

Abstract Title Page

Title: Using Simulation to Understand Consistency in Treatment Effects: An Application to School Choice

Author(s): Spiro Maroulis
Eytan Bakshy
Louis Gomez
Uri Wilensky

Abstract Body

Background / Context:

Research investigating whether students who attend “choice” schools (e.g., private schools, public magnet schools, charter schools) are better off than those who do not has yielded inconsistent findings. Observational studies using a variety of approaches to account for selection bias have resulted in disputed findings, with some finding better educational outcomes for students who attend magnet (Gamoran, 1996) or Catholic schools (Bryk et al., 1993; Coleman et al., 1982; Evans & Schwab, 1995; Morgan, 2001); and others finding little to no effect on academic achievement (Alexander & Pallas, 1983; Goldhaber, 1996; Neal, 1997). Perhaps even more surprisingly, field studies taking advantage of the lotteries put in place to deal with oversubscription to choice programs have not yielded conclusive evidence on the treatment effect for choosers. Randomized field trials of pilot voucher programs in Milwaukee (Greene et al., 1997; Rouse, 1998; Witte et al., 1995), New York City, Dayton, and Washington, DC. (Howell et al., 2002) have resulted in effect sizes ranging from small to modest.

Differences in estimates across such studies could reflect a number of underlying causes, with methodological issues related to the internal validity of the studies serving as the prime suspects. For example, with respect to the randomized field trials, debates have centered on methodological issues pertaining to selective attrition in Milwaukee (Witte, 1997), and subgroup definition in New York City (Krueger & Zhu, 2004). With respect to the observational studies, serious concerns have, not surprisingly, been raised about the approaches used to deal with selection bias (Altonji et al., 2002). But even if internal validity issues could all be satisfactorily settled, school choice programs vary both in their design and in the populations they serve. It is therefore reasonable to want to understand how such contextual features influence the estimates of the aggregate treatment effects, even in the most rigorously designed and implemented of randomized field trials.

Purpose / Objective / Research Question / Focus of Study:

We examine the sensitivity of the school choice treatment effects -- as defined as the difference between participants and non-participants in open enrollment programs -- to differences in i) the underlying student/household preferences of a school districts, and ii) the program participation rates of the district. Data detailed and broad enough to directly estimate these relationships across many districts do not exist. Instead, we use student- and school-level data from Chicago Public Schools to initialize an agent-based, computational model of the transition to public school choice; and then conduct computational experiments with hypothetical districts that would be otherwise difficult or impossible to execute in the field. To be clear, our intent is not to perform a secondary analysis of the Chicago open enrollment program, but instead to gain a better understanding of the connection between the contextual features in which school choice programs are implemented, and the outcome measures used in social experiments that take place in those contexts.

Setting: Chicago Public Schools

Population / Participants / Subjects:

To initialize the model, we used student-level data made available by Chicago Public Schools (CPS). This included achievement, school enrollment, and demographic information on all students enrolled in a CPS school between 2001 and 2003 (See Appendix B.3).

Abstract Body

Background / Context:

Research investigating whether students who attend “choice” schools (e.g., private schools, public magnet schools, charter schools) are better off than those who do not has yielded inconsistent findings. Observational studies using a variety of approaches to account for selection bias have resulted in disputed findings, with some finding better educational outcomes for students who attend magnet (Gamoran, 1996) or Catholic schools (Bryk et al., 1993; Coleman et al., 1982; Evans & Schwab, 1995; Morgan, 2001); and others finding little to no effect on academic achievement (Alexander & Pallas, 1983; Goldhaber, 1996; Neal, 1997). Perhaps even more surprisingly, field studies taking advantage of the lotteries put in place to deal with oversubscription to choice programs have not yielded conclusive evidence on the treatment effect for choosers. Randomized field trials of pilot voucher programs in Milwaukee (Greene et al., 1997; Rouse, 1998; Witte et al., 1995), New York City, Dayton, and Washington, DC. (Howell et al., 2002) have resulted in effect sizes ranging from small to modest.

Differences in estimates across such studies could reflect a number of underlying causes, with methodological issues related to the internal validity of the studies serving as the prime suspects. For example, with respect to the randomized field trials, debates have centered on methodological issues pertaining to selective attrition in Milwaukee (Witte, 1997), and subgroup definition in New York City (Krueger & Zhu, 2004). With respect to the observational studies, serious concerns have, not surprisingly, been raised about the approaches used to deal with selection bias (Altonji et al., 2002). But even if internal validity issues could all be satisfactorily settled, school choice programs vary both in their design and in the populations they serve. It is therefore reasonable to want to understand how such contextual features influence the estimates of the aggregate treatment effects, even in the most rigorously designed and implemented of randomized field trials.

