COULD NEIGHBORHOOD SCHOOLS ALSO BE DIVERSE SCHOOLS? A LINEAR PROGRAMMING APPLICATION FOR CMS STUDENT ASSIGNMENT BOUNDARIES

A Thesis
Presented to the
Center for Interdisciplinary Studies
Davidson College
Davidson, North Carolina

In Partial Fulfillment
of the Requirements for the Degree
Data Science and Educational Studies

by
Lena Parker
May 2017

COMMITTEE MEMBERS

Committee Chair: Hilton Kelly, Ph. D.
Associate Professor of Educational Studies
Davidson College

Committee Member: Raghuram Ramanujan, Ph. D.
Assistant Professor of Mathematics and Computer Science
Davidson College

Committee Member: Eriberto Lozada, Ph. D.
Chair of the Center for Interdisciplinary Studies
Professor of Anthropology and Environmental Studies
Davidson College

ABSTRACT

Though Charlotte-Mecklenburg Schools (CMS) was once at the forefront of school desegregation in the south, CMS schools have become increasingly racially and socioeconomically segregated. As CMS undergoes a review of student assignment for the 2018-2019 school year, I recommend a mathematical optimization approach to rezoning student assignment boundaries to increase socioeconomic diversity in schools. Linear programming is a promising mathematical optimization technique for producing student assignments results which uphold logistical constraints, while optimizing for socioeconomic diversity and commute times. My algorithm uses geographic data from CMS and Mecklenburg County, as well as demographic data from the census. I explore three variances of a linear programming algorithm. A simple optimization which allocates students to their closest school increases socioeconomic diversity compared to the current student assignment. Models that weight for socioeconomic diversity can further increase socioeconomic diversity and reduce commute times from the current model. I find linear programming has the potential to significantly increase socioeconomic diversity within the student assignment boundaries.

Keywords: School segregation, student assignment, linear programming, student assignment boundaries

ACKNOWLEDGEMENTS

Thank you to my advisers, Dr. Hilton Kelly and Dr. Raghu Ramanujan for their guidance and support. Thank you to Dr. Amy Hawn Nelson, Scott McCully, and Michael Alves for their student assignment and CMS expertise. Thank you to Dr. Melinda Adnot, Dr. Hamurabi Mendes, Dr. Laurie Heyer, Dr. David Backus, and Ross Kruse for additional technical support with varying aspect of my thesis. Thank you to Conor Hussey and Rachel Lee for emotional support through the trying times of program crashes and the struggle of working with large data. Lastly, I would like to thank Laurie Parker and Chris Parker their assistance editing this thesis.

TABLE OF CONTENTS

INTRODUCTION

Charlotte-Mecklenburg School (CMS), located in North Carolina, is one of the largest 20 school districts in the US. The level of school segregation in CMS has gained national attention in recent years. The New Yorker published an article which discussed the links between school segregation in CMS and the deep racial tensions which led to rioting in Charlotte during 2016 (Smith, 2016). The popular late-night show, *Last Week Tonight with John Oliver,* spent an entire 20 minute segment focused on school segregation, using CMS as an extreme example (LastWeekTonight, 2016). During the 2010-2011 school year, 44 percent and 47 percent of Latino and Black student respectively attended racially isolated schools in CMS (Ayscue, Woodward, Kucsera, Siegel-Hawley, & Orfield, 2014).

Rewind three decades, and the Charlotte Observer ran a 1984 article titled *You Were Wrong, Mr. President,* responding to Ronald Reagan's critique of Charlotte's desegregation plan; they wrote, "Charlotte-Mecklenburg's proudest achievement of the past 20 years is not the city's impressive new skyline or its strong, growing economy. Its proudest achievement is its fully integrated schools," ("You Were Wrong, Mr. President," 1984). . In the 1989-1990 school year, only one percent of Latino students and three percent Black students attended racially isolated schools in CMS (Ayscue et al., 2014). This public attention is largely due to the notable success of desegregation in CMS during the 1980s in contrast with highly segregated schools today. This return to segregated schools, and reversal of desegregation is termed "resegregation."

CMS spans rural, suburban, and urban areas. CMS now educates significant Hispanic, Black, and white populations, at 23, 39, and 29 percent respectively, and growing populations of other racial and ethnic groups. ("Charlotte-Mecklenburg Schools Month 1, 2016-17 School Diversity Report," 2016). CMS has experienced a large influx of Hispanic students, as well as other immigrant families from many regions of the world, and become more linguistically diverse in the past two decades (Mickelson, Smith, & Nelson, 2015). Despite the diversity of the district as a whole, individual schools in CMS remain largely segregated.

In this introduction, I consider the context of the vast segregation CMS faces, and examine the potential for CMS to diversify its schools. I provide evidence for the importance of having diverse schools, outline the legal framework surrounding segregation and student assignment, and the trends in popular and political support for desegregating student assignment.

History of CMS Segregation and Resegregation

As with all districts, CMS was federally required to desegregate its schools after the 1954 *Brown v. Board of Education* ruling. CMS began desegregating at a small-scale in the 1960s with token desegregation, or placing a very small number of minority students into white schools. The 1971 Supreme Court case *Swann v. Charlotte-Mecklenburg Board of Education* stemmed from CMS community efforts to desegregate. Desegregating student assignment plans, commonly referred to as busing plans were permitted as a means of achieving racial integration. CMS went on to design a robust assignment plan that aimed to desegregate schools more broadly during the 1970s and 1980s. Through various desegregation tactics, CMS managed to lower racial dissimilarity indices of schools, a common measure of desegregation, to an all-time low during the 1980s.

Beginning in the 1990s, the desegregation-focused student assignment plan began to draw considerable opposition, especially from white families who felt as if their children were disadvantaged by the desegregating student assignment plan (Mickelson et al., 2015). *Swann* was eventually overturned by the case *Capacchione v. Charlotte-Mecklenbug Schools* in 1999 (*Capacchione v. Charlotte-Mecklenburg Schools*, 1999). In 2001, the Family Choice Plan replaced the old diversity-focused student assignment plan, and marked the beginning of resegregation in CMS (Mickelson et al., 2015). In 2010 the Black/white dissimilarity index of CMS had returned to a level comparable to the dissimilarity index in 1970, when CMS desegregation had barely begun (Mickelson et al., 2015). Both racial isolation and dissimilarity indices are common measures of racial segregation. By all commonly used measures of desegregation, CMS desegregated its schools successfully, but has resegregated schools since the Family Choice Plan (Mickelson et al., 2015; Siegel-Hawley, 2014).

In regards to terminology, I primarily use the term *desegregation* rather than *integration.* Desegregation typically refers to placing different students in the same school, whereas integration is a step farther and involves the meaningful inclusion of and interaction between different students in a school. Measuring desegregation and trends over time is much easier than measuring the level of true integration within a school, and therefore I primarily discuss desegregation. Furthermore, I use the terms diverse schools and diversify. The phrase diversifies and desegregate can be used almost interchangeably. When discussing a singular type of diversity, such as socioeconomic diversity, I use the term desegregate, since there are common metrics for segregation. The term diversify more broadly encompasses a difficult to measure, yet important understanding of student identities and diversity. Literature around desegregation has shifted towards language of diversity, which allows for the inclusion of more students' identities, as well as more positively frames the dialogue.

School segregation and diversity are complex issues that exist at the intersection of city planning, residential segregation, changing city demographics, and community attitudes. Though it is hard to detangle the many factors that influence segregation of schools, school boards and governments can change few of these parameters. While school boards cannot change the demographics of the city or significantly alter residential segregation, they do determine student assignment plans. These plans determine how students are allocated to schools, and are the most direct lever for changing the composition of schools. Due to the complex nature of diversity and segregation, I choose to focus only on the aspects that are most easily influenced by the school district and schools board: student assignment practices.

CMS Student Assignment Review. As CMS faces vast resegregation of schools and many factors out of their control, CMS is also actively reviewing their student assignment practices, and therefore has a unique opportunity to potentially desegregate their schools. In 2015 CMS announced the beginning of a periodic review of their various student assignment practices. Danielle Holley-Walker's examination of segregation in Southern school districts notes that school districts possess most of the power and responsibility to diversify schools (Holley-Walker, 2010). She finds that intentional efforts on the part of the district is a necessary ingredient for diverse schools. Phase I of this student assignment review led to a plan for school options and magnet schools, and was approved in November of 2016. Phase II aims to redefine student assignment boundaries. These boundaries, which geographically zone students into schools, are a large component of student assignment plans, since they serve as the default assignment for all students.

The Case for Diverse Schools

The *Brown v. Board of Education* decision to desegregate schools in 1954 was based largely on a narrative that segregation harmed black children by isolating them in inferior schools (*Brown v. Board of Education of Topeka*, 1954). Though much of this argument was based on factual and significant resource discrepancies between white schools and colored schools, the narrative that desegregation only benefits disadvantaged students is not the whole story, yet persists today. For this reason, I address the case for diverse schools from both the perspective of the harm done to disadvantaged children, as well as the advantages afforded to all children.

Segregation Harms Disadvantaged Students. A wide body of evidence suggests that heavily segregated schools result in poorer academic outcomes for low income and racial minority students, as well as worse life outcome measures such as post-secondary education and lifetime income (Darling-Hammond, 2010; Giersch, Bottia, Mickelson, & Stearns, 2016; Mickelson, 2008; Orfield & Lee, 2005). Though the exact mechanisms by which these differing outcomes occur may be a compounding effect, disadvantaged students perform far better in diverse schools.

Resource disparities, resulting from desegregated schools, are a significant reason that disadvantaged students are harmed by segregated schools. Schools with high racial and ethnic minority populations usually have worse facilities, less experienced teachers, fewer honors and AP course offerings, and are under-resourced in other ways (Darling-Hammond, 2010, (Giersch et al., 2016; Orfield & Lee, 2005). Segregation of students also leads to disadvantaged students being isolated from socially valuable networks for knowledge sharing and cultural capital, such as knowledge of the college admissions process (Mickelson, 2008). There are likely many other mechanisms by which segregated schools harm disadvantaged students; it is clear that disadvantaged students are harmed by isolated schools and benefit from diverse ones.

In many desegregation narratives, the case for diverse schools ends here. The underlying message is that diversifying schools only serves disadvantaged students within these schools, and even perhaps at the cost of harming more advantaged students. Shifting the case for diverse schools to include all students and dispel narratives that diversity harms advantaged students changes the desegregation narrative.

