Race and Relationships: Analyzing the Presence of Homophily in Notre Dame's Student Body

Introduction

Colleges and universities are hosts to many different people who come from diverse backgrounds. Higher education is the melting pot through which many cultures and backgrounds can interact with each other in an inclusive and supportive environment. For four years, I have had the privilege of being a part of the diverse community that exists at Notre Dame, and I have developed friendships with both in-group and out-group members of my race. The theory of group size effects generally states that members of a minority group will have a higher probability of having relationships with out-group members and members of a majority group will have a higher probability of having relationships with in-group members, and with me being at a predominantly white institution and being a racial minority, the diverse racial composition of my network reflects such a theory. However, I have observed that for some individuals, that is not the case. As I walk around campus, I have generally noticed racial homophily existing within friend groups: Asian people eating with Asian people, Latino people working out with Latino people, and African Americans studying with other African Americans. I always believed that there is an intrinsic benefit to maintaining a diverse network, but some individuals on campus might not share the same belief as me, and I always questioned why such individuals decide to only make connections within their race. However, my observation of homophily on campus is merely conjecture and only based on my observations, and thus my conclusions may be simply based on chance observations of the racial compositions of outside networks. However, with access to the NetHealth data set and the racial data that is combined with the network data, I can test whether a bias towards racial homophily exists within the Notre Dame student population, and the results of such data analysis will either confirm my findings or prove my insights wrong. My overarching research question is: **Does race-based homophily permeate throughout the student population at Notre Dame?** To answer this question, I will use the NetHealth basic survey data and edge list data to compose measures of racial homophily through three different means: analyzing average summary statistics of the nodes based on race, analyzing tie decay between same race and specific race pairs, and analyzing tie formation between same race and specific race pairs. I hope that by the end of this paper, a clearer picture of racial homophily at Notre Dame will be presented.

Literature Review

Sociologists have been studying for decades how and why groups tend to be more homophilous than others. Matheson Currarini and Vega-Redoindo explain that homophily is driven by certain factors that include, but are not limited to socialization costs, preferences for similar partners and meeting biases that are formed by group size effects. They mention that relative population shares and absolute size play distinct roles in influencing meeting biases and homophily patterns and that smaller relative population groups face higher socialization costs if they only restrict themselves to within-group ties (Currarini et al. 18). However, from what I have witnessed, even races such as African-American and Asian-American, who both have small relative group size at Notre Dame, tend to form within-group ties, and this reality goes against the common theory of group size effects that is mentioned within this paper. I hope to provide an answer that runs contrary to the observations of Currarini to visualize how despite group size effects, racial homophily of minority groups permeates at Notre Dame. Similar to Currarini, Kara Joyner and Grace Kao use the theory of group size to explain the existence or lack of same-race relationships within schools, stating that “Adolescents' likelihood of having an interracial friendship increases as the proportion of same-race students in their schools decreases" (Joyner & Kao 810). However, Joyner and Kao used only data from grades 7-12 rather than data from higher education such as undergraduate and graduate, therefore their findings might not be generalizable to a sample of university study and thus their explanation of relative group size effects may not apply towards the questions I want to be answered.

Stephen Antonoplis looks at homophily from a different perspective than Matheson Curriani and Joyner by indicating openness to others (O2) and agreeableness as predictors of same race or diverse relationships, where "High-O2 individuals had an almost equal 1:1 ratio of same-to-different-race network members, whereas low-O2 individuals had 4:1 same-race" (Antonoplis et al. 894). However, the NetHealth Data has no means to determine O2 in the students at Notre Dame, and although openness to others may be a reliable indicator for racial heterophily in networks, the measurements for racial homophily through the NetHealth data do take into account extraneous factors such as agreeableness and openness to others. However, further research into O2 measurements and the NetHealth data may provide insight into which races are more open to others, a factor that may be a driver for racial heterophily in social networks. Furthermore, Lars Leszczensky and Sebastian Pink claim that ethnic identification is the driver for racial homophily, stating that "Friendship formation is necessarily a relational process, which implies that, in addition to the strength of an individual's own ethnic identification, the strength of potential friends' ethnic identification must also be considered" (Leszczensky and Pink 398). Additionally, Julie Park explains in her article on student organizations and interracial friendships that "participation in Greek life, religious groups, and ethnic student organizations were all negatively associated with interracial friendship" (Park 653). Therefore, to intertwine the findings of Leszcensky and Park, organizations in which ethnic identity is prevalent and important can promote the creation of racially homophilous networks. At a university such as Notre Dame, I believe that the existence of student clubs such as the Filipino American Student Organization (FASO), Asian-American Association (AAA), Wabruda, and so on provide avenues for which people can express their ethnic identification, and as I will explain later, I believe that the existence of these student organizations not only create a greater sense of ethnic identification, but provides the social foci for which racially homophilous relationships can form.

