November 2023

Using Social Network Analysis to Measure and Visualize Student Clustering Within Middle and High Schools

Geoffrey David West
University of South Florida

Using Social Network Analysis to Measure and Visualize Student Clustering Within Middle and High Schools

by

Geoffrey David West

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
with a concentration in Measurement and Evaluation
Department of Educational Research and Measurement
College of Education
University of South Florida

Major Professor: Jennifer Wolgemuth, Ph.D.,
John Ferron, Ph.D.,
John Skvoretz, Ph.D.,
David Lamb, Ph.D.,

Date of Approval:
October 27, 2023

Keywords: tracking, ability grouping, sociograms, ERGM, equity

Dedication

To my wife, who inspired me to begin, persevere, and finish my degree.

Thank you Chlöe.

Table of Contents

List of Tables

List of Figures

Abstract

The dominant philosophy of American public schools has been to group students together based on similar characteristics. Known as tracking, high achieving students would take courses on the “college track” while others would take “career track” courses. It was not long until advocates noticed that this process unfairly advantaged affluent and White student over poor and minoritized groups. A new process called “ability grouping” took over where tracking left off, but to the same effect. It is difficult to measure the degree students are grouped together by a certain characteristic, and while a few research papers aim to do so, none of them have used advanced social network techniques. This dissertation introduces Exponential Random Graph Modeling (ERGMs) and sociograms to the study of curricular networks. My aim is to show how these SNA tools add to the literature regarding ability grouping and tracking within schools. This is possible since as students take classes together, they form connections with one another, which can be measured and visualized. Using ERGMs, I show how to measure the degree to which students are clustered together within curricular networks based on grade level, race, special education need, and academic performance. Using Sociograms, I show how to visualize these curricular networks and glean insight to various structural patterns that emerge. In this dissertation, I highlight how ERGMs and sociograms tell different parts of the same story as well as show how they can be used together. My goal is to provide researchers with a way to measure and visualize middle and high school curricular networks in such a way that avoids the downsides of other methods.

Chapter 1: Introduction

Overview of Research Literature

Grouping students based on shared characteristics predominates educational philosophy, policy, and practice in the United States, both historically and today (LaParade, 2011). Occasionally, this looks like grouping students together based on similar ages or similar academic performance (Rubin, 2018). Other times, this takes the form of grouping students of similar races or special education needs together (Archbald, Glutting, Qian, 2009; Karin et al., 2012; Boer et al., 2013). Both explicit policies, and the unintended consequences of those policies, can determine how students are placed in school classes (Riehl, Pallas, & Natriello, 1999).

Studying how students are placed in schools is difficult due to the multifaceted processes that take place at schools. Some studies that attempt this look at the proportion of students taking advanced classes, but these methods cannot adjust for multiple factors simultaneously. Regardless, no study to date has examined how researchers can use *social network analysis* to study the ways students are distributed and grouped within schools, and the degree to which this grouping is associated with student characteristics.

Because students can share multiple classes with one another, this creates measurable connections between students based on those shared courses. One can see schools as curricular structures, where connections between students and the classes they take form a network (Heck, Price, & Thomas, 2004). The way students take courses is not random, but rather follows certain patterns and flows. For example, a high school student is unlikely to take Calculus 2 in their first year; they must first take various prerequisites. Students who take advanced classes in one subject area are likely to take more advanced classes in other subject areas. Rather than being randomly placed into classes, students are sorted into classes based on various student and school characteristics (Rubin, 2018). Because of this, what SNA calls “homophily”, students who are similar to one another are more likely than chance to take classes with other students from similar backgrounds.

Social network analysis is a theoretical framework as much as it is a methodological tool (Carrington, P. & Scott, J., (2016). Social network analysts conceptualize the world as a series of connections (ties) between social actors (nodes) within a broader societal context (network). Networks are constructed with ties based on friendship, advice networks, co-occurrences, and other connections. If two people go to the same party, they are connected via that shared event. If a child nominates a classmate as being their best friend, one can say a tie is “sent” from one child to the other. These ties connect all sorts of social actors, whether those actors are children, people at parties, dolphins, or even multinational corporations.

Researchers can visualize networks by constructing sociograms, or graphical representations of social networks (De Freslon et al., 2019). In sociograms, social actors are represented by nodes. Nodes are connected to one another by lines that represent the social connections between them. One can modify these visualizations in any number of ways, from coloring the nodes based on node characteristic, sizing the edges by the number of ties between two actors, or constructing the layout of the graph to emphasize clustering.

One can not only visualize networks but also measure them. One tool to measure social networks is exponential random graph modeling (ERGM). This statistical tool is akin to logistic regression in that it predicts the presence (1) or absence (0) of a tie between actors i & j based on network, dyad, and node level parameters (Goodreau, Kitts, & Morris, 2009; Sweet, 2016). For example, if actors i and j are part of the same group, this may be associated with a change in the likelihood that they are connected. ERGMs can estimate how different parameters are associated with the likelihood two nodes are connected.

Literature Gap and Research Problem

Social network analysis has strong potential to investigate curricular networks, but few studies have done so. Some studies use clustering analysis to identify within-school clusters of students, and what characteristics those clusters display. However, these studies use less sophisticated statistical methods than ERGMs. By using ERGMs, researchers can adjust for a variety of node, dyad, and network parameters that less advanced clustering methods cannot accomplish. For example, McFarland (2007) showed that a school in Hawaii had seven different clusters. Each cluster had very different proportions of students based on the number of academic vs. career-oriented classes, as well as proportions of ethnoracial groups. The method McFarland (2007) used can provide estimates of how overrepresented or underrepresented students are in each cluster based on ethnoracial groups. It cannot, however, measure clustering on ethnoracial groups while at the same time accounting for clustering based on other student factors such as economic status, age, or academic performance. This is a strength of ERGMs.

Additionally, ERGMs allow us to test a hypothesis while simultaneously controlling for other kinds of network factors (Sweet, 2016). Using EGRMs, I was able to estimate how likely two students are to share a class if both share a common trait, while controlling for other variables. For instance, it may be the case that students with disabilities are more likely to be placed together in classes. While estimating clustering on the basis of special education need, I was able to simultaneously control for the number of classes each student takes. By controlling for other factors such as the number of classes a student takes, I can rule out alternative explanations for why clustering is occurring.

Of the few existing papers using social network analysis to study curricular networks, none utilize sociograms to their full potential. Sociograms provide a unique look at curricular networks in that they can articulate patterns of “where” students are within a network. For example, in a typical middle school, there are three grade levels—sixth, seventh, and eighth. One may expect the sociogram to be primarily made up of three large clusters, one for each grade level. A sociogram would be able to confirm this, but also give more context to the network. It is possible that each cluster is further divided by academic performance, or there is a separate fourth cluster of primarily students with disabilities.

To my knowledge, no academic papers have used ERGMs or sociograms in researching the curricular networks of schools. This dissertation fills an important gap in the literature by showing how social network analytic tools can shed light on how students are distributed within schools. The purpose of this dissertation is to highlight the strengths and weakness of using sociograms and ERGMs to study concepts of ability grouping, tracking, and curricular networks.

Sociograms can depict a school curricular network to show the viewer where students from different groups are placed within the school. These visual aids can tell us, for example, if students with disabilities are clustered together separately from students without disabilities, or if students from different ethnoracial backgrounds are segregated into certain portions of the network. ERGMs can then give a numerical estimate of the degree of clustering, which can inform policy makers or district officials on what network or student characteristics are driving the clustering.

Questions/Hypotheses

This dissertation explores the ways social network analysis can add to the literature about within-school clustering. I used social network analysis to measure within-school clustering and identify which network, student, and dyad characteristics predict co-course enrollment. In addition, I used social network analysis to visualize school curricular networks to glean additional information. I then combined the measurements with the visualizations to determine if there was an additional benefit of simultaneously considering both factors. This dissertation addressed the following five research questions.

- In what ways can SNA **visualize** within-school clustering on the basis of grade level, race, academic performance, and special education need?
- In what ways can SNA **measure** within-school clustering on the basis of grade level, race, academic performance, and special education need?
- What information about within-school clustering do SNA visualizations provide compared to the information provided by SNA measurements?
- What information about within-school clustering do SNA measurements provide compared to the information provided by SNA visualizations?
- What information can be gleaned from using both SNA visualizations and measures in tandem?

Dataset

To answer these questions, I used a dataset compiled from a large school district in Florida. After going through the district's Institutional Review Board, district employees were asked to identify the various sources of data required for this project, and to construct a dataset according to my specifications. This dataset is a composite of several data-files used by the district and by the state of Florida. Over 30 middle and high schools are represented in the data, as well as over 30,000 students. The dataset contains information on what courses students took during the 2015-2016 school year. The dataset also contains student academic and demographic information; students' ethnoracial group, disability/gifted label, grade level, and academic performance on a statewide standardized test.

These student-level data allowed me to test a variety of hypotheses about how different student characteristics are related to whom they are more, or less, likely to take classes with. Because I used ERGMs, I was also able to control for confounding variables. For example, I was able to see if race is predictive of student clustering while also controlling for academic performance of the students, or if the number of overall classmates varies across ethnoracial groups. By including these control variables, I achieved more precise estimates of the clustering parameters.

Significance

One of the major significances of this dissertation is that is it the only study to use ERGMs and sociograms to visualize and measure the curricular networks of middle and high schools. After extensive literature review, I found that no study has used these social network analysis tools to study how students are distributed within schools. ERGMs allow researchers to measure a host of node, dyad, and network parameters at once, while less sophisticated methods may only be able to identify clusters of students and determine if students are over- or underrepresented in those clusters (Robins, 2016). By using ERGMs, one can simultaneously estimate how different student characteristics play a role in clustering students within a curricular network. This research can inform the literature on how students are distributed throughout school curricular networks on the basis of academic performance, race, disability, and other important demographic characteristics such as gender and socioeconomic status.

The ability of ERGMs to measure clustering within curricular networks gives researchers a numeric value of *how clustered* a school is based on student characteristics (Sweet, 2016). ERGM coefficients are useful for a variety of purposes, such as measuring their associations with school-level variables—e.g., Title 1 status, charter school status, magnet school status, and percent of students who are minorities or on free-or-reduced lunch. Researchers can use these measurements to track how clustered schools are over time and to determine if district, state, or federal policies are associated with the ERGM coefficients.

Where ERGMs provide numerical measurements of the degree to which a school is clustered on the basis of student characteristics, sociograms visualize the school's curricular network. These network visualizations provide an additional component of the story, telling the viewer *where* the clustering is within the curricular network. While the concept of place or location is meaningless for most networks, sociograms of curricular networks can highlight where the clustering of students occurs. For example, ERGM results may show that students in a school are highly clustered on the basis of special education placement. A sociogram of the school can show that these students are highly isolated from students in mainstream classes, indicating that these students are in self-contained classrooms. While ERGMs can tell us whether or not there is student clustering on the basis of a characteristic, sociograms can provide additional context.

Lastly, measuring and visualizing within-school clustering based on student characteristics is a valuable tool for student advocacy and measuring if policies or programs have had the desired impact. For example, a middle school principal noticed that students with disabilities were rarely in mainstream classes, and that classes were racially homogeneous. The principal could use the methods described in this dissertation to determine the degree to which their school is clustered on the basis of these factors. Additionally, after implementing various programs to counteract the clustering, they could re-measure the school's curricular network to assess the degree to which these programs worked.

Delimitations

This dissertation contains several important delimitations including. These include; condensing categories of special education placement, omitting cross-school connections, excluded elementary schools, and consolidated the categories of academic performance.

For this dissertation, I chose to condense different categories of special education needs into Gifted, Student with Disability, and Neither. The dataset from the school district only contained the "*primary* exceptionality" of the student. This means that if a student has two disabilities, only the *primary* special education need is listed. Additionally, if a student is labeled gifted and also has a disability, only gifted status is known. Because of the limitations of the data, and some categories of disabilities having too few students to model, I chose to condense them into the three categories.

Occasionally, a student will go to two or more schools over the course of the year. This leads to students' being connected to the curricular networks of two or more different schools. For this dissertation, I chose to focus only on those within-school connections, ignoring the connections that exist between schools. For example, if a student takes eight courses at School X and eight courses at School Y, when visualizing and modeling School X I only consider those eight courses that occur there. I then visualize and model the eight courses at School Y separately. This allows for the network visualizations to be clearer and the statistical modeling to be more precise.

I decided to only look at middle school (grades 6-8) and high school (9-12) curricular networks. Middle and high school students will generally have multiple courses comprised of different students throughout the school day. Unlike middle and high school students, elementary students generally stay with the same group of students throughout the day and are taught by the same teacher. Because of this, students are only ever connected to the same set of students. Put another way, an elementary school with 12 classes of 20 students each will only have 12 clusters of students who are never connected to other students in other clusters and connected to all students within their cluster. In network analysis, these are known as cliques, where all actors in a network are connected to each other actor in the network. I chose to only focus on middle and high school curricular networks because they avoid forming cliques due to the interconnections between grades and cohorts of students. By only focusing on these students, I will better highlight the ways social network analysis can visualize and measure within-school clustering within curricular networks.

For a measure of academic performance, I chose to use the Florida Standards Assessment for English Language Arts (FSA ELA). I chose this measure as an indicator of academic performance for two reasons. First, during the academic school year in which I collected these data, the FSA ELA was administered to all students in second grade all the way to tenth grade (FLDOE, 2019). This is the standardized test that the largest proportion of students take consistently. Whereas FSA Math or History is taken in select grade levels, the FSA ELA is taken in the most grade levels. Second, this measure is standardized to the grade level. When a student takes an FSA exam, they receive a diagnostic scale score as well as an Achievement Level. The range of possible diagnostic scale scores for the exam varies by grade level; whereas third grade's range is 240 to 360, tenth grade's range is 284 to 412. However, within each grade level, these scores are broken into one of the five Achievement Levels. So, while a scale score of 360 in third grade is in level 5, the same score is in level 3 for tenth-grade students. Put another way, an FSA score of 3 is considered "proficient" across grades, while the minimum Diagnostic Scale Score to be considered proficient is different across grade levels (FLDOE, 2019). By using the

Achievement Level as a measure of academic performance, I avoid complications that would arise from using the diagnostic scale score.

This introductory chapter outlined the historical and conceptual ideas of ability grouping within US middle and high schools. It also talked about how grouping students based on shared characteristics has often been problematic along lines of race and disability. It then turned and gave an overview of social network analysis both as theory and as method. Lastly, it presented the various research questions this dissertation will answer. The next chapter explains how this dissertation fits into the broader literature as well as justify the different methodological choices I made throughout the dissertation.

Chapter 2: Literature Review

The purpose of this dissertation was to investigate and evaluate how social network analysis can visualize and measure how students are distributed within schools on the basis of grade level, race, academic ability, and special education need (SEN). After showing how SNA is a valuable tool for both measuring and visualizing student distributions in schools, I discuss how each method has unique strengths and weaknesses. I then discuss the benefits of using both SNA measures and visualizations in tandem.

In this review of the literature, I discuss how students are sorted into classes based on these characteristics. It is important to detail the mechanisms behind how students are sorted into classrooms, as this will provide important context when interpreting the results of social network analysis. I then show how various social network analysis concepts and tools can visualize and measure within-school clustering. I introduce exponential random graph models and sociograms as two specific tools well-suited for the study of student clustering within school curricular networks.

Homogenous Grouping of Students

Historically, educators have believed sorting students into groups of similar peers will have advantageous consequences (Argys, Rees, & Brewer, 1996). Homogenous groups of students reduce the amount of variation within a classroom. This could allow teachers to teach to the average student without alienating students at the ends of the academic distribution. There are various ways schools have sorted students into groups. Schools often sort students based on a variety of factors including academic ability, special education needs (SEN), and race.

One can conceptualize this sorting as being *between* schools or *within* a school. Houtte, Demanet, and Stevens (2012) define between-school grouping as the process in which students are distributed between various schools. An example of this practice is using a standardized entrance exam to determine which school students may attend. Based on this entrance exam, the system may send high-achieving students to "elite" schools while sending others to "lower tier" schools.

Within-school grouping is the practice of sorting students into more or less advanced courses on the basis of some criteria (Steenbergen-Hu, Makel, & Olszewski-Kubilius, 2016). For example, in order to take advanced courses in math, a school may require students to score a certain level on a standardized math test. While the purpose of this dissertation is to measure within-school tracking/clustering, it is worth briefly discussing mechanisms for between-school tracking/clustering.

Between-School Tracking

Between-School Sorting on Academic Performance

Educators have historically believed grouping students together based on shared characteristics is strategy that benefits all students. Sorting students between schools on academic performance is one of those strategies. An example of between-school tracking based on academic performance comes from Jackson (2009). Jackson describes the process by which fifth-grade students in Trinidad and Tobago schools are sorted into secondary schools in sixth grade. All fifth-grade students take an examination at the end of the year and are sorted into schools based on student preference. Schools all determine minimum scores that students must score above to attend the school. If the student does not score high enough for their first-choice school, they are placed in their second-choice school, and so on. In this system, students are sorted between schools based on academic performance as decided by test scores and cut scores that determine which school the student is placed in.

Between-School Sorting on Race

Before 1954 in the United States, schools were legally allowed to segregate students into schools based on the race of the student. Orfield (2004) describes how when President Kennedy first proposed the Civil Rights Act of 1964, 99% of Black students in the South were in completely segregated schools, there were no Black teachers teaching White students, and no White teachers were teaching in predominantly Black schools. After *Brown v. Board of Education* (1954) and the passage of the 1964 Civil Rights Act, schools were no longer to legally segregate based on race. However, this segregation continued in the subsequent years as politicians resisted.

For decades, schools were pushed to desegregate, but this took a turn in 1991 when SCOTUS started to dismantle *Brown v. Board of Education*. *Freeman v. Pitts*, *Missouri v. Jenkins*, and other court cases slowed plans of integration to a halt. Today, we find that schools are often just as *de facto* segregated based on race as they were in the past (US GAO, 2016). Since 1991, over 35 school districts have dismantled their desegregation programs after being declared to have successfully desegregated (Gamoran, 2015).

Between-School Sorting on Labels of Disability and Giftedness

Between-school tracking may also look like schools specifically designed to meet the needs of students with disabilities or gifted students. A school district may deem that the best way to accommodate students with disabilities is to have specialized schools equipped and resourced to provide services. This practice is very rare, however, as 2.74% of students served by the Individuals with Disabilities Education Act (IDEA) are placed in a separate facility (NCD, 2018); the majority of these students are inside general education classes 80% or more of the day.

There is some variation across groups when it comes to which students with disabilities are placed in separate schools (NCD, 2018). Asian and Black students with disabilities are placed in separate schools 3.7% of the time, White and Multiracial 2.7% of the time, and Native Americans 1.4% of the time. There is also variation across disability labels. Students with multiple disabilities, deafness, or blindness are taught in separate schools 17.7% of the time, while those with speech/language disabilities are taught in separate schools only 0.21% of the time.

Gifted students many also attend different schools than non-gifted students via a similar mechanism. A school may be specially designed to meet the needs of gifted students or have specialized programs to attract students. For example, some middle and high schools have a magnet program for gifted students, encouraging parents of gifted children to send their students to these schools. This muddies the distinction of between- and within-school tracking.

Within-School Tracking

The aim of this dissertation is to investigate how social network analysis tools can measure and visualize within-school clustering on the basis of different student characteristics. As such, it is important to understand the mechanisms behind the different sorting processes. This helps in interpreting the results of exponential random graph models as well as the resulting sociograms. For example, if a network has high clustering of students with disabilities and the sociogram shows students with disabilities clustered separately from the main network component, it may indicate that the school has implemented self-contained classrooms.

Arguments for Within-School Tracking

Houtte, Demanet, and Stevens (2012) define within-school grouping as the process in which students are distributed into different curricular “streams” within a school. Within-school grouping based on academic performance has been referred to as “tracking,” where students are placed into vocational or academic “tracks” (Oakes, 2008). This is a common practice across countries’ education systems (Banks et al., 2014; Dumont, Jansen, & Becker, 2017). This was also a common practice in the US prior to the mid-1960s (Lucas & Berends, 2006). High school would assign students to one of two mutually exclusive programs, determining what classes they would take during high school. There was little to no mobility between tracks, constraining what classes one could take. These tracks were often labeled as “career” or “academic,” where students in the latter group were placed in courses to better prepare them for college and the former group into classes to better prepare them for careers in blue-collar jobs.

This is a feature, not a bug, as tracking advocates suggest this method better matches students to careers suited for them. In this view, education aims to meet the demands of the labor force, job training, and integrating students to occupational structures (Mickelson & Everett, 2008). Students who struggle in school, the argument goes, may be most suited for vocational jobs such as mechanic, technician, or other blue-collar jobs. Students who excel in school may be best suited for white-collar jobs requiring college degrees. These students are sorted into advanced academic courses to best prepare them for college, while others are sorted into less challenging courses.

Advocates of tracking also suggest low-achieving peers will “drag down” high-achieving peers (LaPrade, 2011; Argys, Rees, & Brewer, 1996). This fear of “holding back” higher-achieving students comes from the idea that teachers may struggle to teach highly heterogenous classes. If a classroom has a significant amount of variation in academic ability, teachers must teach to both high- and low-performing students. In heterogenous learning environments, students who struggle may be underprepared for the pace required, and students who excel may not be challenged by the material enough to keep their attention. In homogenous learning environments, the variance of academic performance is lowered, and thus teachers can teach to the level their students need.

Ability Grouping

Ability grouping is the process by which students are grouped into different levels of the same course (Miller, 2018). According to Lucas and Berends (2008), prior to the 1960s and 1970s, students were assigned to a program determining their course-taking for their high school career. After this time, many school systems turned away from this style of assigning students to courses and moved towards allowing students to choose classes at different levels. Rather than continuing to practice *de jure* tracking, where students were formally grouped into curriculum streams, schools instead started to place students into differentiated curricula. Put another way, rather than being horizontally differentiated curricula (college-track courses and career-track courses), classes are vertically differentiated (English Language Arts and English Language Arts Honors).

According to the National Center for Education Statistics National Teacher and Principal Survey (2018), 34.1% of schools across the United States assign students to classes based on their academic ability. In public schools, this occurs in 31.6% of schools compared to 40.3% of private schools. There are minor differences between public schools when it comes to the percent of the student body receiving free or reduced lunch as well as minor differences in the proportion of students who are of minoritized groups. Public middle schools have the highest percent of ability grouping (48.5%), while elementary schools have the least (23.9%). However, as the size of the school increases, the more likely the school is to ability group, from 25.9% at small 100-student schools to 50.3% at schools with over 1,000 students.

Steenbergen-Hu, Makel, and Olszewski-Kubilius (2016) discuss this new form of assigning students to courses. This form of course assignment is most often referred to as "ability grouping." In their comprehensive literature review, they classified ability grouping into four categories: between-class ability grouping, within-class ability grouping, cross-grade subject grouping, and grouping for gifted students.

