Socio-Economic Planning Sciences 64 (2018) 1–8

Contents lists available at ScienceDirect

Socio-Economic Planning Sciences

journal homepage: www.elsevier.com/locate/seps

Invited paper

Assigning students to schools to minimize both transportation costs and socioeconomic variation between schools

Elizabeth L. Bouzarth [a,*], Richard Forrester [b], Kevin R. Hutson [a], Lattie Reddoch [a]

[a] Department of Mathematics, Furman University, Greenville, SC 29613, United States
[b] Department of Mathematics and Computer Science, Dickinson College, Carlisle, PA 17013, United States

ARTICLE INFO

Article history:
Received 29 August 2016
Available online 6 September 2017

Keywords:
School districting
Assignment problem
Multiple criteria decision making
Integer programming
Deviation-based objective approach

ABSTRACT

Several studies have found that students' academic achievement is as much determined by the socioeconomic composition of their school as their own socioeconomic status. In this paper we provide a methodology for assigning students to schools so as to balance the socioeconomic compositions of the schools while taking into consideration the total travel distance. Our technique utilizes a biobjective general 0–1 fractional program that is linearized into a mixed 0–1 linear program that can be submitted directly to a standard optimization package. We show how a parametrized model could be utilized to provide a spectrum of different possible assignments so that a decision maker can decide how to balance socioeconomic factors with transportation costs. As a test case for our approach we analyze data from the Greenville County School District in Greenville, South Carolina.

1. Introduction

Does the socioeconomic composition of a student's school matter in their academic success? Several studies have suggested that it does. The Coleman Report [7], which examined 600,000 students in over 4,000 schools, concluded that "the social composition of the student body is more highly related to achievement, independent of the student's own social background, than is any school factor." Rumberger and Palardy [21] found that students attending the most affluent schools receive the greatest academic benefit, presumably benefiting from such process variables as higher parental involvement, greater teacher expectations, and more advanced (college prep) classes, among others. Mickelson and Bottia [18] found an inverse relationship between a school's socioeconomic makeup, as defined by the percentage of students eligible for free or reduced-price lunch, and outcomes in mathematics, irrespective of their age, race, or family's socioeconomic status. A Brooking's study [20] found that nationwide, the average low-income student attends a school that scores at the 42nd percentile on state exams, while the average middle/high-income student attends a school that scores at the 61st percentile on state exams.

Despite the identification of socioeconomic composition as a correlating factor between education and achievement, most public schools still fill their hallways by the attendance zones defined by the county or state officials. More often than not, these zones are primarily based on distance from the student's home to the particular school. While zoning based on distance typically provides convenience for families and lower transportation costs for the schools, it effectively zones by socioeconomic status since people of similar economic backgrounds form the neighborhoods that are zoned to particular schools.

There have been a number of different approaches taken to assigning students to schools so as to balance socioeconomic composition, including a lottery system, aggressive district efforts to integrate schools, and implementation of new zoning policies. The lottery system and redistricting efforts carry with them potentially high transportation costs as well as political consequences [19]. One of the most effective implementations of zoning policies to achieve socioeconomic integration can be seen in Montgomery County, Maryland. Montgomery County ranks among the wealthiest counties in the nation. Its zoning policy allows the public housing authority to purchase one-third of the zoning homes within each subdivision to operate as federally subsidized public housing. Public housing families are randomly assigned to public housing apartments, which prevents families from self-selecting neighborhoods and elementary schools of their choice.

* Corresponding author.
E-mail addresses: liz.bouzarth@furman.edu (E.L. Bouzarth), forrestr@dickinson.edu (R. Forrester), kevin.hutson@furman.edu (K.R. Hutson), lattie.reddoch@furman.edu (L. Reddoch).

http://dx.doi.org/10.1016/j.seps.2017.09.001

In a study of Montgomery County Schools [26], it was found that over a period of five years, children in public housing who attended the school district's most-advantaged schools far outperformed in math and reading those children in public housing who attended the district's least-advantaged elementary schools.

While the outcomes of these studies are encouraging, there are challenges associated with economic integration of schools. Since American housing is relatively segregated by income, assigning students to high schools to create socioeconomic uniformity would require students to travel greater distances. Not only could the associated transportation costs be prohibitive, but the greater travel distances could provide a barrier to active PTA and extra-curricular participation. As home prices and tax structures are often tied to school districts, there are political challenges that must be considered as well—parents who pay high property taxes are often very protective of their access to high-performing schools. Thus, school districts must seek to balance the benefits of socioeconomic uniformity with longer busing distances.

In this article, we develop a methodology to assign students to schools that minimizes socioeconomic variation between the schools while also minimizing the total busing distance. Our technique utilizes a biobjective general 0–1 fractional program, which is recast as a linear mixed 0–1 program. The linearized model has the advantage that it can be solved using standard commercial software. As a test case, we apply our technique to a school assignment problem based on data from the Greenville County School District in Greenville, South Carolina.

The remainder of this paper is organized as follows. The next section reviews the literature related to the use of optimization techniques for assigning students to schools. Then we introduce our school assignment model that minimizes both the socioeconomic variation between the schools and the total busing distance. The results of our test case are presented in the fourth section, where we apply our technique to a data set based on the Greenville County School District. The last section provides our concluding remarks.