Diverse Schools Benefit All Children. Diversifying schools inevitably leads to some degree of pushback from more advantaged families who feel as if their child will be harmed by increasing diversity. Often, this pushback is so severe, districts do not carry through with plans to diversify schools. Despite the pervasive narrative that diversity harms more advantaged children, most research has found no negative impact on the academic outcomes of white and wealthier students in diverse schools. (Giersch et al., 2016; Jayakumar, 2008). On the contrary, some research has found positive effects on academic outcomes for more advantaged students in diverse schools (Wells, Cordova-Cobo, & Fox, 2016).

Diverse schools socially benefit all students. Students in more diverse schools demonstrate reduced implicit biases and negative attitudes towards other races and ethnicities, and more positive intergroup relations and understanding. (Hallinan, 1998; Schofield, 1991). Reducing implicit biases and improving intergroup understanding from a young age has the potential to reduce deeply ingrained racism and classism.

The potential for diverse schools to improve social, educational and long-term life outcomes, paired with currently widespread segregation is cause for concern. Not only may diverse schools benefit students academically and socially, but districts also have a legal obligation to desegregate schools.

Legal Framework

The long and complex history of school segregation, desegregation, and resegregation in US public schools is punctuated and guided by several legal cases. Understanding the effects of these court cases is crucial to understanding the effects and limitations of geographically-based student assignment plans. The legal framework constructed by segregation court cases requires that schools be desegregated, but also limits how districts may desegregate their schools. Most segregation litigation focuses on racial segregation, though the benefits of diverse schools extend to other types of diversity; for example, my analysis considers socioeconomic diversity. This race-focused legal framework does relate to other types of desegregation.

Schools Must Desegregate. The most well-known case concerning school desegregation, *Brown v. Topeka Board of Education* in 1954, decided that “in the field of public education the doctrine of ‘separate but unequal’ has no place” and that racial segregation within public schools violates the US Constitution (*Brown v. Board of Education of Topeka*, 1954). *Brown II,* as the 1955 follow-up case has become known, expanded upon the initial *Brown* decision, specifying that local districts must fully comply with “deliberate speed” (*Brown v. Board of Education of Topeka*, 1954). *Brown I and Brown II* concluded that school districts may not have student assignment plans to separate races. However, the judicial system did not propose specific race-neutral alternative assignment methods. During the 1960s and 1970s when many districts and communities resisted desegregating schools, many racial minority communities and families found that threat of litigation was one of the only ways to hold districts accountable for the *Brown* decision (Douglas, 1995).

Keyes v. School District No. 1 is one such case brought against a district for its assignment plan on behalf of district families. In this 1973 case, courts ruled 7-1 in favor of *Keyes*. This case ruled that student assignment policies with a segregating effect are also illegal under *Brown*, even if the policy does not explicitly segregate based on race (*Keyes v. School District No. 1*, 1973). This case specifically focuses on policies with a segregating effect, but not necessarily explicit racial language in the policy. Segregation which is explicitly written into laws or policies is called *de jure* desegregation, while policies which desegregate but are not explicitly written into policy, are called *de facto* segregation. Drawing geographic boundaries to take advantage of racial residential segregation, as in the *Keyes* case, is considered a violation of *Brown*.

Legal Limits on Desegregation Policy. In response to movements to hold districts accountable for desegregation, other groups pursued legal limitations on what school districts could do to desegregate. *Washington v. Seattle School District No.* and *Parents Involved in Community Schools v. Seattle School District No. 1* are two instances of pushback against desegregating policies. In 1982 *Washington* argued it is unconstitutional to assign students to any other school than their geographically closest school, but judges ruled in favor of *Seattle Schools*, giving districts discretion in their pupil assignment strategies (*Washington v. Seattle School District No. 1*, 1982). The 2007 *Parents Involved* case ruled that certain student assignment practices which considered individual student race violate The Fourteenth Amendment (Parents Involved, 2007). As a result, an individual's race or ethnicity may not be used in student assignment decisions. This has implications for student assignment plans which consider students on an individual basis, such as school choice lotteries.

School districts often replace individual's race with neighborhood demographics or socioeconomic status as proxies, to comply with the from *Parents Involved* (Siegel-Hawley, 2014; Stroub & Richards, 2013). For example, the CMS-set goal to "reduce the number of schools with high concentrations of poor and high-needs children" in the district's new student assignment plan; this is an example of the shift in willingness to prioritize socioeconomic segregation rather than racial segregation (CMS Board of Education, 2016).

In the 1977 case, *Milliken v. Bradley*, the Supreme Court ruled that districts are only responsible for desegregating within the district, not between districts. This limits the possible desegregation, since studies estimate that the majority, more than 75 percent, of segregation in schools is due to between-district segregation rather than within-district segregation (Bischoff, 2008; Clotfelter, 2004; Siegel-Hawley, 2013). District fragmentation has often been a tactic for suburban, wealthier portions of districts to separate and maintain highly concentrated wealth, as well as means for white communities to avoid racial desegregation. Highly homogeneous districts, which are usually smaller, simply do not have the diversity of students to assign students in a way that leads to diverse schools. CMS on the other hand is far from homogenous, and has a large number of schools to assign students to.

These court cases have created a legal framework that school districts and states must work within. In general, any form of *de jure* or intentional *de facto* desegregation is illegal. School districts may face legal intervention if their schools are not sufficiently desegregated. More and more districts have been granted unitary status in the past decades, and instances of legal prosecution for not desegregating schools have drastically decreased. This means fewer schools face losing federal funds for not having desegregated schools.

Conversely, there is litigation risk if school districts attempt to desegregate schools in a way that specifically targets individual student's race or ethnicity. School districts may not use individual student race or ethnicity, and do not need to desegregate across districts, thus limiting desegregation tools available to states and districts. Parallel to the decrease in legal pressure to desegregate, political will has also declined.

Political Framework

Student assignment plans are typically approved by the district school board, which has been elected by the citizens within the district. Because school boards rely upon the support of their constituents, they are responsive to the priorities and attitudes of their constituents. Student assignment plans are one of the most contentious and difficult decisions that school boards make, and often determine the results of the next school board election. For this reason, popular attitudes around student assignment and desegregation influence school board decisions and therefore student assignment plans.

As seen in Charlotte, the community organizing focused on desegregating schools reached a highpoint during the 1970s and 1980s. This political will for desegregation has steadily declined in the decades since. Mickelson et al. attribute a portion of the decrease in Charlottean's pride for their integrated school system to a large influx of newcomers (Mickelson et al., 2015). Nationally, popular will and thus political will for desegregating schools has decreased despite patterns of increasing segregation.

The legal threat if districts involve race in assignment plans, paired with the disappearance of federal funds to incentivize desegregation, has caused districts to not consider race at all in student assignment despite the prevalence of racial segregation (Fiel, 2013). This has resulted in a decrease in desegregation-focused student assignment (An and Gamoran, 2009; Reardon,Grewal, Kalogrides & Greenberg, 2012). The decrease in legal pressure, rise of legal constraints, and changes in social pressure have led to many districts retreating entirely from plans to racially desegregate schools.

Additionally, the decreasing community and political will for desegregated schools has contributed to a new discourse around education reform and desegregating schools: school choice and school excellence. Families increasing value school choice, such as voucher system, charter schools, magnet programs, lottery assignment systems, over equitable schools. The shifting popular will away from desegregation, and towards school choice is evidenced by the CMS Student Assignment Survey; school choice and school excellence were consistently ranked as more important than diverse schools (Charlotte-Mecklenburg Board of Education, 2016). Though some choice advocates argue that school choice can increase within school diversity has, the literature finds this to be the case only if districts intentionally use choice plans to diversify schools; otherwise choice plans typically further segregate schools (Reardon & Yun, 2003; Sohoni & Saporito, 2009). In this paper, I do not focus on school choice for several reasons, though school choice and excellence are an important narrative to mention. First, a large amount of research on school choice and school diversity already exists. Secondly, I believe that non-choice student assignment practices have larger potential to desegregate schools.

The Promise of Optimizing Student Attendance Boundary Rezoning

Though the school choice movement is a dominant debate currently, a National Center for Education survey in 2012 found that 87 percent of students in traditional public schools attend their assigned school (NCES, 2012). The composition of schools is largely determined by assignment boundaries rather than school choice. Recent research suggests that student assignment boundaries have significant consequences for the diversity of schools (Richards, 2014; Saporito & Van Riper, 2016; Siegel-Hawley, 2013; Sohoni & Saporito, 2009).

Broadly, where is there potential to further desegregate, and what tools are available to increase desegregation? Student attendance boundaries emerge as a clear linchpin for altering school diversity.

Literature on Student Assignment Boundaries. Meredith Richards performs a large-scale analysis of all available attendance zones in the US. She examines the effect of their "gerrymandering," and defines attendance zone gerrymandering as the difference between the actual zone and a natural zone. Natural zones are the "logical" zones that would be expected if the attendance boundaries were only based on proximity to a school. Richards finds that the majority of attendance boundary gerrymandering increases the segregation in schools. However, notably, she finds that a very small portion of districts have irregular boundaries that actually increase diversity in schools. A disproportionately high number of these districts are under court desegregation orders, and are likely using irregular boundaries intentionally (Richards, 2014, p. 1143).

Richards' large-scale analysis provides valuable insight into the widespread impact of attendance boundaries and how they typically further segregate schools. However, her study spans so many districts and areas with different contexts and histories. By focusing on just one district, I am able to more fully understand the political climate, history, and local policy around student assignment.

In a narrowly-focused case study similar to my analysis, Genevieve Siegel-Hawley analyzes the potential impact of several proposed attendance boundary plans for Henrico County Public Schools. She finds that Henrico eventually chose a plan that “failed to embrace the growing diversity of the school system,” and would “solidify extreme patterns of racial isolation,” (Siegel-Hawley, 2013, p. 580). In a qualitative study of the political process of rezoning in Richmond schools, one report finds that the voices of white families were privileged over less advantaged families. This led to boundaries that largely served to isolate these white families and furthered segregation in the schools (Siegel-Hawley, Bridges, & Shields, 2017).

Salvatore Saporito and David Van Riper examine attendance boundaries, specifically irregularly shaped zones, to determine their effect on segregation (2016). Saporito and Riper find that most attendance boundaries are fairly compact ,and therefore simply mimic patterns of residential segregation.

Mathematical Optimization to Construct Student Assignment Boundaries

Considering the resegregation in CMS, and the proven negative effects of segregation, the effort to reassign school boundaries in CMS is critical. Given the history of student assignment and desegregation, the widespread use of student assignment boundaries, and CMS’s current student assignment review, I examine CMS’s student assignment boundaries specifically. I would like to test the popular understanding that low commute times and more diverse schools are a trade-off.