The first form of ability grouping is between-class, where students are distributed between different levels of the same course such as "advanced," "honors," "regular," or "remedial." For example, based on prior English achievement, students may be placed in remedial English, honors English, or AP English. While similar to traditional tracking because students' track and coursework are correlated, this form of ability grouping is less rigid than traditional tracking because students can take advanced classes for some subjects and less advanced for other subjects. While traditional tracking focuses on *program* placement, between-class ability grouping focuses on *course* placement (Heck & Mahoe, 2010; McFarland, 2007).

In a traditionally tracked school, all of a student's courses would be from either the "high" or "low" track. Between-class ability grouping would allow a student who struggles in math but excels in science to take remedial math courses but AP science classes. Heck, Price, and Thomas (2004), however, suggest that while there is no formal track, it is difficult to take higher-level classes if a student is initially placed in a lower-level class. For example, if a student takes pre-algebra instead of algebra in eighth grade, they may be less likely to take AP Calculus in high school.

The second type of ability grouping is within-class. This simply involves taking a class of students and breaking them into smaller groups of students with similar academic performance. One can think of this as small-group instruction, where students in the same class work together with a few peers to complete tasks. This type of ability grouping is beyond the scope of this dissertation because it would require observational or classroom-level network information, something not contained in the district transcript files.

The third category of ability grouping involves grouping students across grade levels based on similar academic performance. For example, in a middle school with grades 6, 7, and 8, a high-performing sixth-grade student may be placed in eighth-grade-level courses. Struggling students may be placed with other struggling students in remedial courses regardless of grade level. This dissertation can investigate the degree to which students across grade levels are taking courses together. The methods used in this paper can estimate the degree to which students are connected between grade levels. Additionally, sociograms can show how students are placed within curricular networks on the basis of grade level.

The fourth and final category is special grouping for gifted students. Gifted students are those students who excel in a number of academic aspects and require special educational accommodations. To accommodate these students, some schools pull them out of their regularly scheduled courses and work on special advanced projects with them. Alternatively, some schools have a regularly occurring gifted course for students to attend during their day. While Steenbergen-Hu, Makel, and Olszewski-Kubilius (2016) only discuss the grouping of gifted students, schools often group students on the basis of disability, often having a self-contained classroom or pull-out programs. This dissertation looks at the degree to which students with special education needs (students with disabilities and gifted students) are clustered together on the basis of having a separate self-contained classroom.

To summarize, this dissertation assesses how students are grouped together between classes, between grades, and by special education need. The dataset I am using has information about these three types of groupings, but unfortunately does not contain within-class information.

Arguments Against Ability Grouping

Not everyone agrees that tracking, and by extension ability grouping, is living up to the purpose of education. These advocates show multiple problems with ability grouping, specifically how they reproduce inequality based on race, socioeconomic status, and special education needs. Those who argue against tracking find that it reinforces societal hierarchies rather than providing equitable education across groups (Dornbusch, Glasgow, & Lin, 1996). It is important to better understand the process in which students are placed into classes based on various characteristics so that we may interpret the results of the SNA models and sociograms correctly and provide useful information.

Within-School Tracking and Race

Many researchers believe tracking has less to do with matching students to jobs best suited for them, but rather maintaining racial and social class hierarchies. Oakes and Guiton (1995) forward this idea by showing evidence from a two-year case study of four high schools. The authors discern seven propositions regarding tracking decisions. Propositions 1 and 2 suggest that schools view student academic ability as fixed and aim to accommodate, rather than alter, these student characteristics. Those who are high-achieving receive additional advantages (proposition 3) by being placed in more advanced classes. However, decision makers are not simply using academic performance, but rather allow student race, ethnicity, and class as markers of where students ought to be placed (proposition 4). Course streams are also highly regulated (proposition 5) in that there is often a rigid course progression that is unalterable. Between-school sorting of students also influences what kinds of courses are offered (proposition 6). For example, if a school receives a large number of struggling students, student counselors may be overwhelmed, teachers may have higher turnover rates, and advanced and elective courses may be cut due to funding issues. In the end, there are exceptions and irregularities, but these irregularities benefit already advantaged students (proposition 7). For example, a school that has limited counseling services may focus on students who are already doing well. In summary, school decision makers allocate students within preexisting differentiated hierarchical curricular structures in ways that are influenced by culture and not solely on the basis of merit as they may believe.

Mickelson and Everett (2008) show that track placement is related to not only student socioeconomic status but also student race, with Black and Hispanic students consistently placed into lower tracks. The authors describe "neotracking," a new form of tracking in North Carolina that combines older forms of tracking of students into particular career and college tracks and newer forms of tracking of within-subject area curricular differentiation. They show that racial stratification within schools is larger at schools with a higher proportion of students from racial minority groups.

O'Connor (2002) describes how racial stratification is often very apparent, with Black students often noticing "I'm usually the only Black [student] in my class." The way racial stratification occurs within schools creates noticeably Black spaces and movement within the schools. O'Connor gives an example of a school where courses within the career track are placed in the north wing of the school building. The overrepresentation of Black students in the career track creates the situation where students within the north wing of the school building are predominantly Black.

On the other hand, Black students who are placed in advanced courses face different challenges. O'Connor (2002) describes an interview with Black students placed in Advanced Placement courses. The students describe their experience feeling alone as the only Black student in a class, or how White students look down on them. Others feel pressure to not drop out of a difficult class because of how their White peers may attribute their dropping out to Black people as a whole, putting massive pressure on the shoulders of Black students to perform. While much of the literature on within-school segregation emphasizes how Black students suffer educationally under these systems, O'Connor describes the emotional toll Black students face.

Burris et al. (2006) used a quasi-experimental cohort design to determine the extent to which tracking was limiting students from receiving New York State Regents and IB diplomas. The study outlines a detracking reform that took place in a diverse school in New York State. Students who took a series of eight advanced state exams and met course requirements could receive an advanced Regents diploma. The district eliminated low-track courses that did not follow the Regents curriculum and offered struggling students instructional support. These policies, plus the addition of allowing any student to take honors courses if they wanted to, led to more and more students' taking Regents courses and receiving a Regents diploma. These structural changes led to drastic increases in the probability Latino or African American students received a Regents Diploma regardless of original aptitude levels.

It is important to better understand the process in which students are placed into classes based on race so that we may interpret the results of the SNA models and sociograms correctly and provide useful information. For example, sociograms can tell us "where" students are clustered within a network and how this position correlates with how students are placed via other characteristics. ERGMs then measure the degree to which students are clustered together.

Previous studies have used other metrics to measure within-school clustering, one being the Segregation Index. Conger (2005) describes the Segregation Index as the proportion of students from group B in the typical classroom for a student from group A. In other words, it is the probability that a student from group A comes into contact with a student from group B in their classroom. The Segregation Index normalizes these calculations based on what we would expect the exposure rate to be if classes were random, and then compares what we would expect by chance to what is observed. One can use this index to measure within-school clustering, it is limited to one characteristic at a time. If one wanted to see how a school was segregated by race and academic performance, one would need to do two separate calculations, whereas ERGMs can do both estimations simultaneously.

Within-School Tracking and Socioeconomic Status

Tracking exacerbates not only racial inequality, but also socioeconomic inequality. Epple, Newlon, and Romano (2002) show that the general practice of tracking is increasing inequality on the basis of socioeconomic status. The authors show that private schools have very little amounts of tracking due to their selective nature and often prohibitive tuition costs. Private schools use these mechanisms to determine the composition of the study body, mostly students with high academic performance, reducing the variability of student academic performance and thereby reducing the need for tracking mechanisms. Public schools, on the other hand, have limited control over who is able to join their student body. Thus, the variability of student performance is heightened, increasing the use of tracking.

By using tracking, public schools are able to attract predominantly high-performing students and offer them a challenging curriculum, while lower-performing students are relegated to less rigorous education. Even students placed in a lower track who are comparatively higher-performing and more affluent have the option to switch to private schools, thus further increasing the inequality we see within schools. For low-performing students from poor families, the only option is to remain in schools and classes with low resources and support.

Special Education Needs

There are many different types of special education needs. While some may assume this is only referring to students with disabilities, it can also refer to students who have been labeled "gifted and talented." In the same way that students with various physical, mental, and emotional needs need specially tailored instruction, so too do gifted students, and gifted students with disabilities. This section talks more about the background of how students with disabilities and students labeled as gifted have been sorted within school curricular networks. This assists my interpretation of sociograms, helping to identify the different ways these students are sorted together.

One way of measuring the degree to which students with disabilities are grouped together is measuring the percent of students with disabilities who are included in general education classes. In 2019 in the United States, 65% of all students with disabilities spent at least 80% of their school day within general education courses, and only 13% spend less than 40% of their day inside general education classes (NCES, 2019). This varies significantly by race, as White students with disabilities are most likely (66%) to spend 80% or more of their time in general education classes, whereas 57% of Pacific Islander and Asian students with disabilities spend 80% or more of their time in general education classes. Across disability labels, there is substantial variation on who spends at least 80% of their time inside general education classrooms, from 88% of students with speech or language disabilities to 14% of students with multiple disabilities.

This measure of within-school clustering is limited in that it does not account for other network or student variables. Students with disabilities may take more or fewer courses than students without disabilities, which may explain part of the degree to which students are segregated. Additionally, this measure does not lend itself to graphical representations such as sociograms that can add clarity and context. If students with disabilities are spending more than 80% of their school day in general education classes, but those are only elective or remedial, that information is lost if only using the percent participation of students with disabilities in general education classes. Sociograms, tied with ERGMs, can provide additional context lost by using this measure.

Within-School Tracking and Students with Special Education Needs

In order to use ERGMs and sociograms properly, it is important to know the ideas behind and mechanisms used when talking about the education of students with disabilities. Grouping students based on disability may take the form of self-contained classrooms where students receive limited interaction with students without disabilities. NCD (2018) states that “to the maximum extent appropriate,” students with disabilities must be educated 1) with children without disabilities 2) as close as possible to home and 3) in the school they would attend if they did not have a disability. Removing these students from general education classes must only occur if their education cannot be successfully achieved in these settings.

Elbaum (2002) outlines three different educational within-school placements for students with disabilities. These placements include general education classes, resource rooms, or self-contained classrooms. These classifications correspond to the statistics IDEA requires states to submit (NCD, 2018). Students with disabilities in general education classrooms receive 80% or more of their instruction in classes with students without disabilities. While these students may be given accommodations, they nonetheless remain in general education courses for most of the day. Resource rooms are separate classes where a student receives some instruction. These students spend 40%-79% of the school day with students without disabilities. These students may be "pulled out" of general education classes at certain times of the school day for separate instruction. Lastly, students in self-contained classrooms receive all or most of their school in a class separate from students without disabilities. Participants in this group only spend 40% or less of their day in general education classes.

Reasons for Grouping by Special Education Needs

Due to stigma surrounding disabilities, many students with disabilities face ridicule from peers without disabilities. Chen et al. (2015) show that 74% of students with disabilities were involved in consistent peer victimization through the year and were significantly more likely to be a victim of bullying. The researchers also found that students with disabilities are more likely to be bully victims, those victims of bullying who in turn bully as a coping mechanism. Estell et al. (2008) echo this finding, discussing how students with "mild" disabilities are more likely to be labeled as victims of bullying. Pearl et al. (2018) concurs that students with disabilities are not "well accepted" by their peers in many school settings. By placing students with disabilities into separate classes, they may be spared from being bullied and victimized by their peers.

Additionally, teachers with students who have more severe disabilities may not have the curriculum or support to deliver specialized accommodations for these students. These circumstances lead some to advocate for teaching students in separate classrooms with more teacher support and with less exposure to students without disabilities. For example, the Government Accountability Office (2009) conducted a survey that showed most teacher preparation programs (73% of elementary programs and 67% of secondary programs) required at least one class entirely focused on students with disabilities. Most elementary and secondary teacher preparation programs also required pre-service teachers to have field experience with students with disabilities (58% and 51%, respectively). However, when asked about specific aspects of their field experiences, respondents were less likely to be able to point to specific experiences. For example, only 7% of students in elementary programs interacted with parents of students with disabilities, only 11% participated in IEP meetings, and only 30% said they helped teach students with disabilities. Because the *average* teacher is expected to only have a moderate amount of training and experience with students with disabilities, having specialized classes taught by highly trained teachers may be a better option to support students with their special education needs (Government Accountability Office, 2009).

In favor of grouping students by disability within a school, as opposed to sending them to a separate school, is that teachers in schools specifically dedicated to those students with disabilities are half as likely to be certified to teach. Between 2013 and 2014, 4.5% of teachers at schools specifically dedicated to serving students with disabilities were not certified to teach, as opposed to 1.4% of teachers at traditional schools with medium-sized populations of students with disabilities. Placing students in separate schools increases the likelihood that they are taught by non-certified teachers.

Reasons Against Grouping Based on Special Education Need

Others advocate for inclusion, pointing out that there are academic and non-academic benefits for both students with and without disabilities when they can interact with one another

(Boer et al., 2013). For example, this may reduce the level of prejudice held by peers without disabilities and can teach all students how to appropriately interact with one another.

Boer and Pijl (2016) used social network analysis to show that students with attention deficit hyperactivity disorder (ADHD) and students with autism spectrum disorder (ASD) have significantly fewer friends than “typically developing” peers. When the study population was asked to identify students whom they liked to hang out with, and whom they would rather not hang out with, students with ADHD or ASD received less friendship nominations and more rejection nominations than their typically developing classmates. The authors also found that student attitudes modified this relationship, as typically developing students with positive attitudes towards those with disabilities were less likely to reject students with ADHD or ASD. By separating students with disabilities away from those without, there are fewer opportunities for cross-group interaction, which may lead to fewer friendship formations and more missed opportunities to increase out-group positive attitudes.

Szumski, Smogorzewska, and Grygiel (2020) showed that students without disabilities who were taught in inclusive classrooms had higher levels of positive attitudes towards students with disabilities. In their cross-sectional study, the authors also showed that moral identity, the idea that being a moral person is important for individual identity, was also predictive of positive attitudes towards students with disabilities. The authors are unable to identify this effect as causal due to nonrandom assignment of students to schools. They discuss how it could be the case that parents with positive attitudes towards people with disabilities are more likely to select schools with inclusive classrooms. Nonetheless, this study shows there is a relationship between inclusive education settings and positive attitudes towards people with disabilities, which provides the first step in creating a causal argument.

However, Elbaum (2002) reminds us that inclusive environments are not automatically better for students, but that there are other variables at play. The author conducted a 38-study meta-analysis showing that the self-concept of students with learning disabilities was not statistically significantly different across most educational placements. These results show that there is no inherent relationship between educational placement and self-concept, but rather, there are other contextual factors to consider. For example, the author points out that some parents pursue litigation for *more* restrictive educational environments, while others fight for *less* restrictive educational environments. This lends to the idea that there are other explanatory factors that influence the association between educational placement and self-concept of students with special needs.

These studies highlight the association between positive benefits for students with disabilities and inclusion within the classroom. However, we should take these conclusions cautiously, as it may not be *inherently* beneficial academically, behaviorally, emotionally, socially, etc. to be in an inclusive classroom setting. Rather, one would need to consider other tangible and intangible contextual factors.

Gifted Students and Inclusion

Similar arguments are made concerning gifted students. A major rationale for separating gifted students out from non-gifted peers is that by doing so, they are given differentiated instruction at the level they require. Gifted students are often under-challenged in general education courses; to address this, they are often given additional or more challenging work. Callahan, Moon, and Oh (2017) describe different ways gifted students are taught in schools. As with students with disabilities, gifted students may also be “pulled out” by a special gifted education teacher who can guide the student through a more advanced curriculum in an individual or small-group setting, the most popular program service delivery for elementary school gifted programs. Other schools may also offer specific classes for gifted students to take, or even to the point where there is limited interaction with non-gifted peers. This method is the most-used method in middle schools. A third major way is to place these students in more advanced, but still general education, courses such as Advanced Placement. This is the most-used delivery method in high school gifted programs. These three gifted program delivery methods are common ways to address the needs of gifted students to be challenged in schools (Callahan, Moon, & Oh, 2017).

Barber and Wasson (2015) discuss how gifted students may perceive themselves negatively due to their being different. While acceleration of the student to the next grade may reduce the differences between a gifted student and their peers, one could also view this as a reason to isolate gifted students into separate classes. Separate classes would reduce the amount of "difference" these students see on a daily basis. This may lead to fewer feelings of difference for these students. However, studies seem to agree that gifted students may have equal or lower rates of peer victimization as non-gifted peers. Estell et al. (2008) shows that teachers rarely nominated gifted students as being victims of bullying, while socially isolated students were more likely to be victims. This may lead one to conclude it is not the giftedness per se that is associated with being a victim of bullying, but rather the social isolation (which may be higher for gifted students).

No studies have sought to measure the degree to which students with special education needs are clustered within a school. Additionally, no studies have used ERGMs and sociograms to do this. Using ERGMs provides a unique opportunity to use a clustering mechanism that can control for multiple sources of homophily at once. For example, if we wanted to know the degree to which gifted or disability label predicted co-course enrollment, but also wanted to detangle other student characteristics such as academic performance or race, ERGMs would be able to do so. Additionally, by using sociograms, one can visualize how students are distributed throughout the network, and then overlay other student characteristics. If there are clusters of gifted students in the network, sociograms can identify those students, and then answer other questions about them such as their race, academic performance, or grade level.

Special Case of Gifted Inclusion

It is important to note a special case regarding gifted inclusion: Magnet schools with a specific focus on gifted education start to blur the distinction between within- and between-school placements. Historically, magnet schools arose after desegregation efforts in the 60’s and 70’s were met with substantial opposition (Tefera et al., 2011). One mechanism for desegregating schools was through busing, where students in Black neighborhoods would be bused to schools in majority White neighborhoods, and vice-versa (Cascio, et al., 2007). Opposition to this practice was overwhelming, with the US Supreme Court declaring it unconstitutional.

This is where magnet schools come into play. Magnet schools are typically located in underserved communities and offer unique programs or curricula (Tefera et al., 2011). While all parents are likely to want their children to participate in these programs, they are designed to be particularly attractive to parents from more privileged backgrounds. By providing these programs, magnet schools are better able to attract parents from affluent and White backgrounds, which increases the social and political capital of the school.

These programs can include the arts, STEM, military, and other focuses. Sometimes a magnet school will be specifically designed to house a gifted program. One can think of these magnet gifted programs of as “schools within schools” (Tefera et al., 2011). Imagine a school with 100 students, 50 of whom are labeled as being gifted. At a school level, these students are in an “integrated” school, since the two groups are represented equally. However, these gifted students are taking gifted and more advanced classes than their peers; while the school is integrated at a macro level, it is not integrated at a micro level. This is especially important to note due to anticipated groupings based on sociograms. While one may assume there are three large clusters in a middle school curricular sociogram, it may be the case that there are more if the school is a magnet school due to where “magnet” students are placed within the curricular network.

Overview of Social Network Analysis

In order to accurately measure and visualize school curricular networks, we must first discuss different social network concepts. The way school curricular networks are measured hinges on a number of factors including n-partite-ness, directionality, and edge weights. Without considering how these factors may play a role in measuring and visualizing social networks, we may make mistakes.

Basics of SNA

Social network analysis conceptualizes the world as connections between “actors” that are measurable and visualizable. Within networks, individuals, ideas, groups, or other types of actors are considered “nodes,” which are connected to other nodes via a relationship. These connections between nodes are called edges or ties, and can represent relationships such as friendship, advice-seeking, or co-course enrollment. These relationships may exist within what is called a bounded system. For instance, if I wanted to study the communication patterns of lawyers in a law firm, I would ask employees whom they communicate with *from within the* *office*. This section discusses those aspects of social network analysis that are relevant to this paper, namely, bipartite and unipartite networks, directed and undirected ties, and weighted and binary ties.

Bipartite and Unipartite

Within social network analysis, there is a major distinction between networks that have nodes of only one type and networks with nodes of multiple types. For the purposes of this dissertation, it is most important to discuss the difference between

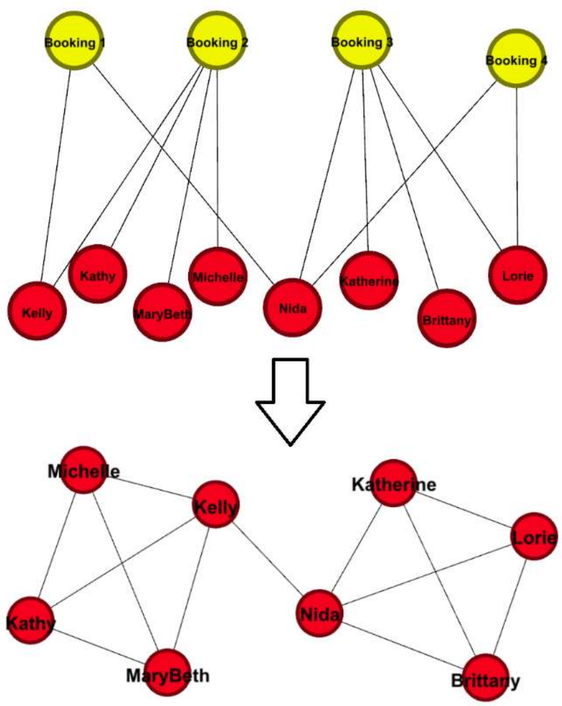


Figure 1: Simulated Bipartite to Unipartite Transformation

unipartite and bipartite networks. Borgatti and Halgin (2016) describe bipartite networks as having the property where nodes are demarcated into two classes, and that all ties occur between classes and not within classes.

Newman, Watts, and Strogatz (2002) provides an example of women who attend parties. They give the image of a matrix, where each row is a different woman, columns are the events that took place, and the intersecting cells are “1” if the woman attended that party, or “0” if they did not. The only connection within such a network is between the women and the parties they attend, or on the converse, parties and the women who attend them.

Unipartite networks, on the other hand, are those networks where all nodes are of the same type. In the above graphic, the bipartite network is transformed into unipartite so that all white nodes are connected to other white nodes if they shared a connection with a black node. To continue with the aforementioned party example, one can use matrix multiplication to transform the network from bipartite to unipartite so now women who attended the same event are now connected *directly* to one another rather than indirectly. This transformation is done by taking the Women X Event Matrix, transposing it across the diagonal, and then multiplying it by the original matrix.

Schaefer (2012) is a good example of using bipartite networks and transforming them into unipartite networks. The purpose of the research is to test the association between youth co-offending and how close the students live to one another. Using data from Maricopa County, Arizona, the author uses a two-mode dataset of juveniles and crimes they have committed, and the juvenile and the census tract they reside in. These two bipartite networks were manipulated to unipartite by matrix algebra, creating matrices if the two juveniles co-offended or lived in the same area.

