The Pennsylvania State University

The Graduate School

College of the Liberal Arts

SAME SCHOOLS, DIFFERENT CLASSES: SOCIO-ECONOMIC STATUS HOMOPHILY IN ADOLESCENT FRIENDSHIP NETWORKS

A Thesis in

Sociology

by

Karen A. Ertrachter

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Masters of Arts

December 2014

The thesis of Karen A. Ertrachter was reviewed and approved* by the following:

Diane Felmlee
Professor of Sociology
Thesis Advisor

John Iceland
Professor of Sociology and Demography
Head of the Department of Sociology

Wayne Osgood
Professor of Criminology and Sociology

Melissa Hardy
Distinguished Professor of Sociology and Demography
Director of the Graduate Program in Sociology

*Signatures on file with the Graduate School

ABSTRACT

Research has found that the preference for associating with those most similar to us in terms of race and gender has been particularly important for shaping the social networks and friendships for American youth. Homophily is defined as the tendency for people to associate with those most similar to themselves within networks (McPherson et al., 2001). However, research on homophily within social networks has not traditionally included socio-economic status in analyses. In this paper, socio-economic status is defined as parental education level, familial income, and type of parental occupation. This paper explores how socio-economic status, in three distinct forms, shapes the composition of adolescent social networks within one school in the National Longitudinal Study of Adolescent Health using three different methods. These methods include centrality analyses and Multiple Regression Quadratic Assignment Procedure. Similarly, the extent of social class homophily among friendship dyads is assessed in comparison to gender, race, and grade homophily using Exponential Random Graph Models. Income, occupational, and educational homophily have separate and distinctive effects, as shown by analyses, which is why including more forms of homophily beyond typical control variables (grade, gender, age, and race) may be helpful for understanding friendship ties between students. Overall, results do not find that one socio-economic status type is most prevalent among students in Jefferson High School. However, results indicate that students may interact with more students outside their own socio-economic groups than expected.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I thank my parents, who have supported me through this and encouraged me even when I doubted myself.

I also dedicate this thesis to my close friends who have supported me through this process and the surrounding chaos as I wrote and edited my thesis. I have to specially thank Victoria Austin for helping me stay motivated.

I cannot thank Jacob Turner enough for his encouragement and late night work parties.

Lastly, I dedicate this thesis to my advisor Diane Felmlee, who has supported me since the beginning of my time here at Penn State.

Introduction

Adolescent friendship networks are important to study since they shape adolescents at a crucial turning point in their lives. More specifically, studying friendship patterns helps researchers gain a better understanding of how adolescent networks are structured and how external aspects related to student's background affect their in-school network. Race and gender have been heavily explored in the social network literature, but research on the impact of socio-economic status has been fairly limited by comparison. This research will explore how socio-economic status influences the friendship ties of high school students and their social position for one high school in the National Longitudinal Study of Adolescent Health (ADD Health). Socio-economic status is traditionally conceptualized as the social and economic standing of an individual or family in comparison to others in their society and defined using information about economic activities and familial background. Unlike previous studies that have used only one measure of socio-economic status to investigate the impact of socio-economic status on networks, this research will use three different types of socio-economic status to evaluate the impact of each type (occupation, education, income) on adolescent social networks. Additionally, this study uses three different network analysis techniques to determine which type of socio-economic status emerges as the most important aspect for adolescent friendships.

THEORY

Socioeconomic Status

Socio-economic status, in its various forms, predicts marriage, nonmarital fertility, family structure as a child, health, and death (Carlson & England 2011; McLanahan & Percheski 2008; Palloni et al. 2009). Due to the importance of socio-economic status for so many outcomes, it is important to explore how different measures of socio-economic status may influence the friendships of youth.

Socio-economic background is a combination of different aspects of a person's life and lifestyle that define where he or she compares to others in their society. Traits that define socio-economic status can include the types of jobs that people have, familial yearly income, consumption patterns, wealth and other factors that contribute to a class identity (Hout 2007). Specifically in the United States, researchers have found five classes exist in the United States: working poor, working class, middle class, upper middle class, and upper class. The lower classes have two parts: the underclass and the working class. The underclass is composed of people who struggle to get by. The working class is composed of individuals who tend to work in manual and low-paying service-sector jobs. The middle class is comprised of individuals who earn average incomes for the United States and often work in jobs that involve white-collar components, such as administrative positions. The upper-middle class is defined by the tendency towards having higher degrees, such as college degrees or graduate degrees. Lastly, there is the upper class, which is typically defined by its extremely high income (typically over $250,000 per year) (Gilbert 1998).

However, some sociologists question whether class boundaries can be as easily drawn as some assert. According to Michael Hout (2007), many Americans are hesitant or unsure of which aspect of their lives defines them. Typically, sociologists use the previous year's income or achieved education to decide where individuals fall in comparison with others in a given country, but simply having higher income does not automatically make one a member of the upper class nor does it make a recent college graduate a member of the working-class. As a result, Hout believes that defining people using the broad socio-economic classes mentioned previously is problematic.

Due to the difficulties defining socio-economic class even at the large scale level, researchers have proposed with other conceptualizations of class boundaries to enable them to include socio-economic class in analyses. One of the most famous approaches is the socio-economic Index, which assigns numeric values for familial income, education and occupational prestige prior to aggregating the three measures into a single index (Duncan 1961). However, Warren, Sheridan and Hauser (1998) find that indices that rank occupations by prestige are "dead" and that meaningful combinations of education, occupation and income may be obsolete in a time in which socio-economic class can be unclear. The authors also assert that all individuals do not share the same beliefs as to which aspect of socio-economic status contributes most to their class identity. Therefore, indices are a problematic way of taking socio-economic status into account when distinctions between social classes are important to the research question. Consequently, Krieger, Williams and Moss (1997) find that analyzing each aspect of socio-economic status separately enables researchers to examine unique contributions of each aspect of socio-economic status. Otherwise, these unique contributions may be masked within aggregate measures of socio-economic status. This is an important point, since previous research on social networks and socio-economic status has focused on only one aspect of socio-economic status or used an aggregate measure. Here we will rely on disaggregated measures of the three main components of socio-economic class, including occupation, education, and income.

Adolescent Friendship Research

Friendship is particularly important within the formative years of adolescence, because adolescents increasingly depend on their friends as their primary form of social support during their critical developmental years (De Goede et al. 2009). This reliance on peer networks can be both positive and problematic since friends can provide crucial emotional support but can also influence adolescents to adopt risky behaviors (La Greca, Prinstein, & Fetter, 2001 and Osgood et al., 2013). However, forming close ties with other teenagers may result in victimization and aggression once the relationship goes sour (Felmlee and Faris 2014).

Scholarship regarding adolescent friendships is facilitating the use of social network analysis. Unlike traditional data, social network data help to more objectively determine what a school looks like socially by asking students about their friends and enemies (Borgatti 2013). Using those kinds of data, researchers are able to determine a student's position in the school hierarchy prior to conducting analyses with more information on a student's background and opinions. Students who are more popular within a given school are considered to be "central" within a school's network, which is why popularity will be referred to as centrality for this paper (Freeman 1978). Centrality is important to understand within social networks because more popular teens are more likely to use substances even after controlling for past substance use (Moody et al. 2011). Lastly, understanding the role of centrality within networks gives researchers a better understanding of how adolescent friendship processes operate within a larger context.

Socio-economic Status in Adolescent Networks

Previous research on socio-economic status in adolescent networks can be traced to Hollingshead and Coleman, who studied how adolescent networks operate and the role of family background in network related phenomena. Hollingshead's *Elmtown's Youth,* which was originally published in 1949, revolutionized how socio-economic status was perceived by social scientists. Hollingshead uncovered strong links between adolescent behavior and adolescents' socio-economic backgrounds. Using information from members of the community about the socio-economic background of adolescents in schools, Hollingshead characterized the friendships by similarity in socio-economic background. Hollingshead found that the composition of friendship groups (or cliques) was strongly associated with his prescribed categories of socio-economic status, and that the majority of friendship ties occurred intra-class. Overwhelmingly, the majority of friendship ties were between students of equivalent classes—and even when friendships crossed class boundaries, a third of friendship ties tended to extend to the proximal class above or below (Hollingshead, 1975). Cohen (1979) later analyzed this same data set using Freeman's index of segregation and found a high degree of segregation (0.6 for boys) by socio-economic class among youth (Freeman 1979). Using Hollingshead's earlier data, a high degree of segregation would indicate that students were more likely to only associate with other students within the same socio-economic class.

