. 444–62. JSTOR, http://www.jstor.org/stable/45176597. Accessed 3 Dec. 2022.

Brewer, Devon D., and Cynthia M. Webster. “Forgetting of Friends and Its Effects on Measuring Friendship Networks.” Social Networks, vol. 21, no. 4, 1999, pp. 361–373., doi:10.1016/s0378-8733(99)00018-0.

Kapucu, Naim, et al. “Social Network Analysis (SNA) Applications in Evaluating MPA Classes.” Journal of Public Affairs Education, vol. 16, no. 4, 2010, pp. 541–63. JSTOR, http://www.jstor.org/stable/20790765. Accessed 3 Dec. 2022.

Wimmer, Andreas, and Kevin Lewis. “Beyond and below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook.” American Journal of Sociology, vol. 116, no. 2, 2010, pp. 583–642., doi:10.1086/653658.