Similarly in this dissertation, the student data was originally in bipartite form; each student connected to the classes they take. The matrix in this scenario had *rows* made up of each student, *columns* made up of the possible courses students could take, and the intersecting *cell* being whether or not the student took the class. In this matrix, students are connected to their courses, which in turn are connected to all the students taking the class. I transformed the network from bipartite to unipartite using the aforementioned matrix manipulation. This became a matrix where students are connected *directly* to their classmates on the basis of shared courses.

Directed and Undirected Networks

Edges are either directed or undirected. In a directed network, a node (ego) is the source of a tie, which is then sent to a different node (alter). This can come in the form of peer nominations where students list who their friends are. In a directed network, the direction of the tie is important information and intelligible; it is important to know how ties are sent as well as received. For example, if Jamal nominates Greg as a friend, but Greg does not reciprocate that friendship tie, this gives us valuable information about reciprocity within the network.

Undirected networks, on the other hand, ignore the direction of the tie. In these networks, it is only important that two ties are connected, not who sent a tie to whom. When bipartite networks are transformed into unipartite graphs, they become undirected networks. This is because this transformation binds together two nodes based on a shared connection, and this

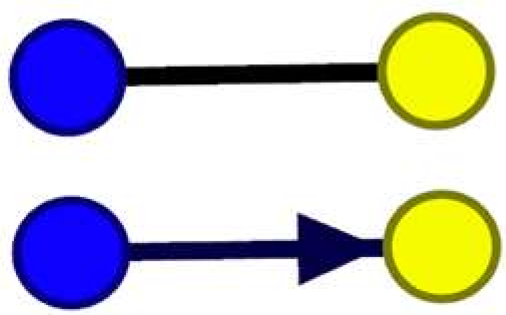


Figure 2: Undirected and Directed Edges

shared connection exists mutually. For example, if Greg and Jamal both go to the same party, you could say "Greg went to the same party as Jamal" or "Jamal went to the same party as Greg." It would be impossible for one to be true and the other false. The image above illustrates the difference between undirected (top) and directed (bottom) edges. The former tells us if the two nodes are connected, while the latter tells us who the sender of the connection is (blue) and who receives the connection (yellow).

This dissertation used undirected rather than directed networks because if two nodes are connected by being in the same course, the relationship is undirected. Additionally, once a network is transformed from bipartite to unipartite, the information about directionality is lost. Similar to the above example, if Rosa and Kevin are both in AP Calculus, the statements "Kevin is taking AP Calculus with Rosa" and its reverse have the same meaning from a network standpoint.

Weighted and Binary Ties

Edges can also be weighted by the number of times the relationship occurs. For example, if I were to ask middle school students whom they sought advice from for 1) teachers, 2) extracurricular activities, and 3) schoolwork, I could aggregate this network so students could have 0, 1, 2, or 3 connections to another student. On the other hand, this network could also be transformed into a binary network based on if a student made any connection to another student. Students who formed 1, 2, or 3 advice relationship ties would be recoded to 1.

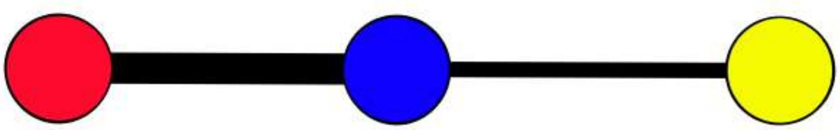


Figure 3: Weighted vs Unweighted Edges

This image reinforces the difference between weighted and binary ties. The tie connecting the red and blue nodes is significantly larger than the tie connecting the blue and yellow node. This is because the width of the edge is tied to the number of times the two are connected, often referred to as edge weight.

This dissertation used binary networks due to the large computational requirements for modeling weighted networks. Regardless of the number of times two students are in the same class, the weight of the edges between the two was transformed to binary. Future research ought to consider weighted networks.

Sociograms

At a basic level, sociograms are graphical representations of networks where individual social actors are represented by dots (nodes) and are connected to one another by lines (edges) based on shared social connection (Carolan, 2014). Jacob Levey Moreno arguably introduced the sociogram as a means of visualizing network data (Carolan, 2014). Moreno’s sociograms introduced using color, shape, and position to highlight not only variation across nodes, but also network characteristics such as density and multiplexity (Freeman, 1996).

Klovdahl (1981) discusses how Moreno introduced the concept of manipulating aspects of the sociogram to convey different kinds of information. Depending on the researcher’s intentions, the researcher may color a node based on some attribute of that node. For example, one may choose to color the nodes based on different levels of anxiety the student reports. A researcher could also manipulate the size of a node to convey importance.

The researcher can also alter other facets of nodes including shape, fill, borders, and opacity. One may also similarly customize edges by style, width, color, and opacity. Additionally, one can represent edges’ directionality or degree of social connection between nodes. Generally, the distance or edge length between nodes is not meaningful; however, most SNA software will integrate graphical distance and network information such as geodesic distance (Sweet, 2016).

Here is a simulated network sociogram. Each node represents a student in a middle school, and the lines represent the friendship connections between one another. Nodes are colored based on the students' gender: boy (blue), girl (red), and nonbinary (yellow).

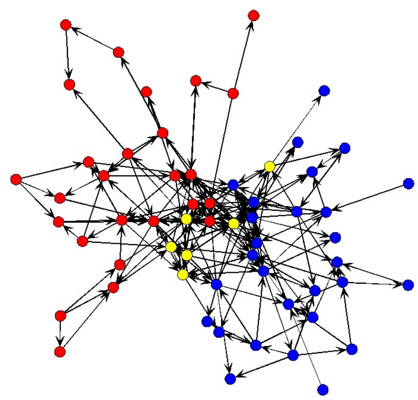


Figure 3: Simulated Network Sociogram 1

While sociograms give insight to the researcher about the nature of the network they are studying, many aspects of the networks are meaningless. For example, generally speaking, the "top," "left," "right," and "bottom" of a sociogram are meaningless artifacts of SNA graphing software algorithms.

One can flip this simulated network upside down or rotated 180 degrees and nothing changes about the interpretation of the network. The fact that boys (blue) are on the right side of the graph is not informative, as if you flipped it to have boys on the left, the information within the graph is not changed.

One can, however, impart meaning onto the orientation of a network to keep interpretations consistent and for ease of readability. Consider the simulated high school course network here. Each node is a student, and they are connected to their classmates if they share a class. If I rotate the graph so that seniors (dark blue) gravitate towards the "top" of the network and freshmen (yellow) gravitate towards the "bottom," one can get a sense of direction using *position* as a visual cue, rather than only using color.

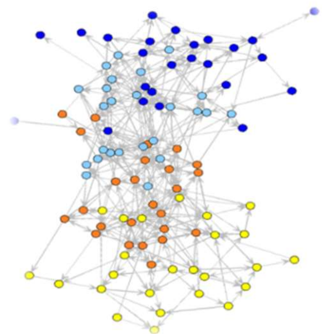


Figure 4: Simulated Network Sociogram 2

In this network, a student located towards the "top" of the graph is more likely to be an upperclassman than a student located towards the "bottom" of the graph. I can flip the sociogram 180 degrees and the only interpretation that changes is where seniors and freshmen gravitate towards in the network.

This is especially useful for this dissertation due to the level of clustering that occurs within schools. Students are highly clustered within schools on the basis of grade level as well as academic performance based on standardized test scores. One can orient these sociograms so that grade level is structured vertically while academic performance is structured horizontally. Using the visualization variable of orientation to represent attributes such as grade level opens other visualization variables to be used elsewhere. For example, it would be difficult to visualize grade level, academic performance, race, and SEN with different colors. However, if grade level is vertical, and academic performance is horizontal, one can use color to represent race or SEN.

Sociogram Issues: Hairballs

Due to the large number of nodes and edges in the sociograms I created, it is important to discuss a common pitfall in creating social network visualizations. Some sociograms are derogatorily referred to as “hairballs” due to their high density of edges, complex structures, and number of nodes (Crnovrsanin et al., 2014). Due to the significant number of edges, it is impossible to follow an edge from one node to another. Likewise, these graphs sometimes make it difficult to identify triangulation, reciprocity, and homophily. These sociograms are unable to convey a significant amount of information because it is difficult to see who is connected to whom.

Consider the below sociogram. Ted Polly (2014) discusses how images such as these display characteristics of a hairball. He notes that while there are a few important aspects of these graphs you *can* glean, the more “granular” aspects of the network are more difficult to see.

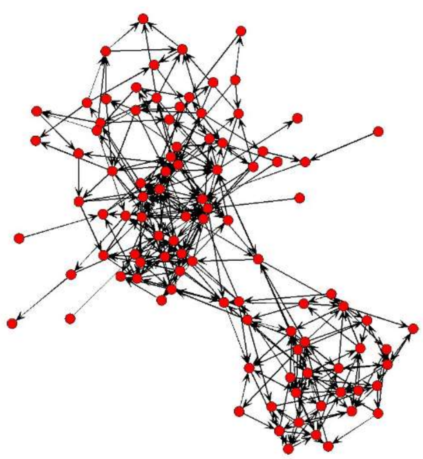


Figure 5: Simulated Hairball Example

For example, one can see nodes are grouped together in two major clusters. You may also see a few nodes in the center that have very many edges, while most other nodes have fewer. However, deciphering details such as reciprocity and triangulation is near impossible in this visualization.

Multiple authors have created techniques to solve the problem of hairballs. Dianati (2016) outlines methods of pruning, where “noisy” nodes that have a significant number of edges are removed from the network to uncover the underlying structures. One can remove nodes that have either too many or too few edges to make the underlying social network more visible. Crnovrsanin et al. (2014) describes bundling edges, using sensitivity cutoffs, and using network layouts. These strategies help simplify these complex network visualizations.

Schools in the selected district range from having a few hundred students to over 1,000. With each student having eight or more courses, the number of edges for even a small school has tens of thousands of edges. These visualizations create the perfect conditions for hairballs to arise. However, because of large amounts of clustering that takes place on the basis of student characteristics such as grade level, academic achievement, race, disability label, etc., these visualizations can still serve as useful tools without the need for pruning. When student nodes are set to represent specific colors based on these characteristics, and the orientation of the network is positioned to convey other characteristics, it is easy to see how much clustering is occurring. So, while network patterns such as different connectivity rates across groups or triangulation may be difficult to see from curricular network sociograms, homophily is still be easily discernable.

Homophily

McPherson, Smith-Lovin, and Cook (2001) discuss homophily as the principle that people from similar groups have more contact with one another than people from different groups. Throughout the literature, authors using social network analysis use the phrase “birds of a feather flock together” to describe this phenomenon. The authors discuss how homophily based on race, gender, age, education, and religion is involved in a variety of social situations. Within schools, there are a number of mechanisms that could be causing homophily to occur, such as grade level, disability label, and race, among others.

Zbigniew (2012) furthers discusses how one must consider two kinds of homophily, "population" and "relational." For population homophily, imagine there is a school with 1,000 students, 50% of whom are White, 30% of whom are Black, and 20% of whom are Asian. If there were no friendship homophily on the basis of race, we would expect the average Black student to have friends who are 50% White, 30% Black, and 20% Asian. Relational homophily looks at how much homophily exists in a network *above* population homophily. If in the same school, the average Black student's friends were 60% Black, there would be in-group homophily on the basis of race since we would expect that number to be 30% due to population homophily.

This is important to consider for this paper because of the asymmetric population sizes of different groups within schools. Some measures of homophily, such as the Freeman Index, do not consider the population homophily in a network. The Freeman Index simply counts the number of ties that are to members outside of the group compared to how many ties go to members within the group. This works for networks with groups of identical sizes, but schools are unlikely to have evenly sized groups. For this reason, it is important to find a social network analytic tool that can account for population homophily when measuring homophily.

Kavaler and Filkov (2017) look at a different aspect of homophily: selection and influence. If people who are similar to one another are more likely to be friends, it could be that people who already previously share a common characteristic are more likely to *become* friends. This mechanism is *selection*, where people who are similar are likely to choose one another for friendship.

On the other hand, it could be that people who are already friends become more like one another. This mechanism is called *social influence*, where people become more like one another if they share a connection. In this proposal, school network homophily is more likely to be caused by selection than influence, since students are sorted into similar classes based on shared characteristics, rather than their shared characteristics being caused by sorting. Many of the variables examined in this paper would not make sense to be "influenced" by peers. Student categories such as race, gender, and disability label are not caused by how the students are sorted. It may be the case that by being originally sorted into a curricular track, students may perform differently academically, but discerning if this is the case goes beyond the scope of this dissertation.

Literature on Measuring Network Clustering

Acknowledging that ties are more likely to form between similar than dissimilar actors, Bojanowski and Corten (2014) systematically review different ways the literature has measured segregation in social networks. They define segregation as "the extent to which groups are exposed to one another by occupying nearby positions." The authors review a number of different measures of segregation. They review the E-I index (Krackhardt & Stern, 1988); assortativity coefficient by Numan and Girvan (2002); Gupta, Anderson, and May's Q (Gupta et al., 1989); odds ratio for within-group ties (Charlres & Grusky, 1995); exponential random graph models; Freeman's Segregation Index; Segregation Matrix Index; and Coleman's Homophily Index (Coleman, 1958).

To measure network clustering within schools, I chose to use exponential random graph models (ERGMs). This is because ERGMs offer significant specification flexibility (Bojanowski & Corten, 2014). Within an ERGM, I can simultaneously model homophily on the basis of grade, race, SEN, and academic performance. One can measure these homophily effects either uniformly or differentially. For example, ERGMs can either measure if the race of students is predictive of being classmates, as well as measuring if students from different races are more or less likely to form classmate ties. Further, ERGMs can measure homophily while controlling for other network effects—for example, measuring if different groups make more connections than others.

Literature on Curricular Networks

There is limited literature that uses social network analysis to measure clustering on the basis of curriculum in schools, and even less that uses ERGMs to do so. The study that most resembles the current project is Heck, Price, and Thomas (2004). They conceptualize tracking as sorting machines, "emergent structures," that resemble pathways through a student's high school career. The authors discuss how most schools track their students, and how the course-taking patterns are "hierarchical and differentiated sociocurricular[ly]," without regard to formal and informal sorting mechanisms. While previous authors have looked at tracks as formal structures, the authors conceptualize them as structures that emerge from "student encounters with courses."

This is, in part, a necessity that comes out of the shift from *de jure* to *de facto* tracking. Previous studies that looked at tracking required students to be *formally* designated into one track or another. In this regime, student data would indicate what track the student was assigned to. With the shifts in tracking mechanisms in previous decades, these formal track designations are not so explicit. To adjust for these changes, Heck, Price, and Thomas (2004) use the connections between students and their courses to infer the track a student is on. For example, if a student is mostly taking AP or other advanced courses, they are in the "higher" track. This helps overcome the limitation of previous literature in identifying what track a student is on since that information is no longer explicitly stated on transcripts.

To connect students to courses, the authors collected data in an urban, grade 9-12 high school in Hawaii with 1,500 students, and focused on the ninth-grade cohort of 402 students.

This school did not officially track its students, but fieldwork showed students neither were randomly assigned to classes nor had complete control of what classes they took. Rather, they were assigned into courses based on students' previous educational attainment. The authors gathered the data from transcripts to create a hypergraph. Heck, Price, and Thomas (2004) define a hypergraph as "a matrix, $H=[h_{ik}]$ with N actors and k events in which h_{ik} is the status of the relationship between actor i and event k." Put another way, the authors' hypergraph was a matrix with each student represented within each row, and each possible course represented across the columns. The intersecting cell represents if the student took the course or not. This overcomes previous limitations in the literature because this method does not assume students and classes are organized in any particular way and "accommodates the absence of formal track designations…"

The authors used a novel criterion for developing course profiles, resulting in seven distinct curricular positions. The authors then describe the seven curricular positions in terms of student academic performance, enrollment in food service, English proficiency, ethnicity, and other demographics. Consistent with previous research, the authors showed that students were placed in various tracks in non-random ways that revealed patterns based on ethnicity, standardized test scores, and other demographic variables. A weakness of this paper is that it only considers "student-course" events instead of "student-course-teacher-time" events. In other words, the authors only consider "Jamal took AP Calculus" rather than "Jamal took AP Calculus with Mrs. Sebring in 7th period during the first semester." This misses how some students may not only be grouped by the courses they take, but also the class period of that course (i.e., AP Calculus is taught during 1st and 2nd period, but students who perform better on assessments are put in 1st period).

Overview of Exponential Random Graph Models

Harris (2014) describes the early years of social network history as avoiding statistical modeling due to the complexities of the underlying data. Rather than using complex statistical modeling, empirical network research had largely relied on graphical representations of social networks and descriptive statistics of those networks. Later, however, two branches of network statistical modeling emerged. The first branch focused on actor-modeling that aims to predict characteristics of actors within a network. This branch would include predicting a person's anxiety level based on characteristics of the person and the anxiety levels of their friends. One can think of this branch as using network information as an independent variable.

The second branch focused on tie-modeling that aimed to explain or predict ties between actors. For example, the tie-focused branch aimed to predict if a tie exists between nodes i and j, and what network, dyad, or node covariates predict that tie existing. Exponential Random Graph Models (ERGMs) fall into this branch of social network analysis.

ERGMs are a SNA tool that allows researchers to estimate various network parameters given the presence or absence of other network structures or node composition (Robins, 2016). One can think of ERGMs as a version of logistic regression, where ERGMs estimate the likelihood two nodes within a network are connected to each other if different network, dyad, or node characteristics are met (Carrington & Scott, 2016). ERGMs estimate and evaluate model parameters to see how well they fit observed network data (van der Hulst, 2016).

Harris (2014) outlines different types of ERGMs, including: simple random graph models, dyadic independence models, and dyadic dependence models. In a random graph, a network is chosen out of the universe of possible networks of the same size and density. Ties in a random graph occur randomly; the likelihood of a tie occurring is purely based on chance. Erdos and Renyi (1959) showed how to model these kinds of graphs using simple random graph models, but these models could only assess the density of the network. Other network or social forces that may influence tie formation are ignored, which makes this form of modeling not very informative.

Later researchers noticed that observed networks often have a larger variation in degree and more reciprocity than random networks. Researchers in turn created dyadic independent models. These models allowed for the use of covariates such as homophily, reciprocity, node degree, and differential attractiveness, rather than only measuring density. These models assume that the absence or presence of a tie between two nodes is independent of any other ties between any other nodes. In many social networks, however, this assumption is untenable. For example, if Carmelo and Meredith are friends, and Meredith and Nick are friends, then it is more likely than chance that Carmelo and Nick are friends. Put another way, the ties between two individual nodes are influenced by ties between other nodes in the network. Dyadic dependence models were created to address this concern by allowing for the modeling of these network structures.

ERGM Estimation

Researchers using social network analysis want to know if the observed network, and network structures, are more or less likely to have occurred than just by chance alone. Throughout the history of ERGMs, statisticians have developed various algorithms to estimate parameter estimates. A pseudo-likelihood estimation technique that worked similarly to logistic regression was originally suggested by Strauss and Ikeda (1990). When change statistics are calculated, each possible tie becomes a case in a typical logistic regression procedure (Robins et al., 2017). However, this technique is at best “approximate” and generally does an inaccurate job of estimating parameters. This is in part because logistic regression assumes independence of observations, which in network analysis is unlikely to be the case.

As an alternative, researchers developed a family of techniques to measure networks. Robins et al. (2007) describes the process for modeling these networks. An *observed* network is the network a researcher has gathered data on and wants to model. The authors note that the observed network is simply one network from a set of *possible* networks with similar densities and numbers of nodes.

First, one approximates model parameter estimates using pseudo maximum likelihood estimation, then simulates a range of possible networks generated to determine a distribution of model statistics (Robins, 2016). Then, using Monte Carlo Markov Chain Maximum Likelihood Estimation (MCMCMLM), one simulates a sample of possible networks based on characteristics of the observed network (Becker et al., 2018). A large sample of possible networks is generated and represented by a probability distribution. The observed network is then placed within the probability distribution, and a maximum likelihood criterion is used to choose the parameter values most likely to occur given the data. Once the means are sufficiently close to the observed network, the estimation has "converged" (Robins, 2016).

ERGM Model Parameters

As previously mentioned, one can think of ERGMs as a version of logistic regression where the model predicts the change in likelihood that two nodes are connected if different conditions are met. This change in likelihood is measured in log-odds. When identifying what predicts if two nodes are connected by an edge, there are a number of model parameters to choose from. One can divide predictors into separate levels including network, dyad, and node levels. As this dissertation aims to explore different ways ERGMs can explore the mechanisms related to student co-course enrollment, it is worth mentioning the different parameters I used in my models.

Network Parameters

The first parameter is a network term within the R-package *statnet*. One can think of the parameter "edges" as the baseline likelihood of any two nodes being connected. When this is the only parameter in a model, it calculates the overall network density of a network. Typically, density is reported as a percentage of *the total number of actualized ties* divided by the *total number of possible ties.* For example, if there are 30 ties in a 10-node directed network, there are 10*10-10, or 90, possible ties. The density of this network would be 0.333…, or "33.3…% of possible ties are actualized." Within the *statnet* package, the edge parameter is given as the log of the odds of two ties being connected. This is easily converted back into density as a percentage by exponentiating the log-odds and using (odds / (1 + odds)).

Dyad Parameters: Homophily

An important dyad-level ERGM term in the *statnet* R package is "nodematch". This parameter measures the change in likelihood that two nodes are connected given that the two nodes share a common characteristic. For example, if two students are in the same grade, the likelihood the two students are connected will likely increase. This is one of the major homophily parameters used throughout this dissertation. I examine how different student characteristics such as grade, race, SEN, and FSA scores are associated with the likelihood of sharing the same classroom.

Nodematch can also be specified for uniform or differential homophily. Uniform homophily measures the change in likelihood that two nodes are connected if they share the same group. This measure assumes each subgroup within a group is as homophilous as the others. In other words, uniform homophily measures the change in likelihood that i and j are connected if i and j share the same group affiliation (same race, same gender, same grade, etc.).

Differential homophily captures how different subgroups may be more or less homophilous than others. Put another way, differential homophily measures the change in likelihood that two nodes are connected if they are both a *specific* category within the attribute. Where uniform homophily refers to "if two students share the same race," differential homophily measures "if two students are both Black" separately from "if two students are both Asian."

The estimate of the nodematch parameter is reported in log-odds, and the comparison group is always the likelihood two *dissimilar* peers are connected. Put another way, the nodematch parameter for grade level homophily may be estimated as "2.45," which means, if two students share the same grade, the likelihood they are connected is 2.45 log-odds higher than students who don't share a grade. Differential homophily is estimated in the same way, except the comparison is between same-sub-group homophily and non-same-sub-group homophily.