James Coleman re-studied Elmtown in his classic book, *The Adolescent Society (1961)*. Unlike Hollingshead, who focused on one specific community, Coleman studied Elmtown along with nine additional high schools taken from various communities throughout the United States. Coleman argued that adolescent cultures are distinct from the larger adult society and that indicators of social status relevant to adolescents (e.g., good looks or possessing a car) may be more important than social status derived from their parents. He does not entirely reject the notion that socio-economic background may be influential, but he does not find a coherent pattern throughout all the schools in his study in terms of socio-economic background and its importance for friendship ties. For certain measures, including socio-metric centrality (an older form of centrality) derived from all the students in 10 schools, Coleman found that youth with college-educated fathers were overrepresented among student leaders in all but one school. Overall, he found that family backgrounds mattered most for sociometric centrality in the peer networks of middle-class schools (Coleman, 1961). However, no study to date has re-examined the role of socio-economic status with centrality analyses specifically.

Recent research on socio-economic status in networks

A handful of studies have examined socio-economic status, more generally, since Hollingshead's and Coleman's early work. Ennett and Bauman (1996) explored how demographic traits manifest within adolescent networks by analyzing cliques within five schools. Their dataset consisted of network and demographic data collected from adolescents enrolled in five schools in one county in North Carolina as part of a 1980-1981 panel study (Ennett and Bauman 1994). The authors categorized students according to their mother's education level and found very similar results to Hollingshead's earlier findings. About half of clique ties were among adolescents from similar socio-economic backgrounds despite only approximately a third of students having similar socio-economic backgrounds. Cliques are defined as groups where all students have ties to each other (Wasserman and Faust 1994). Most of the cliques were composed of students from the same background or contained students whose parents possessed an education level one level below or above their own level. The authors dichotomized the education level of a child's mother into "high" or "low" levels of education without discussing how this affected the results. Since the analyses were largely exploratory, it's not clear how using different types of socio-economic variables, such as occupation or income, would influence analyses.

In spite of the intriguing findings found previously, further research is needed see whether Ennett and Bauman (1996)'s clique findings extend to friendship dyads, or that is, friendship ties between two given students at a school. Friendship dyads are different from cliques since the friendship dyads themselves are the basis of larger groups of students. Using information from multiple friendship dyads, it is possible to aggregate the friendships into groups of students, who are all friends with each other. Since the dyad is the lowest level grouping possible, it is important to examine whether socio-economic status affects friendship patterns at this most basic level.

Previous research has found that race and gender are strong bases for understanding adolescent friendship patterns, yet previous studies do not control for these demographic similarities (Faris and Felmlee 2011; Moody et al. 2001). Additionally, the earlier research does not control for tendencies towards triads, the density of the network or other structural properties common to networks as suggested by McPherson, Smith-Lovin, and Cook (2001).

A recent paper by Bojanowski and Corten (2014) on segregation in social networks used various segregation measures to evaluate classroom networks in Poland. Using socio-economic status index, the authors compare socio-economic status homophily by measuring it using an index of socioeconomic status with two (high v. low) and three categories. Homophily is defined as the tendency for people to associate with those most similar to themselves within networks (McPherson et al., 2001). The authors find that by using a three category variable to represent the socio-economic status of students instead of a dichotomized variable, they identified more socio-economic segregation within the same classrooms. Similarly, Bojanowski and Corten suggest that analyses using three categories results may be more robust. Due to the focus of the paper, Bojanowski and Corten do not explore further implications of this finding nor show analyses beyond segregation measures. However, the authors do find robust effects for socio-economic status as a measure for , within classrooms.

Role of Socio-economic Status in Networks

Since previous research does not elaborate on how socio-economic status works in modern friendship networks beyond potentially being the basis for friendships, here we consider how socio-economic status may manifest itself within the school hierarchy in terms of centrality. There are many forms of network centrality, but we will be focusing on three forms that take into account different aspects of networks: degree, eigenvector, and betweenness. Wasserman and Faust (1994) define degree centrality as the number of ties that a given node (or student) has to other students. This ends up giving higher values to students who have more connections to other students. Eigenvector centrality is calculated through a complex calculation involving the average distance between a student and everyone else in the network. Students with higher values for eigenvector centrality will have more connections with other influential students within the network. Betweenness centrality measures the number of times that a student "falls" on the shortest path between two nodes across the network, which is a unique measure because it takes into account overall network structure (Wasserman and Faust 1994).

If socio-economic status is a positive and highly significant predictor of network centrality, we can consider two potential possibilities as to how socio-economic status may manifest itself within social networks. Karl Marx asserted that class is the foundation for understanding inequality within modern society—and that the bourgeoisie, otherwise known as the upper class, has a tendency towards dominating institutions (Marx and Engels 1978). Based on Marx's work, we would expect privilege, in a number of forms, to manifest itself in the hierarchy of schools themselves. As a result, we may find that students from the highest socio-economic status groups (e.g. students with parents in white-collar occupations, higher family incomes and higher levels of achieved education) would be more central in the school hierarchy, with the assumption that higher socio-economic status benefits students.

On the other hand, socio-economic status also might be influenced by the dominant socio-economic structure within a given area. Based on Hollingshead (1975), we might expect that in a largely working-class school, for example students from working-class backgrounds may be more influential within the school hierarchy. Therefore, we would expect that students from backgrounds representative of the overall socio-economic structure of the community will be overrepresented within the most central positions. Since the majority of the students, in the case of Jefferson High School, come from lower middle class homes where most parents have less than a college degree, work in blue collar occupations, and earn close to the average income for respondents in the survey, it may be that more central students came from lower middle class backgrounds.

Contributions of this paper

Unlike most previous studies that have relied on one social network method to explore socio-economic status as a stratifying mechanism in networks, this paper will use three different social network analysis methods to analyze socio-economic homophily and how it manifests itself within the structure of the school hierarchy. The three techniques used will be Ordinary Least Squares regressions, Quadratic Assignment Procedure analyses, and Exponential Random Graph Models. Ordinary Least Squares regression will be used to account for central positions within the school hierarchy using information about students' backgrounds. The Quadratic Assignment Procedure will be used to predict dyadic friendships using information on how similar pairs of students are to each other. Lastly, exponential random graph models (ERGMs) will be used to predict the network given general network tendencies, controlling for other kinds of homophily, and examine socio-economic status homophily between pairs of students. Previous models examining similarity have not controlled of the tendency for students to resemble each other in terms of age, gender, grade, and race. The QAP model and ERGM models will include controls for similarity on the basis of these traits. Lastly, this analysis will use three different socio-economic status types (parental occupation, family income and parental education) instead of just one type of socio-economic status. All analyses will focus on one lower middle-class community with 880 students from Jefferson High School in the National Longitudinal Study of Adolescent Health (ADD Health) (Mullan Harris 2013).

Hypotheses

This study hopes to examine the relationship between socioeconomic status and adolescent social networks. Based on Coleman (1961), Hollingshead (1975) and Ennett & Bauman (1996), we would expect that students would have a tendency towards homophily on the basis of socioeconomic status. Although researchers have not tested the role of socioeconomic status in dyadic friendships previously, we would expect that adolescents will be more likely to choose friends similar to themselves, which leads us to hypothesis one:

Hypothesis 1: Students at Jefferson High School are more likely to have friends who come from similar socio-economic backgrounds.