Kristen McDonald uses racial homophily to predict friendship stability and quality from sampling 490 young adults, and she concluded that “Cross-race friends were significantly less similar than same-race friends on dimensions of peer-nominated popularity/sociability, exclusion, and victimization” (McDonald et al. 906). From this conclusion, maybe it is not ethnic identification or group size effects that promote racially homophilous friend groups, but maybe the similarity between nodes in terms of socio-behavioural traits that potentially derive from living in a specific racial environment that drives the formation of same-race relationships. I understand well that home environments can differ according to the social norms of each race, and therefore the similarity gained by living in similar home environments may promote ethnically homogeneous friendships. However, within the NetHealth Data, there is no measure of such sociobehavioral traits for each node, and therefore such similarity may account for the prevalence of racial homophily at Notre Dame, but it cannot be measured by the means I will be describing within my paper. Additionally, David Melamed shows in his article that “homophily promotes cooperation, net of two other key determinants of cooperation—dynamics and reputations” (Melamed et al. 1106), and thus the cooperation that is engaged in through the existence of racially homophilous networks may influence tie formation of same-race relationships, as the added camaraderie and support systems that derive from that homophilous social network can provide the incentive to form same-race relationships. This is plausible, as the racially homogeneous clubs on the Notre Dame campus both provide the cooperation between individual nodes and the social support to engage with each other properly, but I will not be specifically diving into the dynamics of racially homogeneous clubs at Notre Dame, only the presence and trends of racial homophily over time.

Cassie McMillan uses exponential random graph models in a study of 153 adolescent friendship networks to show that the strength of a tie may influence the racial homophily of a person's network. She explains that "As the proportion of racial minority students increased, one's odds of naming a same-race peer as a best friend sharply increased, while their odds of naming a same-race peer as a non-best friend slightly declined" (McMillan 295), thus the driver of racial homophily in adolescent networks is the existence of surrounding diversity and the strength of the tie in the first place. In general, the findings from this study indicate that if you are in a more diverse environment, you are more likely to have same-race strong relationships and less likely to have different-race strong relationships, and I believe that this theory holds at Notre Dame, as Notre Dame is a racially diverse environment and of the friend groups I know of, the strongest ties are with racially homogeneous people. However, throughout my paper, I do not use measures of tie strength as an indicator of racial homophily, and thus that factor may be present in the data as an influence on racial homophily but not explicitly outlined.

All of these articles attempt to explain why and how homophilous relationships tend to form. However, the purpose of my paper is to understand whether racially homophilous relationships decay or form over time within the student population at Notre Dame, not so much to describe the means through which these racially homogeneous networks tend to form. I hope through my research that I can show a tendency towards racial homophily over time for certain minority races at Notre Dame, and from the existence of such, I can offer insights into how and why they form. I believe that it is through the existence of social foci at Notre Dame such as ethnically homogeneous clubs that promote the cultivation of racially homophilous networks, but I will go into that in the discussion portion of my paper. Now, I will describe the data and methods I used to attempt to answer the question of whether there is a trend towards race-based homophily over time at Notre Dame.

Data and Methods

To determine how racial homophily changes over time on college campuses, I utilised data from the NetHealth Project, a multi-year longitudinal study of social networks and health administered through the Center for Network Science and Data at the University of Notre Dame. This data contains information on about 700 Notre Dame students who were recruited for the initial two-year study for the NetHealth Project. Within this dataset, each node (student) in the study's basic survey data ranged from basic demographic information such as race and religious affiliation to individual stress levels. Furthermore, there is communication data within the study, which details all communication between each node which is used to construct edge and arc lists for each node. For my paper, I used the NetHealth data from Wave 3 and Wave 5, indicating Spring 2016 and Spring 2017 data to determine how network dynamics and racial homophily change over a year, and in this case, between the end of freshman year and the end of sophomore year for students at the University of Notre Dame. Using the basic survey data, I decided to choose three categories of note: egonet size, egonet clustering coefficient, and egonet race interquartile variation. To determine how such categories change over time, differentiated by race, I used R coding software to take all of the network data, filter each group by race, and then take the mean of these three categories for each race. I did this methodology for both the Wave 3 data and the Wave 5 data, and then I took the difference of the mean of these three categories by race to quantify the change of these three categories over time. The results of such differences are in **Table 1**.