There is a second node-level ERGM term in the *statnet* R package is "nodemix". This also measures homophily but does so in a different way than the nodematch parameter. The nodemix parameter uses a mixing-matrix to calculate the change in likelihood that students are connected between specified groupings in relation to a specified reference group. I have reproduced a nodemix output for reference.

Table 1: Example of Nodemix Estimate Results

Example of Nodemix Estimates			
	Grade 6	Grade 7	Grade 8
Grade 6	-0.064		
Grade 7	-2.843	-0.129	
Grade 8	-4.989	-3.08	-0.886*
* Reference group			

The reference group in this case is the grade 8-grade 8 connection. The parameter estimate between eighth-grade students is taken from the estimated "edge" parameter and acts as the intercept. The interpretation of the intercept is that the log-odds of any two eighth-grade students being connected is -0.886. If we wanted to know the likelihood that a sixth- and eighth-grade student are connected, we would add the intercept and the estimate for the grade 6-grade 8 connection. The likelihood that a sixth-grade student and eighth-grade student are connected decreases from the intercept of -0.886 by 4.989, to -5.875 log-odds.

Nodematch and nodemix can both be included in the same model, but only if the specified attributes are different. Put another way, one can have a nodemix for race and nodematch for grade in the same model. However, one cannot have a nodemix and nodematch for grade in the same model. This dissertation only used nodematch parameters in the models, but future researchers can incorporate nodemix parameters.

Node-Level Parameters

One node-level ERGM term in *statnet* that is important to mention is the "nodefactor" term. This parameter looks at the change in likelihood that two nodes are connected, given that a node is from a particular group. Nodefactor can estimate how students of different races are more or less successful at tie formation than students of other races. This parameter specifies a reference group, and then each estimate is the difference between a specified group and the reference group. If there are three groups, this parameter will estimate two different results. For example, within the dataset I have categorized students into three groups based on their special education needs. There are students with disabilities, gifted students, and students in neither category. The reference group is non-gifted students without disabilities. The nodefactor parameter estimates how much more or less likely students with disabilities and gifted students are to make ties compared to non-gifted students without disabilities.

A second node-level ERGM term in *statnet* worth mentioning is the “nodecov” parameter. While the nodefactor parameter is for categorical node variables, nodecov is for continuous node variables. For my dissertation, this was used to control for variation between students in the total number of classes a student takes. For example, some students may only take half-year courses, while others will take quarter-year classes. It is important to control for this variable to obtain better estimates for other variables.

Network Backbones

To aid in the analysis of ERGM results, I utilized network backbones. Domagalski, Neal, and Sagan (2021) describe a network backbone as a weighted subgraph of the network that only includes the most important edges in the network. Not all edges are equally important. The network backbone identifies those edges that are part of the “core” or “backbone” of the network. One can remove extraneous edges by using a threshold such as “only consider edges with weights of above 3.” A more sophisticated method, however, is creating a statistical test measuring an edge’s observed weight compared to a distribution of possible weights given a null model. If the probability that an observed edge’s weight is beyond a certain p-value, it is unlikely to have occurred by chance, and thus, it is important enough to be included in the network backbone.

I primarily used network backbones as a check for the validity of my analysis. Domagalski, Neal, and Sagan (2021) discuss how multiple challenges arise from analyzing bipartite networks as well as their transformation from bipartite to unipartite. These issues could threaten the internal validity of my dissertation and results.

The authors first discuss how the transformation from a bipartite to unipartite network creates a weighted network, which is just as challenging as analyzing the original bipartite network. Second, the authors talk about how bipartite projections inflate the level of clustering; projections of randomly generated bipartite networks will have artificial clustering.

Third, information about the edges that connect nodes is lost in transforming networks from bipartite to unipartite; a bipartite network can show if two students are connected through either Physical Education courses or AP Calculus, but once the network is transposed, that information is lost. Lastly, the authors discuss how information about the individuals in the network is lost. For example, in the present study, it may look like one student is highly connected within a network, but it may be that this is only the case because of the high number of classes they take. It is possible that they are less connected than we would expect due to their high course load, but with the transformation from bipartite to unipartite, that information is lost.

A fifth consideration is that the projection from bipartite to unipartite can generate an enormous number of additional edges. Imagine a high school with 20 students and a single class. As a bipartite network, this network has 20 edges. However, when transformed, this network will have over 300 edges, each student connected to each other student in the class. In a school with hundreds of students, the number of edges can get into the hundreds of thousands. The network backbone eliminates extraneous ties that may have just occurred by chance. This creates the issue of hairballs, as discussed earlier, where there are too many edges to make sense of the network.

This dissertation avoids the first concern by using a binary network. To avoid the second concern, I used network backbones on selected networks to identify if the amount of clustering that occurred in the network is artificially inflated due to the nature of bipartite projections. By comparing ERGM results and sociograms of the original network to the network backbone, I could determine whether this was a concern. The third concern was not directly addressed in this paper, but future research could create two networks separating core courses (Math/Science/History) and elective courses (Physical Education/Home Economics/Foreign Languages). One could then assess if these networks are correlated with one another, among other research questions. The fourth concern was addressed by calculating the number of courses each student takes and using this as a covariate in the ERGM models. This allowed me to keep the information about the underlying connectivity of each student. The amount of clustering within the sociograms based on student characteristics, the exponential increase in the number of edges is not an issue, as the resulting hairball is a feature of the analysis, not an error.

Sociograms and ERGMs Working Together

The purpose of this dissertation is to explore how sociograms and ERGMs can add to the academic study of curricular networks, especially regarding how students are clustered together based on demographic and academic characteristics. While ERGMs provide an estimate of how clustered students are based on student characteristics, sociograms give context to these estimates. If grade level and academic performance are both highly predictive of student course co-enrollment, ERGMs provide an estimate comparable across schools, and sociograms provide meaningful visualizations. The next chapter of this dissertation explains more about the dataset, how to estimate ERGMs, how to create sociograms, and what models were run.

Chapter 3: Methodology

Social network analysis is both a theoretical framework (i.e., the world is made of connections between actors) and research methodology (that aims to model these connections). The purpose of this dissertation was to explore the ways researchers can use social network analysis tools, specifically ERGMs and sociograms, to characterize within-school clustering on the basis of grade level, academic performance, race, and special education need. Previous methods of measuring curricular networks are inadequate and should be replaced with these more sophisticated tools. Using observational and cross-sectional data, my aim is to use ERGMs and sociograms to answer the following questions:

- "In what ways can SNA **visualize** within-school clustering on the basis of grade level, race, academic performance, and special education need?"
- "In what ways can SNA **measure** within-school clustering on the basis of grade level, race, academic performance, and special education need?"
- "What information about within-school clustering do SNA visualizations provide compared to the information provided by SNA measurements?"
- "What information about within-school clustering do SNA measurements provide compared to the information provided by SNA visualizations?"
- "What information can be gleaned from using both SNA visualizations and measures in tandem?

This portion of the proposal presents the different methodological considerations taken to answer these research questions.

Research Design

The purpose of this dissertation proposal is to explore the different ways ERGMs and sociograms are useful in studying within-school clustering of students within schools' curricular networks. To do this, I used an explanatory correlational research design. Explanatory correlational designs are used for exploring and explaining relationships between variables that are not manipulated by the researcher, or cannot be manipulated (Fitzgerald, Rumrill, & Schenker, 2004). This dissertation uses data from previous school years, which I did not and could not manipulate. Additionally, correlational research is not designed to elicit causal explanations; instead, they merely describe the relationship between variables. When an ERGM is run, for instance, there may be a strong positive association between two students being in the same grade and their likelihood of being in the same classes. While ERGMs may merely describe these associations, visual analysis of the sociograms can add additional explanatory power to the estimated ERGM parameters.

The Location

For this dissertation, I have received data from a large, diverse Florida school district through their Institutional Review Board. I decided to choose a district in Florida for a number of reasons. The first reason is the large size of school districts in Florida. On the list of the 50 largest school districts in the United States, Florida has 10 districts (NCES, Table 215.20, 2018). This is in part due to Florida school districts following local county borders and these counties also being quite large. My analysis requires me to examine a large number of schools in order to get a sense of a variety of different types of schools.

A second reason for choosing Florida is related to the first. Since Florida has such large districts, they also have a larger tax base for collecting property taxes, which fund schools. This additional funding is often then used to develop a research division with the district offices. Small districts with very few schools would be unlikely to afford large research divisions, while many Florida school districts are large and well-funded enough to afford teams of people who conduct evaluations and research within the district. These divisions host large amounts of data and have employees who are data-savvy. These employees can manipulate and distribute data in a safe and secure manner while also maintaining data integrity.

A third reason I chose Florida was because it is a racially diverse state. According to the US Census (Census QuickFacts, 2019), Florida has a population of over 21 million people, 54% of whom are non-Hispanic White, 26% Hispanic, 17% non-Hispanic Black, 3% Asian alone, 2% Multiracial, and less than one percent of a different race. Due to the fact that I am looking at school clustering on the basis of race, it makes sense that I need to find a school district in a state that is racially and ethnically diverse.

The Data Itself

The Florida DOE and local school districts are in constant communication about student and school data. There are multiple standardized datafiles that the State and local school districts share, such as the "INDV" file containing information about student achievement, the "DEMO" file containing student demographic characteristics, and the "COURSE" file containing information about what courses each student is taking. The DEMO, INDV, and COURSE files come from student records from school years 2014-2015 and 2015-2016. One can connect these files to one another through district school ID numbers, state school ID numbers, and Social Security numbers. Because of this, the datasets are remarkably clean and malleable.

The Edge File

For social network analysis, one typically needs two files or sets of information. The first is a dataset of what the "edges" are. Edge files sometimes tell us who is connected to whom, but in the present case, the edge file tells us who is taking which classes. This, in turn, was transformed into a matrix indicating which students are taking classes together.

Information on network edges came from the “COURSE” file. One can think of this file as a transcript file that contains the list of every class each student in the district is taking across the year. Each student has one entry per student per class, so if a student is taking 16 classes across the year, that student will be in the dataset 16 times. I converted the Course file into a matrix through several steps.

The first step was to convert the district’s data from the form which it was stored in into a matrix. The district stored the course file spreadsheet as an “edgelist” where the first column is a student’s ID number, the second column is a course that the student is enrolled in, and each row is a student-course relationship. For example, if Isabella is taking AP Biology, AP Calculus, and Physical Education, Isabella would be in the first column dataset three times, connected to the three courses she is taking (in the second column).

This was then converted to an affiliation matrix where each student appears in one row only, each course is a column, and the intersecting cells indicated if the student is taking the course (1) or not (0). Using matrix algebra, I multiplied this network by the transverse of itself, making the network unipartite. This transformed network had each student appearing in one row only, and then one column only, and the intersecting rows now represent the number of courses the two students take together. This was then made into a binary network. The final matrix is a unipartite network of students connected to their classmates.

This is a similar methodology to Heck, Price, and Thomas (2004) where the authors used “student-course” events to measure curricular networks. I made a major deviation, however, from the methodology of this paper. The authors initially attempted to analyze the similarity of ties between students and “course-teacher-time” events but changed this to simply reflect student-course events. In other words, originally, the authors wanted to do the analysis using

“AP History – Mr. Calcagno – 1st Period” but changed it to simply reflect “AP History.” This was in line with the purpose of their study, which was to characterize different course clusters and to see what students resided in these different clusters. Because this dissertation has a different purpose, I will use the course-teacher-time event to better visualize and measure the co-occurrence of students within classes.

This allowed me to take into consideration the idea that students are not only stratified in courses but may also be tracked into specific classes of the course—i.e., high-performing students get placed into Algebra II in 4th period, while low-performing students get placed into Algebra II in 3rd period. Using course-teacher-time events also allows me to measure homophily more directly within the network. Heck, Price, and Thomas (2004) measure homophily by comparing the percentage of students in a cluster to the percentage of students we would expect to be in that cluster at random. For example, if 36% of a cluster is made up of Hispanic students, but Hispanic students make up 18% of the school’s population, then Hispanic students are overrepresented in that course. This can become more difficult when measuring clustering on the basis of multiple student attributes. Using exponential random graph models allowed me to overcome this issue by controlling for multiple sources of homophily at once.

The Node File

The second dataset is the “node” file. This file tells us information about individual nodes. For example, a node file will tell us the race, grade level, gender, and disability labels of students. Districts within Florida regularly have access to both kinds of data. These data came from two different district datasets, the INDV and DEMO files.

The INDV File

The INDV file is a statewide and standardized dataset that has one entry per student. Each line has information on the student such as their name, date of birth, race, sex, grade, and a few other characteristics. However, the main point of this dataset is to house assessment data. Each assessment across the state is included as a variable name. These assessments include the Florida Standards Assessment (FSA) taken in different grades in English, Math, Algebra, and Geometry, just to name a few. I used this dataset primarily for the assessment data, most specifically English FSA Scores. I used English FSA scores rather than other tests because students in grades 6 through 10 take the English FSA each year, while students only take math, science, or other tests in specified grade levels. By using a test students take on a near-yearly basis, I was able to have a consistent indicator of academic performance for students.

The DEMO File

The DEMO file is found across school districts within Florida and contains a plethora of demographic variables about the students. These variables include race, sex, language spoken at home, if they receive public housing assistance, and their address, among many other variables. These datasets are similar to one another across Florida, but some have different variables included. For example, one difference between two districts I know of are whether they have a variable for "student has internet access at home." However, most districts have the demographic variables of interest to this paper: race, grade, primary exceptionality, and free-or-reduced lunch status.

Data Collection Procedure

Since the data was housed by a local school district, I went through their in-house Institutional Review Board. I requested the data to be specifically structured in a particular way that would allow quick creation of both the edge and node files for later use. District employees identified the necessary data sources and combined them according to specific instructions.

District employees were asked to start with the student COURSE file. They were then asked to left join student characteristics from the DEMO file to the COURSE file. Lastly, they were asked to left join student FSA scores to the COURSE file. This created a large dataset where each line contained a student-course event with information on the student taking the course as well as information on the course.

Below, I have given a simulated version of what the data looks like. For example, student 9022348 is a White ninth-grade student taking math, English, science, and history. The first three columns reside in the COURSE file, and the fourth and fifth come from the DEMO and INDV files, respectively. The DEMO and INDV files were left joined into the COURSE file on the unique student identification number of the student (column 1). The second and third columns indicate information about the course, while the first, fourth, and fifth column are related to the students themselves.

Table 2: Example of Final Dataset Format

Student_ID	Course_ID	Course_Name	Race	FSA_Score
90223	9840179	Math	White	255
90223	4346021	English	White	255
90223	6099678	Science	White	255
90223	8600637	History	White	255
47373	9840179	Math	Black	275
47373	4346021	English	Black	275
47373	5222519	Music	Black	275
47373	8600637	History	Black	275
74906	5222519	Music	Asian	285
74906	6099678	Science	Asian	285
74906	9840179	Math	Asian	285
74906	8600637	History	Asian	285

Due to the way the final dataset was constructed, it was easy for me to separate it into an edge file and a node file. To create the edge file, I simply only kept those columns that resided in the COURSE file. In the above table, this meant deleting the columns for student race and FSA score. To create the node file, I did the reverse. I deleted the columns with information on the specific courses, and only kept student information. After deleting these columns, I deduplicated the rows. Because each row was originally a student-course event and student information was matched to it, by deleting the course information the resulting table is simply a *student* event. The resulting table then had multiple identical rows for students. Using the above table as an example, if I deleted columns 2 and 3, I would be able to deduplicate the 12 rows into three distinct rows.

Variables

Grade Level

The first variable of interest is grade level. In the US, students across the country are typically placed into classes based on the K-12 system. Certain courses are often only available to students who pass pre-requisites, which are sometimes restricted to certain ages of students. Because of this sorting mechanism, it is important to model grade level when examining how students are sorted within a school's curricular network.

This dissertation only looked at the course co-enrollment patterns of middle (grades 6-8) and high school (grades 9-12) students. Elementary school students (K-5) were excluded from analysis. This is because most elementary school students within this district only take one course with the same teacher throughout the day. Put another way, students in elementary school typically only have one set of classmates, which stays static throughout the day. When these networks are transformed from bipartite to unipartite, the networks become star patterns, where every node is connected to each other node with a density of 1. These networks are not informative to how students are connected to one another because these students are connected to all of their classmates.

Florida Standards Assessment

The Florida Standards Assessment (FSA) is Florida's statewide exam to assess student academic performance and growth, as well as assess school-level performance. The FSA has multiple tests including English, Science, and various Mathematics courses. The FSA is a criterion-referenced test that has different diagnostic scale scores depending on the grade level and test the student takes (FLDOE, 2019). These diagnostic scale scores are converted to *achievement levels* on a one to five scale, one being the lowest and five being the highest. A student scoring a level 1 or 2 on the FSA indicates they are below grade level, while scores 3, 4, and 5 indicate the student is at or above grade level. These achievement levels are comparable across grade levels and tests, which allows for cross-year and cross-grade comparisons.

Because FSA levels are on an ordinal scale, there are other methods available for measuring homophily. For example, among other options, I could measure the change in likelihood two students are connected to one another if they share the same exact FSA level, or based on the absolute difference between the FSA levels of two students. I chose to define homophily as whether or not two students are both scoring at grade level or above (passing) or below grade level (not passing) for two reasons. One is ease of interpretation. It is easy to understand that if two students are both passing, that they are more likely to be in the same classes. It is more difficult to explain that as the absolute difference of FSA scores increases, the likelihood of two students are connected decreases. Secondly, school district policy treats the categories of 1 and 2 very distinctly than how it treats the categories of 3, 4, and 5. If a student scores a 1 or a 2, they are often automatically enrolled in remedial courses. Future research should look to investigate the heterogeneity of how different FSA levels may form connections differently.

An additional reason for using the FSA is that students in grades 6 through 10 all take the English Language Arts (ELA) component of the FSA. For students in these grades, I used their previous year's ELA FSA score as a measure of academic performance. Students who do not meet the required passing score in tenth grade must take the exam again in eleventh grade, and if they continue to not pass, again in twelfth grade. Therefore, many students in eleventh and twelfth grades do not have test scores for those years because they had previously passed. For these students, I imputed the passing grade they achieved in tenth grade. In summary, academic performance of students in sixth, seventh, eighth, ninth, and tenth grade is measured by their FSA scores from the previous year; while academic performance of eleventh and twelfth grade students is measured by either their tenth-grade score (if they successfully passed the ELA FSA) or their most recent score.

Ethnoracial Group

Within the county, there are several different ways students can indicate their race. During student enrollment, parents indicate the ethnicity and race of the students. Ethnicity in this district asks if the student is Hispanic or not. Race is a "check all that apply" question, allowing for students to indicate one or multiple racial groups. While the state of Florida and the district typically separate the race and ethnicity of students, there is a district-level "local ethnic code" computed by the district. This is computed by taking the students' race and imputing it into the local ethnic code. However, if a student indicates their ethnicity is Hispanic, it overrides their response to the "race" code. For example, if a student indicates their race is Black and their ethnicity is Hispanic, their local ethnic code is registered as Hispanic.

Local ethnic codes can take one of seven different races: Asian, Black, Hispanic, Multiracial, White, American Indian, and Pacific Islander. For the purposes of this dissertation, American Indian and Pacific Islanders were removed from the dataset due to incredibly low numbers within the networks. There are often only one or zero students from these groups in a school; therefore, this would not give us information concerning the degree to which they are clustered within schools. Since part of this dissertation examines school-level characteristics that explain variation in homophily of these groups, these students were removed from the dataset. Future research taking place in regions with larger numbers of these students should use this dissertation's methodology to analyze how these students are clustered together within curricular networks.

Special Education Needs

The district identifies those students who have special educational needs. These needs include learning, intellectual, speech and language, hearing, visual, physical, and emotional disabilities, among others. Additionally, gifted and talented students are also considered to be students with special education needs. Each special education need has a unique code within the DEMO file. To simplify these categories, the district computes a "primary exceptionality," which is the most influential to how a student learns. This variable now only has one entry for the student's primary exceptionality, rather than listing each special education need of the student. This variable unfortunately ignores twice-exceptional students and students who have a disability as well as are gifted.

In my dissertation, I further simplified the special education needs category of each student into one of three groupings. The first group is gifted students. If a student has a primary exceptionality indicating they receive gifted services, they are placed in the "Gifted" category. The second grouping will be students with disabilities. If a student has a listed primary exceptionality other than gifted, they were placed in this category. Lastly, if a student does not meet any of these conditions, they were placed in a non-SEN category.

This classification system misses a few important aspects of special education needs. First, it misses students with disabilities who are also gifted. This is in part due to using the primary exceptionality code, which privileges giftedness over disability status. Said another way, a student's special education need will be labeled as "gifted" for those students with disabilities and labeled as gifted. Second, this will limit the how students with different disabilities may be sorted differently. By grouping students as either having a disability or not, this overlooks how different groups may experience homophily at different rates. For example, a student who is hard of hearing may be more likely to be integrated into the general school body than a student with severe disabilities who is likely to be placed in a self-contained classroom.

Variables Not Included

The dataset contains several additional variables that I did not include in the analysis. I excluded the variables for Limited English Proficiency label, gender, and free or reduced lunch status. Additionally, course level data included the value-added measure for the teacher of record for the particular class. This is a teacher effective metric, measured by the difference between a student's expected academic performance and their actual academic performance, aggregated to the teacher level. These variables were omitted from the analysis due to being outside the scope of this dissertation. While these omitted variables are important factors to consider, as they can form associations with course taking patterns of students, the other variables included took precedent.

Data Manipulations and Missing Student Data

I conducted a few different manipulations of the data to clean the dataset. The first manipulation was to separate edges of those students who attended two or more schools within the same year. For example, let us say that Meredith goes to East Boston High School in the fall semester, but West Boston High School in the spring. In this scenario, she will have connections to courses at both schools, making modeling the connections within schools more difficult. For students who attended more than one school during the school year, only their connections within the same school were considered. In the previous example, Meredith would have connections within East and West Boston High School, but not between them.

The second data manipulation I conducted was regarding missing student data. Student variables of highest interest were student race, disability label, grade level, and academic performance. As such, students with missing data for any of these categories, except for academic performance, were removed from the dataset.

A higher-than-expected number of students (13.9%) had no test data for their FSA scores, for various reasons. For many students, this is expected, as Florida exempts many students with Limited English Proficiency (LEP) and students with disabilities from assessment tests. An additional explanation is that students transferred from an outside district were not required to take the FSA if they transferred either after the FSA or within a certain time before the FSA. This is evidenced by a large proportion of those students with no FSA data in 2015, having data for 2016’s FSA. Of the 6807 students with no scores on their FSA, 3264 (48%) either had a disability, LEP, or have FSA 2016 data. Regardless, students with no FSA scores were not associated with any other group membership, suggesting no systematic biasing issues.