Secondly, this study hopes to move beyond the current findings by accounting for other types of similarities commonly found in adolescent friendship networks (Felmlee 2014; Moody 2001; McPherson 2001). Without controlling for effects related to gender, race, age and grade, results that find robust socio-economic status homophily effects may prove to be misleading since the omitted variables may be influencing the results. For instance, a strong finding for income homophily may be confounded by the omission of race in the model if racial homophily is the actual basis for similarity. This paper will use a series of commonly used control variables to test whether the relationship between socioeconomic status and friendship remains even with the inclusion of control variables in hypothesis two:

H2: Socio-economic similarity in terms of friendships will remain important even after accounting for other types of similarities common in adolescent friendships (gender, race, grade and age).

Based on Hollingshead's findings on the "leading crowd" in Elmtown, and Coleman's findings regarding centrality in *The Adolescent Society* (1961), we would expect that socioeconomic status is an important predictor of centrality more generally, which leads us to hypothesis three:

Hypothesis 3: *Socio-economic status will be a positive and highly significant predictor of three types of centrality (Degree, Eigenvector, and Betweenness).*

However, it’s not clear how centrality will be influenced by socio-economic status, which results in two competing hypotheses. In the first hypothesis (3a), we would expect that based on Marx (1978), students from more prestigious socioeconomic-groups would also be privileged in the school hierarchy. On the other hand, the location of a school in a primarily lower middle class community may influence the school hierarchy, which may account for hypothesis 3b.

Hypothesis 3a: Students from higher socio-economic status groups will be more central in the school friendship network.

Hypothesis 3b: Students from working-class families will be more central in the school friendship network.

Data, Methods and Measures

Data

This research examines Jefferson High School, which is a school within the National Longitudinal Study of Adolescent Health, more commonly known as ADD Health. The survey was first administered in the 1994-1995 school year, and the survey followed the adolescents from this school for three additional waves. Jefferson High School is situated in a midsize mid-western city. The community is relatively isolated from other cities, and Jefferson is the only public high school in the town. According to Bearman (2004), this high school is pretty typical compared to other ADD Health schools, despite the high prevalence of cigarette and alcohol use. Approximately 880 students attend the school and all of these students were included within ERGM analyses, since exponential random graph models are capable of estimating networks even with missing data. However, not all students had full network data and in-home parental interviews with complete information on the dependent variables. This resulted in a reduction of the sample size for the Quadratic Assignment Procedure and centrality analyses to 540 students. Despite this limitation, at least one socio-economic variable is available for the majority of students at Jefferson.

This school is ideal for investigating the role of socio-economic status in networks because it involves students from a range of socioeconomic backgrounds and is the only public school in the surrounding area. As noted by Frankenberg (2013), school districts are drawn according to geographic boundaries, which are often influenced by the socio-economic characteristics of an area. As a result, many communities have multiple public school districts that may be relatively homophilous in terms of the socio-economic status composition. As mentioned previously, Jefferson is the only public high school in the town, which means that it likely displays more socio-economic status diversity than in many school districts. This aids analysis since it means that there are more students from both ends of the socio-economic spectrum and there may be more opportunities for students to interact with students from different socio-economic backgrounds.

Due to ADD Health sampling, this school compromised a full network since all students were sampled about their friendships with other students in the school. As noted by Hanneman (2005), a full network allows researchers to examine relations between all students in a given school. On the other hand, Borgatti (2013) explains that the collection of full networks comes with many caveats due to boundary issues (e.g., not receiving full participation from the subsample of interest and the need to limit analysis to a small social setting where all interactions can be captured) and survey issues (e.g., limiting the number of friendships that a respondent can list).

Dependent Variables

Three different types of analyses will be conducted, and the dependent variable will vary for each type of model. With the Quadratic Assignment Procedure, the dependent variable will be a dyadic friendship tie between students. With Exponential Random Graph Models, the dependent variable is a simulated network that closely resembles the network of the school. Lastly, for the Ordinary Least Squares regression, the dependent variable will be each of the three centrality measures mentioned previously.

For the Quadratic Assignment Procedure, where we are predicting the friendship tie itself, we are interested in friendships at the dyadic level. Instead of evaluating how many friendships exist within cliques or within the overall school, the relationship of interest is the friendship tie between two given students (student i and student j). In the ADD Health data collection, students were given the opportunity to nominate up to five friends of each gender within his or her school. Not all students reciprocated a friendship nomination and students were limited in the number of nominations within gender. Here we consider the student i and j to be friends if student i claims that student i is their friend. Previous papers using ADD Health with exponential random graph models suggest treating friendship nominations in this way (Goodreau et al. 2008). Faust (2006) also comments that many friendship networks among humans tend to be relatively sparsely populated with ties between actors, which is the case with this dataset. The average density of Jefferson's friendship network is 0.006, which tells us the ratio of realized friendship ties to the number of possible friendship ties. On average, Jefferson students received 11.78 nominations with the most popular student at Jefferson receiving 46 nominations from his/her peers.

Network centrality was the dependent variable in the OLS regression. Wasserman and Faust (1994) define degree centrality as the number of ties that a given node (or student) has to other nodes. This ends up giving higher values to students who have more connections to other students. It is determined by the formula below:

$$C_D^{'}(N_i) = \frac{\sum_{1}^{g} X_{ij} (i \neq j)}{g - 1}$$

In the formula, the numerator is the total number of ties that individual i has within the network while g is the total number of possible ties in the network. It does not take into account the overall position of a student in the school hierarchy or the importance of a student's ties to other popular actors (e.g., prom king.)

Eigenvector centrality is a measure of centrality that more heavily weighs students who are directly connected to other well-connected actors. This measure takes into account the entire structure of the school's social network. Eigenvector centrality may give inflated values for members of the school network who lack connections to other students. It is calculated using the following formula:

$$\lambda c(v_i) = \sum_{j=1}^{n} a_{ij} c(v_j) \quad \forall i$$

In the formula, eigenvector centrality uses a scaled eigenvalue (λ) to weight nodes with more prestigious connections with higher centrality scores. Typically, we take the largest eigenvalues when weighing nodes. The eigenvector centrality of node v_i is the sum of adjacent centralities of i where j is the nearest neighbor (Bonacich 1972).

Freeman's betweenness centrality measures the degree to which a student falls onto the shortest path between two other students. It is calculated using the formula below:

$$C_B(N_i) = \sum_{j<k} g_{jk}(n_i) / g_{jk}$$

Betweenness centrality is the sum of probabilities over all pairs of actors (or students) that a path (e.g. actor j to actor k) takes a specific route, with the assumption that the shortest path will always be chosen. In the formula, g is the number of geodesics between actors j and k. A geodesic is the shortest path between two nodes while a path is a 'line' between two nodes where all nodes and lines are distinct (Wasserman and Faust 1994). The maximum number for betweenness is when the *i*th actor is part of all geodesics for a given social network. Students who bridge together otherwise unconnected friendship cliques will have higher values of betweenness centrality (Hanneman 2005).