2. Literature review

Modeling school assignments to provide racial balance [1,6,11,13,25,27] and low cost transportation [3,14,15,23,28] has a long history in the OR/MS literature. For a good review of the early studies see Church [5]. Malezowski and Jackson [16] provide a general framework for modeling and solving problems of efficient allocation of educational resources. They note that there are really two conflicting criteria at play when modeling these problems: resource efficiency and equity. Resource efficiency usually has involved efficient use of facilities, teachers, buses, etc. For example, objective functions and constraints have been imposed to ensure minimization of the maximum busing distance or total travel distance, to enforce school capacities, or to achieve certain student/teacher ratios. Equity has usually involved achieving equal opportunities for all students, and constraints are used to ensured a racial balance amongst schools or a balanced expenditure per student. Often though, these criteria are in conflict. Schoepfle and Church [25] exhibit an inverse relationship between the total weighted travel (busing) distance and the maximum allowable percent deviation from the community racial balance amongst optimal school assignments. Further, Malezowski and Jackson [16] point out that school administrations stress decisions based on operational efficiency while the public tends to concern itself with equity. Given the competing criteria of balancing socioeconomic factors and transportation costs, researchers have modeled this problem as a multi-criteria decision problem. Solution approaches have generally fallen into either the use of parametric programming or goal programming methods [16].

Traditionally, those studies using parametric programming have taken one of two approaches. Some researchers optimize a single-criteria objective function while constraining other criteria to fall within some specified acceptable range [5,6,8,9,11,12,15,17,24,25,27,28]. Other researchers assign a weight for each criterion function and convert the problem to a single objective function consisting of a combination of weighted objective functions [16]. The parametric programming approach is attractive because it gives the modeler the ability to vary parameters to provide a range of options to the decision maker.

Other authors have employed goal programming [2,13,23,29], which allows the decision maker to specify preferences with respect to solution evaluation criteria, such as travel distance and socioeconomic balance in our case, through establishing an aspiration level for each criteria. The program seeks solutions that are within acceptable ranges around these aspiration levels. The downside of this approach is that it requires criteria preferences to be stated a priori, but methods like the Analytical Hierarchy Process (AHP) [22] can be used to help with this as was demonstrated in Ref. [2].

Our approach to making school assignments utilizes a biobjective model, where the first objective is to minimize the total busing distance and the second objective is to minimize the socioeconomic variation between the schools. Specifically with regards to the second objective, we minimize the sum of the absolute deviations of the average median household income of the students assigned to each school from that of the entire school system. Thus, our model attempts to assign students to schools so that the average median household income is the same for all schools. To the best of our knowledge, this approach has not appeared in the literature before, possibly because of the challenges of formulating and solving a nonlinear model of this form. However, after we apply a novel linearization technique, our model can be submitted directly to a standard optimization package.

3. School assignment model

As is typical in school assignment models, rather than assign individual students to schools we make assignments based on larger subdivisions. We have chosen to use *block groups* as defined by the United States Census Bureau (other possible options include neighborhoods, tracts, etc.). Block groups are statistical divisions of census tracts, typically containing between 600 and 3,000 people, and usually cover a contiguous area. There is an extensive amount of public data on block groups that is collected at least every ten years, including socioeconomic data such as average household income, median household income, and median home value. To balance the socioeconomic backgrounds of the schools we chose to utilize the median household income (MHHI) for two reasons: the median is more resistant to outliers than the average and the median income provides a more accurate reflection of a household's wealth (home value is not necessarily as accurate because home values can fluctuate dramatically based on the real estate market).

The parameters of our model, which are defined in terms of the index sets BG = the set of block groups in the school district and S = the set of possible high schools, are as follows:

n_i = the number of students in block group $i \in BG$;
c_j = the student capacity of high school $j \in S$;
h_i = the MHHI in block group i;
d_{ij} = the distance from the center of block group i to high school j; and
$p = \frac{\sum_{i \in BG} n_i h_i}{\sum_{i \in BG} n_i}$ = average MHHI of the entire school system.

For each $i \in BG$ and $j \in S$ we define the binary variable x_{ij} as

$$x_{ij} = \begin{cases} 1 & \textit{if block group } i \textit{ is assigned to high school } j, \\ 0 & \textit{otherwise}, \end{cases}$$

and for each $j \in S$ we define the continuous "bookkeeping" variable w_j as

$$w_j = \textit{average MHHI of the block groups assigned to high school } j.$$

With these variables defined, the essential constraints of our model are as follows.

$$\sum_{j \in S} x_{ij} = 1 \quad \textit{for} \quad i \in BG \tag{1}$$

$$\sum_{i \in BG} n_i x_{ij} \le c_j \quad \textit{for} \quad j \in S \tag{2}$$

$$w_j = \frac{\sum_{i \in BG} n_i h_i x_{ij}}{\sum_{i \in BG} n_i x_{ij}} \text{ for } j \in S \tag{3}$$

$$x_{ij} \in \{0,1\} \text{ for } i \in BG, j \in S \tag{4}$$

Constraints (1) ensure that each block group i is assigned to exactly one high school, while constraints (2) ensure that block group assignments do not exceed the capacity of high school j. Constraints (3) enforce the definition of our w_j variables, that is, they ensure that w_j equals the average MHHI of the block groups assigned to high school j.

