R E P O R T   R E S U M E S

ED 018 859 EA 001 235

A MODEL FOR THE DETERMINATION OF SCHOOL ATTENDANCE AREAS UNDER SPECIFIED OBJECTIVES AND CONSTRAINTS.

BY- O'BRIEN, RICHARD J.
NATIONAL CENTER FOR EDUCATIONAL STATISTICS (DHEW)

REPORT NUMBER TN-55 PUB DATE 22 JAN 68
EDRS PRICE MF-$0.25 HC-$0.76 17P.

DESCRIPTORS- *URBAN SCHOOLS, *EDUCATIONAL POLICY, *SCHOOL LOCATION, SCHOOL DEMOGRAPHY, *RACIAL COMPOSITION, *STUDENT TRANSPORTATION, METHODOLOGY, SOCIAL DISTRIBUTION, EDUCATIONAL FACILITIES, GEOGRAPHIC DISTRIBUTION, COSTS, SCHOOL SIZE, COMPUTER ORIENTED PROGRAMS, EDUCATIONAL PARKS, BUS TRANSPORTATION, STUDENT ENROLLMENT, SCHOOL SYSTEMS, DISTRICT OF COLUMBIA, URBAN EDUC. MODEL,

THIS TECHNICAL NOTE, ONE OF A SERIES PUBLISHED ON THE URBAN EDUCATION MODEL, PRESENTS A MODEL FOR DETERMINING REQUIRED SCHOOL ATTENDANCE AREAS WHEN RESTRICTIONS HAVE BEEN PLACED ON THE RACIAL AND/OR SOCIAL COMPOSITION OF EACH SCHOOL PLANT. THESE ATTENDANCE AREAS ARE GENERATED IN A MANNER INSURING THE ASSIGNMENT OF STUDENTS WHICH MINIMIZES THE TOTAL "DISTANCE" TRAVELED BY ALL STUDENTS. THE METHODOLOGY ALLOWS FOR SYSTEMATIC STUDY OF THE RELATIONSHIP BETWEEN SCHOOL LOCATION DECISIONS, RACIAL AND SOCIAL COMPOSITIONS OF SCHOOLS, AND OBJECTIVES SUCH AS THE MINIMIZATION OF TOTAL STUDENT TRAVEL TIME. INPUTS REQUIRED TO CONDUCT SUCH ANALYSES INCLUDE THE EXISTING DISTRIBUTION OF SCHOOL PLANTS, THE LOCATION OF PROPOSED PLANTS, AND THE GEOGRAPHIC DISTRIBUTION OF STUDENTS DEFINED BY THEIR RACIAL, SOCIAL, AND AGE CHARACTERISTICS. THE ANALYSIS CONSISTS OF (1) THE DEFINITION OF THE PROPOSED AND EXISTING SCHOOL PLANT(S) BY LOCATION, AGE GROUP, AND CAPACITY, (2) DATA THAT DEFINE THE STUDENT POPULATION CROSS-CLASSIFIED ACCORDING TO AREAL UNIT, RACIAL GROUP, SOCIAL CLASS, AND AGE GROUP, (3) PROPORTIONS OF RACIAL AND SOCIAL MIXES FOR THE INDIVIDUAL SCHOOL PLANTS, (4) AN OBJECTIVE FUNCTION SUCH AS MINIMIZING THE TOTAL DISTANCE TRAVELED BY ALL STUDENTS, AND (5) SOLVING THE MODEL TO YIELD THE ASSIGNMENT OF STUDENTS TO SCHOOLS. EA 001 236 IS A RELATED DOCUMENT. (HW)

ED018859

NATIONAL CENTER FOR EDUCATIONAL STATISTICS
Division of Operations Analysis

A MODEL FOR THE DETERMINATION OF SCHOOL ATTENDANCE AREAS UNDER SPECIFIED OBJECTIVES AND CONSTRAINTS

by

Richard J. O'Brien

TECHNICAL NOTE
NUMBER 55

January 22, 1968

EA 001 235

OFFICE OF EDUCATION/U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
John W. Gardner, Secretary

OFFICE OF EDUCATION
Harold Howe II, Commissioner

NATIONAL CENTER FOR EDUCATIONAL STATISTICS
Francis C. Nassetta, Acting Assistant Commissioner

DIVISION OF OPERATIONS ANALYSIS
David S. Stoller, Director

A MODEL FOR THE DETERMINATION OF SCHOOL ATTENDANCE AREAS UNDER SPECIFIED OBJECTIVES AND CONSTRAINTS

INTRODUCTION

One of the factors under consideration in the determination of the location and size of schools is the racial and/or social composition of the school attendance area that results from such a decision. One of the advantages advanced for the educational park concept, for example, is the degree of racial integration that can be achieved by the large school system when it replaces existing school systems that are largely of one racial group.