To quantify how racial homophily may decrease over time between the same race and each pair race individuals between the Wave 3 and Wave 5 datasets, I used the measure of tie decay to determine the likelihood of ties breaking over the years given a certain racial combination (same_race, pair_race, each race pair). Ethan helped create a dataset that took the Spring 2016 edge data, included each interaction between one node with another, determined if the nodes were the same race or not, included the race pair between each node, and used the Spring 2017 edge data to determine whether that tie decayed or not over the year. From this dataset, I grouped the data by whether the node pair was the same race or not and took the mean of the tie decay variable to quantify the difference in tie decay by whether the nodes were the same race or not. The results of such are expressed in **Table 2**, with TRUE under the same_race column indicating that the node pair were the same race and FALSE indicating that the node pair were different races. To quantify tie decay for each race pair, I grouped the data by each race pair and then took the mean tie decay to see how tie decay rates differ between each race pair. In total, within this dataset, there were over 15 race pairs with some being excluded due to representing the same relationship (Asian American - White and White - Asian American are counted as one race pair). The results of this methodology are expressed in **Table 3**, detailing the mean percent decay of each race pair, the mean tie decay rate overall, and the highest and lowest tie decay percentages of the race pairs.

To determine whether a node pair being the same race or not can be a predictor of whether that tie will decay or not, I ran a generalized linear model regression with same_race as the predictor variable for tie decay, controlling for gender, first-generation student, egonet clustering coefficient, and egonet size. I decided to control for these variables because of the various effects they can have on predicting tie decay. For example, for egonet size, the larger a person's egonet size is, the more likely they are to experience tie decay overtime as they have to pick and choose which relationships to devote the most time to, and ultimately, these nodes cannot maintain the substantial effort required to maintain each relationship if their network size is large. Controlling for egonet size, in this case, would allow me to isolate the effect that the ego network size has on tie decay, meaning that the inferences I can draw on whether same-race relationships can predict tie decay or not are more statistically accurate. This same logic applies to the variables gender, first-generation student, and clustering coefficient. The coefficients drawn from the regression are in the form of log-likelihood, and this form of the data does not provide the most insight into the relationship I am trying to determine between same-race relationships and tie decay. Therefore, to get the coefficients out of the form of log-likelihood, I take each coefficient from the regression, take the coefficient to the power of e, subtract by 1, and then multiply each coefficient by 100 to get the percentage change of the variable, and this percent change of the coefficient provides a way to quantify whether same race pair is a good predictor of tie decay. The results of the regression are displayed in **Table 4**, with the variable column displaying the same_race predictor and all of the control variables, the coefficients expressed in percent change, and the statistical significance level of each variable.

To measure whether a specific pair of a race is a predictor of tie decay, I ran the same generalized linear model regression, but with pair_race as the predictor instead of same_race, controlling for all of the aforementioned variables. For interpretive purposes, I had to declare the race pair "White-White" as a factor through which the regression of pair_race on tie decay can take place. Doing such will allow me to have a point of reference to compare the percentage change of each race pairs' tie decay to. After running the regression, controlling for certain variables, I took the coefficient of each predictor variable (in this case every race pair excluding "White-White"), transformed it out of log-likelihood into percentage change, and created a table with each race pair, the percentage change of tie decay in reference to "White-White," and the corresponding significance levels associated with the coefficients. The results of such are expressed in **Table 5.**

On the other side of the coin, to quantify how racial homophily may increase over time between certain combinations of race pairs or by same-race relationships, I used tie formation between Wave 3 and Wave 5 to measure the likelihood of a tie forming over two semesters. Ethan also curated another dataset to measure tie formation, which includes the same node and race information as the tie decay dataset, but to come up with the formation numbers, he used the edge list of the Spring 2017 dataset, created 218747 different observations representing the number of possible ties that could form over the year, and indicated, for each possible tie that could form, whether they formed over the year. The final number of actual ties that were formed from the nodes in the dataset between Wave 3 and Wave 5 was 1874. In a similar methodology to the tie decay calculations, I filtered the observations by whether the nodes were the or not and determined the mean tie formation rates based on whether the node pair was racially homophilous or not. The results are displayed in **Table 6**, which indicates both the same_race variable and the percentage of ties formed over the years. I performed this same methodology, but instead of the same race as the filter, I filtered by race pair and calculated the average tie formation rate for each race pair. The results of this methodology are displayed in **Table 7.** The percentages displayed in **Table 7** are very small, but that is due to the number of observations being determined by the number of possible ties between the nodes. To account for such small figures, I calculated the percent change between the average tie formation of each race pair and the average tie formation of the "White and White" race pair to create figures that are easier to interpret. The manipulation of such figures is displayed in **Table 8.**