Our model is a biobjective program in which the first objective is to minimize the total busing distance, which is computed as

$$f_1(\mathbf{x}) = \sum_{i \in BG} \sum_{j \in S} d_{ij} x_{ij}.$$

The second objective of our model is to minimize the socio-economic variation between the schools, which we accomplish using the function

$$f_2(\mathbf{x}) = \sum_{j \in S} |w_j - p|.$$

Note that $f_2(\mathbf{x})$ is not written explicitly in terms of x_{ij}, but implicitly in terms of x_{ij} by constraints (3). Here, $|w_j - p|$ quantifies the absolute deviation of the average median household income of the students assigned to high school j from that of the entire school system. Thus, by minimizing $f_2(\mathbf{x})$ we are attempting to balance the socioeconomic backgrounds of the schools. Utilizing $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$, our biobjective program takes the following form.

$$\text{FAP : minimize} \quad \{f_1(\mathbf{x}), f_2(\mathbf{x}) : (1) - (4)\}$$

Unfortunately, the nonlinearities of FAP prevent us from directly utilizing standard optimization software. However, we can transform FAP into an equivalent *linear* biobjective program that can be solved using a standard optimizer. To linearize the absolute value terms in $f_2(\mathbf{x})$ we perform the substitution

$$|w_j - p| = w_j^+ + w_j^-,$$

where w_j^+ and w_j^- are nonnegative continuous variables, and then add the following restrictions.

$$w_j - p = w_j^+ - w_j^- \quad \textit{for} \quad j \in S \tag{5}$$

$$w_j^+ \le M b_j \quad \textit{for} \quad j \in S \tag{6}$$

$$w_j^- \le M(1 - b_j) \quad \textit{for} \quad j \in S \tag{7}$$

$$w_j^+, w_j^- \ge 0 \quad \textit{for} \quad j \in S \tag{8}$$

$$b_j \in \{0,1\} \quad \textit{for} \quad j \in S \tag{9}$$

Note that M is a large number and that b_j is a binary indicator variable, which enforces that at least one of the values w_j^+ and w_j^- are zero by constraints (6)–(8). This in turn ensures that $|w_j - p| = w_j^+ + w_j^-$ since $w_j - p = w_j^+$ when $w_j - p \ge 0$, and $w_j - p = -w_j^-$ when $w_j - p \le 0$. We can now rewrite Problem FAP as

$$\text{FAP : minimize} \quad \left\{f_1(\mathbf{x}), \bar{f}_2(\mathbf{x}) : (1) - (9)\right\},$$

where

$$\bar{f}_2(\mathbf{x}) = \sum_{j \in S} \left(w_j^+ + w_j^-\right).$$

To linearize the fractional constraints (3), we first rearrange (3) as

$$\sum_{i \in BG} n_i x_{ij} w_j = \sum_{i \in BG} n_i h_i x_{ij} \text{ for } j \in S. \tag{10}$$

Now we can replace each mixed 0–1 term $x_{ij} w_j$ in (10) with a continuous variable y_{ij} and then add in additional restrictions to ensure that $y_{ij} = x_{ij} w_j$ for any binary value x_{ij}. Specifically, we can replace (10) with the following restrictions, where U_j is an upper-bound on w_j.

$$\sum_{i \in BG} n_i y_{ij} = \sum_{i \in BG} n_i h_i x_{ij} \text{ for } j \in S \tag{11}$$

$$y_{ij} \le U_j x_{ij} \quad \textit{for} \quad i \in BG, j \in S \tag{12}$$

$$y_{ij} \le w_j \quad \textit{for} \quad i \in BG, j \in S \tag{13}$$

$$y_{ij} \ge w_j - U_j(1 - x_{ij}) \quad \textit{for} \quad i \in BG, j \in S \tag{14}$$

$$y_{ij} \ge 0 \quad \textit{for} \quad i \in BG, j \in S \tag{15}$$

Note that when $x_{ij} = 0$ then (12) and (15) force $y_{ij} = 0$, while (13) and (14) are redundant. When $x_{ij} = 1$ then (13) and (14) force $y_{ij} = w_j$, while (12) and (15) are redundant. Thus, (12)–(15) ensure that $y_{ij} = x_{ij} w_j$ for any binary x_{ij}. We computed the upper bounds U_j on w_j as follows.

$$U_j = \frac{\max\left\{\sum_{i \in BG} n_i h_i x_{ij} : (1), (2), x_{ij} \in \{0,1\} \text{ for } i \in BG\right\}}{\min_{i \in BG}\{n_i\}}$$

Thus, Problem FAP is equivalent to the following linear biobjective mixed 0–1 program.

$$\text{LFAP : minimize} \quad \left\{f_1(\mathbf{x}), \bar{f}_2(\mathbf{x}) : (1) - (9), (11) - (15)\right\}$$

Our goal is to simultaneously minimize $f_1(\mathbf{x})$ and $\bar{f}_2(\mathbf{x})$ subject to (1)–(9), (11)–(15), which we will accomplish using the ε-constraint method introduced in Ref. [10] and later formalized in Ref. [4]. In this approach, the decision maker chooses one of the two objectives to minimize, while constraining the other objective to be less than or equal to a given target value. We chose to utilize this well-known technique because it is easy to implement and is well-suited for discrete biobjective optimization problems since, depending on the choice of ε, it can capture any solution along the Pareto frontier (as opposed to other classical techniques, such as weighted-sums, which can fail to recover portions of the Pareto frontier due to non-convexities). Our specific implementation utilizes Utopia and Nadir values within the actual ε-constraint. The Utopia point $\mathbf{z}^U = (z_1^U, z_2^U)$ is obtained by minimizing each of the objective functions subject to the model constraints, that is,

$$z_1^U = f_1(\mathbf{x}_1^*) \quad where \quad \mathbf{x}_1^* = argmin_x\{f_1(\mathbf{x}) : (1)-(9),(11)-(15)\} \quad and$$
$$z_2^U = \bar{f}_2(\mathbf{x}_2^*) \quad where \quad \mathbf{x}_2^* = argmin_x\left\{\bar{f}_2(\mathbf{x}) : (1)-(9),(11)-(15)\right\}.$$