The methodology present in this paper allows for the systematic study of the relationship between school location decisions, racial and social composition of school attendance areas and objectives such as the minimization of total student transportation time or cost or distance.

The model defines the location of the students by geographical units such as census tracts and the region containing the school system is defined as the aggregate of these geographical units. It thus allows for the investigation of regions on the school system level (as approximated by the aggregate of census tracts for example), central city level or SMSA (Standard Metropolitan Statistical Area) level. In its application to the school system level, the model serves as an evaluation technique in the location of school plants. In higher aggregate levels, where it is assumed that several governmental units and authorities are involved, the model serves as an analytical tool in assessing the potential gains and costs achieved through regional planning of school locations.

Questions of interest that may be answered are: what size schools are required to achieve given racial and/or social composition levels; what geographical units (and therefore governmental units) are required to cooperate in such planning; what resource use in terms of student travel time or travel cost is required to achieve given social or racial compositions; what are the trade-offs or exchange rates between student travel time or cost or distance and racial or social composition; and finally what is the assignment of students to schools, under stated constraints on the composition of the school attendance areas, that will result in minimum student travel time or cost or distance.

Inputs required to conduct the above analyses are: (1) the proposed school plant(s) defined by location, capacity and age group serviced; (2) the existing inventory of school plants similarly defined as above; (3) the distribution of students defined by their racial, social and age characteristics and geographical location, at a level of an areal unit such as a census tract.

The output is the assignment of students to schools in a way that satisfies constraints on the resulting school attendance areas and that satisfies the stated objectives in an optimal manner.

No attempt is made at this time to define social class or for that matter, racial group. The method is general to permit any definition of these terms. What is required, however, is the data in number of students corresponding to the classification scheme selected.

MODEL

A. Definitions:

1. Let X_{ijkmp} = the number of students in school location i, from areal unit j, of racial group k, social class m and age group p

where

$i = 1, 2, --- I$	school locations
$j = 1, 2, --- J$	areal units
$k = 1, 2, --- K$	racial groups
$m = 1, 2, --- M$	social classes
$p = 1, 2, --- P$	age groups

2. A school is defined by its location, capacity and age group serviced. Its capacity is r_{ip} = capacity of the school located at i for the p^{th} age group

3. Measures associated with the transportation of a student from an areal unit j to school location i:

t_{ij} = time to move from areal unit j to school location i

d_{ij} = distance from areal unit j to school location i

c_{ij} = cost of transportation from areal unit j to school location i

It is assumed that these values are fixed for a given areal unit, that is, they are applicable to all students in the given areal unit.

4. Measures associated with the constraints on the composition of the school enrollment:

a_{ikp} = minimum proportion desired in school (i,p) of racial group k

b_{ikp} = maximum proportion desired in school (i,p) of racial group k

c_{imp} = minimum proportion desired in school (i,p) of social class m

d_{imp} = maximum proportion desired in school (i,p) of social class m

B. Constraints or Requirements of Solution:

1. School capacity is defined for each school, defined by its location and age group serviced,

$$(1) \quad \sum_j \sum_k \sum_m X_{ijkmp} \leq r_{ip}$$

which states that the number of students assigned to a given school can not exceed a given capacity.

2. School composition of racial group k is to be between the proportion a_{ikp} and b_{ikp} of the total school plant enrollment,

(2) $a_{ikp} \sum_j \sum_k \sum_m X_{ijkmp} - \sum_j \sum_m X_{ijkmp} \leq 0$

(3) $\sum_j \sum_m X_{ijkmp} - b_{ikp} \sum_j \sum_k \sum_m X_{ijkmp} \leq 0$

These equations state that the number of students of a given racial group k cannot be less than or exceed given proportions of the total school plant enrollment.

3. School composition of social group m is to be between the proportion c_{imp} and d_{imp} of the total school plant enrollment,

(4) $c_{imp} \sum_j \sum_k \sum_m X_{ijkmp} - \sum_j \sum_k X_{ijkmp} \leq 0$

(5) $\sum_j \sum_k X_{ijkmp} - d_{imp} \sum_j \sum_k \sum_m X_{ijkmp} \leq 0$

These equations state that the number of students of a given social class m cannot be less than or exceed given proportions of the total school plant enrollment,

4. Each student is assigned to a school,

$\sum_i X_{ijkmp} = N_{jkmp}$

where

N_{jkmp} = the given number of students of racial group k, social class m, age group p residing in the areal unit j. This equation assures the assignment of all students to a school.