To do so, I explore mathematical optimization to systematically construct assignment boundaries in a manner that intentionally diversifies schools. GIS software and mathematical optimizations have long been paired to solve planning problems in other sectors. As long as there is political will, a linear programming approach shows strong promise as an option to significantly increase diversity in schools.

I found no literature on American school districts actually implementing an optimized student assignment boundary plan. I found this gap surprising given the extensive literature on the effects of choice systems on school diversity. Many districts have moved away from explicit desegregation plans, assuming that choice systems offer paths to desegregation, despite the lack of evidence for this. Without focused efforts to ensure diverse families enter choice systems, or somehow promoting diversity in choice systems, choice systems seem to further segregate schools. I re-examine the potential for student assignment boundaries to diversify schools, and address the perception that diverse schools cannot come from reasonable geographic student assignment boundaries. Specifically, I explore how a mathematical optimization approach to rezoning student attendance boundaries in CMS might significantly increase diversity within schools.

METHODS

Overall, this linear programming based approach uses road segments as the basic unit of assignment for this optimization. Each road segment is given an estimated population of students who live on this road segment and an estimated socioeconomic status measure. I aim to have a model that assigns all road segments, and therefore the students who live on these roads segments, to schools in a way that minimizes commutes times and SES imbalance, while not assigning more students than a school's capacity.

I aligned the methods of this optimization and analysis with the priorities and needs of the current CMS Student Assignment review. For guidance on the specifics of mathematical optimization, I use previous optimization research on student assignment. Lastly, and by necessity, I make some assumptions to fit the data that is available to me, and simplify this problem where necessary.

CMS Priorities

In their Student Assignment Plan goals, CMS prioritizes "constructing attendance boundaries [ ] that contribute to a socioeconomically diverse student population," and not other types of diversity (Charlotte Mecklenburg Schools, 2015). For this reason, I focus only on SES diversity, and do not include racial, ethnic, or linguistic diversity in my model. The goals and priorities of Phase II also include keeping commute times to schools low, and enrolling schools close to capacity. These are also facets of my optimization algorithm.

Literature on Mathematical Modeling of Student Assignment.

Beginning in the 1960s and into the 1970s, researchers from the management science field proposed various math modeling approaches to assign students to attendance boundaries (McNamara, 1971). Several researchers, such as Laila Heckman and Howard Taylor proposed linear programming as a solution to racially integrating schools through the design of student assignment boundaries (Heckman & Taylor, 1969). Others have proposed a different modeling approach, called a network flow model, to optimize for racial desegregation (Belford & Ratliff, 1972).

In 1980, a study was published which constructed potential assignment boundaries for a school district in Odense, Denmark. With a sample of 46 schools and 176 census blocks, these researchers are able to better assign students to keep enrollments within the capacities of schools (Jennergren & Obel, 1980). This linear programming example uses quite large blocks of assignment and therefore is not very granular in its conception of commute time or student population.

The general objective of these linear programming and other math modeling approaches is to assign students and clusters of students to schools in a way that best solves a set of goals, and this requires an algorithm to assess many possible solutions. Many of these optimizations were initially implemented on small scales, such as the assignment of only one census tract, or the assignment of students to only a couple of schools. This is due to computer having less processing ability in the 1960s and 1970s.

More recently, a group of researchers proposed a tool that incorporates linear programming optimization and human tweaking of boundaries to achieve new boundaries (Caro, Shirabe, Guignard, & Weintraub, 2004). Other than this, few researchers after 1980 address math modeling approaches to student assignment boundaries. This could be due to a decreasing political will for desegregation, an increase in the study of magnet schools and a preference for controlled choice plans to desegregate schools. Overall, there have been few studies which examine the potential for linear programming to improve the process of geographic student assignment, and fewer that examine it on a granular scale.

Data Sources

Data used for this problem are publically available from The Census Bureau, CMS, and Open Mapping Mecklenburg. Open Mapping Mecklenburg is an open source spatial data collection run by Mecklenburg County. Table 1, below, summarizes the data used for this analysis. All data either describe the roads and the students which live on these roads, or the schools in CMS. These are the basic units of assignment for this optimization.

Table 1
Data Sources and Descriptions

Data	**Description of Data**	**Source**
CMS School Locations	CMS elementary school locations for the 2015-2016 school year in a shape file	Open Mapping Mecklenburg
CMS School Attendance Boundaries	CMS elementary school boundaries for the 2015-2016 school year in a shape file	Open Mapping Mecklenburg
School Capacities	School capacities for all schools in sample, I used the “Student Core Adjusted CR (Classroom) Capacity” to account for school capacity including all classrooms (mobile and traditional)	CMS
CMS Schools SES Demographics	Current student population SES percentages, CMS provided this student population breakdown in Spring 2016.	CMS

Roads	Shape file of all Mecklenburg County streets with speed limits, road name, road type, and other data about the network of schools	Open Mapping Mecklenburg
Block Groups Boundaries	The shape file for the Mecklenburg County block groups	Open Mapping Mecklenburg
Population Estimates	The 2015 ACS population estimates for each census block group. I restrict population to individuals between five and nine years old.	American Community Survey (ACS), US Census Bureau
SES Category	Categorization of low, medium or high SES was obtained from CMS and the image below showing the classification of each of 555 census block groups in Mecklenburg. These values are based on a compound measure using five Census-collected variables.	CMS (based upon ACS data)

Student Population Estimates. Data to estimate the student population on each road segment comes from the 2015 American Community Survey. The ACS collects population data in age brackets, and the bracket from ages from five to nine most closely aligns with the age range of elementary students. I use the ACS 2015 five to nine year old population to approximate the number of elementary school aged children in CMS, however, it is an imperfect measure for two main reasons (United States Census Bureau - American Fact Finder, 2015). First, elementary schools serve children in a broader age range than five to nine, including students who can be as old as 10 or 11. Therefore the ACS population estimates exclude these older students, and are therefore potentially an underestimate of the elementary age population in CMS. Second, ACS data include all students in Mecklenburg County, which include students who do not attend CMS schools, leading to an overestimate of the number of five to nine year old students in CMS schools.

The total ACS Mecklenburg County population between five and nine, 71,098, is close to the total school capacity of all 94 home elementary schools in CMS, 71,821. Because this ACS population measure is close enough to the CMS school capacities total, I used the ACS population estimates in my optimization.

Socioeconomic Status Compound Measure. Charlotte Mecklenburg Schools provides the compound measure of socioeconomic status for each Census block group for all of Mecklenburg County. Household income, parent educational attainment, English language ability, single parent household status, and home ownership rates are the five factors used to assign a SES value to each of 548 block groups. CMS uses this measure to prioritize magnet school admission based on the home residence of applicant students. To parallel CMS's multi-dimensional understanding of SES, I rely on this SES measure for my optimization of SES diversity.

Each block group is given a compound measure of SES and then ranked 1 – 548. These block groups are classified into three categories: low SES, medium SES, and high SES, in order of least to most socioeconomically advantaged. There is a near equal breakdown of CMS students in the low SES (36 percent), medium SES (35 percent), and high SES (29 percent) categories.

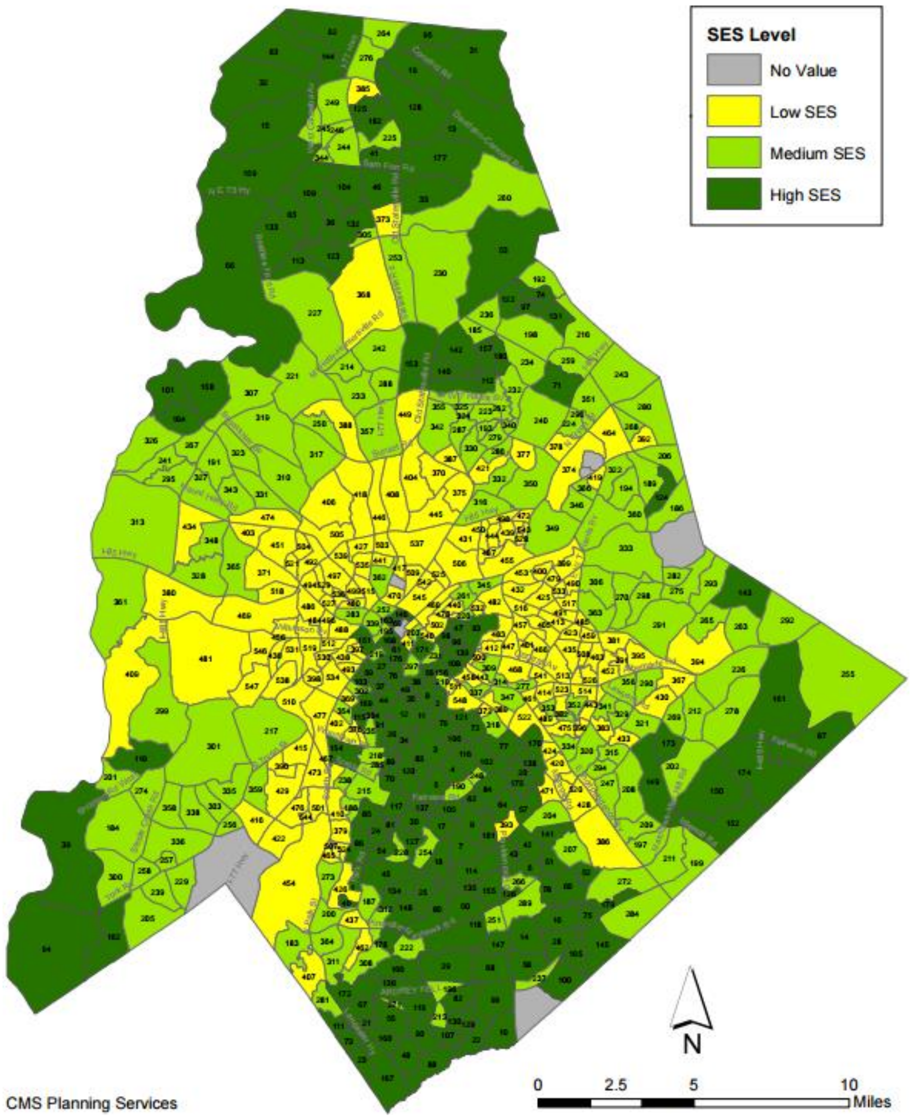

Figure 1. The SES measure across CMS census block groups.