Purpose / Objective / Research Question / Focus of Study:

We examine the sensitivity of the school choice treatment effects -- as defined as the difference between participants and non-participants in open enrollment programs -- to differences in i) the underlying student/household preferences of a school districts, and ii) the program participation rates of the district. Data detailed and broad enough to directly estimate these relationships across many districts do not exist. Instead, we use student- and school-level data from Chicago Public Schools to initialize an agent-based, computational model of the transition to public school choice; and then conduct computational experiments with hypothetical districts that would be otherwise difficult or impossible to execute in the field. To be clear, our intent is not to perform a secondary analysis of the Chicago open enrollment program, but instead to gain a better understanding of the connection between the contextual features in which school choice programs are implemented, and the outcome measures used in social experiments that take place in those contexts.

Setting: Chicago Public Schools

Population / Participants / Subjects:

To initialize the model, we used student-level data made available by Chicago Public Schools (CPS). This included achievement, school enrollment, and demographic information on all students enrolled in a CPS school between 2001 and 2003 (See Appendix B.3).

Intervention / Program / Practice:

We focus on open enrollment, a popular form of public school choice where households can choose among existing public schools in the district, but do not receive vouchers to go outside of the public system.

Research Design:

Simulation

Data Collection and Analysis:

Our analysis proceeded in three broad steps. First, we developed a simulation comprised of two sets of agents – students and high schools – who operate on a landscape that represents the geography of Chicago (Appendix B, Figure 5). Students vary in ability and background. Schools vary in quality and building capacity. In each time period of the simulation incoming 9th graders rank high schools using a preference function based on the mean achievement and geographic proximity of the school. If choice is allowed, the students (in random order) attempt to attend their top ranked school; if choice is not allowed, the students attend their assigned neighborhood school. Acceptance to a school depends only on capacity constraints. If there are no available spaces at the student's top choice, the student attempts to attend the next school on their ranked-list, and continues to try schools until finding one with room. Regardless of availability, a student's assigned neighborhood school must accept them. Upon enrollment, students update their academic achievement based on a combination of individual traits and the "value-added" parameter of the school estimated from CPS data. Schools that do not meet a minimum threshold of enrollment are permanently closed (See Appendix B.3 for model details).

Second, we used data from Chicago Public Schools to initialize the students, schools, and census blocks in the simulation. Students were sampled directly from the CPS data and placed on the appropriate census block. To obtain the achievement growth for a student attending a particular school in the simulation, we estimated a hierarchical linear model of student achievement that nests students inside of schools. The resulting equation was used to predict achievement for each student. The second-level residuals of the estimation were used to initialize the value-added of each particular school. (Please see Appendix B.1 for details.). Building capacity was based on the actual design capacity of each school building.

Third, we used the model to run computational experiments. The primary outcome of interest was the treatment effect of a public choice program – i.e., the difference in achievement between choosers and non-choosers attributable to the being able to attend a school of their own choosing. The key independent variables were the weight, α , placed on school quality relative to geographic proximity in the school choice decision rule used by the choosers, and the percentage of students who take advantage of the ability to choose, *pctChoosers*. The data were generated by running the model for twenty time periods under systematically chosen combinations of independent variables. Since each run of the model represents one instantiation of a stochastic process, each unique combination of parameters was repeated twenty times to create the distributions of outcomes presented below.

Findings / Results:

Figure 2 in Appendix B plots the final mean achievement across all students versus the percent of them who choose, at several different values of α . Figure 3 in Appendix B plots the treatment effect at the completion of the simulation given the same conditions. Since choosers

are randomly selected, the treatment effect for any given run of the model can simply be calculated by taking the difference in mean achievement between choosers and non-choosers at the completions of the run. Each point represents the average across the twenty runs for that particular combination of parameters, and mean achievement is measured in standard deviations.

When comparing the mean achievement in each of the scenarios in Figure 2, there were no surprises. The more that individual students favored achievement relative to geographic proximity (i.e., high α), the higher the mean achievement. Since mean achievement and value-added are positively correlated, the gain comes straightforwardly from student attending higher value-added schools. Also, with the exception of the case where students heavily favor geographic proximity ($\alpha = 0.2$), the more people who choose, the higher the mean achievement. However, when comparing the treatment effects across scenarios (Figure 3) the treatment effect goes down as participation increases – the exact opposite relationship than what is observed when one calculated the overall mean achievement of the students under each counterfactual condition. This underlying reasons for this highlights the importance of considering the micro-level processes of a system when estimating causal effects: When there is very low participation and some excess capacity at the better schools, all participants looking for a new school find a spot at one of the highest value-added schools. As more and more people take advantage of the open enrollment option, and the top schools reach capacity, the choosers necessarily have to go to schools that do not have as high value-added, but likely still better than the schools from which they came. (See Appendix B, Figure 4 for verification.) Consequently, the mean achievement value of the choosers is lower when more students choose.