Finally, in the same manner that the tie decay regressions were performed, I ran 2 generalized linear model regressions with same_race and pair_race as the predictor variables, but instead of tie decay being predicted, tie formation was predicted. Using the same controls and manipulating the coefficients in the same fashion, both regression outputs provided the likelihood of tie formation based on the specific predictor variables, and the results of such regressions are displayed in **Tables 9 and 10**. Through the methods detailed above and the data accumulated through regressions and R data manipulation, I hope to show how the racial composition of students' networks at Notre Dame changes over time and how individual race characteristics may influence such change.

Results

Average Summary Statistics

Looking at **Table 1**, the difference in mean egonet size between Spring 2016 and Spring 2017 for each race seemed to decrease over the years, with the African-American student group having the highest differential of the mean egonet size of -2.86. The overarching theme of egonet size reduction between the two years is expected. In my experience and my friends' experiences at Notre Dame, as you increase in grade throughout University, you learn which ties you want to keep and which ties are not worth the effort, and thus your egonet size decreases as a result of this selection process that happens as you mature and increase in grade. In terms of the mean clustering coefficient between the races, the ones I would like to highlight are the Asian-American mean clustering coefficient differential of - .0322 and the Foreign Student mean clustering coefficient differential of .0366. On average, the nodes in the network of the Asian-American students sampled in the NetHealth Data became 3.22% less clustered over time and the nodes in the network of the Foreign students sampled became 3.66% more clustered. There can be many reasons for the increase or decrease in the clustering coefficient, such as an increase or decrease in overall network size or an increase or decrease in the number of people in the network who have fewer ties. However, if you look at **Table 1** under mean egonet race IQV, the average interquartile variation of race for each race, where 0 indicates a fully homophilous network and 1 indicates a fully heterophilous network, the differential for Asian-Americans is -.0455 and the differential for Foreign Students is .0451. On average, the nodes in the network of Asian Americans sampled became 4.55% more racially homogeneous and the nodes in the network of Foreign Students sample became 4.51% more racially diverse. Therefore, it might be the case that on average for Asian Americans at Notre Dame, as their networks became more racially homophilous, their clustering coefficient decreased as a result of these homophilous ties not being as clustered together, whereas on average for Foreign Students at Notre Dame, as their networks become more racially heterogeneous, their clustering coefficient increased as a result of these diverse ties being more internally connected in their network. Therefore, racial homophily and heterophily may account for variation in the clustering coefficient of different races at Notre Dame. Furthermore, the only race that the mean egonet race IQV decreased and thus the only networks that became more racially homogeneous between Wave 3 and Wave 5 were Asian Americans, meaning that as time passed, only the Asian American networks at Notre Dame tended toward being more racially homophilous whereas the rest of the races in the dataset became more racially diverse. The result coincides with my observations, as all over campus, I witness Asian-American groups sticking together and only sometimes see Asian-American students grouping with heterophilous individuals. Therefore, on average, Asian-American student networks become more racially homophilous whereas every other race gravitates towards racial diversity over time to a certain extent.

Tie Decay

Tables 2 and 3 indicate the average tie decay rates for the same race/not the same race and each racial pair. Looking at **Table 2**, when the same race = TRUE, meaning that the node pair is of the same race, on average, their ties decayed by 55.78% between the Wave 3 and the Wave 5 Dataset. When same_race = FALSE, meaning that the node pair consists of two nodes of different races, on average, their ties decayed by 61.14% over a year. Therefore, based on the data above, same-race relationships compared to different-race relationships were on average 5.36% less likely to decay, meaning that same-race relationships were 5.36% more likely to persist over time. Thus, homophilous relationships are less likely to decay than heterophilous relationships at Notre Dame. Moving on to **Table 3**, the highest average tie decay percentage was between Foreign Students and White Students with a decay percentage of about 70% and the lowest average tie decay percentage was between African-Americans and Asian-Americans with a percentage of about 48%. Additionally, it is apparent that for same-race relationships, every race besides Foreign Student has a drastically lower percentage of tie decay for a year compared to some of the different race ties such as "Foreign Student and White" and "African American and White," who have tie decay percentages hovering around 70%. Therefore, in general, the table demonstrates the tendency for same-race relationships to persist more often than different-race relationships.