SES Diversity Measure. CMS released data about the distribution of students across these SES categories for each school, with a number and percentage of students who fall into each category. I use this data to weight the relative value of assigning each road to a school. For example, if a school has very few high SES students, a road segment in the high SES category would be relatively more valuable to add to this school than other road segment.

ArcMap Calculations and Data Formatting

I used ESRI's Geographic Information Systems (GIS) software, ArcMap, to format data needed for this optimization.

Figure 2. This map shows a section of Mecklenburg County with three schools, a network of roads and their midpoints. Schools and roads are overlaid on census block groups with varying populations and SES measures.

Restricting Optimization Units. In ArcMap, I restricted the sample of all schools to the 94 elementary schools (serving at least kindergarten through 5^{th} grade) that are home schools. Home schools have a unique geographic attendance boundary to assign students to this school;

this excludes several full magnet schools and specialized schools which draw students from a wide geographic area. I also restricted the road segments used in my assignment. I excluded road segments that are highways with speed limits of 65 mph or above, as well as high on-ramps. This is to limit to roads that are likely to have students living on them for this assignment. Several of the roads in *Figure 2* do not have marked midpoints. This is because these roads are not considered in my model due to the above restrictions.

Estimating Road Segment SES and Population. I used ArcMap to spatially join SES measures to each road segment based on which census block group the midpoint of the road segment fell within. In *Figure 2,* block groups are shown as the blue and green tone background polygons. Each of these has a different, CMS measured, SES score. I assign an SES value (low, medium, or high) to each road segment (signifying the estimated SES of the students living on that road) based on which census block group this road segment's midpoint falls within. I assign a medium SES value to the seven census block groups which do not have CMS assigned SES value due to the lack of CMS students living in the census blocks. This choice does not have a large impact on my optimization since a very small number of children are estimated to live within these "nonresidential" census block groups.

The estimated population living on each road segment was also determined using block groups and spatial joining. The population estimated to live on a road segment is based on which block group this road is within and how long this road on. Each road is given the total population of the block group weighted by its proportion of the total road length. For example, consider a block group which contains 10 total miles of roads, and 100 students. A road of length two miles would have and estimated 20 students living on it, since this road comprises 20 percent of the total road length. The age 5 to age 9 population living within each block group ranged from 0 to

666, with a mean of 128 students. Each road segment had a range of 0 to 63.71 students estimated, and an average of 1.51 students per road segment.

Many census block groups are divided along road lines, and therefore the midpoint of roads lay across the block group's boundary. In this scenario, it is difficult to decide which block group attributes should be assigned to a boundary road. In this case I assign this boundary road the attributes of the block group this road's midpoint falls within. Since the roads data and block group data are from different data sets there is a slight misalignment and the road midpoint falls on one side of the line. I assume that this is a near-random way of deciding which attributes to assign to roads which are likely similar to both block groups that it borders.

Calculating Commute Times. ArcMap calculated the commute times, in minutes, between every road segment and its five closest schools, assuming it would be unreasonable to assign students to a school that is more than the fifth closest school. Commute times were based upon road segment lengths and listed speed limits. ArcMap produced a cost matrix of the commute times between all roads and their closest five schools. I artificially assign an arbitrarily large number (99,999 minutes) to the commute time of a road segment to every other school. This is to ensure that a road segment is not assigned to a school farther than its fifth closest school by imposing an unreasonably high commute cost for such an assignment. The commute times appear low because ArcMap commute calculations do not include traffic, stoplights, or other impediments that a driver encounters.

Optimization Model

Based on the requirements of this problem, and the above data, I have chosen a linear programming (LP) approach, since all constraints can be specified linearly and LP fits well with allocation problems such as this one. Linear programming is a mathematical optimization technique that solves for a number of variables for a set of linear constraints and costs. In this problem, the variables being solved for are the assignment of road segments to schools. See Appendix B for an introduction to LP and a simple example to illustrate the basics of LP.

Logically, the assignment of road segments to schools would be binary; either a road segment is or is not assigned to a road. LP where variables are constrained to integers is called integer linear programming (ILP); and when variables are further constrained to be binary, it is called binary integer linear programming. Constraining variables to be integers increases the number of branch and bound options, and therefore increases the time it takes for an algorithm to find an optimal solution exponentially.

Assigning over 47,000 road segments to 94 home elementary schools is far too large of an integer programming problem, given the computational resources at my disposal. Linear programming instead allows for an approximate solution to this assignment problem. Binary integer programming would allow for more logical results, but given the computational limitations, a linear program produces comparable results in a feasible amount of time.

Costs and Constraints of this Model. Optimization solutions are constrained to school capacities, and assigning each road segment to exactly one school. My objective function is minimized and considers two costs: commute times and SES diversity. I minimize the total commute times for road segments to schools, but weight each commute time by a socioeconomic diversity weight. This SES weight can be understood as manipulating the commutes times to appear larger or smaller based on how “desirable” it would be assign a given road to a given school. If the SES category of a road is underrepresented at a school, the commute between this road and this school is weighted downwards and appears to be closer to that school. This makes the algorithm more likely to assign this school to this road. Alternatively, a road with an SES category that is overrepresented at a school would have a weight that artificially increase the commute of this school. The algorithm would therefore be less likely to assign this road to this school.

Comparing Various Assignment Solutions

I examined three possible models that vary the weight of SES diversity alongside commute times. Different SES weight magnitudes alter the relative importance of SES diversity and commute times, leading to slightly different assignment solutions. For example, if the SES weights are relatively larger, the algorithm will increasingly accept farther away roads if they increase SES diversity.

The three models below are identical except for the weight values for SES. Model (1), the commute only model, does not weight for SES diversity at all, but still includes weights of zero to keep the model consistent. Model (2), the moderate diversity model, and (3), the strong diversity model, do factor in SES diversity.

The scalars below are referenced in the model below.

$R = number\ of\ road\ segments, indexed\ from\ 1\ to\ i$
$S = number\ of\ home\ elementary\ schools, indexed\ from\ 1\ to\ j$
$G = number\ of\ SES\ groups\ (three\ in\ this\ problem), indexed\ from\ 1\ to\ k$
$M = the\ number\ of\ models, from\ 1\ to\ 3$

All data for this optimization are summarized in the below variables.

$A_{ij} = binary\ values\ signifying\ the\ assignment\ of\ road\ segment\ i\ to\ school\ j$

$T_{ij} = the\ commute\ time, in\ minutees, between\ road\ segment\ i\ and\ school\ j$

$P_i = the\ estimated\ number\ of\ children\ living\ on\ a\ road\ segment\ i$

$E_{ik} = the\ binary\ membership\ of\ a\ road\ segment\ i\ in\ one\ of\ three\ SES\ levels, k$

$C_j = \ the\ total\ student\ capacity\ for\ each\ school\ j$

$S_{jk} = \ the\ proportion\ of\ students\ fro\ SES\ category, k, at\ school, j$

The following formulas define the binary integer programming problem. The model determines the optimal assignment matrix, *A*. This optimal assignment is subject to the following constraints:

$$\forall i, j \ \ A_{ij} = 0\ or\ 1 \tag{1.1}$$

All assignments for all road segments *i* and schools *j* must be either 0 or 1

$$\forall i \ \ \sum_{j=1}^{S} A_{ij} = 1 \tag{1.2}$$

For all road segments, i, each may be assigned to exactly one school

$$\forall j \ \left[\sum_{i=1}^{R} A_{ij} P_i\right] \leq (C_j \times 1.25) \tag{1.3}$$

For all schools, j, the sum of the population of all assigned roads, or the total number of students assigned to a school, must be less than the 25 percent more than the capacity of that school[1].

The below function is the objective function for this model.

$$minimize \sum_{j=1}^{S} \sum_{i=1}^{R} A_{ij} T_{ij} P_i W_{ijm} \tag{2}$$

This function minimizes the total commute time for all students to schools, weighted by a diversity weight, W_{ij}. The transit times T_{ij} are multiplied by the

[1] I find that simply constraining by capacity leads to issues finding feasible assignment solutions. This is due to the fact that many areas are overcrowded and this requires that schools accept slightly more students than their official capacity. This constraint could be tightened or loosened.

population P_i to account for the cost of multiple students commuting from more heavily populated road segments[2]. This is summed for all schools.

$$W_{ij1} = 1 \tag{3.1}$$

Model (1) only considers commutes and thus the SES diversity weights are all 1, leading to no weight placed on SES diversity.

$$W_{ij2} = \frac{\left(\frac{1}{3} - SES_{ij}\right)}{4} + 1 \tag{3.2}$$

Model (2) incorporates these moderate weights, which find the distance from having a third of this SES group represented at the school, divides by four to dampen this weight, and adds one to normalize this weight. SES_{ij} is the proportion of students from the SES group of this road segment, *i*.

$$W_{ij2} = \left(\frac{1}{3} - SES_{ij}\right) + 1 \tag{3.3}$$

Model (3) includes strong weights. The weights included are the same as in model (2) but are not dampened by dividing the distance from one third.

See Appendix A for the MATLAB code for running this optimization.

Size Constraints and Partitioning the Problem

The full problem leads to extremely large matrices, which are too large to be held in memory for the optimization. Therefore, I partition the problem into four subproblems. CMS has partitioned the district into violet, grey, green and blue zones, as shown in *Figure 3.* I run the optimization separately for each of these zones. Ideally, I would run this optimization all at once

[2] I have chosen a fairly student-centric understanding of the cost of commute time. Though a school district may view the cost of a trip to one area the same, or even cheaper, if multiple students live there, I account for the total cost of student commutes. I view the length of commute for each student as a cost for that student and sum commute multiplied by number of students on that road.

with all data, so that road segments are not constrained to being assigned to schools within their zone.