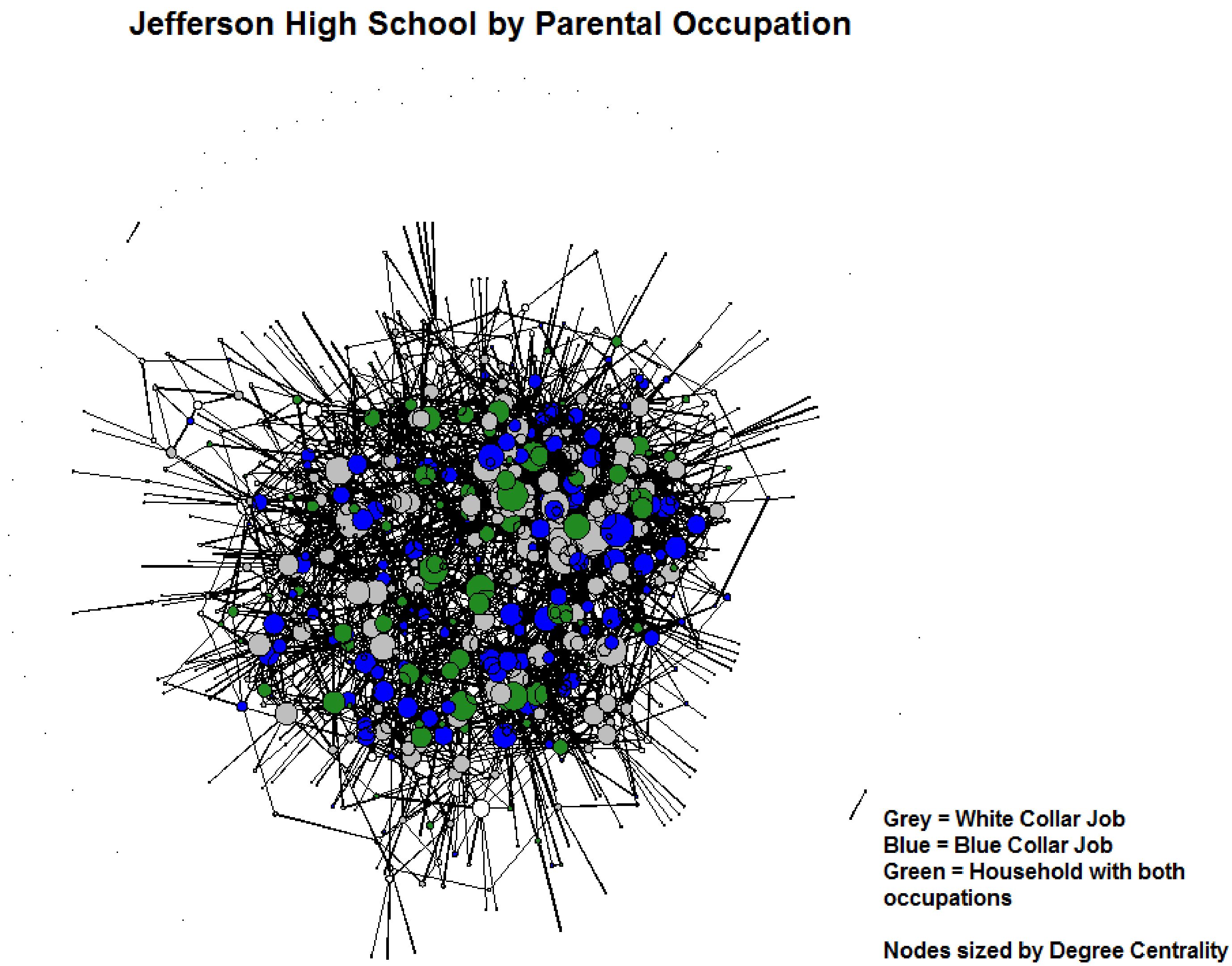
Independent Variables

All information on socio-economic status came from the Wave One In-Home Parental Questionnaire, which utilized the interviews with adolescents' parents. All three socio-economic variables are considered as separate dependent variables, because they assess different aspects of socio-economic status as suggested by Warren, Sheridan and Hauser (1998).

The coding for parental occupation was based on other articles on occupational status that used the National Longitudinal Study of Adolescent Health data, including Balsa et al. (2004) and Humensky (2010). Since not every parent worked within two-parent households, and there are households with only one parent, the coding focused on the parent within the household who was employed. White-collar work was defined by both parents working (or one parent if the student lived in a single-parent home) in professional, managerial, technical, office, or sales jobs. Blue-collar work was defined by one (if the student resided in a single-parent home) or both parents working in food service, personal service, military, security, construction, transportation, a factory, a farm, a craft, or the fish industry. The third category, which was referred to as a mixed-occupation household, represented students in which one parent worked in a blue-collar industry and the other parent worked in a white-collar industry.

Figure 1 shows Jefferson High school by parental occupation. The graph is sized according to degree centrality, where larger nodes are students who are more prominent in the school network. Gray represents students with white-collar backgrounds while blue represents students from blue-collar backgrounds. Green represents students who are from mixed-occupation households. The graph shows that within Jefferson high school, there is a significant mixing of students from different socio-economic status, as defined by occupation. There isn't a clear picture that emerges based on the graph.

Figure 1: Jefferson High School by Occupation



Parental education was coded into five categories focusing on the highest level of education achieved by a parent in a given household: less than high school; high school degree; associate's degree (or some college); college Bachelor's degree; and an advanced degree. Ennett and Bauman (1996) had a similar coding scheme, although they collapsed the categories into a low and high education category. In the interest of retaining more information on students from more educated households, the author has opted to keep these categories distinct. However, there were too few cases in the highest categories to use more fine-grained categories.

Household income was coded with the 25th percentile as a cut-off for each category. Alternative approaches were considered, however, too few cases were in the highest category to allow matching on the basis of socio-economic status. Therefore, the variable is coded as ranging from 0th-25th percentile, 26th-50th percentile, 51-75th percentile and 76th-100th percentile. The percentiles were created in comparison with the rest of the in-home ADD Health sample, which is representative of the United States population.

Control Variables

Based on previous research, the author includes a number of control variables that have been found to influence adolescent friendship networks (Moody, 2001; McPherson et al. 2001; Faris & Felmlee 2011). These include age, grade in school, race, and gender. Since some of these variables may influence friendship ties, we must control for common forms of homophily to fully understand adolescent friendship networks. Adolescents increasingly depend on their friends for social support as they grow older, which is why age is an important control variable (De Goede et al. 2009). Due to overlap in classes and being in the same cohort at the beginning of high school, researchers have found that students are more likely to be friends with others in the same grade (McPherson, Smith-Lovin & Cook 2001). Additionally, Moody (2001) finds that racial segregation remains a reality of adolescent friendship networks, which is why it is important to account for racial homophily. Last, gender is included as a control variable, because gender differences are common in informal school friendship networks (Felmlee and Faris, 2014).

Methods

QAP (Quadratic Assignment Procedure)

Social network data have special attributes that make it difficult to apply conventional data analysis techniques. The main difference lies in the violation of the independence assumption, which requires that observations are independent from other observations. However, in social network analysis, the data are relational and therefore dependent on other observations. The Quadratic Assignment Procedure (QAP) utilizes the relational structure of social networks and enables researchers to produce significance tests that are unbiased by structural autocorrelation between pairs of students, otherwise known as dyads (Krackhardt 1988). This technique will be used to test hypotheses 1 and 2.

A strength of applying the Quadratic Assignment Procedure to the topic addressed here is that it takes into account information about the potential student dyads prior to predicting friendships among dyads. This approach goes beyond traditional forms of analysis that only use individuals' traits to predict friendships. Instead of just considering a student's own background, for example, the QAP allows consideration of the *similarities* in background between students that may influence whether students are friends.

The dependent variable in a Quadratic Assignment Procedure is the symmetric matrix of ties between actors. Here the dependent variable consists of the friendship ties within Jefferson High School with the same number of rows and columns as the number of students. If a friendship exists between student i and student j, a value of 1 will be in the ith row and jth column. Otherwise, a value of 0 will represent the lack of a friendship between these two students. The independent variables are matrices representing whether students resemble each other on the basis of several traits, including gender, race, grade, and each of the socio-economic status variables. For instance, if student i and student j are members of the same family income category, then a value of one will be included in the ith row and jth column of the matrix that represents income as an independent variable; otherwise the value will be zero.

First, the QAP regression regresses the matrix of friendships onto the matrices for the independent variables using Ordinary Least Squares. The initial step provides the estimates for the regression coefficients reported for the QAP regression. In the next step, the rows and columns of the dependent variable matrix, friendship, are permuted. Then, the regression coefficients are recalculated and the permutation procedure is repeated. This aspect of the regression resembles bootstrapping and is the basis for the p-values associated with the regression coefficients. Statistical significance represents the proportion of results from random permutations of matrices that generate OLS coefficients that are as high as those obtained on the basis of the original dependent variable matrix. Here we use 5,000 permutations via Double-Dekker Semi-Partialling, which is a particularly useful permutation method under the conditions of possible skewness and spuriousness (Dekker, Krackhardt, and Snijders 2007).