Analysis Strategy for Exponential Random Graph Modeling

As the purpose of this dissertation is to identify what network structural and compositional factors are associated with the presence or absence of classmate ties, I used exponential random graph modeling.

Dependent Variable

Similar to logistic regression, ERGMs estimate the association between the presence (1) or absence (0) of a tie existing between two nodes and different network statistics. This is the dependent variable of each model being run: whether or not a possible tie between two students is present or absent. As mentioned previously, the tie between students is based on whether or not they share the same "course-teacher-time" event. If two students are both in AP Calculus with Mrs. Gilmore during second period, the tie betwixt them is considered present. Instead of saying "these students both take the same class", the framing would change to "these students are classmates".

Independent Variables

The most important independent variables in this dissertation were the various measures of homophily, the tendency of actors to be more connected to other actors who share similarities. This dissertation aimed to answer the question of how homophily based on grade level, academic performance, race, and special education need is associated with the existence of classmate ties. This was primarily measured with the nodematch parameter in R. If the nodematch parameter was positive and statistically significant, it means that there is homophily based on that characteristic. Additional independent variables include other parameters such as nodecov, and nodefactor (defined in Chapter 1). Because this dissertation is exploratory, I ran several models with these different parameters. There are, however, consistent patterns for each model.

Modeling Round One

The first round of modeling created the baseline model simply measuring and predicting the likelihood that any two nodes in a network are connected by a tie. This model only estimated the *edges* parameter, acting as the intercept of the model—i.e., the baseline likelihood that any two nodes are connected by a tie. As mentioned previously, this is equivalent to how dense the network is. This baseline model also served as the initial model when conducting measures of model fit.

Modeling Round Two

The second round of modeling added two control variables: nodematch based on grade level and nodecov for number of classes. This adjusted the estimates for how grade level serves as the most influential mechanism for school clustering, and how number of *classes taken* is associated with shared classes. These parameters were used in all subsequent models.

Modeling Round Three

The third round of modeling introduced the nodematch parameter on race and created two separate models: one model for uniform and differential racial homophily. These models included an estimate for the nodefactor parameter for race, to see if students from different racial backgrounds are more likely to form ties in general.

Modeling Round Four

The fourth round of modeling simply altered the nodematch and nodefactor parameters in the third round of modeling to measure FSA homophily rather than racial homophily. I did this for both the uniform and differential models.

Modeling Round Five

The Fifth round of modeling simply altered the nodematch and nodefactor parameters in the fourth round of modeling to be Special Education Need homophily rather than FSA. I did this for both the uniform and differential models.

Modeling Round Six

The final round of modeling created models for uniform and differential homophily that incorporated nodematch and nodefactor parameters for race, FSA, and SPEN.

ERGM Checking Model Fit

To make sure the models fit the data, I used AIC and BIC to compare the original model to the subsequent models. If the model fit improved, it was evidence that the subsequent models reduce misfit and are truer to the data. Second, for model 6 (uniform and differential) I utilized the goodness-of-fit procedures in the R package *statnet* to show how the observed networks compared to the simulated networks. If the observed networks were within the bounds of the simulated networks, one would say the network is a strong fit for the data.

Additionally, for multiple networks, I compared ERGM results of the full networks to the network backbone. I did this to address internal validity concerns resulting from how projections of bipartite networks sometimes add extraneous clustering. By comparing ERGM results of the original and backbone networks, I was more confident that the results would be due to actual clustering and not an artifact of the network projections.

ERGM Multicollinearity

A common assumption of many statistical models is the absence of multicollinearity. Multicollinearity occurs when one or more predictor variables are too highly correlated with another predictor variable. In ERGMs, multicollinearity can lead to model degeneracy, non-convergence, and biased parameter and variance estimates.

A typical approach for identifying multicollinearity in regression models is to calculate the variance inflation factor (VIF). In regression models, this is calculated by regressing each term in the model by all other terms in the model, calculating the R-Squared, and dividing 1 by (1 - R-Squared). A VIF of above 10 is considered problematic, and indicative of major multicollinearity.

Calculating VIF in ERGMs is more complex for multiple reasons. The first reason is that the outcome variable is not typically a continuous variable and therefore does not have "variance" in the way OLS regression would. This is a similar reason to why authors have created pseudo-R-Squared values for logistic regression. The second reason is because observations in ERGM networks are not independent from one another; the fact that two people are connected is influenced by the connections around them.

Duxbury (2018) devised a five-step procedure for calculating VIF in ERGM contexts. The first two steps are to simply fit an ERGM, and then simulate a large distribution of networks from the ERGM parameters. This step creates hundreds of simulated networks similar to the original network and calculates their parameters. The third step is to calculate the bivariate correlation matrix from the sufficient statistics of these simulated networks. To calculate R-squared of y, as in OLS VIF calculations, one treats each explanatory variable as y, and regresses against it all remaining explanatory variables. This is then used to calculate VIF using the previously mentioned equation. I used Duxbury's method to identify multicollinearity within my ERG models for model 6 (uniform and differential).

ERGM Pseudo-R-Squared

In linear regressions, an R-squared value is an indication of model fit. On a scale from 0 to 1, it calculates the percent of the total variation in the outcome variable that is explained by the predictor variable(s). For OLS regression, this is straightforward because the dependent variable is continuous and has variance in the typical sense. However, for many statistical models such as logistic, Poisson, and multinomial regressions, the outcome variable is not continuous.

Many authors have developed pseudo-R-squared values for these models. For logistic regression, a common alternative to R-squared values is McFadden's (1974) pseudo-R-squared. Like ERGMs, logistic regression is estimated using maximum likelihood estimation. This estimation method chooses the parameter estimates that maximize the likelihood of observing the underlying data, while also calculating the level of misfit. One can turn this information into a measure of model misfit, for example, by taking the natural log of the observed maximum likelihood value and multiplying it by -2, often denoted by -2LL. The lower the -2LL, the better the fit of the model to the data. McFadden's pseudo-R-Squared takes the -2LL of the full logistic regression model and divides it by a logistic regression model that only has the intercept. It then subtracts this number from 1. Other measures of misfit could also be used, such as AIC and BIC.

This number, like traditional OLS R-squared, is on a scale from 0 to 1, with 1 being perfect fit. Because I estimated my ERGMs with maximum likelihood estimation, I could use McFadden's pseudo-R-squared as a measure of model fit. This is a novel way of estimating goodness of fit for ERGMs, as I have not been able to find any paper that applies McFadden's pseudo-R-Squared to ERGMs.

Data Analysis Strategy for Sociograms

Creating and Analyzing Sociograms

To create sociograms of the different schools' curricular networks, I used Gephi 9.2, an open-source social network analysis software tool. Powered by JavaScript, Gephi has faster computational capacity than R when it comes to creating large complex sociograms. For example, imagine a school with 1,000 students, where each student has 10 courses, and each course has 20 students. Each student could have upwards of 200 ties to others, and, when multiplied by 1,000 students, you are quickly generating a network with 200,000 edges. Where this is computationally difficult for R, Gephi does it with ease. I made use of position and color, to illustrate how students are distributed within the networks.

Using Gephi, I used node position to inform the viewer of the grade level and typical academic performance of students. Due to ability grouping and sorting students by grade level, these networks display a form of symmetry along two imaginary axes; one axis serves as a line dividing students by grade level and the other by academic performance.

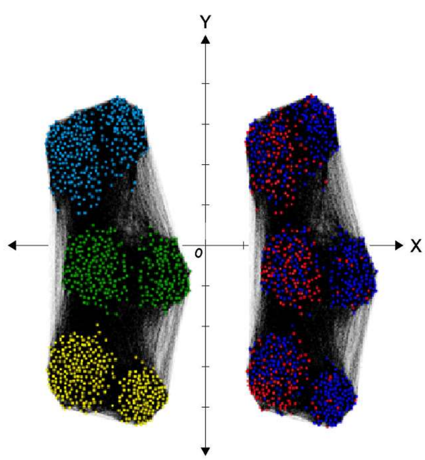


Figure 6: Middle School Sociogram

I reproduced two versions of a middle school's curricular network here: the left colored according to grade level, the right colored by academic performance, wherein each dot represents a student. Vertically, they are clustered by grade level: sixth graders in yellow at the bottom, and eighth graders in blue at the top. When I color the graph by academic performance, one can see that those who are below grade level (red) are more likely to be on the left of the graph while those who are on or above grade level (blue) are on the right.

I constructed all sociograms so that grade level descends along the vertical axis while academic performance increases along the horizontal axis from left to right. With this in mind, I added other variables such as SEN and race to the nodes while maintaining the information that students tend to gravitate towards one another on the basis of academic performance (horizontally) and grade level (vertically). For example, in the above graph, students in eighth grade who are at or above grade level on the FSA are most likely to be in the top right corner of the sociogram. If I color nodes by special education need or race, the information on where eighth-grade students at or above grade level are in the sociogram is left unchanged.

I constructed separate sociograms for selected schools. Keeping the positioning and orientation of the sociograms the same, I colored the nodes by grade level, race, FSA score, and special education need. The Gephi layout algorithm I used for each graph is called Force Atlas 2. This algorithm is computationally the fastest among the options Gephi offers. It also has the option to “dissuade hubs” and to “prevent overlap”. These options prevent nodes from hiding behind one another, which would obscure how the visuals display clustering. By preventing overlaps, it is easier to see how clustering occurs because no nodes are hidden from view.

Analyzing Sociograms

For each school, I created four sociograms, one for grade level, special education need, race, and academic performance. Each sociogram would be colored according to the variable of interest, and aligned so that academic performance was along the horizontal axis and grade level along the vertical axis. I compared these to one another and examined how clustering is varied based on different student characteristics. I made note of times student clustering based on one variable coincided with groupings based on other student characteristics. I asked questions such as “where are students with disabilities located?”, “where are Black students located?”, and “are the locations of students with disabilities similar to the locations of Black students?”.

Then, for each variable of interest, I looked at all the school networks together to compare how students may be clustered similarly based on the same student characteristics. For example, for each sociogram I colored the nodes based on student race. With academic performance along the horizontal axis and grade level along the vertical axis, I looked across the different sociograms to see if there were similarities in how students were clustered according to race. This let me see whether students were clustered together similarly based on demographic characteristics within different curricular networks.

Combining Analysis of ERGMs and Sociograms

The final component of my analysis strategy was to blend the analysis from exponential random graph modeling with the analysis of the sociograms. I did this by analyzing each school’s ERGM results alongside their sociograms. For example, to examine how each school clusters students based on special education needs, I first looked to see “where” clustering occurs based on the sociograms and “to what degree” based on the ERGM results. It could be the case that students with disabilities are clustered together away from the general student population. Sociograms can easily identify this phenomenon. If one colors a sociogram’s nodes by the disability label of the student, then one can immediately see where they are in the school’s curricular network. However, there is no accompanying measure of how much clustering there is. The nodematch parameter of an ERGM, on the other hand, can estimate the change in likelihood that a tie exists between two students, based on if they both share a common characteristic. Therefore, while ERGMs can estimate the degree of clustering in a network, they cannot tell us "Where" the clustering occurs in the network. In the above example, if a school has high levels of clustering based on special education needs, it cannot tell us if those students are being removed from the general student population or if they are simply more likely to be placed into certain kinds of classes. Using both sociograms and ERGMs in tandem sheds light on different aspects of how students are placed in the curricular network of their schools.

This chapter of this dissertation outlined how the dataset was sourced, created, and manipulated into a usable format for ERGMs and sociograms. This chapter also showed how to estimate ERGMs, how to create sociograms, and described what models were run. The next chapter discusses the results of these models and sociograms. It does this by describing the strengths of the separate approaches, and then what benefits exist by using them in tandem.

Chapter Four: Results

The purpose of this dissertation was to investigate and evaluate how social network analysis (SNA) can visualize and measure how students are distributed within schools based on grade level, race, academic ability, and special education need (SEN).

Before showing how SNA can measure and visualize student distributions in schools, I will first review the demographic characteristics of the school district studied.

Demographic Summary: Aggregated to District Wide Level

Across the school district, after applying filtering constraints listed in the previous chapter, my final sample included 48,959 students. These students were predominantly White (57.9%), followed by Black and Hispanic (17.4% and 15.8%), and Asian and Multiracial students accounting for 5% and 4% respectively. There were slightly more male students than female students (50.8%).

More students passed the ELA FSA than not (45.1% to 41%) while 13.9% did not have a FSA score. These students are likely students with disabilities who are exempt, or transfers from homeschool or other states without the FSA. Most students were neither labeled as gifted nor had a disability (84.8%), but 5.6% had a disability while 9.6% of students were labeled as gifted. Across grade level, most grades accounted for between 12.44% to 16.5% of students. Middle school grades had fewer students while high school grades had more students with a notable decline from 9th grade to 12th grades. Table Three shows this in more detail.

Table 3: Descriptive Statistics of Student Body

Descriptive Statistics of School District Student Composition			
Student Grouping		Total Students	Percent of Sample
Total Students		48,959	100.00%
Race	Asian	2,438	5.00%
	Black	8,512	17.40%
	Hispanic	7,714	15.80%
	Multiracial	1,968	4.00%
	White	28,327	57.90%
Gender	Female	24,091	49.20%
	Male	24,868	50.80%
FSA Score	FSA_Fail	20,088	41.00%
	FSA_Pass	22,064	45.10%
	FSA_No_Test	6,807	13.90%
Special Education Need	Disability	2,745	5.60%
	Gifted	4,715	9.60%
	Neither	41,499	84.80%
Grade Level	Sixth	6,569	13.42%
	Seventh	6,559	13.40%
	Eighth	6,381	13.03%
	Ninth	8,077	16.50%
	Tenth	7,630	15.58%
	Eleventh	7,654	15.63%
	Twelfth	6,089	12.44%

Demographic Summary: Aggregated to School Level

Certain patterns occurred after aggregating student demographic information to the school level. The median school had 1246 students, while the largest had 2510 and the smallest had 272. Schools varied significantly between each other on how students of different races, FSA scores, and special education needs, Table Four shows this in more detail.

Table 4: Descriptive Statistics of School Composition

Descriptive Statistics of School Composition

Student Grouping		Mean	Median	Max	Min	SD	IQR	Skew	Kurtosis
School Size	n_total	1360.0	1246	2510	272	526.38	702.25	0.19	2.47
Totals by Race	n_asian	67.7	52	223	8	51.61	54.25	1.20	3.78
	n_black	236.4	182.5	841	5	204.74	221.50	1.48	4.72
	n_hispanic	214.3	184.5	603	30	127.50	132.00	1.08	4.21
	n_multiracial	54.7	50.5	91	14	18.73	27.00	0.13	2.68
	n_white	786.9	681.5	1962	94	443.89	584.75	0.72	3.09
Percentages by Race	p_asian	4.9%	3.9%	12.2%	1.0%	0.03	0.04	1.034	3.202
	p_black	19.6%	12.9%	62.7%	1.3%	0.18	0.11	1.253	3.318
	p_hispanic	15.6%	13.3%	35.5%	5.7%	0.07	0.09	0.878	3.219
	p_multiracial	4.2%	4.2%	6.1%	2.7%	0.01	0.01	0.496	3.499
	p_white	55.7%	55.2%	82.0%	23.3%	0.17	0.24	-0.198	2.107
Totals by Gender	n_sex_f	669.2	601	1396	126	279.24	436.50	0.36	2.79
	n_sex_m	690.8	640	1271	146	262.18	337.50	0.20	2.58
Percentages by Gender	p_sex_f	48.8%	48.0%	58.5%	38.3%	0.05	0.05	0.108	2.899
	p_sex_m	51.2%	52.0%	61.7%	41.5%	0.05	0.05	-0.108	2.899
Totals by FSA Score	n_FSA_Fail	558.0	553.5	978	86	217.02	232.00	-0.10	2.68
	n_FSA_Pass	612.9	568.5	1609	6	324.23	338.75	1.03	4.40
	n_No_Test	189.1	164.5	358	20	87.31	127.25	0.28	2.26
Percentages by FSA Score	p_FSA_Fail	42.1%	43.8%	61.8%	20.0%	0.11	0.14	-0.460	2.438
	p_FSA_Pass	43.8%	42.5%	77.6%	2.2%	0.14	0.13	0.015	4.296
	p_No_Test	14.2%	13.9%	39.0%	2.5%	0.05	0.03	2.499	14.451
Total by SEN	n_disability	76.3	40	413	0	99.82	97.50	1.86	6.07
	n_gifted	131.0	130	223	13	51.41	64.25	-0.26	2.63
	n_neither	1152.8	1024.5	2300	216	526.45	828.25	0.32	2.23
Percentages by SEN	p_disability	7.2%	3.2%	43.4%	0.0%	0.10	0.10	1.948	6.335
	p_gifted	10.0%	9.9%	20.6%	3.3%	0.03	0.03	0.511	4.524
	p_neither	82.8%	85.7%	92.3%	46.8%	0.10	0.10	-1.943	6.891

Regarding race, Black students comprised 12.9% of the median school, and ranged from 1.3% to 62.7%. White students comprised 55.2% of the median school, but the range was from 23.3% to 82%. If students were distributed across the schools evenly, we would expect these numbers to be closer to the district level measures for Black and White students (17.4% and 54.9% respectively).

Regarding FSA Scores, the median school has 43.8% of students passing, and 42.5% of students failing. However, the range of percent passing ranged from 2.2% to 77.6% and the range of percent not passing goes from 20% to 61.8%.

Students were also distributed unevenly across schools based on SEN. While the median school had 3.2% students with disabilities, and 9.9% gifted, these percentages ranged upwards of 43.4% and 20.6% respectively at the maximum school. Some schools had as many as 53.2% of their students being either gifted or having a disability.

Exponential Random Graph Models (ERGMs)

Model One

The first model run was the null model which calculated the baseline likelihood that two nodes are connected simply by both nodes existing in the same network. These models only had the ERGM edge parameter incorporated into the model. For the 34 schools in the dataset, the average edge parameter was -2.127, with a maximum of -1.157 and a minimum of -2.725. Across all schools, the estimate was statistically significant, where p was less than or equal to .001. This shows that, at a typical school, any two students have a very low likelihood of being classmates at random.

An ERGM edge parameter value of 0 would indicate a 50/50 chance of being classmates. Turning the edge parameter values for mean, max, and min into percentages, students would have a 10.7% chance, 23.7% chance, and 6.01% chance, respectively.

Model Two

The second model added a grade level node-factor and nodematch as well as a node-covariate for the number of classes taken. Intuitively, two students sharing a grade level would be associated with their likelihood of being classmates, as would taking more classes than others.

For the edge parameter, the mean, maximum, and minimum estimates were -5.17, -2.26, and -8.04, respectively. For nodematch on grade level, the average estimate was 2.28, 3.6, and 0.48. All estimates for nodematch on grade level were statistically significant at the p = .001 level. The node-covariate of number of classes had mean, max, and min parameter estimates of 0.11, .25, and -.04, and was statistically significant at 32 out of 34 schools (p = .001). McFadden's R-Squared for model 2 ranged from 0.016 to 0.375, with an average of .164 log odds. This suggests that parameters in model 2 explain around 16.4% of the total variance.

Model Three

Each subsequent model was divided into both a uniform node-match model (3-U) and a differential node-match model (3-D). As previously discussed, uniform models look at the change in likelihood based on if two people share the same group. Both types of models aim to measure the change in the likelihood two nodes are connected based on if they share the same group, but uniform assumes all groups have similar homophily, while differential models drop that assumption. Subsequent models also introduced a node-factor estimation to control for students from different groups being more or less likely to be classmates with other students.

In models 3-U and 3-D, I introduced the first parameters most interesting to this study, a node-match estimate based on the race of students. Model 3-U estimated that if two students in a school share the same race, there is an associated increase of 0.151 log-odds to the likelihood they are classmates. This effect had a dramatic range, from a maximum of .72, to a minimum of -.029. For uniform homophily based on race, 85% of schools have statistically significant estimates at the p<.001 level.

Model 3-D showed good reason to assume different racial groups have different levels of homophily. This estimate was strongest for Black students, the average homophily estimate was

0.307, and was statistically significant for 71% of schools. The estimate was weakest for multiracial students, with a homophily estimate of -0.05 and only 6% of schools were statistically significant.

Model Four

Models 4-U and 4-D focused on how similar academically performing students, as measured by the FSA, are associated with a change in the likelihood of two students being classmates. This was operationalized by identifying when two students both passed (by receiving a score of 4 or 5) or did not pass (by receiving a score of 1, 2, or 3).

When assuming uniform homophily for model 4-U, the average parameter estimate across the school district for nodematch based on FSA was .505, with 97% of estimates being statistically significant. After dropping the assumption of uniform homophily, model 4-D teased out how students who passed and students who do not pass may have different tendencies to form homophilous bonds. In fact, if two students are both passing the FSA, on average there was an associated .855 increase in the log-odds they are classmates. This is much higher than the associated increase for non-passing students, with an average of .312. These were statistically significant differences across schools 84% and 88% respectively.

Model Five

Models 5-U and 5-D focused on how students with similar special education needs may be more or less likely to form classmate ties with similar peers. As previously summarized, I condensed the different categories of special education needs into the categories of Gifted, Student with Disability, or Neither because the dataset only listed one “primary” special education need, and many disability categories were too small to model. This unfortunately removes the possibility of identifying students; with two or more disabilities, a disability and gifted, or how students with different disabilities may be tracked throughout school curricular networks.

Model 5-U showed the average parameter estimate for special education was 6.11, with 100% of schools being statistically significant. Among uniform homophily based on racial group, academic performance, and special education needs, only the parameters for special education needs had 100% of estimates being statistically significant at the $p<.001$ level. Model 5-D showed that differential homophily for students with disabilities and students labeled as gifted have significant positive associations with being classmates. The average school's parameter estimate for differential homophily for students with disabilities was 1.29, and for gifted it was 2.35. These estimates were statistically significant for 97% and 100% of schools, respectively. For non-gifted students without disabilities, the associated change in likelihood of forming classmate bonds with similar students was -0.299, which was statistically significant at 91% of schools.

Model Six

The final two models, models 6-U and 6-D, incorporated all three variables of interest into the calculations. They added a node-match parameter for race, FSA, and special education needs. They also included node-factor controls for grade level, race, FSA, and special education needs. Because these were the final models, I include Tables five and six to show additional descriptive statistics regarding the parameter estimates.

Model Six – Uniform Homophily

Across the district, the mean edge parameter and node-covariate for number of classes taken was -5.52 and .122, respectively. This means any two students in the same school taking the average number of classes (8.5 classes) would have a baseline likelihood of 1% to be classmates. If a student shares the same grade with another student, but no other characteristics, this likelihood changes to 9%.

Extrapolating further, I did calculations separately for the main variables of interest; the likelihood students in the same grade - taking the average number of classes - are classmates with other students who are also of the same race, same FSA, and same SPEN, goes from 9% to 10%, 14%, and 16% respectively. Substituting the average parameter estimates for the maximum parameter estimates gives us changes in likelihood from 9% to 16% for race, 19% for FSA, and 64% for SPEN.