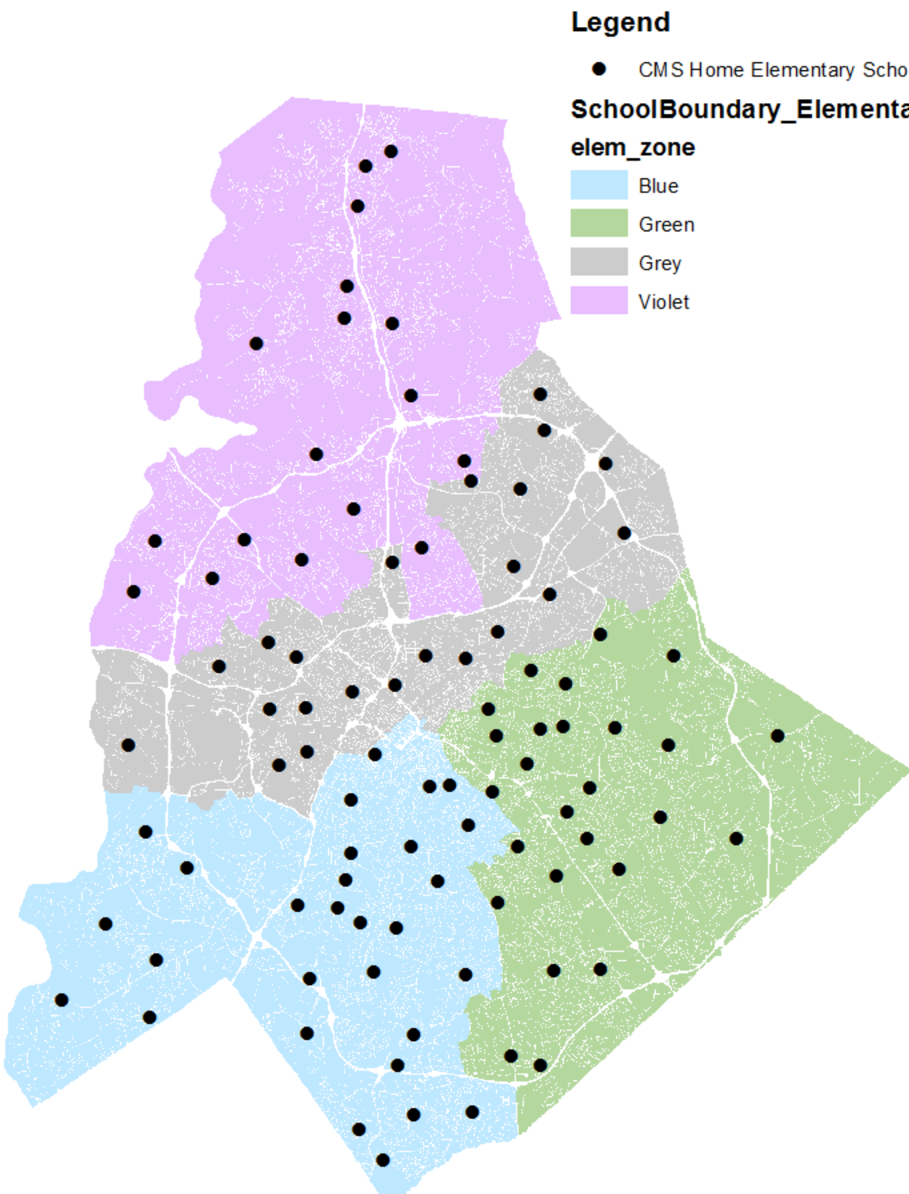

Figure 3. This map shows the four zones into which schools and portions of the county are partitioned.

Limitations and Potential Extensions of this Model

For this analysis the, the limitations and potential extensions of the model fit closely. Many improvements upon this model could be made given improved data.
Though I have attempted to tailor my linear programming application to CMS specifically, this application is laden with many assumptions imposed by that data available.

First, census block group estimates only allow me to make estimations about each road segment. I do not actually have data about the students who both attend CMS and live on that road segment. Secondly, ACS data only provides a measure of the average SES of a given census block. CMS has an extensive student information system which tracks this data, not estimates of student population.

CMS also collects self-reported SES data from families when they enter the lottery process. Though I only use data available through the census, self-reported data may give a more nuanced understanding of residential pattern of socioeconomic segregation.

Beyond the limitations of the data, there are other factors that could be incorporated into this optimization. This model could easily be expanded to include a broader understanding of diversity. Though I focus on SES diversity, racial diversity could be incorporated into this model, by adding a cost weight for how racially diverse assignment solutions are. Legally, CMS may consider widespread patterns of race in CMS to assign students to schools. Therefore incorporating racial and ethnic data into a model could lead to decreases in racial and ethnic segregation in schools.

Additionally, the commute times of this model are a simplified understanding of the actual process of transporting students to schools. Incorporating a complex transportation cost, such as the one used by T. Bektaş and Seda Elmastaş, could better account for the true cost of transport for assignment zones (2007). Bektaş and Elmastaş include transportation networks cost, which models busing routes, to optimize assignment boundaries.

Similarly to the maximum capacity constraint, a minimum capacity constraint could be imposed to insure no schools would operate under-enrolled. Lastly, I use current population estimates for this model. Projected population growth estimates could be used to account for the population growth that would affect student assignments.

Lastly, linear programming only provides a partial and near-optimal solution for this assignment. To understand how this optimization has increased diversity, I simply round the maximum assignment value to a full assignment. A binary integer programming model would better solve this problem. However, this problem is too large to solve with that method.

For now, the given data and optimization technique provide a solid foundation for the potential of optimization to better assign students to schools, but not a ready-to-implement application.

FINDINGS

Overall Findings

The assignment solutions from these optimization models result in various levels of increased SES diversity and decreased commute times, as compared to the current assignment boundaries. The commute only model leads to a reduction in average commute time, as well as a increase in SES diversity, even though this model does not include SES diversity as a factor. Even more diverse school assignments can be achieved with the other two models, at the expense of increasing commute times. All three optimization models reduce commute times and increase SES diversity from the current assignment.

Criteria for Successful Assignments

To evaluate the "success" of this optimization, I measure several outcomes, and compare these to the current assignments, as well as other hypothetical optimal assignments. Table 2, below, shows the measures of success which I use to evaluate how much improvement optimized boundaries might offer CMS in terms of commute and diversity.

Table 2
Optimization Measures of Success

	Description of Measures	Calculation
Average Total Commute Time for All Students	The measure is the total commute time, in minutes, for all students assigned to all schools.	For a given school $$\sum_{i=1}^{R} \frac{A_{ij}P_iC_{ij}}{A_{ij}P_i}$$
Average Commute Time by School	This measure is the commute time for all students assigned to each of the 94 elementary schools.	$$\sum_{j=1}^{S}\sum_{i=1}^{R} \frac{A_{ij}P_iC_{ij}}{A_{ij}P_i}$$
SES Proportions by school	This summarizes the SES distribution for each school, and the proportion of students assigned to each school that fall into each SES category.	For a given school and SES group $$\sum_{j=1}^{S} A_{ij}P_iE_{ik}$$
SES Imbalance	This is a basic measure of imbalance of SES within a school. It can be thought of as the distance from equal representation of each SES group. The smaller this imbalance measure, the more diverse this school is. Values of imbalance range from 0 to 4/3.	For a given school, where P_k is the proportion of a certain SES group at that school. $$\sum_{k=1}^{3} abs\left(P_k - \frac{1}{3}\right)$$
High Concentration of Poverty Schools	A total count of the number of schools that have a high concentration of low SES students assigned.	Counts schools from above measure where this proportion sum is greater than 90% for the low SES category.

Evaluating Optimization Success on Small-Scale Optimization

As the optimization is run on a larger and larger scale, the results become increasingly difficult to visualize with ArcMap. I calculate summary statistics to measure the success of the models at a larger scale and present those. Before the full district summary, I present a small-scale optimization of four schools' assignments to demonstrate the three optimization models as compared to the current assignment model. I select four schools and the road segments in the region currently assigned to one of these four schools. Table 3 shows the summary of measure of success for this small optimization of four schools and 2,253 road segments.

Table 3
Comparison of Current and Potential Assignments and Measures of Success for a Small Optimization

Measure	**Current Assignment**	**Model (1): Commute Only Assignment**	**Model (2): Moderate Diversity Assignment**	**Model (3): Strong Diversity Assignment**
Total Average Commute (minutes)	2.941	2.719	2.769	2.952
Average SES Imbalance Score	0.783	0.716	0.676	0.661
Count of High Poverty Concentration Schools	0	0	0	0

In Table 3, the average commute time decrease from Model (1) to Model (3), as expected by the modes. Both Model (1) and Model (2) reduce average commute time. Model (3), however has 0.0037 percent increase in commute time from the current assignment. Also as expected, the average imbalance of schools decreases from Model (1) to Model (3). Model (3) manages to reduce the average imbalance of schools to 0.661 from the current assignment average school imbalance of 0.783. Model (1), the commute only model which does not weight for SES diversity, actually reduces the average SES imbalance measure from the current assignment. This suggests that simply optimizing students with the commute time as the only cost would lead to an SES diversity increase in schools. Though there are not schools with high concentrations of poverty in this small-scale optimization, this measure is important when looking at the entire district optimization.

Figures 4 through 7 map the assignments of road segments to schools produced by the current assignment (*Figure 4*) and the three optimization models (*Figures 5, 6, and 7*). As can be seen in *Figure 7*, the optimization algorithm chooses to assign a cluster of farther away road segments to Providence Spring Elementary because they increase diversity enough. This Model (3) assignment allows for the assignment of this cluster of roads because these roads increase the diversity at Providence Spring enough to warrant the farther commute. The overall changes in boundaries from Model (1) in *Figure 5* to Model (2) in *Figure 6* and Model (3) in *Figure 7* are not that dramatic. However, from Table 3, we see that assignment differences lead to real differences in the commute times and SES diversity of schools.