However, neither the QAP correlation nor MR-QAP takes into account structural properties that may influence networks beyond tendencies towards similarity at the dyadic level. Additionally, the QAP correlation and MR-QAP procedure ignore all cases with missing data, which brings the number of valid cases from 880 students to only 540 students.

ERGMS

An exponential random graph model (ERGM) is a class of models that simulate real-life patterns of association by generating random graphs of possible networks with the same network attributes, which allows researchers to examine both the structure of the networks and the attributes of the actors (or people) in the networks. The ERG models, the outcome of interest is the tie (or friendship, in this study) between two actors (or students). This is a dichotomous outcome, since friendships are treated as existing or not existing between students. For this reason, ERG models resemble logistic regression models. In standard logistic regression models, however, observations are required to be independent of each other, which is not the case with any social network data. Instead, ERG models assume that all observations (or ties) are dependent on other observations. The ties that are predicted by ERG models are conditional on the other aspects of the network (Hunter et al. 2008). The advantage of ERG models over traditional methods of analysis is that they take into account the dependent nature of social network data.

The probability of a tie between actors can be represented through the formula shown below (Goodreau 2007):

$$P(Y = y) = (1/\kappa)exp\left(\sum \eta_A \, g_A(y)\right)$$

where A is the set of actors, Y is the sociomatrix with the same number of rows and columns as the number of students, Kappa (κ) is a normalizing constant, and Eta $(\eta)_A$ represents the parameters specific to the user-specified terms included in the ERG model, which are estimated using Markov Chain Monte Carlo (MCMC) approximation. Lastly, $g_A(y)$ represents the user specified terms within the network that the ERG model is simulating. This can include information about the number of friendships in the network and information about the students. In this study, the user specified term focuses primarily on homophily on the basis of gender, grade, race, household income, parental education level, and parental occupation type. The inclusion of data about both networks and the ties allows ERG models to generate random networks with the same attributes as the network it is attempting to simulate, and compare the simulated networks to the observed network using statistics generated by maximum likelihood and MCMC methods.

In the case of this study, the friendships are non-directed, which changes the meaning of specific terms in the formula. As noted by Hunter et al. (2008), non-directed network ties force models to have dyadic independence, which means that the probability of a friendship between two students is only conditional on both students in the dyad and their attributes. As a result of dyadic independence, the likelihood function is much simpler, which prevents dyadic independent ERG models from suffering from degeneracy. Degeneracy is when simulated models of the network are not a good fit to the real social network. As a result, ERG coefficients may become biased and the model may not even converge. This is a relatively common issue with most ERG models. However, dyadic independent models are simpler to estimate and therefore, less likely to encounter degeneracy. More about degeneracy can be found in Hunter et al. (2008).

Unlike other common social network analysis methods that do not approximate missing data, ERG models calculate the probability that information about respondents and friendships is missing. Based on both probabilities, ERGMs will correct for missingness when calculating network statistics (Handcock and Gile 2010). Since it is relatively uncommon to have complete information about all actors and their ties, this is a significant advantage of ERG models over other traditional forms of analysis and common algorithms to estimate Multiple Regression Quadratic Assignment Procedures (such as that employed in UCINET), which do not account for missing data. For this reason, the ERG models in this study utilizes data from all 880 students, even those with missing data who are not included in other analyses.

In addition, ERGMs allow researchers to account for similarity in terms of traits that actors may share. Homophily, or similarity, is determined at the dyadic level for ERGMs. This means that for homophily to exist, the dyad needs to share similar traits in terms of a single attribute (e.g. same gender or level of parental education). In the case of this paper, homophily effects include gender, race, grade, parental education, household income, and parental occupation. The two terms included to capture homophily effects with the ERG model are "nodematch" and "nodefactor." Nodematch measures the general tendency towards homophily, whereas nodefactor measures the tendency for members of a specific attribute group to become friends (Morris et al. 2008). For example, there may not be tendencies towards racial homophily more generally, however African American students may be more likely to choose friends similar to themselves in terms of race. Similarly, there may be homophily in terms of parental education and students may be more likely to choose friends whose parents have similar levels of education to their own. For this reason, ERGMs will be used to test hypotheses one and two.

OLS

For analyses predicting centrality using degree, betweenness and eigenvector centrality, Ordinary Least Square Regression analysis was used. As detailed by Alison (1999), multiple regression is appropriate for continuous numeric dependent variables, which is the case with centrality measures. This analysis will be used to test hypotheses 3, 3a, and 3b on SES and centrality.

Results

Descriptive Statistics

The descriptive statistics are presented in Table 1. The school is split fairly evenly between girls and boys (46% female). Similarly, little racial diversity exists within the school, since only 6% of the school is part of an underrepresented ethnic or racial group [non-white or Hispanic]. Despite the lack of racial diversity, Jefferson has a fair amount of socio-economic diversity, ranging from the least advantaged backgrounds to the most advantaged. Most of the respondents, however, come from lower- to middle-class socio-economic backgrounds, falling within the 25-75th income percentiles, which would resemble Hollingshead's Elmtown. In terms of occupation, there is a split between types of parental backgrounds, with 48% of students from white-collar backgrounds, 30% of students from blue-collar backgrounds, and 9% of students coming from homes where parents work in both blue- and white-collar jobs.

Table 1: Descriptive Statistics

Variable	mean	SD	min	Max
Degree Centrality	0.009	0.006	0	0.033
Betweenness Centrality	0.006	0.007	0	0.042544
Eigenvector Centrality	0.032	0.052	0	0.344658
Number of ties (undirected)	11.786	9.084	0	46
Gender	0.473	0.500	0	1
Minority Race/Ethnicity	0.060	0.238	0	1
Grade	10.293	1.089	9	12
Age	16.337	1.195	14	19
Parent Blue Collar Occupation	0.467		0	1
Parent White Collar Occupation	0.311		0	1
Parent Mixed Household Occupation	0.192		0	1
Parent Other Occupation	0.040		0	1
Parent Less than HS Education	0.035		0	1
Parent HS Degree Education	0.277		0	1
Parent Some College Education	0.356		0	1
Parent College B.A. Education	0.210		0	1
Parent Graduate School Education	0.121		0	1
Parent 0-25th percentile Income	0.130		0	1
Parent 26th-50th percentile Income	0.303		0	1
Parent 51th-75th percentile Income	0.333		0	1
Parent 76-100th percentile Income	0.235		0	1

Quadratic Assignment Procedure Correlations

To test if socio-economic homophily exists in Jefferson High School, we used the QAP correlation procedure for the friendship matrix and socio-economic mixing matrices. A significant positive correlation between the friendship ties of students at Jefferson High School and the socio-economic similarity of peers would indicate that socio-economic homophily does exist. Using the occupational similarity matrix and the friendship matrix, there is a significant positive relationship between the matrices using a one-tailed test (p-value of 0.04). This result partially confirms the first hypothesis that one form of socio-economic homophily exists at Jefferson High School, suggesting that youth are significantly more likely to befriend a student from a similar parental occupational background as themselves. QAP correlations were used with the other socio-economic status matrices, but neither income nor education similarity was significant.

Centrality Analyses

An Ordinal Logistic Squares Regression was used to test which student attributes significantly predicted the likelihood that a student was a central member of the Jefferson school hierarchy. Since this topic is exploratory, three common forms of centrality found in the social networks literature were used to compare the robustness of socio-economic status. Here, we conduct analyses on degree centrality, eigenvector centrality, and betweenness centrality. It should be noted that the centrality measures are highly skewed compared to standard distribution.