Table 5: Model 6-U Parameter Estimates

Descriptive Statistics for Parameter Estimates of Model 6-Unform Homophily

Parameter	Median	Mean	Maximum	Minimum	Percent Statistically Significant
Edges	-5.52	-5.72	-2.26	-9.38	100%
nodecov.NClasses	0.124	0.122	0.248	-0.033	97%
nodematch.grade	2.18	2.32	3.83	0.513	100%
nodematch.race	0.069	0.107	0.585	-0.0284	82%
nodematch.FSA	0.472	0.463	0.711	0.0284	97%
nodematch.SEN	0.502	0.57	1.94	0.185	100%

Special education needs appeared to be a major driver of homophily within curricular networks, though FSA was a close second. Uniform homophily based on race is heavily varied across schools, with heavy outliers pulling the average much higher than the median.

McFadden Pseudo R-Squared ranged from .019 to .384, with an average of .180. This suggests the parameters in this model explain about 18% of the total variance, a slight increase from 16.4% in model 2.

I checked the Variance Inflation Factor (VIF) of all estimated parameters, looking for those parameters exhibiting elevated (VIF >= 20) or very concerning (VIF >= 100) levels of multicollinearity. No parameters exhibited this amount of VIF, suggesting the measures do not display problematic levels of multicollinearity.

Model Six – Differential Homophily

Like the uniform model (6-U), the differential model (6-D) covariates for number of classes, grade level, and edges were relatively similar. This reinforces the idea that these are strong baseline control variables that help each model.

Black students had the highest mean and median parameter estimates for racial homophily, though the maximum parameter was higher for Asian students. The percentage of schools with statistically significant parameters widely varied across racial groups, from a low of 9% to a high of 68%. This may suggest that few schools are driving homophilous results with heavy outliers.

There was a strong positive association between two students passing the FSA and forming classmate bonds. On average, students who have both passed the FSA have an associated increase of 0.761 log-odds of forming a classroom tie, compared to 0.259 for students who did not pass. These results were statistically significant at 94% and 85% of schools, respectively.

Lastly, there was a very strong positive association between two students both having a special education need and forming classmate ties with similar students. If both students have a disability, or are both labeled gifted, there is an average associated increase of .957 and 1.76 log-odds, respectively. These are statistically significant at over 95% of schools. These are the largest of the estimated parameters, suggesting that special educational needs are a major driver of homophily across schools.

Table 6: Model 6-D Parameter Estimates

Descriptive Statistics for Parameter Estimates of Model 6-Differential Homophily

Parameter	Median	Mean	Maximum	Minimum	Percent Statistically Significant
Edges	-5.96	-6.18	-2.7	-9.57	100%
nodecov.NClasses	0.123	0.123	0.261	-0.033	97%
nodematch.grade	2.18	2.33	3.87	0.513	100%
nodematch.race.A	0.106	0.142	1.28	-0.488	29%
nodematch.race.B	0.137	0.213	1.25	-0.192	68%
nodematch.race.H	0.036	0.019	0.171	-0.369	44%
nodematch.race.M	-0.035	-0.038	0.294	-0.433	9%
nodematch.race.W	0.032	0.064	0.436	-0.051	41%
nodematch.FSA.Pass	0.761	0.736	1.2	-0.767	94%
nodematch.FSA.Fail	0.259	0.293	1.0	-0.011	85%
nodematch.FSA.No_Test	-0.003	0.0044	1.23	-0.603	65%
nodematch.SEN.Disability	0.957	1.08	2.61	0.285	97%
nodematch.SEN.ND	-0.177	-0.187	0.468	-2.07	77%
nodematch.SEN.Gifted	1.76	2.03	4.85	0.615	96%

Checking the Variance Inflation Factor (VIF) for the differential model parameters proved more difficult, specifically the node-factor parameter. Many schools had too few students from different groups to have correctly estimated parameters. For example, if a school has zero gifted students, or zero multiracial students, these parameters cannot be estimated, creating a "null" value. The program to calculate VIF does not accept null values, and crashed when it reaches these errors. For those schools where the VIF calculation did not fail, there was not enough significant multicollinearity to cause concern based on the recommended limits.

McFadden Pseudo R-Squared for the differential homophily model ranged from .018 to .394, with an average of .183, only slightly better than the uniform homophily model. For reasons of parsimony, this slight increase in model fit may not be worth the additional 7 parameters this model estimates that the uniform homophily model does not (model 6-U estimates 10 and 6-D estimates 17).

I randomly chose 10% of the schools in my dataset to rerun Model 6-D using a network backbone. Taking the original parameter estimates, I compared them to estimates from the backbone models. For these schools, no parameter estimates changed dramatically enough to cause concern.

Distribution Visualizations of different statistical parameters

There was a great deal of variation across schools for each ERGM parameter estimate. I call out a few notable examples below. The first parameter of interest is the nodematch parameter for homophily for Black students, shown by Figure 8.

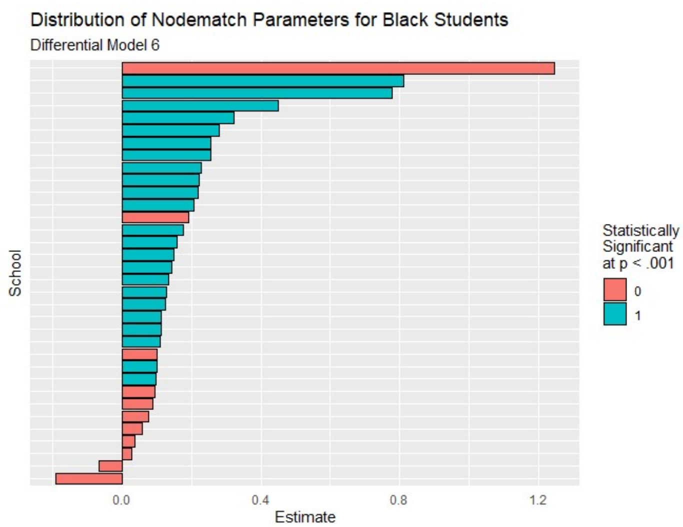


Figure 7: Nodematch Parameters for Black Students

Across the district, 68% of these parameters were statistically significant. It also appears that there were a few outliers pulling the average upwards. This could be a useful method for identifying those schools with the most racial homophily for Black students, and then using qualitative tools to identify what mechanisms are likely responsible for this trend. After the mechanism is determined, these outlying schools could be good candidates for interventions to reduce racial homophily.

In almost every school, there was a statistically significant amount of homophily based on academic achievement according to FSA scores. Figure 9 shows the distribution of nodematch parameter estimates for non-passing students.

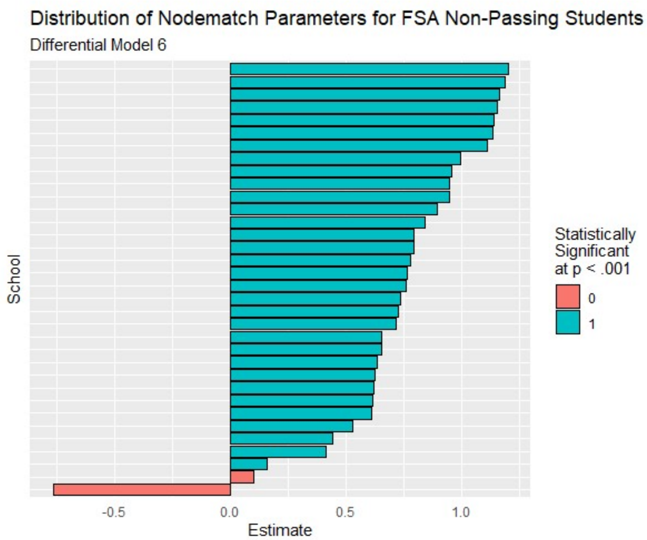


Figure 8:Nodematch Parameters for FSA Non-Passing Students

No major outliers appeared were statistically significant. These data could then be used to compare different districts to one another. If another district has less academic performance homophily, we can incorporate those measures in regression models to predict other school level variables such as the percentage of students passing a standardized test.

Measuring Clustering by Exponential Random Graph Models Wrap Up

These results show there are significant opportunities in using ERGMS to measure clustering of school curricular networks using ERGMs. Researchers can use the parameter estimates from the ERGM results as covariates in other regression models for outcome measures for quasi-experimental models such as regression discontinuity design, tracking change over time, or identifying schools with high levels of homophily for interventions.

Sociograms

High School Sociograms

Sociograms are powerful tools for visualizing social networks. While they are often described as "hairballs" with little informative value, the following sociograms of curricular networks (when modified correctly) can provide insights into how students are distributed across demographic groups.

Take for example school 3062, a high school with a majority black population. When colored by grade level (top left), the sociogram appears how one may expect; there is tight clustering within grade levels, but also plenty of overlap between grades. This clustering of

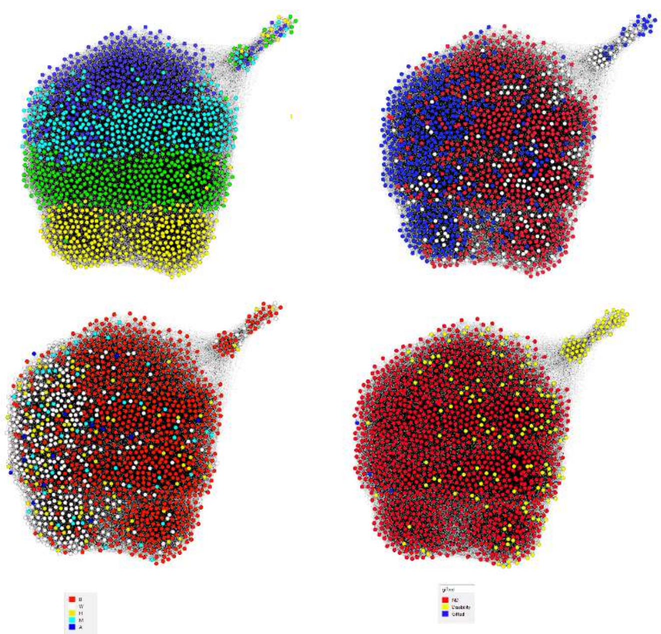


Figure 9:High School Sociogram

groups shows that 9th grade students are more likely to form classmate ties with 10th grade students, rather than 12th grade students. For high schools, one can often use the convention that grade level acts as a "vertical axis" for the graph. Many sociograms do not use position/location as meaningful orientations within their visualization (one can rotate or flip many sociograms and the meaning stays the same); however, by implementing grade level as the vertical axis and treating sociogram orientation graphically, we can compare the sociograms of different schools to one another to elicit meaningful relationships between them.

School 3062 also appeared to have significant clustering around FSA performance (top right) Students who passed the FSA (blue) are presented as clustered together on the left side of the graph, whereas students who did not pass appeared more on the right side of the graph. This is a second convention by which we can graphically position and define data distribution within school sociograms, as most schools experience this phenomenon. While grade level can orient us North to South, academic performance can orient us East to West.

School 3061 appeared to have a significant amount of clustering based on ethnoracial groups (bottom left). White students tended to be on the left side of the graph, while Black students tend to be located on the right side of the graph. Black students also tend to be located in the "tail" of the network sociogram on the top right side of the graph. However, this distribution is more likely due to the influence of the 4th graph (bottom right), where nodes are colored by special education needs; the tail of the graph is completely comprised of students with disabilities. This is likely due to there being a self-contained classroom. We can also zoom into the tail to see how there appear to be two clusters. This shows some students with disabilities are connected to the major network component, while others are less connected to it. This is due to some students with disabilities taking physical education courses with students without disabilities, while others do not. This is why it's important to construct sociograms in layers and include these layers in reporting and results. The different layers of sociograms can give insight to how the clustering of different groups relate to one another.

Middle School Sociograms

School 2562's sociogram looks like a typical middle school within the school district. Regarding grade level (left), there are three major clusters, representing sixth (bottom), seventh (middle), and eighth grade (top), splitting the network vertically. Students are again clustered by academic performance within each grade, splitting the network into an East and West side.

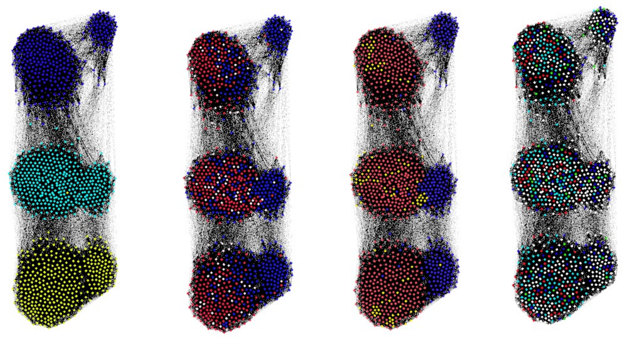


Figure 10: Middle School Sociogram

The sociogram for this school showed students with disabilities (green, center right) were largely clustered with one another within clustered grade levels, indicating they are integrated into the broader network. However, this school is a magnet school for gifted students (blue, center right), which reflected here by there being a second cluster within each grade level. This is especially visible in 8th grade, where the cluster of gifted students are further away than their peers than in sixth or seventh grade. In this graph, it is difficult to see how students may be clustered across the network on the basis of race. This is a limitation of just using sociograms, only homophily at high levels that form their own clusters are truly visible using the network graphs alone.

Ecological Fallacy and Sociograms

When analyzing these graphs, it is easy to fall into the ecological fallacy. For example, students with disabilities tend to be clustered within the sociogram for 3061 on the right side of the graph. Students who do not pass the FSA also tend to be located on the right side of the graph. With these two premises, one should be hesitant to conclude that most students with disabilities do not pass the FSA. There are many passing students on the right side of the graph where students with disabilities are located. It could be the case that all those passing students are the ones with disabilities! To make claims about individual students purely based on where they tend to be located in the graph can lead to erroneous conclusions. For people who want to draw conclusions such as these, one should use other methods.

Common Traits of High School Sociograms

As mentioned in the last section, many school sociograms tend to develop similar patterns. Clusters based on grade level develop in a way that separates students by academic performance and grade level. Graphing the sociogram vertically with grade level and academic performance horizontally makes the interpretation of the networks uniform.

Take for example two schools 7842 (top row) and 1502 (bottom row). These schools exhibit these characteristics, grade level clusters form North to South, and academic performance tends to separate East to West. School 7842 shows this very clearly.

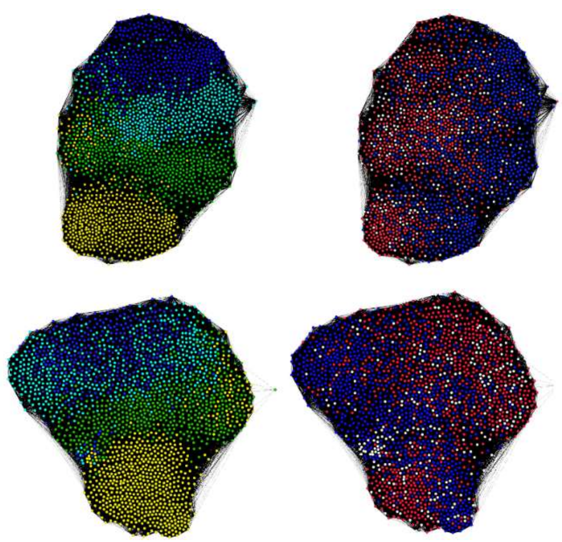


Figure 11: Two High School Sociograms

There is a clear inflection point about halfway across the graph where one starts seeing more of one color than the other. While there are also many high achieving freshmen (yellow) that cluster with the 10th grade students (green), freshmen tend to be at the bottom of the graph.

However, looking at the 9th grade cluster for 1502, students who are not passing the FSA are on the left side of the graph, rather than the right where 10th, 11th, and 12th grade non-passing students are located. This challenges the idea that these network graphs will always have clear patterns. While for most sociograms there is a pattern of grade level and academic performance clustering that goes vertically and horizontally (respectively), this is not always the case.

High Schools Within High Schools

There is a second convention of high school sociograms that is worth mentioning, specifically, there are many schools that have "schools within schools". Take school 4168 for example. There is a giant component, with four major offshoots. The giant component appears to be vertically demarcated by grade level and horizontally by academic performance.

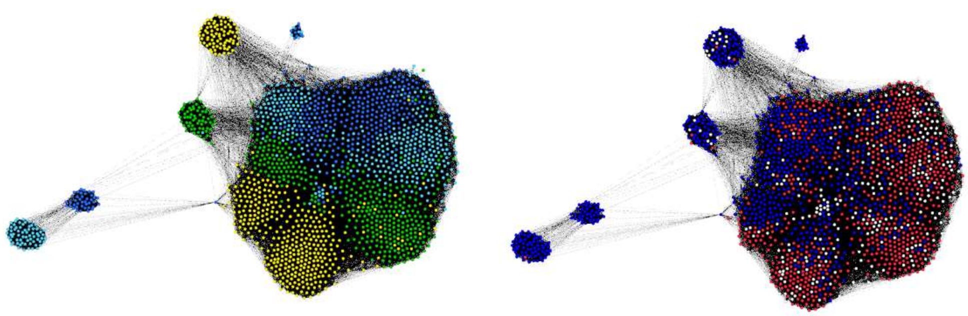


Figure 12: IB High School Sociogram

However, the four major offshoots all appear to be homogenous based on both grade level and academic performance. This is because the school has an International Baccalaureate (IB) program. You can see this playing out by underclassmen in these clusters taking classes with high achieving upperclassmen. This means that many 9th and 10th grade IB students (yellow and green) are taking classes with high achieving 11th and 12th grade students (sky-blue and blue) in the regular academic track of the school. Upperclassmen in the IB program take a decreasing number of classes with students in the regular track.

Something similar occurs in school 9362. Below are two sociograms, the left for grade level, and the right for academic performance. This school also has an International Baccalaureate program. There is a large cluster of students, with four large offshoots each corresponding to a grade level. These students also tend to pass the FSA (blue).

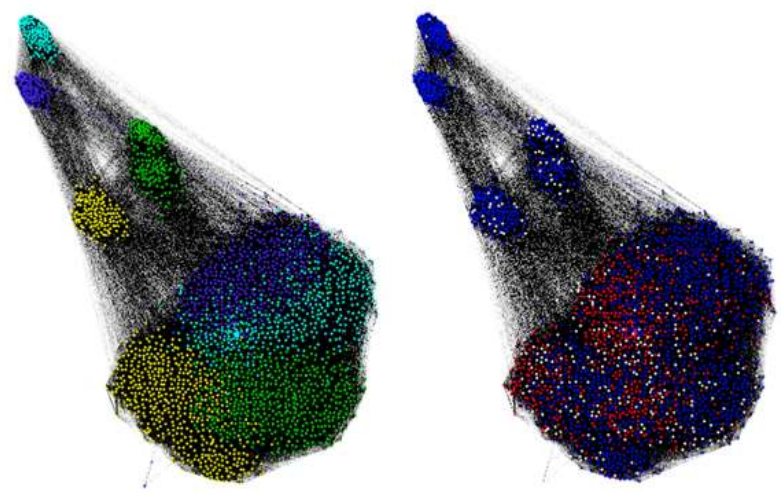


Figure 13: Second IB High School Sociogram

In the main cluster, there is somewhat of a vertical demarcation of clusters based on grade level, though horizontal demarcation by FSA score is less visible. Compared to the school 4162 previously, it appears as though IB freshmen at 9362 are taking more classes with their 9th grade (yellow) regular track peers. Here, it is 10th grade students (green) that are more likely to be taking classes with upperclassmen (sky-blue and blue) in the school's regular track. This could be an indicator of different policies on tracking or different class offerings of the two schools. Researchers could be interested in using these network diagrams in tandem with on the ground qualitative research.

Common Traits of Middle School Sociograms

Middle school curricular networks display similar characteristics to those of high schools. There are high levels of homophily based on grade level (left) as well as academic performance (right). Grade level homophily is even more pronounced than for high school networks, as fewer inter-grade ties are formed.

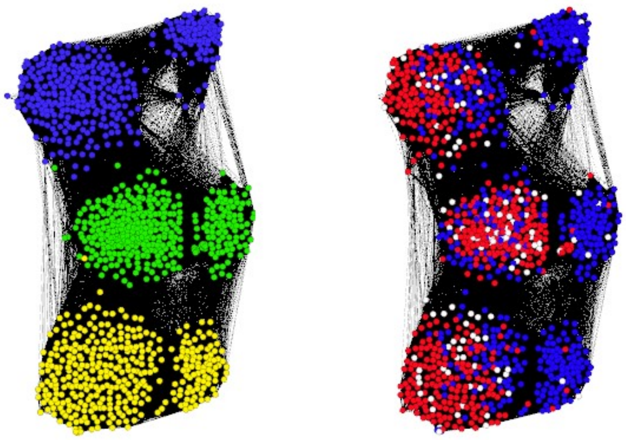


Figure 14: Middle School Sociogram

Whereas grade levels in high school are closely connected to one another due to students from different grades taking multiple classes with students from other grades, middle school students tend to cluster with only those students in their own grade. Within each grade level, there are also horizontal patterns separating students by academic performance; students take classes with similarly academically performing peers. These clusters are often more pronounced than in high school curricular networks. It appears that middle schools have higher levels of academic performance homophily than for high schools. However, the sociograms cannot give us a measured parameter we can use to make such conclusions about grade level of academic homophily; we would need to incorporate ERGM parameter estimates to make such a case.

Visualizing Clustering with Sociograms Wrap-Up

The sociograms of middle and high school networks are giant "hairballs". They are initially messy webs of meaninglessness. Once we add information and directional structure to them, however, these "hairballs" have the capacity to tell us a great deal of information. Where ERGM's tell us clustering exists, sociograms show where this clustering occurs and how it is forming. They are also visually striking and have the capacity to illustrate and communicate data to both researchers and curious laypersons.

It is one thing to say, "This school exhibits racial homophily - if two students share the same race, there is an associated increase of 0.55 log-odds they are classmates". It is an entirely different thing to visualize how White kids and Black kids, or students with and without disabilities, are highly concentrated into different sections of the network.

Whereas very few people would understand a change in log-odds, many more would understand a graphical representation of a school with colored nodes that signify the race of students. This could improve messaging of the need for de-tracking measures, encouraging school and district administrators to make changes in policies surrounding tracking.

ERGMs and Sociograms Working Together

High School Network

We can look at school 3062's ERGM results alongside its sociograms for an example of how we can use the two in tandem to glean even more information about the school. The grade level parameter is the strongest ERGM estimate, which corresponds to the sociogram's visual representation of where students are located within the network across grade levels (top left). The ERGM estimate tells us that this association between homophilous grade levels and forming classmate ties is very strong, and the sociogram also adds information by telling us where the clusters are forming. Looking from just the ERGM estimate, one could not tell that grade levels are stacked atop one another.