Conclusions:

Analysis of the model finds that treatment effects calculated by comparing choosers to non-choosers are highly dependent on both the household participation rates in the program and the distribution of available capacity across schools. In particular, as participation rates rise, the magnitude of the treatment effect falls, because capacity constraints increasingly limit the amount of choosers who are able to attend the highest value-added schools. From a policy perspective, this finding highlights the importance of connecting an understanding of the mechanisms in each context that give rise to aggregate school choice outcomes. From a research perspective, the most direct implication of this finding is that one must account for the amount of “better” capacity available to choosers when using treatment effects estimated from existing programs to either a) project the impact of a larger scale program, or b) synthesize effect sizes estimated across programs. The measure we developed to characterize the better available capacity of the district in the model, *bac*, could be applied to program data to aid with both these purposes. Future work should include a more refined understanding of student preferences is required. More specifically, the current model only partially addresses heterogeneity in the decision-making rules of households. Additional heterogeneity could come in the form of categories of agents weighing elements of the existing preference function differently, or in the form of additional and varied criteria on which to judge schools that go beyond mean achievement and geographic proximity.

APPENDIX A

References

- Alexander, K. L., & Pallas, A. M. (1983). School sector and cognitive performance: When is a little a little? *Sociology of Education*, 58(2), 115-128.
- Altonji, J., Elder, T., & Taber, C. (2002). An evaluation of instrumental variable strategies for estimating the effect of catholic schools (No. 9358). National Bureau of Economic Research.
- Belfield, C. R., & Levin, H. M. (2002). The effects of competition between schools on educational outcomes: A review for the united states. *Review of Educational Research*, 72(2), 279-341.
- Borland, M., & Howson, R. (1992). Students' academic achievement and the degree of market concentration in education. *Economics of Education Review*, 11, 31-39.
- Bryk, A. S., Lee, V. E., & Holland, P. B. (1993). *Catholic schools and the common good*. Cambridge: Harvard University Press.
- Coleman, J. S., Hoffer, T., & Kilgore, S. (1982). *High school achievement: Public, catholic, and private schools compared*. New York: Basic Books.
- Evans, W. N., & Schwab, R. M. (1995). Finishing high school and starting college: Do catholic schools make a difference? *The Quarterly Journal of Economics*, 110, 941-974.
- Gamoran, A. (1996). Student achievement in public magnet, public comprehensive, and private city high schools. *Educational Evaluation and Policy Analysis*, 14(3), 185-204.
- Goldhaber, D. D. (1996). Public and private high schools: Is school choice an answer to the productivity problem? *Economics of Education Review*, 15(2), 93-109.
- Greene, J. P., Peterson, P., & Du, J. (1997). Effectiveness of school choice: The Milwaukee experiment.
- Howell, W., Wolfe, P., Campbell, D., & Peterson, P. (2002). School vouchers and academic performance: Results from three randomized field trials. *Journal of Policy Analysis and Management*, 21(2), 191-217.
- Morgan, S. L. (2001). Counterfactuals, causal effect heterogeneity, and the catholic school effect on learning. *Sociology of Education*, 74, 341-374.
- Neal, D. (1997). The effects of catholic secondary schooling on educational achievement. *Journal of Labor Economics*, 15(1), 98-123.
- Wilensky, U. (1999). NetLogo. <http://ccl.northwestern.edu/netlogo>. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.
- Witte, J. F. (1997). Achievement effects of the milwaukee voucher program. In American economics association annual meeting. New Orleans.
- Witte, J. F., Sterr, T., & Thorne, C. (1995). Fifth-year report: Milwaukee parental choice program.