As defined by Wasserman and Faust (1994), degree centrality refers to the number of ties that a given student has with other students after taking into account the largest number of ties possible within the school's network. After controlling for gender, grade, parental income, race, and parental education, the only significant predictors of a student having high degree centrality are the occupations of a student's parents. A student from a blue-collar background is significantly less likely to have higher degree centrality (p-value of .018) than students from white-collar backgrounds. Results are shown in Table 2. As a result, we find support for hypothesis 3b, which suggests that privilege may manifest itself within the school hierarchy.

Table 2: Degree Centrality Regression with Income, Education, Occupation, Gender, and Race

Variables	Coef.	Std. Err.	P>t
Parental Income			
26-50th Percentile	-0.0019	0.0015	0.2240
51-75th Percentile	-0.0009	0.0016	0.5870
76th-100th Percentile	0.0002	0.0016	0.9030
Parental Education			
HS Degree	0.0052	0.0028	0.0610
Some College	0.0029	0.0028	0.2900
College B.A.	0.0042	0.0028	0.1390
Graduate Degree	0.0028	0.0029	0.3470
Parental Occupation			
Blue Collar Background	-0.002*	0.001	0.0350
Mixed Occupation Household	-0.0010	0.001	0.3390
Other Occupations	-0.0029	0.0021	0.1670
Gender	-0.0003	0.0008	0.6630
Race	0.0019	0.0019	0.3000
Constant	0.0082	0.0029	0.0040

^p<.05 (one-tail test); * p< .05; ** p< .01; *** p < .001 (two-tail test)
Reference group: Male, white, non-Hispanic, parent(s) dropped out from high school, lowest familial income percentile, and parent(s) in white-collar occupation.

Next, analyses were conducted for eigenvector centrality, as shown in Table 3. This type of centrality gives preference to students who are more connected to other well-connected actors (Hanneman 2005). Compared to students in the lowest income categories, students from the second highest income quintile (25-50th percentile) and the middle income percentile (51-75th percentile) were less likely to be popular ($p<.05$). We also see marginal effects for friendship homophily in terms of minority students and students whose parents have Bachelor's degrees. This result provides support for hypothesis three and hypothesis 3b since we see that socio-economic status by income matters for centrality even after controlling for other variables.

Table 3: Eigenvector Centrality Regression with Income, Education, Occupation, Gender, and Race

Variables	Coef.	Std. Err.	P>t
Parental Income			
26-50th Percentile	-0.022*	0.011	0.044
51-75th Percentile	-0.026*	0.011	0.020
76th-100th Percentile	-0.011	0.012	0.363
Parental Education			
HS Degree	0.026	0.020	0.186
Some College	0.022	0.020	0.253
College B.A.	0.037^	0.020	0.066
Graduate Degree	0.018	0.021	0.377
Parental Occupation			
Blue Collar Background	-0.009	0.007	0.211
Mixed Occupation Household	-0.001	0.007	0.935
Other Occupations	-0.019	0.015	0.194
Gender	0.001	0.005	0.919
Race	0.024^	0.013	0.067
Constant	0.026	0.020	0.195

^p<.05 (one-tail test); * p< .05; ** p< .01; *** p < .001 (two-tail test)
Reference group: Male, white, non-Hispanic, parent(s) dropped out from high school, lowest familial income percentile, and parent(s) in white-collar occupation.

Betweenness centrality is a form of centrality that more heavily takes into account connectedness between students and gives preference to students who connect otherwise disconnected groups (Hanneman 2005). In Table 4, an Ordinary Least Squares regression is used to predict betweenness centrality. In the case of betweenness centrality, there is not a clear pattern that emerges for predicting centrality, at least based on the demographic traits of students.

Table 4: Betweenness Centrality Regression with Income, Education, Occupation, Gender, and Race

Variables	Coef.	Std. Err.	P>t
Parental Income			
26-50th Percentile	-0.001	0.002	0.382
51-75th Percentile	-0.001	0.002	0.709
76th-100th Percentile	-0.001	0.002	0.743
Parental Education			
HS Degree	0.003	0.003	0.231
Some College	0.002	0.003	0.546
College B.A.	0.002	0.003	0.488
Graduate Degree	0.002	0.003	0.548
Parental Occupation			
Blue Collar Background	-0.001	0.001	0.187
Mixed Occupation Household	0.000	0.001	0.968
Other	0.000	0.002	0.989
Gender	0.000	0.001	0.584
Race	0.001	0.002	0.544
Constant	0.005	0.003	0.108

^p<.05 (one-tail test); * p< .05; ** p< .01; *** p < .001 (two-tail test)
Reference group: Male, white, non-Hispanic, parent(s) dropped out from high school, lowest familial income percentile, and parent(s) in white-collar occupation.

Based on the centrality analyses, we find mixed support for hypotheses 3a and 3b although socio-economic status is predictive of centrality in two out of three models. Besides finding marginal effects for race in one model, we do not find robust effects for gender or race, as we would expect based on the previous literature. Considering the important role of race and gender homophily in adolescent networks, we would expect that being a minority or female may affect a student's centrality. However, we do not find that gender or race seems to play an important role in understanding centrality. As noted previously, these results only use a subsample of adolescents that provided full data on all variables, and missing data might be biasing these results. For this reason, it may be appropriate to use a social network analysis technique that can ignore missing data, such as MR-QAP or ERGMs.

Multiple Regression Quadratic Assignment Procedure

Multiple Regression Quadratic Assignment Procedure was used to compare the effects of race, income, gender, grade, and occupation similarity on friendship ties. As noted previously, the MR-QAP procedure effectively scrambles the matrices in order to obtain unbiased significance tests. . Results (see Table 5) can be interpreted like a logistic regression. Students in the same grade are significantly more likely to be friends ($p<.001$). This is an important finding since grade similarity among friends is common to most adolescent friendship networks. Most importantly for this paper, income similarity (grouped by percentiles) positively predicts friendship ties between students in similar income categories after controlling for race, gender, grade, and occupation similarity ($p<.04$ using a one tailed test). These results suggest that familial income similarity plays an important role in determining the friendship ties that exist at Jefferson High School as suggested, which provides evidence to support hypothesis 1, which focuses on students choosing friends similar to themselves in terms of socio-economic status similarity. Based on previous research on racial and gender homophily, we would expect that the matrices representing similarity in terms of minority status and gender would be significant in the regression; however, they are not significant within the MR-QAP. It should be noted that parental occupation is marginally significant when specified differently within models, which suggests that income and occupation may measure different facets of socio-economic status, which fits with findings by Warren, Sheridan and Hauser (1998) on how including the different socio-economic status variables have unique components.

Table 5: Multiple Regression QAP (MR-QAP) Predicting Friendship Ties

	Coefficient	Standard Error
Income Similarity	0.001^	0.000
Gender Similarity	0.000	0.000
Grade Similarity	0.005***	0.002
Race Similarity	0.000	0.000
Occupation Similarity	0.000	0.000

^p<.05 (one-tail test); * p< .05; ** p< .01; *** p < .001 (two-tail test)

ERGMS

Exponential Random Graph models allow for the estimation of models that include better mechanisms at understanding similarity within networks beyond similarities. Since we include different versions of homophily, we can take account the relative size of each group within a school and tendencies towards within-group preference (Hunter 2006; 2008).

The first ERGM model only includes the socio-economic status variables and a term for edges. The edges term, which is a necessary term in ERGM model, is similar to an intercept in a linear regression model. It can be interpreted as the propensity towards ties in a given network (Morris et al. 2008). As shown in Table 6, all the education homophily terms are positive and significant at the 1% level. This indicates a tendency towards choosing peers from similar parental education, regardless of level. For example, a student with parents who attended graduate school are more likely to be friends with another student with the same background. For income, we do not see a trend toward familial income-based homophily more generally. However, students from the most well-off families (76^{th}-100^{th} percentile) have a tendency towards within-group selection (p<.001). We also see an overall tendency towards homophily on the basis of occupation (p<.001), although the effect is not as strong within occupation types (p<.05). This analysis supports Hypothesis 1.