C. Objectives of Solution

1. Under the constraints defined above, assign students to schools that will minimize one of the following:

a. $\sum_i \sum_j \sum_k \sum_m \sum_p d_{ij} X_{ijkmp}$, aggregate distance traveled by all students

b. $\sum_i \sum_j \sum_k \sum_m \sum_p c_{ij} X_{ijkmp}$, aggregate cost of transportation for all students

c. $\sum_i \sum_j \sum_k \sum_m \sum_p t_{ij} X_{ijkmp}$, aggregate time traveled by all students

D. Solution

The above structure of the problem is in the form of a general linear programming problem and may be solved by standard computer programs. There will be a solution for each of the objectives defined in C above. The solution will yield,

X^o_{ijkmp} = the number of students of areal unit j, racial group k, social class m, and age group p that are assigned to school i.

This output will define the composition of the individual schools and the school attendance area.

Procedure to Evaluate Policy

The methodology may be used to evaluate proposed school plant locations and sizes according to the criteria and measures defined above. In summary, the analysis will consist of:

(1) the definition of the proposed school plant(s) by location, age group and capacity (i.e. i, p, r_{ip})

(2) the definition of the existing school plants by location, age group and capacity. While these will represent actual physical plants their conversion from one age group to another may be

examined. For example the conversion of an elementary school to a secondary school may be considered.

(3) data are acquired that define the student population cross classified according to areal unit, racial group, social class, and age group.

(4) proportions of racial and social mixes are selected for the individual school plants. The overall average of these selections will, of course, be equal to the average for the region. Considerable latitude will be possible for the individual school plants however. This measure may also be examined by varying the size of the region. In the application to a metropolitan region, for example, where the central city school system is largely of one racial group, it may be of interest to investigate the region size (and size and location of schools) required to achieve racial and/or social integration. By changing the level of the constraints on the racial and/or social mix. the relationships among region size, school size, and location, and objectives of cost, time and/or distance may be examined.

(5) an objective function is selected, such as the objective that the total distance traveled by all students be as small as possible. These objectives insure the assignment of students to schools by a "closeness" measure. A particular objective may

be more desirable for one region than another, depending on the geographical features, transportation system and so forth. Whether one objective is or is not more desirable may only become evident after examining all three (or variants of the three, such as t_{ij}^2 instead of t_{ij}).

(6) the model is solved to yield the assignment of students to schools. Standard computer programs are available to solve this model.

An illustration of the application of this model is shown in the Appendix.

Time Dependent Formulation

One could implicitly introduce time in the previous model by letting the values N_{jkmp} represent average (or maximum) values over some appropriate interval of time. The solution then would approximate the average condition experienced during the interval (where the interval corresponds to the planning time frame). A more explicite approach would be to formulate the previous relationships as a function of time. This will be done by considering a variable, say t, which is defined at discrete intervals of time. For example t=1,2,3 could represent 15 years of time where t=1 represents the first 5 years, t=2 the next five years and t=3 the remaining five years.

This extension would require the following data:

$N^{(t)}_{jkmp}$ = the number of students of racial group k, social class m and age group p residing in areal unit j during the time t.

The development of the constraining equations and objective function is a straightforward extension of the previous formulation. Only a few equations will be shown to demonstrate the extension.

Consider the equations constraining the capacity of the individual school plants,

$$\sum_j \sum_k \sum_m X_{ijkmp} \leq r_{ip}$$

This will now be written as a function of time,

$$\sum_j \sum_k \sum_m X^{(t)}_{ijkmp} \leq r_{ip}$$

where there is now an equation for each defined value of t. For example, for t=1,2 these constraints would be written,*

$$\sum_j \sum_k \sum_m X^{(1)}_{ijkmp} \leq r_{ip} \; ; \; \sum_j \sum_k \sum_m X^{(2)}_{jkmp} \leq r_{ip}$$

where $X^{(1)}$ would indicate the assignment during the first interval of time and $X^{(2)}$ the assignment during the second interval of time. Similarly the remaining constraints would be written with superscript t.

The objective function would now contain the additional summation over t,

$$\sum_t \sum_i \sum_j \sum_k \sum_m \sum_p d_{ij} \; X^{(t)}_{ijkmp}$$

*One could also define the capacity of each school as a function of time i.e., $r^{(t)}_{ip}$.

APPENDIX: ILLUSTRATION OF MODEL

The number of variables required to define a realistic problem in the application of this technique would be on the order of several hundred, with a comparable number of constraints. There are linear programming programs available that can handle several hundred constraints and almost an unlimited number of variables (see for example Gass, Saul I, Linear Programming, McGraw-Hill, pg. 131 ff).