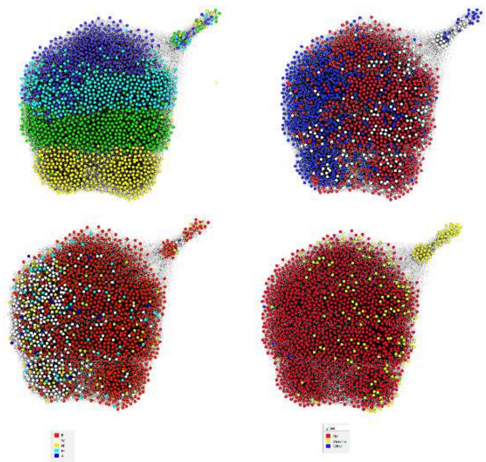


Figure 15: High School Sociogram (3062)

Also notable is the ERGM estimate for special education needs. While the ERGM does support the idea that students are clustered throughout the network on the basis of SPEN, only the sociogram (bottom right) could tell us that there is a "tail" in the network comprised of students with disabilities. Furthermore, it could not tell us that there are two clusters of students within that tail. If we only had the ERGM estimate, we could not tell what the clustering on SPEN looked like.

Table 7: Model 6-U Estimates for School 3062

Parameter	Estimate	SE	Statistical Significance
edges	-4.750	0.091	***
nodecov.NClasses	0.050	0.003	***
nodematch.grade	1.690	0.005	***
nodematch.race	0.550	0.007	***
nodematch.FSA	0.438	0.006	***
nodematch.SEN	0.476	0.010	***

The parameter estimates for race and FSA both have strong and positive associations. Students who share the same race or share the same FSA score are more likely than chance to be classmates with one another. However, the sociogram for race (bottom left) can tell us where this clustering is occurring, that they are clustered East to West. If a school was 75% Black, comprising 100% of the 9th grade class, there would be homophily based on race and grade level. The ERGM nodematch estimates for both grade level and race would be high, but would be across grades, rather than within.

Opportunity gaps across the district lead to race and FSA scores to be associated with one another. While the sociograms show that Black students and students who do not pass the FSA tend to be on the right side of the graph, we cannot infer that they are one in the same. How students from different groups do on the FSA is not a research question answerable by sociograms.

Middle School Network

Turning our attention to model 6-U for school 2562, we can see how the results from the school's ERGMs and sociograms can add nuance to our understanding of the curricular network. Just as in high school curricular networks, grade level is a major predictor of forming classmate ties, however, it is much more evident in middle schools. All nodematch-grade parameters of middle schools are higher than those of high schools. Additionally, FSA score is also a major driver of homophily.

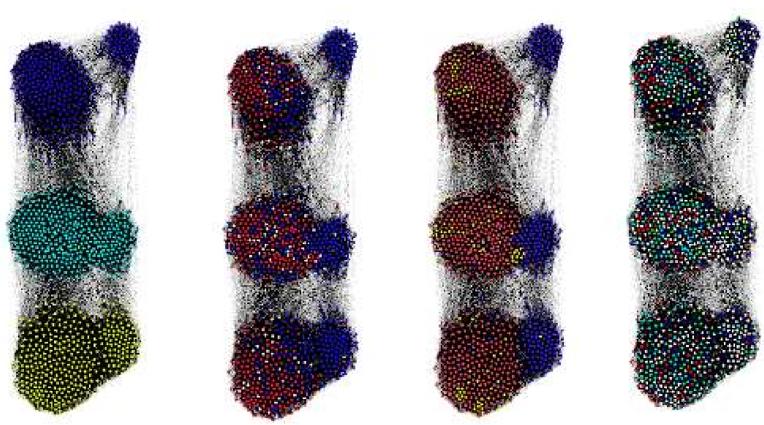


Figure 16: Middle School Sociogram (2562)

Using our compass convention for analyzing curricular network sociograms, grade level serves as a vertical axis while academic performance acts as a horizontal axis. The third sociogram shows where students labeled as gifted tend to be in the sociogram. They are forming their own clusters within each grade level, so much so, that for the differential homophily model, the parameter estimate for gifted students is 3.94, which is the second highest estimate for nodematch on gifted label across the district.

From looking at the fourth sociogram, one may conclude there is no homophily based on race. However, the model 6-uniform models show there is statistically significant clustering based on race, though the estimate is small (.087). This would not be evident from just looking at the sociogram. Further, after dropping the assumption of uniform homophily, the differential homophily model shows that there is strong positive homophily among Asian, Black, Hispanic, and Multiracial students, and it is only White students who do not have statistically significant levels of homophily. This is where ERGMs have an edge over sociograms. The sociogram for race does not immediately show racial homophily, however, ERGMs show precisely how much homophily occurs.

Differences between Middle and High Schools

From looking at the sociograms, it appeared as though homophily based on grade level and academic performance were more pronounced in middle schools than in high schools. The sociograms, while useful, do not give us a well estimated measure of homophily. To support the idea that homophily on grade and FSA are higher in middle schools, we can use the ERGM parameters to estimate them. It turns out that these assumptions were correct.

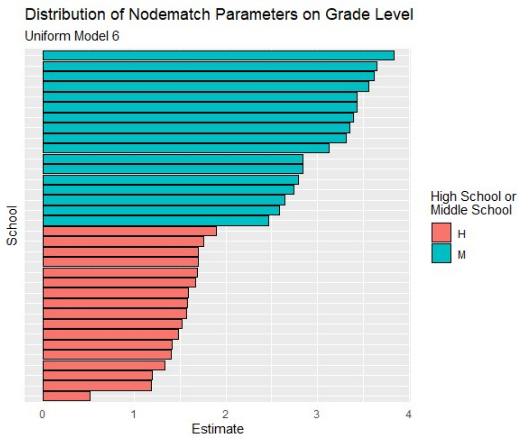


Figure 17: Grade Level Nodematch Estimates by School

When sorting the grade nodematch parameters from least to greatest, we see that each middle school's nodematch-grade parameter is higher than each high school's nodematch parameter.

Using a boxplot to compare nodematch parameters between middle and high schools, we can see that middle schools tend to be more homophilous based on education performance.

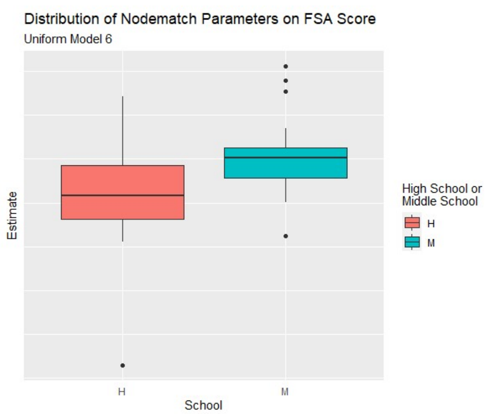


Figure 18: FSA Nodematch Parameter by FSA Score

How the sociograms presented themselves lead to a hypothesis that there were differences between middle and high schools. Looking at sociograms qualitatively can conjure ideas for quantitative research. Using the two in tandem could be a form of exploratory mixed methods.

The opposite is also a possibility, using results from ERGMs to identify strong clustering patterns, and then turning to the sociograms to get a sense of where this clustering is occurring and why. Take for example school 3061. This school has the second highest level of statistically significant nodematch-Black homophily in the district, with an ERGM estimate of 0.77. However, there is no context as to why or where this homophily is occurring. If every 9th grade student was Black, it could be the case that the school would have high homophily for black students purely as an artifact of demographic shifts of the school. A sociogram can provide important contextual information.

According to the sociogram, it appears that there are two major drivers for why school 3062 has high levels of homophily for Black students. The first is academic tracking.

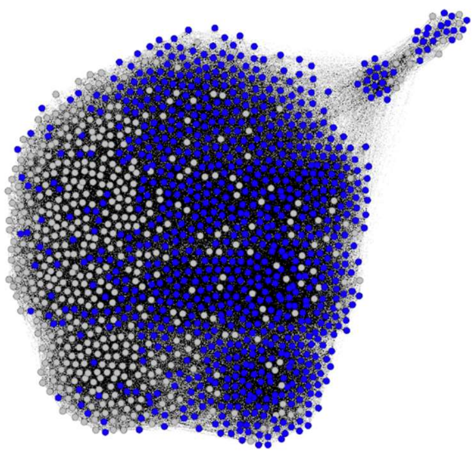


Figure 19: Sociogram of School 3062

Due to inequal educational opportunities, Black students (Blue) fair worse on the FSA than their non-Black peers (Grey). Though even after controlling for FSA homophily, there is still a large amount of homophily for black students, suggesting that there are other mechanisms in place.

The second driver of this homophily is that the "tail" of the network made up of students with disabilities is also made up of majority Black students. It is common for high school networks to have a sort of "tail" for students with disabilities, either due to self-contained classes, or students with severe disabilities having a unique course they take with one another.

This, coupled with Black students being overrepresented in disability categories, leads to additional homophily. If students with disabilities are more likely to be clustered with one another, and Black students are more likely to be labeled as disabled, then it makes sense that there is increased homophily for Black students.

However, the initial ERGM parameter for Black homophily was statistically significant while controlling for homophily on FSA score as well as disability. This suggests there is a third driver occurring, one that may need a deeper dive into the data, or qualitative observations at the school.

This chapter highlighted many of the different ways researchers can use ERGMs and sociograms to study curricular structures within middle and high schools. From ERGMs providing clustering estimates, to sociograms showing where the homophily is taking place. While both methods are useful in describing homophily based on student characteristics, using them together adds vital context that increases the usefulness of each method. The next chapter takes these results and further discusses the strengths of the different approaches, and how they are applicable in future research projects.

Chapter 5: Discussion and Implications

The aim of this dissertation was to answer five separate questions about using the social network analysis tools of exponential graph models and sociograms to measure and visualize within-school clustering of students within curricular networks based on grade level, race, academic performance, and special education needs.

- "In what ways can SNA **visualize** within-school clustering on the basis of grade level, race, academic performance, and special education need?"
- "In what ways can SNA **measure** within-school clustering on the basis of grade level, race, academic performance, and special education need?"
- "What information about within-school clustering do SNA **visualizations** provide compared to the information provided by SNA measurements?"
- "What information about within-school clustering do SNA **measurements** provide compared to the information provided by SNA visualizations?"
- "What information can be gleaned from using both SNA visualizations and measures in tandem?

Research Question 1: Visualizing Curricular Networks

*"In what ways can SNA **visualize** within-school clustering on the basis of grade level, race, academic performance, and special education need?"*

Sociograms provide a unique graphical representation of within school clustering that numeric parameter estimates cannot give. For example, if a researcher is studying within-school clustering based on special education need, they may use any number of segregation indices or clustering mechanisms. These will provide a parameter estimate as a numeric value, and the main way to display these results is a bar chart of the aggregated data. Sociograms of curricular networks, on the other hand, show each student as a unique node in a web of students. Showing someone how students with special education needs are clustered away, isolated from the main network, has an emotional appeal, asking the viewer to imagine themselves as an isolated student rather than just presenting a bar graph.

Visualizing curricular networks via sociograms also provides the observer insight into not only groups, but individuals within groups. As each node represents a student, one can identify particular nodes in interesting positions within the network. For example, finding a 6th grade student who takes more classes with 8th grade students, identifying an Asian student who is surrounded by White students in the curricular network, or identifying students who passed the FSA but are located near students who did not pass. Researchers could be interested in interviewing these individuals about their lived experiences at school. How is their experience different than other students located elsewhere? This could lead to interesting qualitative research projects.

Research Question 2: Measuring Curricular Networks

"In what ways can SNA ***measure*** *within-school clustering on the basis of grade level, race, academic performance, and special education need?"*

Using exponential random graph models (ERGMs) is a method to measure homophily of curricular networks in a way that controls for other variables such as other forms of homophily, network parameters, and unequal group sizes. I gave a presentation of preliminary results to the school district. When I presented the information on clustering based on race, one audience member suggested that the reason for race-based clustering is because of how students are clustered on academic performance. Since Black students underperform on the FSA compared to White students, their argument was that it is actually FSA scores driving homophily for Black students. While other clustering and segregation methodologies could not tease apart out race and academic performance homophily, ERGMs can do so. I showed the audience member that the parameter estimates for Black homophily was still positive and statistically significant, even after including FSA homophily in the model, suggesting that race is still an important factor in clustering, above and beyond FSA scores.

It is also the case that, if one group of students simply takes more classes or has more classmates than others, then this could influence homophily estimates. For example, if the average student with SEN has 25 classmates, while those without SEN have 100 classmates, this could influence how homophilous ties are created. The nodefactor estimates within ERGMs allow for this to be controlled for in ways that other indices of segregation cannot.

Lastly, several segregation indices struggle to accommodate for uneven sub-group sizes, especially when there are more than two groups. Take for instance the I-E index, which measures the number of "IN" ties and compares them to the number of "OUT" ties. A small sub-group in a network would likely have less ties to others in the same group, simply because there are fewer of those students. Because ERGMs serve as a type of logistic regression, uneven group sizes do not influence the results. ERGMs do not need additional adjustments to parameters to account for multiple groups of uneven sizes.

Research Question 3: Information from Social Network Visualizations

"What information about within-school clustering do SNA ***visualizations*** *provide compared to the information provided by SNA measurements?"*

Visualizing school curricular networks using sociograms can provide researchers with a number of tools to investigate topics such as equitable access to advanced classes and academic tracking. One can orient sociograms so that grade level serves as a vertical axis while academic performance acts as a horizontal axis within each graph. While most schools exhibit a more straightforward vertical and horizontal orientation, other schools, such as magnet and IB schools, exhibit different structures. These schools often have underclassmen in IB taking advanced coursework with upperclassmen not in the IB program. By only using ERGM parameter estimates to analyze grade level homophily, researchers may miss the additional context provided by sociograms.

Many schools that initially appear to be racially integrated may actually display high levels of racial homophily. Some such schools are magnet schools, those schools developed to encourage integration by providing advanced coursework or an educational focus (such as art, music, STEM, etc). These schools are often located in historically Black neighborhoods and aim to encourage non-neighborhood parents to send their children there.

While these schools may appear integrated from just looking at the number of students from each racial group, these students may rarely interact with one another in classes. School 342 is one of these schools, where just over half the student population is Black (Blue). One would expect that if each student was randomly distributed in the network, that Black and non-Black students would form similar amounts of in-group and out-group ties. This is certainly not the case, as within each grade level, Black students are isolated from their non-Black peers.

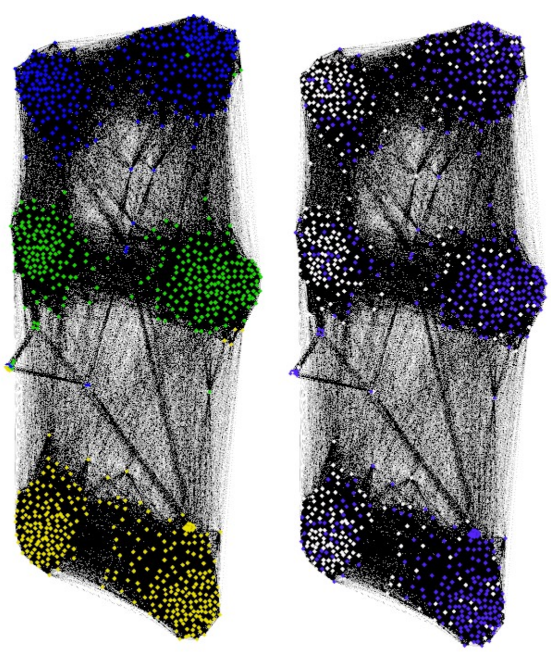


Figure 20: School 342 Sociograms

ERGMs could tell you that there is significant clustering based on race, but it would not be able to tell you what the clustering looks like. ERGMs would not be able to differentiate the case where students were segregated within each grade, vs being segregated by grade. Put differently, if all Black students just happen to be 9th graders, and all 9th graders were Black, then we would see high levels of racial homophily, but it would not be directly related to class networks, but more likely an outside rezoning policy.

Sociograms can also point to students of interest, those who are clustered in a way that is unexpected. In school 342 there are three seventh grade students (green) taking classes primarily with 8th grade students. These three students are all Black, are not passing the FSA, and only one has a disability. These students may be of unique interest to researchers for interviews of their experiences.

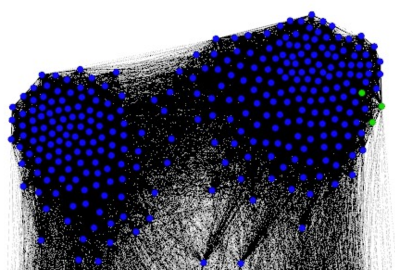


Figure 21: Sociogram of School 342, Zoomed on 8th Grade.

Network visualizations of curricular networks can also show how students from different groups are clustered together, and lead to asking questions about the underlying mechanisms causing the clustering. School 342 has two interesting clusters of students with disabilities (yellow). The leftmost cluster is likely a hub of students in a self-contained classroom. However, this cluster is connected to the main group by just 6 individuals with disabilities. Without these students, and their classmate ties, the yellow cluster would become detached. Do these students act as ambassadors between worlds? Do they form friendship ties with students with and without disabilities? These students could be interviewed to answer these questions. The second cluster on the right is of students with disabilities clustered tightly with other 6th grade students. This

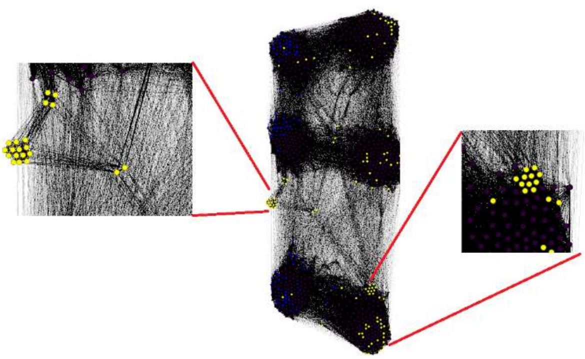


Figure 22: Sociogram of School 342, Zoomed on Students with Disabilities.

could be evidence of a unique class for 6th grade students with disabilities to acclimatize them to the school. Sociograms raise these kinds of questions whereas ERGMs cannot.

Research Question 4: Information from Social Network Measurements

"What information about within-school clustering do SNA measurements provide compared to the information provided by SNA visualizations?"

Researchers can use the parameter estimates from the ERGM results in a number of ways. First, researchers can use these estimates as outcome measures or covariates in other regression models. Take for example a hypothetical school that notices that the school is internally segregated by race when it comes to curricular networks. If this school decides to de-track its curriculum, it can calculate the race-nodematch parameter for each year, before and after the de-tracking program starts. Using regression discontinuity design, ERGM estimates can measure the progress of the program and see if it is successful in reducing levels of racial homophily.

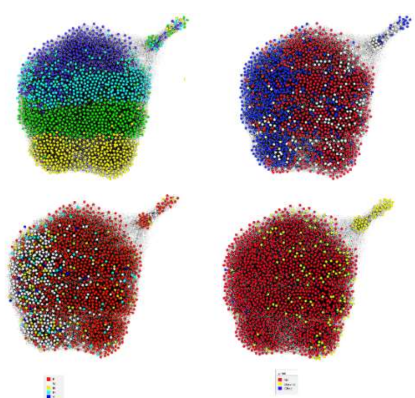
Using the ERGM estimates as covariates, researchers can measure how curricular network clustering is associated with other outcome measures, such as achievement gaps, teacher retention, or school climate. Nodematch parameters may be associated with other school level and student level variables. ERGM parameter estimates can provide a unique opportunity to measure an additional variable to include in multilevel models, with and without interaction effects with student level variables (cross level effects).

Qualitative researchers can use ERGM estimates to identify schools with high levels of clustering. This could be a useful method for identifying those schools with the most racial homophily, and then using qualitative tools to identify what mechanisms are likely responsible. After the mechanism is determined, these outlying schools could be good candidates for interventions to reduce racial homophily. Students from highly clustered schools may also be good candidates for interviews and focus groups, to understand the lived experiences of those students in segregated curricular networks. After finding common themes, researchers can turn and create surveys to see how widespread these experiences are.

Research Question 5: Information from Social Network Measurements *and* Visualizations

"What information can be gleaned from using both SNA visualizations and measures in tandem?

Using ERGMs and sociograms together can provide several insights that they could not do separately. We can return our attention to school 3062. ERGM results show that, like all other schools, grade level homophily is the greatest predictor of being classmates. However, this estimate does not tell us how the different grade levels are connected to one another.



Parameter	Estimate	SE	P-Value
edges	-4.750	0.091	P < .0001
nodecov.NClasses	0.050	0.003	P < .0001
nodematch.grade	1.690	0.005	P < .0001
nodematch.race	0.550	0.007	P < .0001
nodematch.FSA	0.438	0.006	P < .0001
nodematch.SEN	0.476	0.010	P < .0001

Figure 23: School 3062 Suite of Sociograms with corresponding ERGM Parameter Estimates

Sociograms on the other hand show clearly that each subsequent grade is more connected to one another than grades that are more than one grade level apart (top left). Put another way, grades 9 and 10 (yellow and green) are closer together than grades 9 and 12 (sky blue and blue). While expected, sociograms put a graphical representation of this phenomenon, and ERGMs tell researchers how strong this association is.

This is especially the case when discussing students with disabilities (green, bottom right). Special education need is consistently a large predictor of being classmates based on ERGM estimates, but this could be due to multiple reasons. Schools may cluster students with disabilities in the center of the network, clustered alongside 9th grade students, or clustered with academically proficient students. ERGMs cannot tell us where these students are located, but sociograms can.

Tracking how both sociograms and ERGM estimates change over time can provide more insight into curricular networks than either method on its own. For example, if a school removes certain prerequisites for advanced courses, the sociograms and ERGM estimates could reflect this change. The school's sociograms before and after the change could have less visible patterns of academic performance homophily, as students with lower grades start taking more advanced classes. The school's ERGM estimates would also likely change, with estimates becoming smaller on homophily parameters of academic performance.

Limitations

General Limitations

This school district is certainly not the only district to have curricular network homophily based on race, special education need, grade level, and academic performance. However, as with any research study, generalizing to other districts with different demographics and social processes should be done with caution. The school district where this data came from, while representative of the state at large, is not representative of all school districts in all locations. One should not generalize the results of this study to other districts, many of which have different socioeconomic demographics or state and local policies. Additionally, one should not generalize findings backward or forward in history; the data are merely a snapshot in time. Future research should use these social network methods to study a large swath of school districts in various contexts to see what factors are associated with the network metrics as well as track these measures longitudinally.

Due to the IRB agreement with the school district, the identity of the school district and schools within must be kept anonymous. This limits discussion of the historical and social processes that could identify the district. However, the IRB encourages researchers who use its data to present their findings at district meetings with administrators, which I have done in the past.