Table 4 displays the generalized linear model regression results with the same race as a predictor of tie decay, controlling for gender, first-generation student, egonet_clustering, and ego network degree. Controlling for the aforementioned variables, when the same race is equal to TRUE, ties are 14.89% less likely to decay with a significance level of $p < .01$, meaning that the coefficient is statistically significant. It is also apparent from the regression results the importance of controls, as both male and egonet_clustering are both statistically significant at the $p < .001$ level, meaning that the bias in these variables to influence tie decay was taken out of the calculation of the coefficient of the same race = TRUE, meaning that this variable is more statistically accurate. With the significance level of same race = TRUE being statistically significant at the $p < .01$ level, we reject the null hypothesis that a node having a tie with another node of the same race has no impact on the tie decaying, and therefore from the results of this regression, it is apparent that same race relationships at Notre Dame have a lower likelihood of decaying, leading towards the persistence of homophilous relationships over time.

Table 5 runs the same generalized linear model regression results but with pair race as the predictor of tie decay. The statistically significant pairs were the African-American and White pair, the Foreign Student and White pair, and the Latino/a and White pair, with percentage decay likelihoods of 63.14%, 68.38%, and 21.13% respectively. With the African-American and White pair as an example, compared to the White to White pairs, African-American and White pairs are 63% more likely to decay overtime with a significance level of $p < .001$, meaning that we reject the null hypothesis that an African-American node having a relationship with a White node has no impact of whether the tie will decay over time. The same logical conclusion can be made for the Foreign Student and White pair and the Latino/a and White pair, and from the results of this regression, it is apparent that for some diverse nodes at Notre Dame, their ties with White nodes tend to decay at a statistically significant level.

Tie Formation

Moving on to what I believe is a better indicator of racial homophily over time, tie formation statistics between same race/not same race and each race pair are indicated in **Tables 6 and 7**. Because of the number of observations included in the tie formation dataset, the numbers in Table 6 are really small, with a tie formation percentage for same_race = TRUE of .0095 and tie formation percentage for same_race = FALSE of .0077. However, taking the percentage change between the two figures, the number amounts to a 22.9% change. Therefore, if the nodes are the same race, compared to nodes that are a different race, the tie is on average 22.9% more likely to form. Thus, homophilous relationships are more likely to form at Notre Dame versus heterophilous relationships. **Table 7** displays the percentage of tie formation for each of the race pairs, but because such numbers are very small, I converted the race pair tie formation percentages to the percentage change of that number from the White to White tie formation percentage. What I get when I express the percentage change for all race pairs in **Table 8**. Some statistics of note are the African-American and African-American tie formation percent change, the Asian-American and Asian-American tie formation percent change, and the Foreign Student and Foreign Student tie formation percent change. With African-American same-race ties, compared to White and White tie formation percentage, African-American same-race ties are 208% more likely to form over two semesters. The same interpretation can be utilized for Asian-American same-race ties and Foreign Student same race ties with both having a 26% increased likelihood of forming a tie compared to White to White ties. Therefore, it is clear from this table that many same-race connections, and especially African American same-race ties, form at a higher likelihood over time, meaning that these same-race nodes are seeking out racially homophilous nodes with a higher likelihood than heterophilous nodes. Thus, at Notre Dame, same-race relationships tend to form at a higher rate than different-race relationships for specific races, leading towards racial homophily.

Tables 9 and 10 display the regression results of same race and pair race as a predictor of tie formation, controlling for gender, first-generation student, clustering coefficient, and network degree. Looking at **Table 9**, when Same Race = TRUE and controlling for the aforementioned variables in the data, ties have an 18.25% chance of forming throughout the year with a statistical significance at the $p < .001$ level, meaning that we reject the null hypothesis that being the same race as someone else has no impact on whether that tie forms or not. Other controls that had statistical significance were gender and degree with both variables being statistically significant at the $p < .001$ level, meaning that both variables have some impact on whether a tie forms or not. Therefore, the results of this regression indicate that a student being the same race as another student has a significant impact on whether that tie will form or not. **Table 10** indicates the regression results of each race pair as the predictor of a tie forming. A pair of nodes that I would like to highlight is the African-American same-race pair with an astounding 193% higher chance of a tie forming throughout the year compared to the tie formation likelihood of White to White pairs with a significance level of $p < .001$. With this significance level, we reject the null hypothesis that both nodes being African-American have no impact on whether that tie forms and therefore it is apparent that the racial network composition of African-American students has a higher likelihood of being homophilous as time passes. This same idea persists for same-race Asian-American ties with a significance level of $< .1$ as well. What is also interesting is the statistically significant negative percentage change of Foreign Student and White nodes and African American and White nodes, with each having a negative percentage chance of a tie forming of 36.11% and 29.76%, respectively, compared to White to White tie formation. These figures highlight how for some race combinations, there is a decreased chance of that tie actually forming, moving away from racial heterophily and towards racial homophily. Therefore, from the regression results of the same race and the pair race as a predictor of tie formation, it is shown that some races at Notre Dame tend to form racially homophilous ties over time and, in the alternative, stray away from racially heterophilous ties.