The problem presented in this Appendix was run using an existing linear programming program available through a computer timesharing system. The limitations on the size of program (number of constraints and number of variables) was such that a completely realistic problem could not be formulated. These results are, therefore, only presented as illustrative of the model.

Consider the following example. Let there be two school sites at each of which are located an elementary and secondary school. Let there also be two areal units in which there are the following distribution of students:

Areal Unit 1

N_{111} = 10 white elementary students

N_{112} = 10 white secondary students

N_{121} = 250 nonwhite elementary students

N_{122} = 100 nonwhite secondary students

Areal Unit 2

N_{211} = 100 white elementary students

N_{212} = 120 white secondary students

N_{221} = 30 nonwhite elementary students

N_{222} = 5 nonwhite secondary students

The subscripts on the above N values are based on the assumption that only racial data is available on the students (therefore, the subscript m is omitted) and the general term is N_{jkp}, where:

j = 1 (areal unit 1), 2 (areal unit 2)

k = 1 (white), 2 (nonwhite)

p = 1 (elementary), 2 (secondary)

The remaining subscript is the school location where:

i = 1 (location 1), 2 (location 2)

The capacities of the schools are defined as r_{ip} where:

r_{11} = 100, capacity of the elementary school at location 1

r_{12} = 200, capacity of the secondary school at location 1

r_{21} = 300, capacity of the elementary school at location 2

r_{22} = 50, capacity of the secondary school at location 2

These may be existing schools or proposed schools.

The constraints on the racial composition of the schools are given by the values a_{ikp} and b_{ikp}. In this problem the constraints are defined for the nonwhite population. Thus

a_{i2p} = the minimum proportion of nonwhite students desired at school location i and level p

b_{i2p} = the maximum proportion of nonwhite students desired at school location i and level p

The specific values for the four schools are:

(1) $a_{121} = .30$, $b_{121} = .50$

(2) $a_{122} = .40$, $b_{122} = .60$

(3) $a_{221} = .60$, $b_{221} = .80$

(4) $a_{222} = .50$, $b_{222} = .60$

The distances, d_{ij}, between the i^{th} school location and the j^{th} areal unit are given by the values:

$d_{11} = 10$, $d_{12} = 15$

$d_{21} = 5$, $d_{22} = 20$

(What perhaps would be of more interest would be some function of distance such as time or cost.)

The objective then is, given the existing distribution of students, the locations and capacities of the schools, and the constraints on the composition of the schools, to assign students to schools that will minimize the total distance traveled by all students. The solution yields the following school compositions:

(1) Elementary School at Location 1

50 white students from areal unit 2

10 nonwhite students from areal unit 1

30 nonwhite students from areal unit 2

(2) Secondary School at Location 1

110 white students from areal unit 2

70 nonwhite students from areal unit 1

5 nonwhite students from areal unit 2

(3) Elementary School at Location 2

10 white students from areal unit 1

50 white students from areal unit 2

240 nonwhite students from areal unit 1

(4) Secondary School at Location 2

10 white students from areal unit 1

10 white students from areal unit 2

30 nonwhite students from areal unit 1

The total distance traveled by all students is 6,375 units (e.g. miles) which is the minimum distance possible. Excess capacity is seen to occur at school (1), ten spaces, and at school (2), fifteen spaces.

Modifications of this solution are, of course, possible but not at any shorter total distance. For example if 10 nonwhite elementary students that originate in areal unit 1 were shifted from school (3) to school (1) then each student would travel 5 more units per trip (d_{11} - d_{21} = 10-5 = 5) or a total increase of 50 units per trip. Thus a measure is defined for a possible evaluation of assignments that depart from the "optimal" solution.

Other outputs are directly available from the solution to the general linear programming problem. The final numerical calculation (tableau) contains information that answers such questions as, if increased capacity is planned, where should it be allocated to achieve the largest reduction in distance, and similarly, if constraints on the racial mix are to be changed, where should it be altered to achieve the largest reduction in distance. While the numerical calculations will not be shown, some examples of the types of analyses that may be performed will be given. For example, if the capacity of school (3) could be increased one student that would result in the reduction of the total distance traveled by all students by 3 units (increasing the capacity of school (1) or (2) would yield no reduction while school (4) would result in the reduction of 1 unit). Similarly if one more non-white student were admitted to either school (3) or school (4) this would result in the reduction of total distance traveled by all students by 10 units. A similar change in school (1) or school (2) would yield no reduction.

This type of analysis points the way to modifications of the original selection, which may be rerun and similarly analyzed.