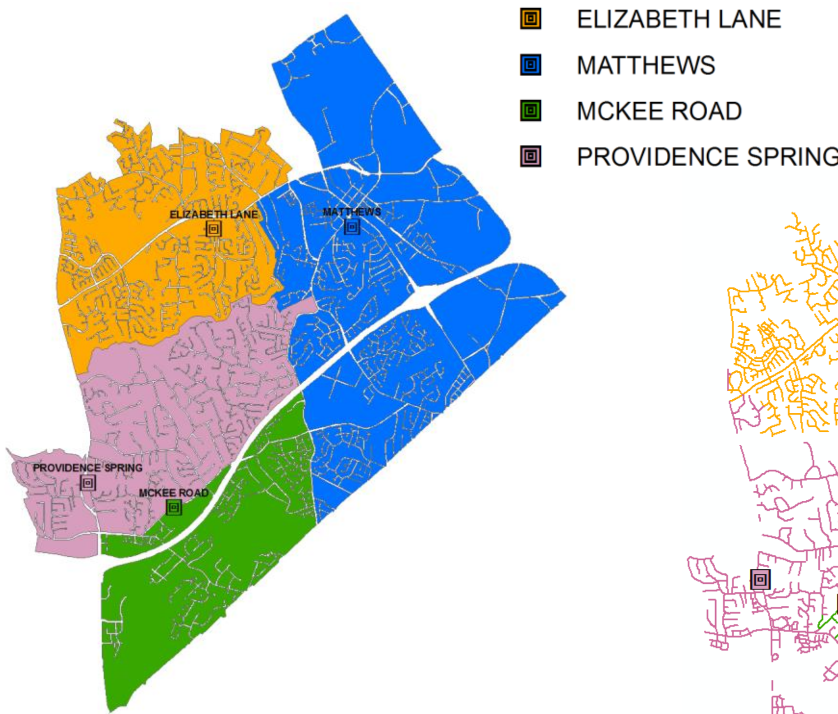

Figure 4. The current assignment boundaries for these four elementary schools

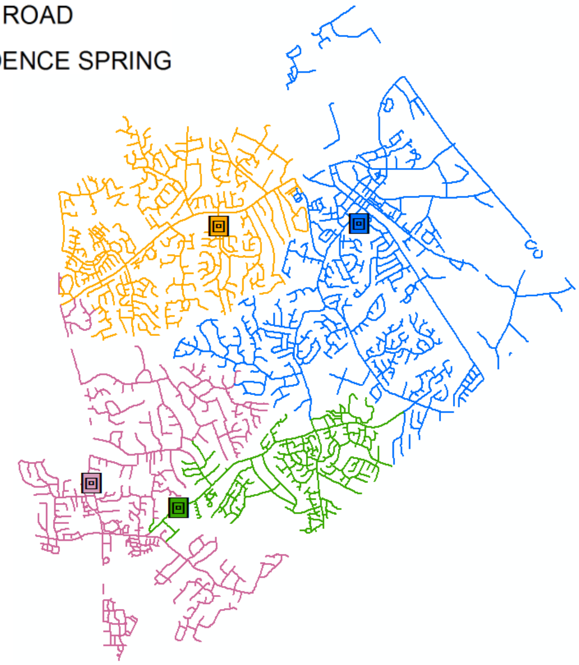

Figure 5. The output assignments for Model (1), the commute-only optimization

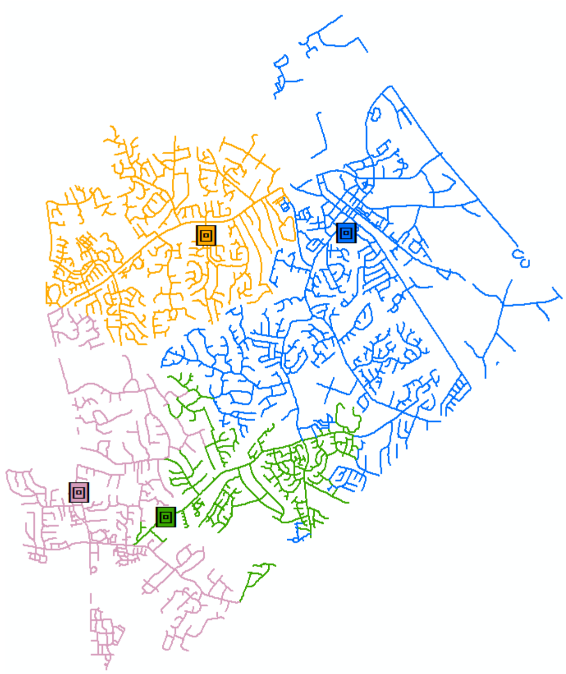

Figure 6. The output assignments for Model (2), the moderate SES diversity optimization

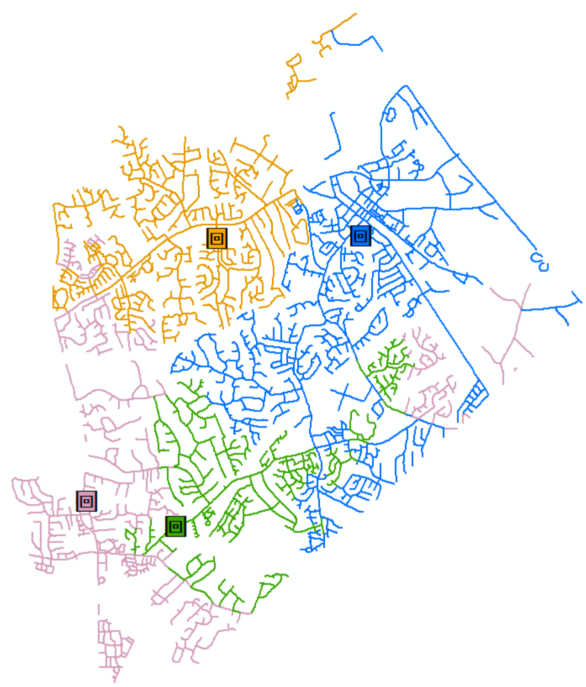

Figure 7. The output assignments for Model (3), the strong SES diversity optimization

Summary of Full District Assignments

Table 4

Comparison of Current and Potential Assignments and Measures of Success Full District

Measure	**Current Assignment**	**Model (1): Commute Only Assignment**	**Model (2): Moderate Diversity Assignment**	**Model (3): Strong Diversity Assignment**
Total Average Commute (minutes)	2.741	2.497	2.688	2.901
Average SES Imbalance Score	0.888	0.813	0.764	0.710
Count of High Poverty Concentration Schools	17	15	8	6

Table 4 results for the full district parallel the results from Table 3, for the small scale example. The models generally have results that are expected. Average commute and SES imbalance decreases from Model (1) to Model (3). The full district actually has a count of school with a high concentration of poverty. Of the 94 home elementary schools, the current assignment leads to 17 which are currently above 90 percent low SES students. As one would hope, the number of high poverty concentration schools decreases in all assignment models. Model (3), the strong diversity assignment reduces the number of high poverty schools to only 6. Model (1), which only uses commute as a cost, manage to decrease the count of high poverty schools from the current assignment from 17 to 15 schools. Model (1) also decreases the imbalance measure to 0.813 from 0.888, in the current assignment. The fact that Model (1), which is seeks only to minimize commutes, actually increase diversity in schools would support the hypothesis that the current boundaries actually serve to further segregate schools. Model (2) also manages to reduce average commute time from the current assignment, from 2.741 minutes to 2.688 minutes, and increase diversity on both measures of SES diversity. Both Model (1) and Model (2) contradict the community attitude that increasing diversity will require an increase in commute times.

CONCLUSIONS

A process, such as optimizing student assignment boundaries remains a legal option for school districts since it only considers wide-spread patterns of demographics. Though political feasibility is more difficult to gauge, I believe that the CMS community could be convinced of the benefits of diverse schools, and thus the benefits of a more systematic and optimal method of student assignment.

A linear programming approach to student assignment in CMS has the potential to increase SES diversity in schools while maintaining commute times. Even without optimizing for SES diversity, this analysis shows that further diversity could be achieved with sacrificing commute times. When including SES diversity in a linear programming algorithm, SES diversity can be increased further, while keeping commutes at a lower level than the current assignment.

CMS has already incorporated a measure of SES diversity into Phase I of their student assignment effort, and made SES diversity a priority for rezoning student assignment boundaries. If CMS intends seriously to desegregate schools, linear programming could be an effective way for CMS to include a focus on SES diversity in Phase II of their student assignment plan.

A linear programming approach to student assignment boundaries has benefits beyond meeting the goals of the assignment plan. The drawing of student assignment boundaries is almost always extremely contentious, but agreeing on a set of priorities to optimize could be significantly less contentious. Linear programming could assign schools in a more impartial manner, based on optimizing selected criteria. Computing boundaries this way can be more systematic and less subjective, and has the potential to reduce or counter some of the unequal power dynamics and biases that can be present with other methods of setting boundaries. Additionally, this method is a more systematic way or CMS to ensure they are meeting their own goals, and most improving education for thousands of students and families.

Although I encountered limitations with this optimization, such as access to certain private data and limitations in the computing power available to me, my preliminary findings are consistent with other math optimization approaches to increasing diversity in schools. My hope is that school districts, including CMS, will seriously prioritize school diversity and consider the value of systematic and optimizing approaches such as linear programming as they continue to work on their student assignment policies.

REFERENCES

Ayscue, J., Woodward, B., Kucsera, J., Siegel-Hawley, G., & Orfield, G. (2014). *Segregation Again: North Caolina's Transition from Leading Desegregation Then to Accepting Segregation Now* (6). The Civil Rights Project.

Bektaş, T., & Elmastaş, S. (2007). Solving School Bus Routing Problems through Integer Programming. *The Journal of the Operational Research Society*, *58*(12), 1599–1604.

Belford, P. C., & Ratliff, H. D. (1972). A Network-Flow Model for Racially Balancing Schools. *Operations Research*, *20*(3), 619–628.

Bischoff, K. (2008). School District Fragmentation and Racial Residential Segregation: How Do Boundaries Matter? *Urban Affairs Review*, *44*(2), 182–217.

Brown v. Board of Education of Topeka, 347US 483 (Supreme Court of the United States 1954).

Capacchione v. Charlotte-Mecklenburg Schools (W.D.N.C. 1999).

Caro, F., Shirabe, T., Guignard, M., & Weintraub, A. (2004). School Redistricting: Embedding GIS Tools with Integer Programming. *The Journal of the Operational Research Society*, *55*(8), 836–849.

Charlotte Mecklenburg Schools. (2015). Goals and Guiding Principles for Student Assignment. Retrieved March 12, 2017, from http://www.cms.k12.nc.us/boe/Pages/2010%20GuidingPrinciplesforStudentAssignment.aspx

Charlotte-Mecklenburg Board of Education. (2016). *2017-2018 Comprehensive Student Assignment Review: Student Assignment Survey Results* (pp. 1–87). Charlotte, NC.

Charlotte-Mecklenburg Schools Month 1, 2016-17 School Diversity Report. (2016). Charlotte-Mecklenburg Schools. Retrieved from http://www.cms.k12.nc.us/cmsdepartments/StudentPlacement/PlanningServices/Documents/PMR%20Month%2001bySchool_2016_Diversity-final.pdf

Clotfelter, C. T. (2004). *After "Brown": The Rise and Retreat of School Desegregation*. Princeton University Press. Retrieved from http://www.jstor.org/stable/j.ctt7sxjd

CMS Board of Education. (2016, April 26). Student assignment principles Approved 4-26-16.pdf. Retrieved May 2, 2016, from http://www.cms.k12.nc.us/cmsdepartments/StudentPlacement/PlanningServices/20172018StuAsgnReview/Documents/Student%20assignment%20principles%20Approved%204-26-16.pdf

Darling-Hammond, L. (2010). *The flat world and education: how America's commitment to equity will determine our future*. New York: Teachers College Press.