Because this study used observational data, one should avoid causal interpretations of coefficients. While there are some social network analysis tools and research methods that can provide causal interpretations, these are not found within this dissertation. The purpose of this paper is to introduce using the social network analytic tools of ERGMs and sociograms, and to provide a methodological method to measure school clustering associated with different student characteristics.

Methodological Limitations

This dissertation also did not consider various network parameters in estimating ERGM models. Parameters such as mutuality and triangulation were not used. The reason for this is that some ERGM parameters such as GWESP (geometrically weighted edgewise shared partners) require substantially large computing power, more than I have available to me. Future research ought to take advantage of computing tools such as cluster computing to incorporate these network parameters into modeling curricular networks.

Additionally, advances in ERGM estimation allow for both weighted and bipartite models. Transforming from bipartite to unipartite reduces the amount of information about the kinds of courses students take, for example, removing the course type/subject/prerequisites, etc. These could be important factors that this dissertation does not address. Similarly, transforming the networks from weighted to unweighted ignores important information. This transformation ignores how some students may not only make one connection with another student, but could be in multiple classes at once. This could lead to underestimates of different ERGM parameters. The reason these are not discussed in this dissertation is because they were more computationally costly than I had resources available to me. Because the purpose of this dissertation is to look at how ERGMs and sociograms work together, I only created sociograms for the networks I could also model. Future research can incorporate these limitations into their study designs.

This dissertation also does not investigate what school factors are related to these patterns. While ERGM estimates are orthogonal to student composition, the composition of the students may be associated with the amount of clustering that occurs. For example, it could be the case that the larger the proportion of Black students in a school, the higher the nodematch parameter for Black students would be. Title One schools, charter schools, or magnet schools may have different levels of clustering based on special education need, race, and academic performance. These school level variables are ignored in this dissertation, but future research ought to look at how ERGM parameters are associated with these higher order variables.

Future Research

Additional Student Demographic Characteristics

Researchers interested in other topics related to student clustering may also find sociograms and ERGM estimations interesting. One can investigate how gender, English language proficiency, migrant status, homelessness status, poverty, and other characteristics are associated with forming classmate ties among students. While this dissertation is limited in its discussion of students with special education needs, other researchers can take a deeper look into different categories of disability labels and see how they are connected to school homophily. While the school district in this dissertation has very few Native American and Pacific Islander students, other districts have larger populations from these groups. Researchers can use the methods described in this dissertation to investigate how race is associated with homophily differently for these groups. This dissertation also does not look directly at interaction effects, for example, students who are Black and have a disability may be clustered differently due to the interaction of the two groupings. This is another avenue for future research, to investigate intersectionality.

Additional ERGM Methods

This dissertation is limited to unweighted unipartite networks of curricular networks in a snapshot in time. ERGMs are not, however, limited to these modalities. There are methods of estimating ERGMs of bipartite networks, weighted networks, as well as networks over time. Sociograms can also be visualized as bipartite, weighted, and longitudinal. These methods may add additional insight on how curricular networks are formed. A bipartite curricular network, for example, could include course level information such as course type, level, subject, or teacher characteristics. Certain class types, such as elective courses, could contribute to homophily of the network differently than other subjects. Modeling these networks this way could provide additional information about how student homophily is created and help explain the phenomenon in more detail.

There are also a host of additional ERGM parameters that researchers can use to study curricular networks. For example, the nodemix parameter estimates how likely ties are formed within a group versus across different groups. Unlike the nodematch parameter, nodemix can estimate how more or less likely an 8^{th} grade student is to form classmate ties with a 7^{th} grade vs 6^{th} grade student. Network parameters such as different kinds of triangulation and edge covariates could also play a part in how curricular networks are formed. These are outside the scope of this dissertation, but they are available research topics for the future.

Additional Research Designs

As discussed previously, both ERGM and sociogram results could be of interest to qualitative researchers. Researchers could identify and interview individuals within a sociogram who are in interesting locations within a network, for example, a 6^{th} grade student clustered with 8^{th} grade students. Sociograms also display unique formations within networks, for example, students with disabilities are often located in the "tail" of a network. Researchers could conduct focus groups with these students and compare their experiences to those who are not located in the tail.

As many districts aim to reduce the amount of unequal access to advanced classes, districts may relax certain prerequisite requirements. Researchers can track the ERGM estimates for academic performance homophily before and after these changes. If this estimate decreases, and there is less homophily based on academic performance, one could say that the district was successful in improving access to advanced classes. If the district rolled out the changes across multiple years, it could be seen as a multiple baseline single case design study.

There are likely a host of school level and district level variables that are associated with how clustered students are based on their demographic characteristics. It is possible that, as the proportion of Black students increases, so does the level of homophily for Black students. Median incomes of districts may be associated with homophily based on academic performance. Putting these district and school level variables into models would lend itself to studies using multilevel modeling. While this dissertation only looked at a single district, and did not consider school level variables, future researchers can easily use the methods described to study these.

Implications for Policy and Practice

While there are other tools available to school and district administrators, sociograms and ERGMs can provide unique insight on their schools' curricular networks. If a superintendent of a school district had a goal of increasing the proportion of minority students in advanced courses, they would be able to identify which schools in their district had the highest level of race based curricular homophily. Then, at those schools with the highest levels of racial homophily, the district could implement strategies to de-track the school by reexamining the process for which students are placed in advanced classes. The district could then, using ERGM estimates, measure the levels of racial homophily before and after different implementing different de-tracking measures. If successful, these policies could be reapplied elsewhere in other schools experiencing racial homophily.

Sociograms and ERGMs, used together, are useful for highlighting the prevalence of clustering in curricular networks. Sociograms can visualize the depth of the clustering, and ERGMs can measure how strong the associations are. ERGMs also can control several factors that opponents to de-tracking may suggest as alternative explanations for measures such as racial homophily. For example, when I presented this data to the district in 2018, some respondents suggested that academic performance, not race, was the primary driver of homophily. ERGMs already control for this, so while academic performance does impact homophily, race is still a salient factor in how students are clustered throughout a curricular network.

Final Thoughts

If the goal of a district is to improve equitable access to advanced course work, one must first know how equitable access is to begin with. Measuring this is difficult, while there are methods, they are often confounded by several factors mentioned in this dissertation. Social network analysis provides two powerful tools to measure clustering within curricular networks, ERGMs and sociograms. These tools overcome the limitations of other methods, allowing researchers to accurately measure the association between group membership and being classmates, and allow one to visualize the clustering in an accessible format. These are effective tools in the toolbox of those who aim to de-track districts and improve access to advanced courses.

References

Archbald, D., Glutting, J., & Qian, X. (2009). Getting Into Honors or Not: An Analysis of the Relative Influence of Grades, Test Scores, and Race on Track Placement in a Comprehensive High School. *American Secondary Education*, *37*(2).

Argys, L., Rees, D., Brewer, D. (1996). Detracking America's Schools: Equity at Zero Cost? *Journal of Policy Analysis and Management*, *15*(4), 612-645.

Banks, J., Byrne, D., McCoy, S., & Smyth, E. (2014). Bottom of the class? The leaving certificate applied programme and track placement in the Republic of Ireland. *Irish Educational Studies*, *33*(4), 367-381. http://dx.doi.org/10.1080/03323315.2014.960148

Barber, C. & Wasson, J. (2015). A Comparison of Adolescents' Friendship Networks by Advanced Coursework Participation Status. *Gifted Child Quarterly*, *59*(1), 23-37. doi: 10.1177/0016986214559639

Becker, K. R., Stojek, M. M., Clifton, A., & Miller, J. D. (2018). Disordered eating in college sorority women: A social network analysis of a subset of members from a single sorority chapter. *Appetite*, *128*, 180-187.

Boer, A. & Pijl, S. (2016). The acceptance and rejection of peers with ADHD and ASD in general secondary education. *Journal of Educational Research*, *109*(3). 325-332.

Boer, A., Pijl, S., Post, W., & Minnaert, A. (2013). Peer Acceptance and Friendships of Students with Disabilities in General Education: The Role of Child, Peer, and Classroom Variables. *Social Development*, *22*(3), 831-844.

Bojanowski, M. & Rense, C. (2014). Measuring Segregation in Social Networks. *Social Networks*, vol. 39, pp. 14–32., doi:10.1016/j.socnet.2014.04.001.

Borgatti, S. P., & Halgin, D. S. (2016). Analyzing Affiliation Networks. In John Scott & Peter Carrington (Eds.), *The SAGE Handbook of Social Network Analysis* (pp. 417-433). London, SAGE Publications Ltd.

Burris, C. C., Heubert, J. P., & Levin, H. M. (2006). Accelerating Mathematics Achievement Using Heterogeneous Grouping. *American Educational Research Journal*, *43*(1), 137–154. https://doi.org/10.3102/00028312043001105

Callahan, C. M., Moon, T. R., & Oh, S. (2017). Describing the Status of Programs for the Gifted: A Call for Action. *Journal for the Education of the Gifted*, *40*(1), 20–49. https://doi.org/10.1177/0162353216686215

Carolan, B. V. (2014). *Social network analysis and education: Theory, methods & applications*. SAGE.

Carrington, P.J. & Scott, J. (2016). Introduction. In John Scott & Peter Carrington (Eds.), *The SAGE Handbook of Social Network Analysis* (pp. 417-433). London, SAGE Publications Ltd.

Cascio, E., Gordon, N., Lewis, E., & Reber, S. (2007) From Brown to Busing. *National Bureau of Economic Research*; Working Paper No. 13279.

Chen, C.C., Hamm, J.V., Farmer, T.W., Lambert, K., & Mehtaji, M. (2015). "Exceptionality and Peer Victimization Involvement in Late Childhood." *Remedial and Special Education*, vol. 36, no. 5, 2015, pp. 312–324., doi:10.1177/0741932515579242.

Conger, D. (2005). "Within-School Segregation in an Urban School District." Educational Evaluation and Policy Analysis 27: 225–44.

Crnovrsanin, T., Muelder, C.W., Faris, R., Felmlee, D., & Ma, K.L. (2014). Visualization Techniques for categorical analysis of social networks with multiple edge sets. *Social Networks*, *37*, 56-64.

De Freslon, I., Martinez-Lopez, B., Belkhiria, J., Strappini, A., & Monti, G. (2019). Use of social network analysis to improve the understanding of social behavior in dairy cattle and its impact on disease transmission. *Applied Animal Behavior Science*, *213*, 47-54.

Dianati, N. (2016). "Unwinding the Hairball Graph: Pruning Algorithms for Weighted Complex Networks." *Physical Review E*, vol. 93, no. 1, 2016, doi:10.1103/physreve.93.012304.

Domagalski, R., Neal, Z.P., Sagan, B. (2021). "Backbone: An R Package for Extracting the Backbone of Bipartite Projections." *PLOS ONE*, vol. 16, no. 1, 2021, doi:10.1371/journal.pone.0244363.

Dornbusch, S. M., Glasgow, K. L., & Lin, I. C. (1996). The social structure of schooling. Annual Review of Psychology, 47, 401-429. doi:10.1146/annurev.psych.47.1.401

Dumont, H., Protsch, P., Jansen, M., & Becker, M. (2017). Fish Swimming into the Ocean: How Tracking Relates to Students' Self-Beliefs and School Disengagement at the End of Schooling. Journal of Educational Psychology.

Duxbury, S. W. (2021). Diagnosing Multicollinearity in Exponential Random Graph Models. Sociological Methods & Research, 50(2), 491–530. https://doi.org/10.1177/0049124118782543

Epple, D., Newlon, E., & Romano, R. (2002). "Ability Tracking, School Competition, And The Distribution Of Educational Benefits," Journal of Public Economics, 2002, v83(1,Jan), 1-48.

Erdos, P. & Renyi, A. (1959) On Random Graphs I, Publicationes Mathematicae Debrecen (1959)

Estell, D.B., Farmer, T.W., Irvin, M.J., Crowther, A., Akos, P., Boudah, D.J. (2008). Students with Exceptionalities and the Peer Group Context of Bullying and Victimization in Late Elementary School. Journal of Child and Family Studies. 18. 136-150. 10.1007/s10826-008-9214-1.

Fitzgerald, S., Rumrill Jr., P.D., & Schenker, J. (2004). Correlational designs in rehabilitation. Journal of Vocational Rehabilitation. 20. 143-150.

Florida Department of Education. (2019). *Understanding Florida Standards Assessments Reports*. https://www.fldoe.org/core/fileparse.php/5663/urlt/UnderFSARpt19.pdf.

Freeman, L.C. (1996). *Visualizing Social Networks*. University of California, Irvine. https://www.cmu.edu/joss/content/articles/volume1/Freeman.html

Gamoran, A. (2015). The Future of Educational Inequality in the United States: What Went Wrong, and How Can We Fix It? William T. Grant Foundation.

Goodreau, S, Kittis, J., & Morris, M. (2009). Birds of a Feather or Friend of a Friend? Using Exponential Random Graph Models to Investigate Adolescent Social Networks. Demography, 46(1), 103-125.

Goodreau, S.M., Kitts, J.A. & Morris, M. Birds of a feather, or friend of a friend? using exponential random graph models to investigate adolescent social networks*. *Demography* **46,** 103–125 (2009). https://doi.org/10.1353/dem.0.0045

Government Accountability Office (2009) Multiple Federal Education Offices Support Teacher Preparation for Instructing Students with Disabilities and English Language Learners, but Systematic Departmentwide Coordination Could Enhance This Assistance. United States

Government Accountability Office (2016) Better Use of Information Could Help Agencies Identify Disparities and Address Racial Discrimination. United States

Harris, J. (2014) An Introduction to Exponential Random Graph Modeling. *Quantitative Applications in the Social Sciences*. https://dx.doi.org/10.4135/9781452270135

Heck, R. & Mahoe, R. (2010). Student course taking and teacher quality: their effects on achievement and growth. International Journal of Educational Management, 24(1), 56-72. https://doi.org/10.1108/09513541011013051

Heck, R., Price, C., & Thomas, S. (2004). Tracks as Emergent Structures: A Network Analysis of Student Differentiation in a High School. American Journal of Education, 110(4), 321-353.

Houtte, M. V., Demanet, J., & Stevens, P. A. (2012). Self-esteem of academic and vocational students: Does within-school tracking sharpen the difference? Acta Sociologica, 55(1), 73–89. https://doi.org/10.1177/0001699311431595

Jackson, C. K. (2010). Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago. Economic Journal. 120. 1399-1429. 10.1111/j.1468-0297.2010.02371.x.

Kaloeti, D., Manalu, R., Kristiana, I. F., & Bidzan, M. (2021). The Role of Social Media Use in Peer Bullying Victimization and Onset of Anxiety Among Indonesian Elementary School Children. *Frontiers in psychology*, *12*, 635725. https://doi.org/10.3389/fpsyg.2021.635725

Karin, H., Ellen, V., Evelien, C., Mieke, H., & Katja, P. (2012). Don't pull me out!? Preliminary findings of a systematic review of qualitative evidence on experiences of pupils with special educational needs in inclusive education. International Conference on Education and Educational Psychology. 69, 1709-1713.

Kavaler, D. & Filkov, V. (2017) Stochastic actor-oriented modeling for studying homophily and social influence in OSS projects. Empirical Software Engineering, 22(1):407–435, February, 2017.

Klovdahl, A. (1981). A note on images of networks. *Social Networks, 3*, 197-214.

LaPrade, K. (2011). Removing Instructional Barriers: One Track at a Time. Education, 131(4).

Lucas, S. & Berends, M. (2002). Sociodemographic Diversity, Correlated Achievement, and De Facto Tracking. Sociology of Education, 75(4), 328-348.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), Frontiers in econometrics (pp. 104-142). New York: Academic Press.

McFarland, V. (2007). Middle Level Grade Configuration: Impact on Hawai'i's Schools. Hawai'I Educational Policy Center

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

Mickelson, R. A., & Everett, B. J. (2008). Neotracking in North Carolina: How high school courses of study reproduce race and class-based stratification. *Teachers College Record, 110*(3), 535–570.

Miller, V. (2018) Access Denied: Tracking as a Modern Roadblock to Equal Educational Opportunity. New York University Law Review 93:903. https://www.nyulawreview.org/wp-content/uploads/2018/10/NYULawReview-93-4-Miller.pdf

National Center for Education Statistics. (2018, February). *National Teacher and Principal Survey*. https://nces.ed.gov/surveys/ntps/tables/ntps1718_21022301_s12n.asp

National Center for Educational Statistics, Table 215.30. (2018). *Digest of Education Statistics.* https://nces.ed.gov/programs/digest/d18/tables/dt18_215.30.asp?current=yes

National Center on Disability. (2018, February). *The Segregation of Students with Disabilities*. https://ncd.gov/sites/default/files/NCD_Segregation-SWD_508.pdf.

Newman, M. E. J., Watts, D. J., Strogatz, S. H. (2002) Random graph models of social networks Proceedings of the National Academy of Sciences Feb 2002, 99 (suppl 1) 2566-2572; DOI: 10.1073/pnas.012582999

Oakes, J. (2008). Keeping Track: Structuring Equality and Inequality in an Era of Accountability. Teachers College Record, 110(3), 700-712.

Oakes, J., & Guiton, G. (1995). Matchmaking: The Dynamics of High School Tracking Decisions. American Educational Research Journal, 32(1), 3–33. https://doi.org/10.3102/00028312032001003

O'Connor, C. (2002) "I'm Usually the Only Black in My Class": The Human and Social Costs of Within-School Segregation, 8 MICH. J. RACE & L. 221 (2002). Available at: https://repository.law.umich.edu/mjrl/vol8/iss1/6

Polly, T. (2014). Networks: How I learned to stop worrying and love the hairball. http://www.ulib.iupui.edu/digitalscholarship/blog/networks-how-i-learned-stop-worrying-and-love-hairball

Riehl, C., Pallas, A., & Natriello, G. (1999). “Rites and Wrongs: Institutional Explanations for the Student Course-Scheduling Process in Urban High Schools.” American Journal of Education, 107(2), 116-154.

Robins, G. (2016).Exponential Random Graph Models for Social Networks. In John Scott & Peter Carrington (Eds.), *The SAGE Handbook of Social Network Analysis* (pp. 417-433). London, SAGE Publications Ltd.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007) An introduction to exponential random graph (p*) models for social networks, Social Networks,Volume 29, Issue 2, https://doi.org/10.1016/j.socnet.2006.08.002.

Rubin, B. (2018). Tracking and Detracking: Debates, Evidence, and Best Practices for a Heterogeneous World.

Schaefer, D., R. (2012). Youth co-offending networks: An investigation of social and spatial effects. Social Networks, 34, 141-149.

Steenbergen-Hu, S., Makel, M., & Olszewski-Kubilius, P. (2016). What One Hundred Years of Research Says About the Effects of Ability Grouping and Acceleration on K-12 Students' Academic Achievement: Findings of Two Second-Order Meta-Analyses. Review of Educational Research, 86(4), 849-899.

Strauss, D., & Ikeda, M. (1990). Pseudolikelihood Estimation for Social Networks. *Journal of the American Statistical Association, 85*(409), 204-212. doi:10.2307/2289546

Sweet, T. M. (2016). Social Network Methods for the Educational and Psychological Sciences. Educational Psychologist, 51(3-4), 381-394.

Szumski, G., Smogorzewska, J., & Karwowski, M. (2017). Academic achievement of students without special educational needs in inclusive classrooms: A meta-analysis. Educational Research Review. 21, 33-54

Tefera, A., Frankenberg, E., Siegel-Hawley, G., & Chirichigno, G. (2011). Integrating Suburban Schools: How to Benefit from Growing Diversity and Avoid Segregation. UCLA, Civil Rights Project.

U.S. Census Bureau (2019). Florida Quickfacts. Retrieved from https://www.census.gov/quickfacts/FL.

Van der Hulst, R.C. (2016). Terrorist Networks: The Threat of Connectivity. In John Scott & Peter Carrington (Eds.), *The SAGE Handbook of Social Network Analysis* (pp. 256-270). London, SAGE Publications Ltd.

Vanfossen, B.E., Jones, J. D., & Spade, J.Z. (1987). Curriculum tracking and status maintenance. Sociology of Education, 60, 104–122

Zbigniew, K. (2012) Heterogeneity, Consolidation, and Intergroup Relations, International Journal of Sociology, 42(4), 8-32, DOI: 10.2753/IJS0020-7659420401

Appendices

Appendix A: Definitions of Terms

Ability grouping – the practice of placing students of similar academic performance into the same courses.

Between-school tracking – the practice of placing students of similar characteristics into separate schools from one another.

Curricular network – the network that is formed by the co-enrollment of students in the same class.

Directed/undirected edges – characteristic of edges based on if the direction of the tie (from sender to receiver) is meaningful (directed) or meaningless (undirected).

Edge (file) – the dataset file that contains information about connections between actors.

Edges (ERGM parameter) – the baseline likelihood that two actors are connected by an edge.

Edges (ties) – the connections that connect actors to one another.

Exponential Random Graph Models (ERGMs) – similar to logistic regression, statistical tool that estimates the likelihood that two actors are connected by a tie based on node, dyad, and network parameters.

Gephi – a social network analysis graphing software.

Gifted and Talented – programs designed to develop and instruct students who exhibit extraordinary academics, creativity, or talents.

Hairball – pejorative term used to describe sociograms that have too many ties among actors to provide meaningful information.

Homophily – the tendency for similar actors to associate with one another above what we would expect by chance alone.

Hypergraph – data structure where actors of one type are displayed as rows, actors of a second type are displayed as columns, and intersecting cells indicate if there is a connection (1) or not (0).

Network backbone – a weighted subgraph of a unipartite network that has been transformed from a bipartite network.

Node file – the dataset file that contains information about the actors themselves.

Nodecov (ERGM parameter) – change in likelihood that two actors are connected based on a continuous variable.

Nodefactor (ERGM parameter) – change in likelihood that two actors are connected based on a categorical variable.

Nodematch (ERGM parameter) – change in likelihood that two actors are connected based on if they share a common characteristic.

Nodemix (ERGM parameter) – change in likelihood that two actors are connected based on if they belong to two different groups, estimating each cross-group likelihood.

Nodes – the actors within a network.

R-Studio – open-source data analysis software.

Social Network Analysis – the theoretical ideas and methodological tools associated with the study of social connections between social actors.

Sociogram – a graphical representation of a network, consisting of edges and nodes.

Special education needs – overarching term identifying those students who need additional accommodations for them to be successful in schools.

Unipartite/bipartite – characteristic of a network based on if the connections between actors are between actors of the same type (unipartite) or different type (bipartite).

Unweighted/weighted edges – characteristic of edges based on if actors can have multiple connects between them (weighted) or only one connection (unweighted).

Within-school tracking – the practice of placing students within the same school into differentiated tracks, often “career” or “academic” placements.