Discussion and Conclusion

I believe that the findings expressed through the average summary statistics of the networks, the mean tie decay and tie formation rates of the same race and each race pair, and the statistically significant coefficients from the regressions of same race and pair race as predictors of the tie decay and tie formation show this overarching tendency for specific races to become more racially homophilous overtime at Notre Dame. More specifically, the decrease in mean race IQV of Asian-Americans who participated in the NetHealth study at Notre Dame shows that they became more racially homogeneous between Spring 2016 and Spring 2017. Furthermore, the decrease in average tie decay and increase in average tie formation between same-race ties and non same-race ties show that there is something intrinsic about same-race relationships that not only result in increased tie persistence over time, but also tie formation throughout the year. Additionally, specific race pair figures such as the 208% increase in tie formation of same-race African-Americans compared to White to White tie formation rates and a 193% statistically significant increased likelihood of an African American same race tie forming over the year display that for some races, this tendency toward racial homophily is prominent. These two huge figures of African American tie formation are also corroborated by outside research, as Elizabeth McClintock explains that "Black students demonstrate a particularly strong tendency for racially homophilous friendship choices" (McClintock 49). Additionally, Tamara Gilkes Borr explains in her research paper "The Strategic Pursuit of Black Homophily on a Predominantly White Campus" that "black students manipulated place and time on the predominantly White campus to find Black networks" (Borr 340) and "same-race relationships provide academic/social safety and serve as a source for positive racial identity development" (Borr 336). Therefore, it may be due to the increased academic and social safety provided by homophilous networks that compel African American students at Notre Dame to spend much more time and effort towards curating a racially homophilous network. Thus, outside research on the topic of black homophily provides reasons as to why black homophily is so prevalent on college campuses, but such research does not extend much to other races such as Asian Americans and Latino/a individuals.

I truly believe that this tendency for some races to become more racially homophilous over time at Notre Dame is due to the prevalence and popularity of ethnically homogeneous social foci at Notre Dame, where over a year, a Freshman going into Sophomore year can join these clubs which provide the opportunity to form racially homogenous friendships. However, my research does not dive into the racial dynamics of clubs, only the existence of network racial composition, tie decay and tie formation between different race and same-race nodes. Therefore, for future research into this topic, I would try to use different predictor variables in my previous regressions to decipher whether the racial composition of clubs can have an impact on same-race and different-race ties over time at Notre Dame. In doing so, I can determine the reasons why racially homogeneous friendships tend to grow over time. However, before doing this research, I would like to solve some of the limitations of this study, which include the ethically ambiguous nature of the "Foreign Student" variable. The foreign student race variable can technically be included within each of the racial categories, and if there was a way to consolidate that variable into the other race categories, then the results would more more statistically accurate. Another limitation of this study is the sample size. The NetHealth data only covers the students at Notre Dame, and therefore for further research, I would extrapolate my methodology over a larger student body size with the inclusion of different schools to determine whether this trend towards racial homophily exists at other schools or only at Notre Dame. I hope that further studies into this topic of racial homophily can provide more insight as to why these racially homophilous relationships form and persevere on college campuses, and I hope my paper took a positive first step towards providing a concrete answer to such a question.