The Nadir point $\mathbf{z}^N = (z_1^N, z_2^N)$ provides upper bounds of the objective functions in the Pareto set, where

$$z_1^N = max \quad \{f_1(\mathbf{x}_1^*), f_1(\mathbf{x}_2^*)\} = f_1(\mathbf{x}_2^*) \quad and$$
$$z_2^N = max \quad \left\{\bar{f}_2(\mathbf{x}_1^*), \bar{f}_2(\mathbf{x}_2^*)\right\} = \bar{f}_2(\mathbf{x}_1^*).$$

In our application of the ε-constraint method, we chose to minimize the distance objective $f_1(\mathbf{x})$ while constraining the socioeconomic balancing objective $\bar{f}_2(\mathbf{x})$, which allows us to directly control how much of the optimal value of $\bar{f}_2(\mathbf{x})$ we are willing to sacrifice to improve $f_1(\mathbf{x})$ (if the decision maker was interested in controlling $f_1(\mathbf{x})$ instead we could simply reverse the functions). This gives rise to the following model.

$$\text{AP}: \quad \text{minimize} \quad f_1(\mathbf{x}) \quad \text{subject to}$$

$$\bar{f}_2(\mathbf{x}) \le z_2^U + \varepsilon\left(z_2^N - z_2^U\right) \tag{16}$$

$$(1)-(9),(11)-(15)$$

Note that AP is a single-objective mixed 0–1 linear program that can be directly submitted to a standard optimization solver for any specific value of ε. To find Pareto optimal solutions, we solve AP over a spectrum of values of ε, where $0 \le \varepsilon \le 1$. As ε is incrementally decreased from 1 to 0, the ε-constraint (16) restricts $\bar{f}_2(\mathbf{x})$ to be between z_2^N and z_2^U, which is the interval of variation of $\bar{f}_2(\mathbf{x})$ over the Pareto optimal set.

4. Results of case study

To demonstrate the usefulness of our approach, we tested it with data from the Greenville County School District in Greenville, South Carolina. This particular district is the 46th largest school district in the United States in terms of numbers of students (over 72,000 in 2015) and covers over 800 square miles. This mixture of population size and district area makes it an appropriate school district to test our approach. Furthermore, Furman University is located in Greenville County, so we had the convenience of location and the knowledge of parents and students who attend the Greenville County public schools.

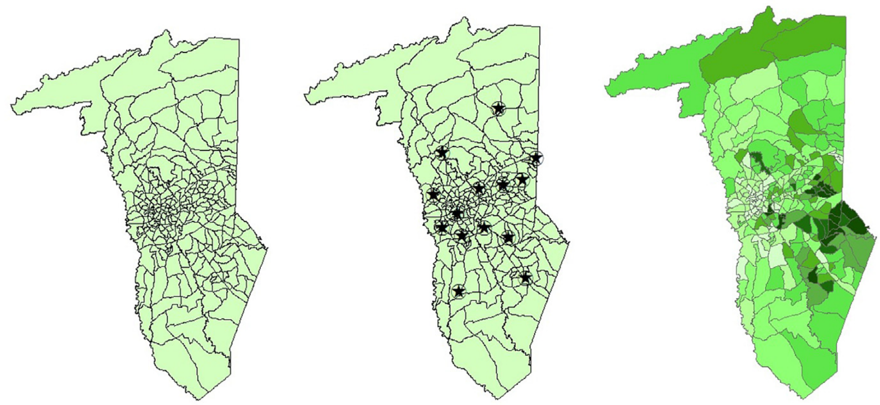
As mentioned in the previous section, we utilize block groups to divide Greenville County into smaller units for assigning students to the 14 existing public high schools. Using Business Analyst, a GIS program with built-in census data, we found population and socioeconomic data from 2013 for each of the 255 block groups in Greenville County. We note that the block groups generally had anywhere from 0 to 200 high school students per block group. Fig. 1 shows Greenville County partitioned into the 255 block groups and the location of the 14 high schools within the county. In addition, the figure shows the MHHI level for each block group where the darker colors represent groups with higher MHHI levels.

Interestingly, the total number of high school aged students in Greenville County is greater than the total capacity of the 14 public high schools because of private schools, home schooling, dropouts, and other circumstances that keep students out of public schools. Therefore, we had to scale our population per block group to ensure feasibility of the high school capacity constraints. Based on the most recent population data released from Greenville County public schools, the current number of high school students in the public school system (20,435 students as of the 15th day of school in the 2015–2016 school year) can be reached by scaling the 14–17 age group in 2013 by 86.7%. Distances from each block group to the different public high schools were determined using ArcMap. Specifically, we created a centroid for each of the 255 block groups, and then calculated the straight-line distance from each centroid to each high school location.

We formulated the optimization model AP with $\varepsilon = 0$ to $\varepsilon = 1$ in increments of 0.1 for the Greenville School District data using the Mosel algebraic modeling language. Each problem instance was optimized with a time limit of 30 min using the Xpress solver 8.0 on a Dell Precision T5610 workstation equipped with two 2.60 GHz Intel Xeon Processors and 32GB RAM running Windows 7. When solving for the Utopia value z_2^U associated with the socioeconomic objective $\bar{f}_2(\mathbf{x})$ we instructed the solver to record all the integer solutions found during the optimization process. These integer solutions were then loaded into Xpress prior to solving model AP so as to provide a warm start. We found that this approach dramatically improved the optimization process, especially for smaller values of ε as the tightness of the ε-constraint (16) made it challenging for the solver to find integer solutions that satisfied the constraint. We also mention that we utilized a 60 min time limit when solving for z_2^U, and thus z_2^U is an *approximate* Utopia value (solving for z_1^U was essentially trivial for Xpress).