Table 6: Exponential Random Graph Model with Socio-economic Status Variables Only

Coefficient	Estimate		Standard Error
Edges	-5.523	***	0.045
HS Degree Node Factor	0.280	***	0.048
Some College Node Factor	0.227	***	0.048
College B.A. Node Factor	0.186	***	0.055
Grad School Node Factor	0.256	***	0.064
Education Node Match	0.069		0.068
Income 26-50th Percentile Node Factor	0.096	*	0.046
Income 51-75th Percentile Node Factor	0.194	***	0.046
Income 76th-100th Percentile Node Factor	0.067		0.053
Income Node Match	0.143	*	0.064
Blue Collar Node Factor	0.090	*	0.039
Mixed Occupation Node Factor	0.181	*	0.043
Occupation Node Match	0.159	***	0.060

^p<.05 (one-tail test); * p< .05; ** p< .01; *** p < .001 (two-tail test)
Reference group: Parent(s) dropped out from High School, lowest familial income percentile, and parent(s) in white-collar occupation.

Next, we create a model with traditional homophily variables and occupational homophily. In Table 7, we see that students with parents in both blue- and white-collar occupations are significantly more likely to be friends with each other (p<.05). These findings fit the previous centrality and MR-QAP results since it shows that, in addition to homophily tendencies typically found in adolescent networks (grade, race and gender), we see a robust effect for parental occupation as a form of differential homophily. These results support hypothesis 2, which suggests that socio-economic homophily is important even when we control for other common forms of homophily.

Table 7: Exponential Random Graph Model with Occupation Only

Coefficient	Estimate		Standard Error
Edges	-5.561	***	0.043
Grade 10 Node Factor	0.034		0.043
Grade 11 Node Factor	0.041		0.045
Grade 12 Node Factor	0.113	**	0.039
Grade Node Match	0.091	^	0.054
Minority Node Factor	0.402	***	0.072
Race Node Match	0.633	***	0.060
Female Node Factor	0.073	*	0.030
Gender Node Match	0.075		0.047
Blue Collar Node Factor	0.000		0.038
Mixed Occupation Node Factor	0.119	**	0.043
Occupation Node Match	0.018		0.060

^p<.05 (one-tail test); * p< .05; ** p< .01; *** p < .001 (two-tail test)
Reference group: Male, Grade 9, non-Hispanic and white, and parents in white-collar occupation.

In Table 8, we add the traditional homophily variables to the socio-economic status homophily variables from the first ERGM model (Table 6). Many of the effects associated with socio-economic status homophily disappear in favor of strong tendencies towards racial and gender homophily (p<.001 and p<.01). Fitting with the Warren, Sheridan, Hauser (2008), we see distinct effects associated with different types of socio-economic status. There are positive and significant effects (at the 5% level) for mixed parental occupation homophily, familial income homophily at the 75th-100th percentile, and parental education homophily for children with parents with some college education.

Table 8: Exponential Random Graph Model with Occupation, Income and Education

	Estimate		Standard Error
Edges	-5.566	***	0.050
Grade 10 Node Factor	-0.030		0.038
Grade 11 Node Factor	-0.033		0.043
Grade 12 Node Factor	-0.032		0.045
Grade Node Match	0.095	^	0.054
Minority Node Factor	0.393	***	0.074
Race Node Match	0.604	***	0.067
Female Node Factor	0.086	**	0.030
Gender Node Match	0.073		0.047
HS Degree Node Factor	0.115		0.049
Some College Node Factor	0.064	*	0.049
College B.A. Node Factor	0.113		0.055
Grad School Node Factor	0.065		0.064
Education Node Match	0.042		0.068
Income 26-50th Percentile Node Factor	-0.003		0.046
Income 51-75th Percentile Node Factor	0.092	*	0.046
Income 76th-100th Percentile Node Factor	-0.029		0.053
Income Node Match	0.093		0.064
Blue Collar Node Factor	-0.012		0.040
Mixed Occupation Node Factor	0.113	*	0.439
Occupation Node Match	0.026		0.060

^p<.05 (one-tail test); * p< .05; ** p< .01; *** p < .001 (two-tail test)
Reference group: Male, grade 9, Non-Hispanic, white, Parent(s) that dropped out from high school, lowest familial income percentile, and parents in white-collar occupation.

Discussion and Conclusions

Using three different social network analysis techniques, we find that socio-economic status homophily is a potential stratifying mechanism for Jefferson High School. Based on the results from the ERGM models, centrality analyses and MR-QAP results, it may be beneficial to include at least one measure of socio-economic homophily in analyses predicting friendships among high school students. It seems that income, occupation, and education homophily have separate and distinctive effects, as shown by the centrality and ERGM analyses, which is why including more forms of homophily beyond typical control variables (grade, gender, age, and race) may be helpful for understanding friendship ties between students. As suggested previously, the inclusion of different facets of socio-economic status in network analyses may provide insight into how socio-economic status may not be a one-dimensional variable.

The findings more generally point to how socio-economic status manifests itself in unique and specific ways when parental education, income and occupation are included. Socio-economic status may differ from other demographic variables included because it can be more subtle. Unlike race or gender, there may not be an external way to know another child's socio-economic status beyond explicitly asking for the occupations, familial income, education level, and wealth level of another student's parents. Similarly, Americans do not fully identify with one social class based on one aspect of socio-economic status. Findings from ERGM analyses (Table 8) support the idea, because students do not select solely on the basis of one socio-economic trait. Part of this finding may be affected by various differential tendencies for families to define their class by different aspects. The tendency towards homophily for students from families with the highest incomes to befriend students with similar backgrounds may be related to these families defining their class by family income. Based on Iceland and Wilkes (2006), families with higher incomes may live closer to each other, and this may provide more opportunities for socialization outside of school. Additionally, children from the highest income percentile and the lowest income percentile were highest in terms of eigenvector centrality. It may be that having more disposable income allows these students to gain status within Jefferson High School.

We also see a tendency towards homophily for students with parents who work in blue-collar and white-collar occupations. It may be that these children have unique family situations that enable them to bond with each other or this may be a byproduct of "traditional" dual-earner homes where a father works in a blue-collar field and a mother works in a white-collar field. However, more nuanced occupational data are required to further explore how living in a dual-earner family with these unique traits may affect a child's friendships.

We do not find total class matching across the three categories of socio-economic status. Public schools were established in order to provide an education for all children within a given geographic area, regardless of socio-economic status. In theory, we would expect that public schools provide an opportunity for students to interact with other students outside of their own socio-economic realm. We do find support for this hypothesis in Jefferson High School, which suggests that there may be more contact between students of different socio-economic backgrounds.

The tendency to socialize with those different from ourselves may have far-reaching impacts on the adolescents as they mature into young adults. The ability to be comfortable around students from different socio-economic contexts may positively affect students as they grow older by expanding their empathy for those who are less fortunate and the size of their social networks. Perhaps, these individuals may be more likely to date someone who comes from a background different from their own and live in an area that is socio-economically diverse. However, future research must be conducted to find long-term effects.

Limitations

Analyses using MR-QAP and centrality are limited to the 540 students at Jefferson High School that had full network and attribute data, which may mean that results cannot be generalized beyond this sample. Additionally, the term for socio-economic similarity used in this paper for Quadratic Assignment Procedure is not a true measure of homophily as it does not control for structural tendencies in adolescent friendship networks. Lastly, this analysis cannot be generalized to friendship formation patterns since analyses only predict existing friendship ties using one wave of network data.

Future Research

Scholars should examine the social networks of the other ADD Health schools where students nominated at least two friendship ties on average. During the administration of the ADD Health social network survey, there was an error with the computer system where students were limited to one friend of each gender (Mullan Harris 2013). As a result, researchers typically focus on schools where students had the option of nominating up to ten names. Focusing on the broader ADD Health network would allow researchers to investigate schools with more racial diversity and different network structures. The use of schools with more racial diversity will allow a more detailed comparison of racial homophily with socio-economic homophily. Lastly, the role of subgroups and centrality should be more fully explored, in conjunction with substance use (alcohol, cigarettes, marijuana) outcomes.