Table 1: Average Summary Statistics of Wave 3 and Wave 5 Net Health Data

	Spring 2016	**Spring 2017**	**Differential**
Mean Egonet Size			
White	14.75055	12.08725	-2.663300
Latino/a	12.21591	10.19767	-2.018240
Asian-American	12.90323	11.27869	-1.624540
African-American	14.53659	11.66667	-2.869920
Foreign Student	8.680851	7.525	-1.155851
Mean Egonet Clustering Coefficient			
White	0.2300835	0.2313515	0.001268
Latino/a	0.2041601	0.2040907	-0.000069
Asian-American	0.1832147	0.150983	**-0.032232**
African-American	0.1923509	0.1829346	-0.009416
Foreign Student	0.2146333	0.2512971	**0.036664**
Mean Egonet Race IQV			
White	0.178998	0.2032369	0.024239
Latino/a	0.2075594	0.2362647	0.028705
Asian-American	0.2403649	0.1947946	**-0.045570**
African-American	0.1687641	0.2024751	0.033711
Foreign Student	0.2306287	0.2757683	**0.045140**

Table 2: Tie Decay Rates for Same Race Between Wave 3 and Wave 5

	same_race	percent_decay
1	FALSE	0.611474585
2	TRUE	0.557890855

Table 3: Tie Decay Rates for Each Race Pair Between Wave 3 and Wave 5

	pair_race	percent_decay
1	African-American and African-American	0.554216867
2	African-American and Asian-American	0.479166667
3	African-American and Foreign Student	0.653846154
4	African-American and Latino/a	0.629213483
5	African-American and White	0.684615385
6	Asian-American and Asian-American	0.547169811
7	Asian-American and Foreign Student	0.59375
8	Asian-American and Latino/a	0.505747126
9	Asian-American and White	0.55666004
10	Foreign Student and Foreign Student	0.652173913
11	Foreign Student and Latino/a	0.617021277
12	Foreign Student and White	0.703296703
13	Latino/a and Latino/a	0.542857143
14	Latino/a and White	0.617176128
15	White and White	0.557792992
	Mean Percent Tie Decay	0.592980246
	Race Pair w/ Highest Tie Decay %	
	Foreign Student and White	0.703296703
	Race Pair w/ Lowest Tie Decay %	
	African-American and Asian-American	0.479166667

Table 4: Generalized Linear Model Regression Results for Same Race Tie Decay, Controlling for Gender, First Generation Student, Clustering Coefficient, and Egonet Size

	variable	percentage_change	significance_of_coefs	p_value
1	Same Race = TRUE	-14.8959156	**	<.01
2	Male	-19.8906016	***	<.001
3	First Generation	26.6759062		>.1
4	Egonet_Clustering	-82.2301088	***	<.001
5	Degree	0.1241771		>.1

Table 5: Generalized Linear Model Regression Results for Each Race Pair Tie Decay, Controlling for Gender, First Generation Student, Clustering Coefficient, and Egonet Size

	pair	percentage_change	significance_of_coefs	p_value
1	African-American and African-American	-19.1501389		>.1
2	African-American and Asian-American	-36.217933		>.1
3	African-American and Foreign Student	19.2778152		>.1
4	African-American and Latino/a	15.2499196		>.1
5	African-American and White	63.1412875	***	<.001
6	Asian-American and Asian-American	-10.701906		>.1
7	Asian-American and Foreign Student	-16.0900425		>.1
8	Asian-American and Latino/a	-22.9741776		>.1
9	Asian-American and White	-3.0032655		>.1
10	Foreign Student and Foreign Student	50.9573109		>.1
11	Foreign Student and Latino/a	2.2495639		>.1
12	Foreign Student and White	68.3873551	**	<.01
13	Latino/a and Latino/a	-8.6458636		>.1
14	Latino/a and White	21.1377519	*	<.05

Table 6: Tie Formation Rates for Same Race Between Wave 3 and Wave 5

	same_race	percent_formation
1	FALSE	0.007732849
2	TRUE	0.009504208

Table 7: Tie Formation Rates for Each Race Pair Between Wave 3 and Wave 5

	pair_race	percent_formation
1	African-American and African-American	0.02887538
2	African-American and Asian-American	0.004719005
3	African-American and Foreign Student	0.011888112
4	African-American and Latino/a	0.010413476
5	African-American and White	0.007078093
6	Asian-American and Asian-American	0.01181767
7	Asian-American and Foreign Student	0.007874016
8	Asian-American and Latino/a	0.00581508
9	Asian-American and White	0.008783124
10	Foreign Student and Foreign Student	0.011764706
11	Foreign Student and Latino/a	0.008072027
12	Foreign Student and White	0.005798661
13	Latino/a and Latino/a	0.008089261
14	Latino/a and White	0.008168587
15	White and White	0.009365882
	Mean Percent Tie Formation	0.009901539
	Race Pair w/ Highest Tie Formation %	
	African-American and African-American	0.02887538
	Race Pair w/ Lowest Tie Formation %	
	African-American and Asian-American	0.004719005