The results of model AP for the eleven different values of ε are listed in the table of Fig. 2. The first column indicates the value of ε, while the second and third columns list the percentages that the objective values z^* of the assignments are above the Utopia values z_i^U, computed as $(z^* - z_i^U)/z_i^U$, for the objectives $f_1(\mathbf{x})$ and $\bar{f}_2(\mathbf{x})$, respectively.

One of the main observations that can be made from Fig. 2 is that our two objectives, minimizing total busing distance and balancing the socioeconomic compositions of the schools, are in direct conflict with each other. When $\varepsilon = 0$ the ε-constraint (16) requires that the solution to AP has $\bar{f}_2(\mathbf{x})$ achieve its Utopia value, that is, the model focuses exclusively on the socioeconomic



(a) Census block groups (b) High school locations (c) MMHI levels (darker colors represent higher MHHI)

Fig. 1. Greenville County block groups and high schools.

ε	Percent Above Distance Utopia	Percent Above Socioeconomic Utopia
0.0	274.4%	0%
0.1	214.4%	8.0%
0.2	206.1%	12.6%
0.3	171.5%	91.7%
0.4	132.4%	110.9%
0.5	111.0%	189.2%
0.6	87.6%	247.0%
0.7	27.4%	314.1%
0.8	20.9%	343.1%
0.9	1.6%	382.1%
1.0	0%	449.0%

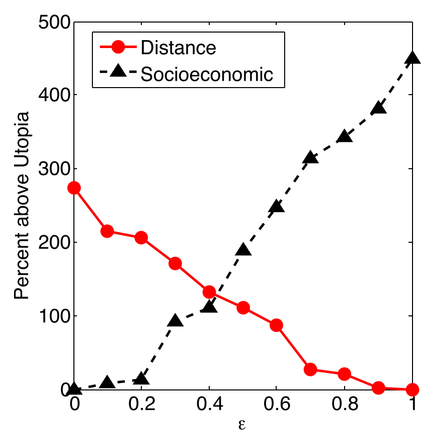
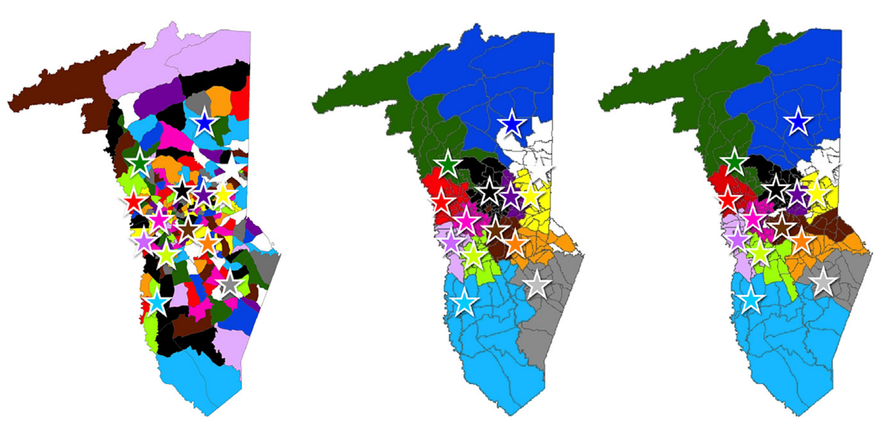


Fig. 2. Percentages above Utopia.



(a) $\varepsilon = 0$ (b) $\varepsilon = 1.0$ (c) Current assignment (est.)

Fig. 3. Assignments corresponding to Utopia points.

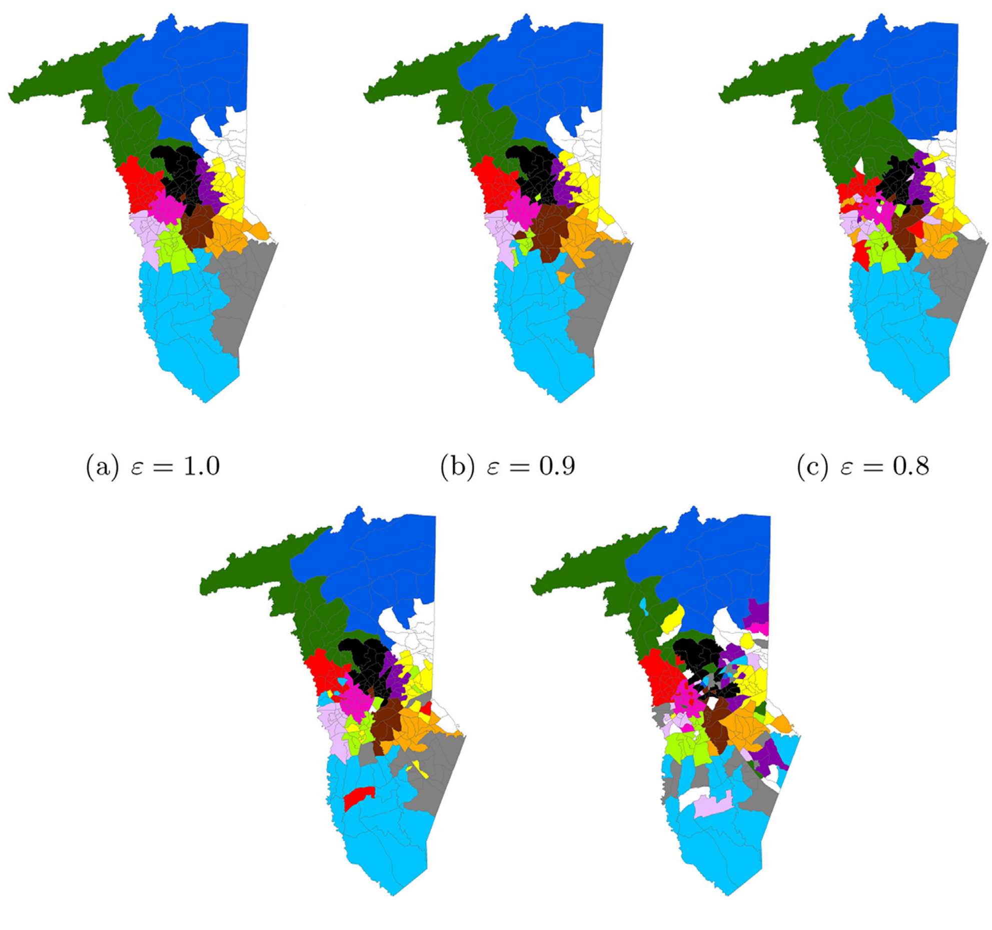
balancing criteria. As ε increases the model focuses more on the distance criteria with that being the sole focus once $\varepsilon = 1$. Note that when $\varepsilon = 0$ the percentage above the busing distance Utopia is 274.4% and then decreases as ε approaches 1. The reverse is true for the socioeconomic Utopia in that the percentage above the Utopia point is 449.0% when $\varepsilon = 1$ and decreases as ε decreases.