References

Alison, P. D. (1999). Multiple regression. Thousand Oaks, CA: Pine Forge Press.

Balsa, A., French, M., & Regan, T. (2013, April). *Relative Deprivation and Risky Behaviors*. Retrieved April 11, 2014, from www.um.edu.uy/docs/working_paper_um_cee_2013_04.pdf

Bearman, P., Moody, J., & Stovel, K. (2004). Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks. *American Journal of Sociology, 110*(1), 44-91.

Bojanowski M. & Corten, R. (2014, October). Measuring Segregation in Social Networks. *Social Networks*, *39*, 14-32.

Bonacich, P. (1972). Factoring and Weighing Approaches to status scores and clique identification. *Journal of Mathematical Sociology*. 2, 113-120.

Borgatti, S., Everett M., & Johnson, J. (2013). *Analyzing Social Networks*. Thousand Oaks, CA: Sage.

Carlson, Marcia J. and Paula England, editors. (2011.) Social Class and Changing Families in an Unequal America. Stanford, CA: Stanford University Press.

Cohen, J. (1979). Socio-economic Status and High-school Friendship Choice: Elmtown's Youth Revisited. *Social Networks, 2*(1), 65-74.

Coleman, J. (1961). *The Adolescent Society*. New York: Free Press.

Dekker, D., Krackhardt, D., & Snijders, T. A. B. (2007). Sensitivity of MRQAP Tests to Collinearity and Autocorrelation Conditions. *Psychometrika*, *72*(4), 563-581.

De Goede, I. H. A., Branje, S. J. T., Delsing, M. J. M. H., & Meeus, W. H. J. (2009). Linkages over time between adolescents' relationships with parents and friends. *Journal of Youth and Adolescence*, 38, 1304–1315.

Duncan, O.D. (1961). A Socioeconomic Index for All Occupations. *Occupations and Social Status*. Ed. Reiss, Albert, et. al. New York: Free Press of Glencoe. 109-138.

Ennett, S.T. & Bauman, K.E. (1994). The contribution of influence and selection to adolescent peer group homogeneity: the case of adolescent cigarette smoking. *Journal of Personality and*

Social Psychology, 67, 653-663.

Ennett, S. & Bauman, K. (1996). Adolescent social networks: school, demographic, and longitudinal considerations. *Journal of Adolescent Research, 11*, 194-215.

Faust, K. (2006). Comparing social networks: Size, density and local structure. *Metodološki Zvezki, Advances in Methodology and Statistics*, *3*(2), 185-216.

Faris, R. and Felmlee, D. (2011). Status Struggles: Network Centrality and Gender Segregation in Same- and Cross-Gender Aggression. *American Sociological Review*, 76, 48-73.

Faris, R. and Felmlee, D. (2014). Casualties of Social Combat: School Networks of Peer Victimization and their Consequences. *American Sociological Review,* 79(3), 228-257.

Frankenberg, E. (2013). The Role of Residential Segregation in Contemporary School Segregation. *Education and Urban Society, 45*(5), 548-570.

Freeman, L.C.. (1979). Centrality in networks: I. Conceptual clarification. *Social Networks*, 1, 215–239.

Gilbert, D. (1998) The American Class Structure: Wadsworth Publishing Company, California.

Goodreau, S. M. (2007). Advances in exponential random graph (p *) models applied to a large social network. *Social Networks,* 29(2), 231-248.

Goodreau, S. M., Hancock, M. S., Hunter, D. R., Carter, B. T., Morris, M. (2008). A statnet tutorial. *Journal of Statistical Software, 24*(9). Retrieved September 11, 2014, from http://www.jstatsoft.org/v24/i09/paper

Handcock, M. and Gile, K. (2010). Modeling social networks from sample data. *Annuals of Applied Statistics*, 4, 5-25.

Hanneman, R. (2005) *Introduction to Social Networks*. Retrieved February 11, 2014, from http://faculty.ucr.edu/~hanneman/nettext

Hollingshead, A. (1975). *Elmtown's Youth and Elmtown Revisited*. New York: Wiley.

Hout, M. (2008). How Class Works in Popular Conception: Most Americans Identify with the Class

Their Income, Occupation, and Education Implies for Them. In A. Lareau & D. Conley *Social Class: How Does It Work?* (Pp. 25-64) New York: Russell Sage Foundation.

Humensky, J. (2010). Are adolescents with high socioeconomic status more likely to engage in alcohol and illicit drug use in early adulthood? *Substance Abuse Treatment, Prevention, and Policy, 5*(19), 1-10.

Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). ERGM: A package to fit, simulate, and diagnose exponential-family models for networks. *Journal of Statistical Software, 24*(3). Retrieved April 11, 2014, from http://www.jstatsoft.org/v24/i03/paper

Hunter, D. R., & Handcock, M. S. (2006). Inference in curved exponential family models for networks. *Journal of Computational and Graphic Statistics, 15*(3), 565–583. Retrieved April 11, 2014, from http://sites.stat.psu.edu/~dhunter/papers/cef.jcgs.pdf.

Iceland, J., & Wilkes R. (2006). Does Socioeconomic Status Matter? Race, Class and Residential Segregation. *Social Problems*. 52(2): 248-273.

Krackhardt, D. (1988). "Predicting with networks: Nonparametric multiple regression analysis of dyadic data." *Social Network*s. 10: 359-381.

Krieger, N., Williams, D. R., & Moss, N. (1997). Measuring Social Class in U.S. Public Health Research: Concepts, Methodologies, and Guidelines. *Annual Review of Public Health*. 18: 341-378.

La Greca, A., Prinstein, M., & Fetter, M. (2001). Adolescent Peer Crowd Affiliation: Linkages with Health-risk Behaviors and Close Friendships. *Journal of Pediatric Psychology: 26*(3), 131-143.

Marx, K, and Engels, F. (1978.) Manifesto of the Communist Party. *The Marx-Engels Reader, Second Edition*. Ed. Tucker, Robert C. New York: W. W. Norton. 469-500.

McPherson, M., Smith-Lovin, L., & Cook, J. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology, 27*(1), 415-44.

McLanahan, S., and C. Percheski. (2008). Family structure and the reproduction of inequalities. *Annual*

Review of Sociology 34: 257–76.

Moody, J. (2001). Race, School Integration, and Friendship Segregation in America. *American Journal of Sociology, 107*(3): 679-716.

Moody, J., Brynildsen, W.D., Osgood, D. W. and Feinberg, M. (2011). Popularity Trajectories and Substance Use in Early Adolescence. *Social Networks*. 33: 101-112.

Morris M., Handcock M.S., Hunter D.R. (2008). Specification of Exponential-Family Random Graph Models: Terms and Computational Aspects. Journal of Statistical Software, 24(4): 1-24.

Mullan Harris, K. (2013). *Design Features of Add Health, The University of North Carolina at Chapel Hill*. Retrieved January 6 11, 2014 from http://www.cpc.unc.edu/projects/addhealth/faqs/aboutdata

Osgood, D. W., Ragan D., Wallace L., Gest S.D., Feinberg M.E., and Moody J. (2013). Peers and the Emergence of Alcohol Use: Influence and Selection Processes in Adolescent Friendship Networks. *Journal of Research on Adolescence*, 23(3): 500-512.

Palloni, A., Milesi, C., White R.G. and Turner A. (2009). Early Childhood Health, Reproduction of Economic Inequalities and the Persistence of Health and Mortality Differentials. *Social Science & Medicine*. 68(9):1574-82.

Snijders, T. (2011). Statistical models for social networks. *Annual Review of Sociology*, *37*, 131-153.

Warren, J. R., Sheridan, J. T., & Hauser, R. M. (1998). Choosing a measure of occupational standing: How useful are composite measures in analysis of gender inequality in occupational attainment? *Sociological Methods and Research.* 27(1): 3-76.

Wasserman, S. & Faust, K. (1994). *Social Network Analysis: Methods and Applications.* New York, NY: Cambridge Press.