Douglas, D. M. (1995). *Reading, Writing & Race: The Desegregation of the Charlotte Schools*. Chapel Hill: University of North Carolina Press.

Fiel, J. E. (2013). Decomposing school resegregation: Social closure, racial imbalance, and racial isolation. *American Sociological Review*, *78*(5), 828–848. https://doi.org/10.1177/0003122413496252

Giersch, J., Bottia, M. C., Mickelson, R. A., & Stearns, E. (2016). Exposure to School and Classroom Racial Segregation in Charlotte-Mecklenburg High Schools and Students' College Achievement., *24*(32/33), 1–24.

Hallinan, M. T. (1998). Diversity Effects on Students Outcomes: Social Science Evidence. *Ohio State Law Journal*, *59*(3), 733–754.

Heckman, L. B., & Taylor, H. M. (1969). School Rezoning to Achieve Racial Balance: A Linear Programming Approach. *Socio-Economic Planning Sciences*, *3*, 127–133.

Holley-Walker, D. (2010). After Unitary Status: Examining Voluntary Integration Strategies for Southern School Districts. *North Carolina Law Review*, *88*(3), 877–910.

Jennergren, L. P., & Obel, B. (1980). A Study in the Use of Linear Programming for School Planning in Odense. *The Journal of the Operational Research Society*, *31*(9), 791–799. https://doi.org/10.2307/2581893

Keyes v. School District No. 1, 413US 189 (1973).

LastWeekTonight. (2016). *School Segregation: Last Week Tonight with John Oliver (HBO)*. Retrieved from https://www.youtube.com/watch?v=o8yiYCHMAlM

McNamara, J. F. (1971). Mathematical Programming Models in Educational Planning. *Review of Educational Research*, *41*(5), 419–446. https://doi.org/10.2307/1169885

Mickelson, R. A. (2008). Twenty-First Century Social Science on Racial Diversity and Educational Outcomes. *Ohio State Law Journal*, *69*(6), 1173–1227.

Mickelson, R. A., Smith, S. S., & Nelson, A. H. (Eds.). (2015). *Yesterday, today, and tomorrow: school desegregation and resegregation in Charlotte*. Cambridge, Massachusetts: Harvard Education Press.

NCES. (2012). Digest of Education Statistics, 2015 [IES - National Center for Education Statistics]. Retrieved April 11, 2017, from https://nces.ed.gov/programs/digest/d15/ch_2.asp

Orfield, G., & Lee, C. (2005). Why Segregation Matters: Poverty and Educational Inequality. *eScholarship*. Retrieved from http://escholarship.org/uc/item/4xr8z4wb

Parents Involved in Community schools v. Seattle School Distict No. 1 (2007). Retrieved from http://caselaw.findlaw.com/us-supreme-court/551/701.html

Reardon, S. F., & Yun, J. T. (2003). Integrating neighborhoods, segregating schools: the retreat from school desegregation in the South, 1990-2000. *North Carolina Law Review*, *81*(4), 1563–1596.

Richards, M. (2014). The Gerrymandering of School Attendance Zones and the Segregation of Public Schools: A Geospatial Analysis. *American Educational Research Journal*, *51*(6), 1119–1157.

Saporito, S., & Van Riper, D. (2016). Do Irregularly Shaped School Attendance Zones Contribute to Racial Segregation or Integration? *Social Currents*, *3*(1), 64–83. https://doi.org/10.1177/2329496515604637

Schofield, J. W. (1991). School Desegregation and Intergroup Relations: A Review of the Literature. *Review of Research in Education*, *17*, 335–409. https://doi.org/10.2307/1167335

Siegel-Hawley, G. (2013). Educational Gerrymandering? Race and Attendance Boundaries in a Demographically Changing Suburb. *Harvard Educational Review*, *83*(4), 580–612. https://doi.org/10.17763/haer.83.4.k385375245677131

Siegel-Hawley, G. (2014). Mitigating “Milliken”? School District Boundary Lines and Desegregation Policy in Four Southern Metropolitan Areas, 1990-2010. *American Journal of Education*, *120*(3), 391–433.

Siegel-Hawley, G., Bridges, K., & Shields, T. J. (2017). Solidifying Segregation or Promoting Diversity?: School Closure and Rezoning in an Urban District. *Educational Administration Quarterly*, *53*(1), 107–141. https://doi.org/10.1177/0013161X16659346

Smith, C. (2016, October 3). The Desegregation and Resegregation of Charlotte's Schools. Retrieved May 13, 2017, from http://www.newyorker.com/news/news-desk/the-desegregation-and-resegregation-of-charlottes-schools

Sohoni, D., & Saporito, S. (2009). Mapping School Segregation: Using GIS to Explore Racial Segregation between Schools and Their Corresponding Attendance Areas. *American Journal of Education*, *115*(4), 569–600. https://doi.org/10.1086/599782

Stroub, K. J., & Richards, M. P. (2013). From resegregation to reintegration: Trends in the racial/ethnic segregation of metropolitan public schools, 1993–2009. *American Educational Research Journal*, *50*(3), 497–531. https://doi.org/10.3102/0002831213478462

United States Census Bureau - American Fact Finder. (2015). *B01001 Population By Age* (2015 American Community Survey). Mecklenburg County, North Carolina.

Washington v. Seattle School District No. 1, 458US 457 (1982).

Wells, A. S., Cordova-Cobo, D., & Fox, L. (2016). *How Racially Diverse Schools and Classrooms Can Benefit All Students*. The Century Foundation. Retrieved from http://www.issuelab.org/resource/how_racially_diverse_schools_and_classrooms_can_benefit_all_students

You Were Wrong, Mr. President. (1984, October 9). *Charlotte Observer*.

APPENDIX A: INTRODUCTION TO LINEAR PROGRAMMING

I arrive at linear programming as the mathematical optimization method for the problem. Linear programming finds an optimal solution given restraints which are linearly defined. Given these linear constraints and costs, a polyhedron of feasible solutions is found, and the algorithm finds the optimal solution along the corners of this polyhedron.

Linear programming problems are difficult to visualize in multiple dimensions, so I provide a simple two-dimensional example to demonstrate how linear programming works. This example is from Chapter 2 of John W. Chinneck's *Practical Optimizations: a Gentle Introduction* Textbook.

If a bicycle company may want to maximize profit by best allocating what products they create. First, the variables must be defined. In this example, the company can produce mountain bikes and racers.

x_1 = the number of mountain bikes and.

x_2 = the number of racers produced.

Next, we define the constraints of this problem:

(1) The company cannot produce a negative number of mountain bikes or racers.

$$x_1 \geq 0,\ x_2 \geq 0$$

(2) The company can only produce a maximum of two mountain bikes in a day.

$$x_1 \leq 2$$

(3) The company can only produce a maximum of two racers in a day.

$$x_2 \leq 3$$

(4) A total of four bikes can be made in one day because of constraints the metal finishing machine needed to complete a bike.

$$x_1 + x_2 \leq 4$$

Finally, we maximize (or minimize) an objective function. This function is the cost we care about either keeping as low or as high as possible. In this case, the objective function maximizes daily profit:

$$\text{mazimize}(Z = 15x_1 + 10x_2)$$

The graph in *Figure 8* shows the constraints, graphed in red, with the feasible solutions region shaded in grey. Any point within this grey region would be a solution that satisfied the four constraints above. The algorithm checks for the optimal point by exploring the corners of this polyhedron, since optimal solutions are always found at the corners.

The green lines represent various levels of profit, Z. The profit function line where profit is fixed at 30 intersects at a solution of three racers and zero mountain bikes. This would result in a profit of $30 and satisfy all constraints of this problem, and is a corner of the polyhedron. This is not, however, the optimal solution, since other solutions can yield higher profits. The actual optimal point is found at a solution of two mountain bikes and two racers, with a maximum profit of $50. No other corner within the polyhedron of feasible solutions results in a larger profit than this solution.

A problem with more than two variables cannot be visualized on a two-dimensional coordinate plane. For the student assignment optimization defined in this paper, there are several million variables. Each variable is the decision to either assign or not assign a road segment to a school. These constraints creates a multiple dimensional polyhedron of feasible solutions. Therefore there are many more “corners to be explored” by a linear programming algorithm to find the optimal solution.