Table 8: Tie Formation Percent Difference from "White and White" for Each Race Pair Between Wave 3 and Wave 5

	pair_race	percent_difference from White and White
1	African-American and African-American	208%
2	African-American and Asian-American	-50%
3	African-American and Foreign Student	27%
4	African-American and Latino/a	11%
5	African-American and White	-24%
6	Asian-American and Asian-American	26%
7	Asian-American and Foreign Student	-16%
8	Asian-American and Latino/a	-38%
9	Asian-American and White	-6%
10	Foreign Student and Foreign Student	26%
11	Foreign Student and Latino/a	-14%
12	Foreign Student and White	-38%
13	Latino/a and Latino/a	-14%
14	Latino/a and White	-13%
15	White and White	0%
	Mean Percent Change Tie Formation	6%
	Race Pair w/ Highest Tie Formation %	
	African-American and African-American	208%
	Race Pair w/ Lowest Tie Formation %	
	African-American and Asian-American	-50%

Table 9: Generalized Linear Model Regression Results for Same Race Tie Formation, Controlling for Gender, First Generation Student, Clustering Coefficient, and Egonet Size

	variable	percentage_change	significance_of_coefs	p_value
1	Same Race = TRUE	18.250577	***	<.001
2	Male	-2.932507	***	<.001
3	First Generation	14.425654		>.1
4	Egonet_Clustering	-60.3093		>.1
5	Degree	6.887709	***	<.001

Table 10: Generalized Linear Model Regression Results for Each Race Pair Tie Formation, Controlling for Gender, First Generation Student, Clustering Coefficient, and Egonet Size

	pair	percentage_change	significance_of_coefs	p_value
1	African-American and African-American	193.3743892	***	<.001
2	African-American and Asian-American	-50.1847385	*	<.05
3	African-American and Foreign Student	28.3964637		>.1
4	African-American and Latino/a	14.7961476		>.1
5	African-American and White	-29.7697157	***	<.001
6	Asian-American and Asian-American	45.2451259	.	<.1
7	Asian-American and Foreign Student	-5.3527198		>.1
8	Asian-American and Latino/a	-27.5285043	.	<.1
9	Asian-American and White	0.5561724		>.1
10	Foreign Student and Foreign Student	48.0015561		>.1
11	Foreign Student and Latino/a	-1.2031032		>.1
12	Foreign Student and White	-36.1107455	***	<.001
13	Latino/a and Latino/a	2.5008591		>.1
14	Latino/a and White	-5.3789941		>.1

Bibliography

Antonoplis, Stephen, and Oliver P. John. "Who has different-race friends, and does it depend on context? Openness (to others), but not agreeableness, predicts lower racial homophily in friendship networks." *Journal of Personality and Social Psychology* 122.5 (2022): 894.

Currarini, Sergio, Jesse Matheson, and Fernando Vega-Redondo. "A simple model of homophily in social networks." *European Economic Review* 90 (2016): 18-39.

Gilkes Borr, Tamara. "The strategic pursuit of Black homophily on a predominantly White campus." *The Journal of Higher Education* 90.2 (2019): 322-346.

Joyner, Kara, and Grace Kao. "School racial composition and adolescent racial homophily." *Social Science Quarterly* (2000): 810-825.

Leszczensky, Lars, and Sebastian Pink. "What drives ethnic homophily? A relational approach on how ethnic identification moderates preferences for same-ethnic friends." *American Sociological Review* 84.3 (2019): 394-419.

McClintock, Elizabeth Aura. “When Does Race Matter? Race, Sex, and Dating at an Elite University.” *Journal of Marriage and Family*, vol. 72, no. 1, 2010, pp. 45–72, https://doi.org/10.1111/j.1741-3737.2009.00683.x.

McDonald, Kristina L., et al. "Contributions of racial and sociobehavioral homophily to friendship stability and quality among same-race and cross-race friends." *The Journal of Early Adolescence* 33.7 (2013): 897-919.

Melamed, David, et al. "Homophily and segregation in cooperative networks." *American Journal of Sociology* 125.4 (2020): 1084-1127.

McMillan, Cassie. "Strong and weak tie homophily in adolescent friendship networks: An analysis of same-race and same-gender ties." *Network Science* 10.3 (2022): 283-300.

Park, Julie J. "Clubs and the campus racial climate: Student organizations and interracial friendship in college." *Journal of College Student Development* 55.7 (2014): 641-660.