Fig. 3 shows the assignments of block groups to high schools when $\varepsilon = 0$ and $\varepsilon = 1$. The actual high school locations have been placed in each figure and are indicated by a star colored to match the block groups assigned to it. Greenville, SC and its suburbs are located in the center of the county and 10 of the 14 high schools are located in this heavily-populated area. As expected, when $\varepsilon = 1$ the block groups have been assigned high schools within a small proximity around the high school, especially in the populated areas. When $\varepsilon = 0$, there is a much wider distribution of assignments of block groups to high schools, reflecting that the county has a large income disparity among the block groups. Many of the block groups in the suburbs just north and east of downtown Greenville have income levels in the upper quartile of the county while block groups to the south and extreme northern regions of the county have income levels in the lower quartile of the county. Fig. 3c shows a rough estimate of the current assignment of block groups to high schools. This is an estimate since the actual assignment uses a finer partition of Greenville County than block groups. However, as the figure shows, the current assignment is close to the assignment when $\varepsilon = 1$.

Figs. 4 and 5 show the evolution of the assignments as ε is decreased by 0.1 starting from $\varepsilon = 1$. While there is little variation between the graphs for $\varepsilon = 1$ and $\varepsilon = 0.9$, we see that as ε is decremented from 0.9 to 0.7 many of the block groups that were assigned to schools in areas of high MHHI (for example those colored white, yellow, and black in Fig. 3b) get assigned to schools in regions associated with lower MHHI (light green, light purple, grey, and red in Fig. 3b). Further, in areas of lower MHHI (for example light green and red in Fig. 3b) get assigned to schools in areas with higher MHHI (white, yellow, orange, dark purple, and black in Fig. 3b). As ε continues to decrease, this phenomenon becomes more and more prevalent so that the high school assignments are no longer based on proximity but rather block groups must travel greater distances to attend the assigned high school.

Fig. 6 presents the average MHHI and total busing distance as ε varies. Notice in Fig. 6a the total busing distance decreases as ε increases, signifying the increased importance of minimizing distance. Additionally, Fig. 6b shows the average MHHI of the block groups assigned to each school (w_j from equation (3)). Notice that for $\varepsilon = 1$, the range of extreme average MHHI values is approximately \$50,000, but as ε decreases to 0, the range is approximately \$10,000.