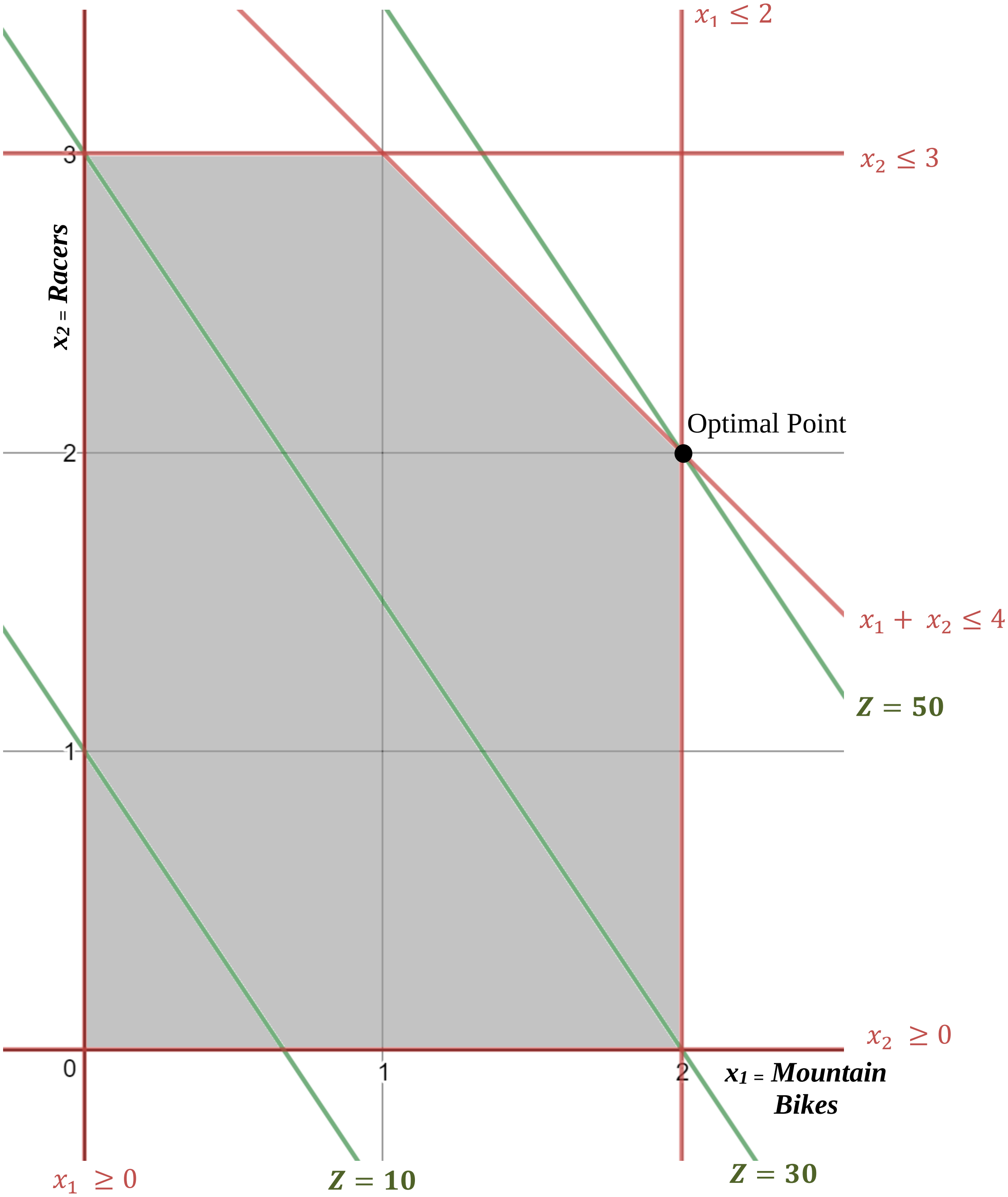

Figure 8. This graph depicts the constraints and feasible region of an example linear programming problem.

APPENDIX B: MATLAB CODE

```
% Initial Data Read
zone = 'blue';
schools = csvread(strcat('schools', zone, '.csv'));
roads = csvread(strcat('roads', zone, '.csv'));
commutes = csvread(strcat('commutes',zone,'.csv'));

roads_id = roads(:,1);
schools_id = schools(:,1);

% Formatting Data into Matrices-----------------------------------------

% E is binary matrix of roads and which SES category they fall in, R x 3
E = roads(:,4:6);
% Pr is the population of each road segment
Pr = roads(:,2);
% C is the school capacity for each school
C = (schools(:,5)) * 1.25;
%Pm is the proportions of each SES group in schools
Pm =schools(:,2:4);

%Dimensions scalars for number of roads, number of schools, and number of
%commutes
R = size(Pr,1);
S = size(C,1);
D = size(commutes,1);

% T is the transit times in a R x S matrix
% Pivot commutes into matrix
T = ones(R,S)*99999;
for i = 1:R
    for j = 1:S
        for k = 1:D
            if (commutes(k,1) == roads_id(i)) && (commutes(k,2) == schools_id(j));
                T(i,j) = commutes(k,3);
            end
        end
    end
end


%Use for moderate SES weight
Wmm = (((1/3)-Pm)/4)+1;
%Use for strong SES weight
Wms = ((1/3)-Pm)+1;
%Use for no SES weight - commute only
Wmn = ones(S,3);
Wused = 's';
Wm = strcat('Wm', Wused);

% Matrix Manipulations for later
Pr3 = repmat(Pr,1,3);
PrPrime = repmat(Pr,1,R);
%-----------------------------------------------------------------------


%A and b creation-------------------------------------------------------
A = zeros([S,(R*S)]);
```

```
for i = 1:S
    for j = 1:R
        A(i,((i-1)*R)+j) = Pr(j);
    end
end
b = C;
%---------------------------------------------------------------------

%Aeq and beq creation-------------------------------------------------
IR = eye(R);
Aeq = repmat(IR,1,S);
%beq
one = [1];
beq = repmat(one,[1,R]);
%---------------------------------------------------------------------


%upper and lower bound------------------------------------------------
x = 1:(R*S);
lb = zeros([R*S,1]);
ub = ones([R*S,1]);
%---------------------------------------------------------------------


%Objective Function ---------------------------------------------------
%W tiled so that rows along the diagonal are each schools 3 weight values
W = zeros((S*3),S);
for i = 1:S
    for j = 1:3
        W(((i-1)*3) + j, i) = Wm(i,j);
    end
end
% Etile of the E matrix repeated S times along the columns)
Etile = repmat(E,1,S);
Tthrees = reshape(repmat(reshape(T',[],1),1,3)',[],size(T,1))';
ET = Etile.*Tthrees;
ETW = ET*W;
PrrepS = repmat(Pr,1,S);
F = ETW.*PrrepS;
ObjF = zeros([1,R*S]);
for i = 1:S
    for j = 1:R
        ObjF(1,((i-1)*R) + j) = F(j,i);
    end
end
%---------------------------------------------------------------------


% Linear Programming Call---------------------------------------------
%assignments = intlinprog(ObjF,x, A,b, Aeq, beq,lb, ub);
%options = optimoptions('linprog','Algorithm','dual-simplex');
%[x,fval,exitflag,output] = linprog(ObjF,A,b,Aeq,beq,lb,ub,options);
assignments = linprog(ObjF, A, b, Aeq, beq, lb, ub);
%---------------------------------------------------------------------

%re-formatting assignments output-------------------------------------
ass_matrix_lp = zeros([R,S]);
for i = 1:S
    for j = 1:R
        ass_matrix_lp(j,i) = (assignments(((i-1)*R)+j));
    end
```

```
end

% find the maximum assignment of values
ass_matrix= zeros([R,S]);
% finds the minimum commute for a road segment
[TMin, TminInd] = min(T, [], 2);
for i = 1:S
    for j = 1:R
        if(ass_matrix_lp(j,i) > 0.5)
            ass_matrix(j,i) = 1;
        end
    end
end
assign_sum = sum(ass_matrix, 2);
for j = 1:R
    if assign_sum(j) == 0
        ass_matrix(j,TminInd(j)) = 1;
    end
end
id_ass_num = ass_matrix * schools_id;

csvwrite(strcat('output-id', zone, '-', Wused, '.csv'), id_ass_num);
%-----------------------------------------------------------------------


%Functions to calculate summary statistics on the assignment-----------
clear enrollments;
clear commute_avgs;
clear commute_avg;
clear SES_props;
clear high_pov_count;
%calculates the assigned enrollment to each school
enrollments = enrollments(ass_matrix, Pr, S);
%average commute time for each school
commute_avgs = commute_avgs(ass_matrix, T, Pr, S);
%total average commute time
commute_avg = commute_avg(ass_matrix, T, Pr, S);
SES_props = SES_props(ass_matrix, E, Pr, S);
%counts the number of schools that still have over 90% of their students
    %in the low SES catate
high_pov_count = high_pov_count(ass_matrix, E, Pr, S );
%-----------------------------------------------------------------------

Summary Statistics Functions:

function [enrollments_out] = enrollments( ass_matrix, Pr, S)
%enrollments gives a column vector of the enrollments by school for a given
%assignment
 enrollments_out = transpose(sum(repmat(Pr,1,S).*ass_matrix,1));
end

function [commute_avgs_out] = commute_avgs( ass_matrix, T, Pr, S)
%commute_avgs - calculates average commute time for those assigned to each
%school.
%   Detailed explanation goes here
R = size(T,1);
S = size(T,2);
Tfiltered = zeros(R,S);
for i = 1:R
    for j = 1:S
        if T(i,j) < 100
```

```
            Tfiltered(i,j) = T(i,j);
        end
    end
end

commute_tot = sum(ass_matrix.*(T.*(repmat(Pr,1,S))),1);
enrollmentsCol = sum(repmat(Pr,1,S).*ass_matrix,1);
commute_avgs_out = transpose(commute_tot./enrollmentsCol);


end

function [commute_avg] = commute_avg( ass_matrix, T, Pr, S)
%commute_avg - calculates average commute time for all students
%   Detailed explanation goes here



R = size(T,1);
S = size(T,2);
Tfiltered = zeros(R,S);
for i = 1:R
    for j = 1:S
        if T(i,j) < 100
            Tfiltered(i,j) = T(i,j);
        end
    end
end

commute_tot = sum(ass_matrix.*(T.*(repmat(Pr,1,S))),1);
enrollmentsCol = sum(repmat(Pr,1,S).*ass_matrix,1);
commute_avg = (sum(commute_tot./enrollmentsCol))/S;

end


function [SES_props_out, imbalance] = SES_props( ass_matrix, E, Pr, S)
%SES_props - reports a S X 3 matrix showing the proportion of L, M, H SES
%students assigned to each of S schools, as a proportion
PrRep = repmat(Pr,1,S);
ERep = repmat(E,1,S);
pop = ass_matrix.*PrRep;
poprep = repelem(pop,1,3);
pop_ses_ass = poprep.*ERep;
sum_pop_ses_ass = sum(pop_ses_ass,1);
enrollmentsCol = sum(repmat(Pr,1,S).*ass_matrix,1);
SES_props_out = zeros([S,3]);
for i = 1:S
    for k = 1:3
        SES_props_out(i,k)=(sum_pop_ses_ass(((i-1)*3)+k)/enrollmentsCol(i));
    end
%end
%imbalances = zeros(1, S);
%for i = 1:S
    %imbalances(i) = ((abs((1/3)-SES_props_out(i,1))+(abs((1/3)-
SES_props_out(i,2))+(abs((1/3)-SES_props_out(i,3)));
%end

%imbalance = mean2(imbalances);
end
```

```
function [high_pov_count_out] = high_pov_count(  ass_matrix, E, Pr, S )
%high_pov_count returns the number of schools with over 90% of children
%assigned to that school in the low SES category
PrRep = repmat(Pr,1,S);
ERep = repmat(E,1,S);
pop = ass_matrix.*PrRep;
poprep = repelem(pop,1,3);
pop_ses_ass = poprep.*ERep;
sum_pop_ses_ass = sum(pop_ses_ass,1);
enrollmentsCol = sum(repmat(Pr,1,S).*ass_matrix,1);
props = zeros([S,3]);
for i = 1:S
    for k = 1:3
        props(i,k)=(sum_pop_ses_ass(((i-1)*3)+k)/enrollmentsCol(i));
    end
end
high_pov_count_out = 0;
for i = 1:S
    % 0.9 indicates high poverty setting of 90% low SES
    if props(i,1) > 0.9
        high_pov_count_out = high_pov_count_out + 1;
    end
end
end
```