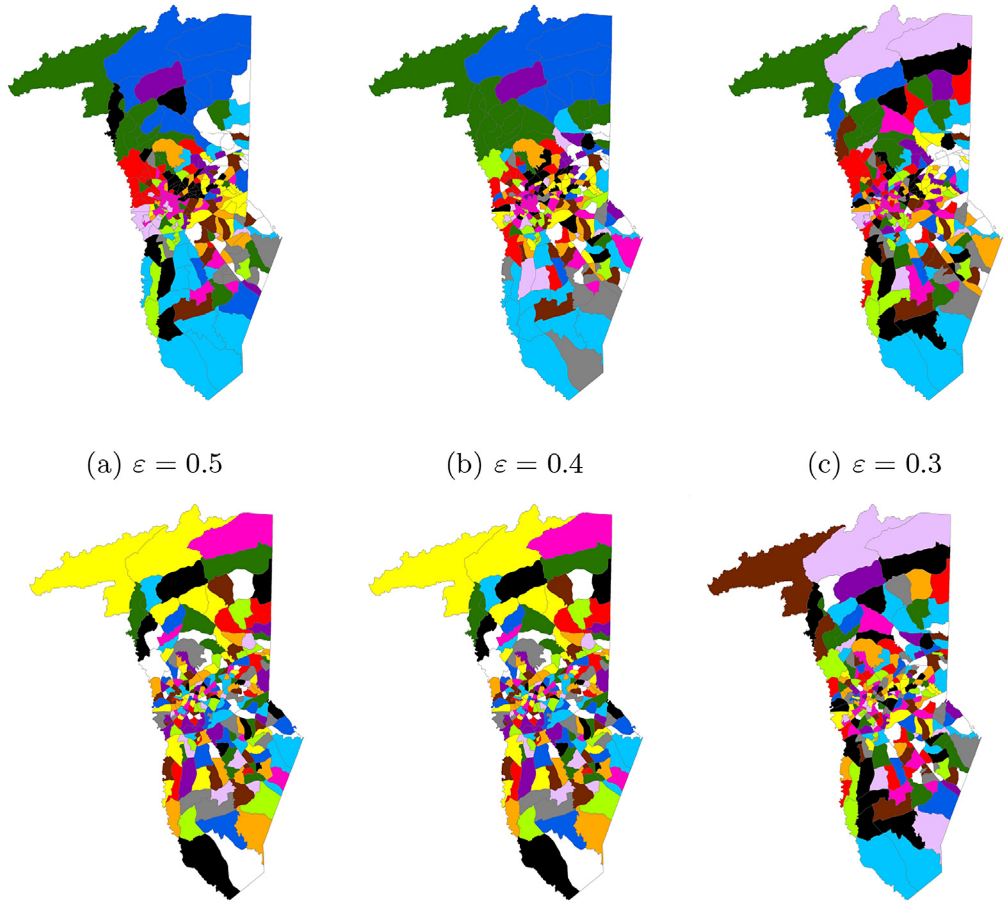
In summary, the assignments generated for the eleven different values of ε provide a variety of options for the Greenville County School District to consider when assigning students to schools. Thus, the district could increase the socioeconomic diversity of the schools while considering the tradeoff with the increased travel distance. There are a number of assignments that seem to provide a



(a) $\varepsilon = 1.0$ (b) $\varepsilon = 0.9$ (c) $\varepsilon = 0.8$

(d) $\varepsilon = 0.7$ (e) $\varepsilon = 0.6$

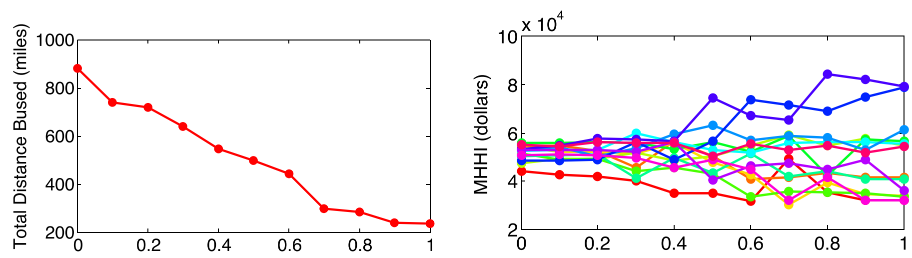
Fig. 4. High school assignments for $\varepsilon = 1, 0.9, \ldots, 0.6$.



(a) $\varepsilon = 0.5$ (b) $\varepsilon = 0.4$ (c) $\varepsilon = 0.3$

(d) $\varepsilon = 0.2$ (e) $\varepsilon = 0.1$ (f) $\varepsilon = 0$

Fig. 5. High school assignments for $\varepsilon = 0.5, 0.4, \ldots, 0$.



(a) Total distance bused (b) Average MHHI (w_j) of high schools

Fig. 6. Total distance bused and average MHHI (per high school) versus ε.

nice balance between total busing distance and balancing socioeconomic compositions of the schools where the solutions are not too far above their Utopia values, such as $\varepsilon = 0.4$. Ultimately, the school district would need to consider the tradeoffs between travel distance and balancing socioeconomic compositions along with the political ramifications involved.

5. Conclusion

This article has described a model that can be used to assign students to high schools so as to balance the socioeconomic differences between the schools while minimizing the total busing distance. The motivation for our study comes from the large body of research that has shown that providing students from lower socioeconomic classes an opportunity to attend a school that has a mix of socioeconomic participants can greatly improve their academic achievement. Our model is a biobjective general fractional program that is linearized into a mixed 0–1 linear program that can be submitted to any optimization software package. To the best of our knowledge, attempting to assign students so that the schools have the same average MHHI has not been utilized in the literature before. As a test case, we applied our methodology to data from the Greenville County School District in South Carolina. We showed how the ε-constraint method could be utilized to provide a spectrum of different possible assignments so that a decision maker could make the best choice between balancing socioeconomic composition and total busing distance.

Future research could focus on the inclusion of additional constraints such as applying a limit on worst-case busing distance for any one block group, taking into account capacity by grade, and prevention of boundaries that cross geographic obstacles such as railroads, rivers, or streets with heavy traffic, etc. Additionally, the use of census blocks instead of block groups is desirable if enough population data is published, which would refine the optimization framework. From a user perspective, various districting authorities may want to approach their school zones different ways. In Greenville County, for example, the high schools are concentrated in the central region of the county. The county could use a version of the model that weights socioeconomic status heavily in the middle of the county, since everything is much closer together and busing distance will be less costly in such a dense region. However, in the northern and southern ends of the county, a distance-weighted model would be more practical because population density is much lower and students have to travel much farther to get to their respective schools. Creativity in this way (e.g., using the model appropriately for differently-shaped counties), is feasible and would be a very practical application.

Acknowledgments

The authors thank Mike Winiski of Furman University for his assistance in utilizing GIS tools and software for data collection. Additionally, Bouzarth, Hutson, and Reddoch thank the Furman Advantage Research Fellowship Program and the Furman University Department of Mathematics for providing summer research funding.

References

[1] Belford PC, Ratliff D. A network-flow model for racially balancing schools. Oper Res 1972;20:619–28.
[2] Brown PJ. Trade-off and compromise in secondary school reorganisation schemes: a goal programming approach. Pap Reg Sci Assoc 1987;61:115–30.
[3] Caro F, Shirabe T, Guignard M, Weintraub A. School redistricting: embedding GIS tools with integer programming. J Oper Res Soc 2004;55:836–49.
[4] Chankong V, Haimes Y. Multiobjective decision making: theory and methodology. New York: Elsevier Science Publishing; 1983.
[5] Church RL, Murray AT. Modeling school utilization and consolidation. J Urban Plan Dev 1993;119:23–38.
[6] Clarke S, Surkis J. An operations research approach to racial desegregation of school systems. Soc Econ Plan Sci 1968;1:259–72.
[7] Coleman JC. Equality of educational opportunity. US Dep Health Educ Welf 1966;22.
[8] Diamond JT, Wright JR. Multiobjective analysis of public school consolidation. J Urban Plan Dev ASCE 1987;113:1–18.
[9] Franklin AD, Koenigsberg E. Computed school assignments in a large district. Oper Res 1973;21:413–26.
[10] Haimes Y, Lasdon L, Wismer D. On a bicriterion formulation of the problems of integrated system identification and system optimization. IEEE Trans Syst Man Cybern 1971;1:296–7.
[11] Heckman LB, Taylor MM. School rezoning to achieve racial balance: a linear programming approach. Soc Econ Plan Sci 1969;3:127–33.
[12] Kahlenberg RD. From all walks of life: new hope for school integration. Am Educ 2012:1–14.
[13] Knotson DL, Marquis LM, Ricchiute DM, Saunders GJ. A goal programming model for achieving racial balance in public schools. Soc Econ Plan Sci 1980;14:109–16.
[14] Koenigsberg E. Mathematical analysis applied to schoolattendance areas. Soc Econ Plan Sci 1968;1:465–75.
[15] Lee SM, Moore LJ. Multi-criteria school bussing models. Manag Sci 1977;23:703–15.
[16] Malczewski J, Jackson M. Multicriteria spatial allocation of educational resources: an overview. Soc Econ Plan Sci 2000;34:219–35.
[17] McKeown P, Workman B. A study in using linear programming to assign students to schools. Interfaces 1976;6:96–101.
[18] Mickelson RA, Bottia M. Integrating education and mathematics outcomes: a synthesis of social science research. N C Law Rev 2010;88(3):993–1089.
[19] Rotherham A. Does income-based school integration work? Time October 2010. http://content.time.com/time/nation/article/0,8599,2027858,00.html.
[20] Rothwell J. Housing costs, zoning, and access to high-scoring schools. Brookings Metropolitan Policy Program Report. 2012. p. 1–30.
[21] Rumberger RW, Palardy GJ. Does segregation still matter? the impact of student composition on academic achievement in high school. Teach Coll Rec 2005;107(9):993–1089.
[22] Saaty TL. The analytic Hierarchy process. New York: McGraw-Hill; 1980.
[23] Saunders G. An application of goal programming to the desegregation bussing problem. Soc Econ Plan Sci 1981;15:291–3.
[24] Schoepfle OB. The modeling of schools under declining and fluctuating enrollments. PhD Thesis. University of California at Santa Barbara; 1988.
[25] Schoepfle OB, Church RL. A fast, network-based, hybrid heuristic for the assignment of students to schools. J Oper Res Soc 1989;40:1029–40.
[26] Schwartz H. Housing policy is school policy: economically integrative housing promotes academic success in Montgomery county, Maryland. New York: Century Foundation; 2010.
[27] Silverstein A. The application of multi-objective linear programming to the school desegregation problem. Master's thesis. The John's Hopkins University; 1979.
[28] Stimson DH, Thompson RP. Linear programming bussing and education administration. Soc Econ Plan Sci 1974;8:195–200.
[29] Sutcliffe C, Board J, Cheshire P. Goal programming and allocating children to secondary schools in reading. J Oper Res Soc 1984;35:719–30.

Elizabeth Bouzarth is an Associate Professor of Mathematics at Furman University. She earned a bachelor's degree in Mathematics and Physics from Dickinson College in 2003 and a Ph.D. in Mathematics from the University of North Carolina at Chapel Hill in 2008. As an applied mathematician, her research interests vary widely and include computational fluid dynamics, optimization problems with community applications, and sports analytics.

Richard Forrester is a Professor of Mathematics at Dickinson College. He received his Ph.D. in Mathematical Sciences in 2002 from Clemson University. His scholarship is at the interface of Operations Research and Computer Science, and has typically centered around techniques for solving nonlinear 0-1 programs. He has published in a number of journals, including *Discrete Optimization*, *Naval Research Logistics*, *Operations Research Letters*, and *Optimization*.

Kevin Hutson is a Professor of Mathematics at Furman University. He earned his Ph.D. in Mathematical Sciences from Clemson University in 2002. His research interests are in network optimization, ranking algorithms, and sports analytics. He has published in a variety of journals, including *Operations Research*, *Annals of Operations Research*, *Networks*, and *Computers & OR*. He also consults in the area of sports analytics to the NCAA and to ESPN.

Lattie Reddoch started this research as an undergraduate student at Furman University. She graduated in May 2017 with a bachelor's degree in Applied Mathematics and Piano Performance. She is a Furman Fellow for the class of 2017 based on her commitment to the Furman community, including her study and application of operations research in 2015 and her internship with the Department of Defense in the summer of 2016. She is pursuing a mathematics career in the intelligence community.