APPENDIX B

B.1 Estimating Achievement Growth By School

To obtain an estimate of achievement growth for a student attending a particular school in the simulation, we estimate a hierarchical linear model of student achievement that nests students inside of schools. More specifically, we estimated the following model that predicts 11th Prarie State Achievement Examination scores for the all students used in the simulation, using the 8th grade Iowa Test of Basic Skills scores and student-level demographics of those students as the independent variables:

$$achiev_{2ij} = \beta_{0j} + \beta_{1j}achiev_{1i} + \beta_{2j}white_i + \beta_{3j}male_i + \beta_{4j}poverty_i + va_j + r_{ij} \quad (1)$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (2)$$

where $r_{ij} \sim N(0, \sigma^2)$ and $u_{0j} \sim N(0, \tau_{00})$

Table 1 presents the results of the HLM estimate for both math and reading scores. Substituting Equation 2 into Equation 1, and replacing the β s with the estimated coefficients, yields the following equation used as the achievement growth rule in the simulation:

$achiev_{2ij} =$

$$-0.0956 + 0.6794 * achiev_{1i} + 0.1567 * white_i + 0.1151 * male_i - 0.0629 * poverty_i + va_j + r_{ij}$$

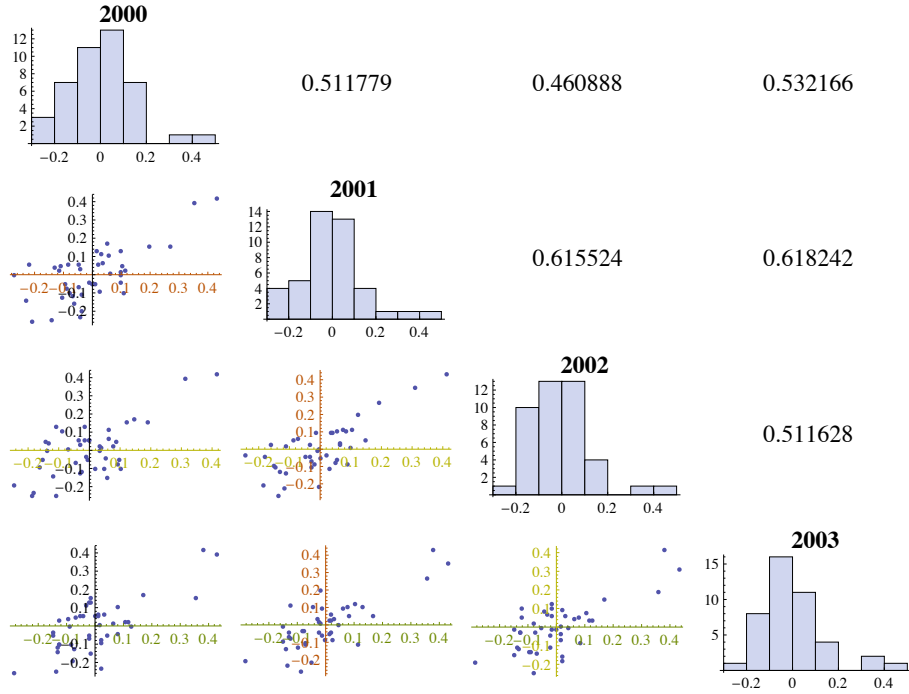
For each school j , the school-level residual u_{0j} , is used as an estimate of the value-added, va_j ; r_{ij} is a random draw from $N(0, 0.3921)$ every time the achievement growth equation is calculated.

Such an approach assumes that the value-added for each school is relatively stable year over year. To evaluate the stability of the value-added estimates, we also estimate the model separately for each incoming cohort of 8th grade students, and examine the year over year association between the school-level residuals (as opposed to Table 1 which generates the estimate by using all the cohorts). Figure 1 shows that all the year-over-year correlations are strong and positive.

Table 1: HLM Estimates of 11th Grade Test Scores
(17,131 students in 43 schools)

		Math		Reading	
		Estimate	S.E.	Estimate	S.E.
<i>Fixed Effects</i>					
Intercept	(γ_{00})	-0.0956	0.0301	-0.0025	0.0247
8th Grade Iowa Score	(β_{1j})	0.6794	0.0053	0.5951	0.0062
White	(β_{2j})	0.1567	0.0183	0.1798	0.0211
Male	(β_{3j})	0.1151	0.0097	-0.1180	0.0112
Poverty	(β_{4j})	-0.0629	0.0080	-0.0159	0.0091
<i>Random Effects</i>					
School-level residual variance	(τ_{00})	0.0363		0.0230	
Student-level residual variance	(σ^2)	0.3921		0.5252	

Figure 1: Association Matrix of School Level Residuals. The upper half of the matrix contains Spearman rank correlation coefficients between the school-level residuals estimated for each cohort of incoming freshman; the lower half of the matrix shows the same association as a scatterplot; the diagonal contains the distribution of school-level residuals in a given year.



B.2 Results from Model

Figure 2: Mean Achievement vs. Percent Choosers

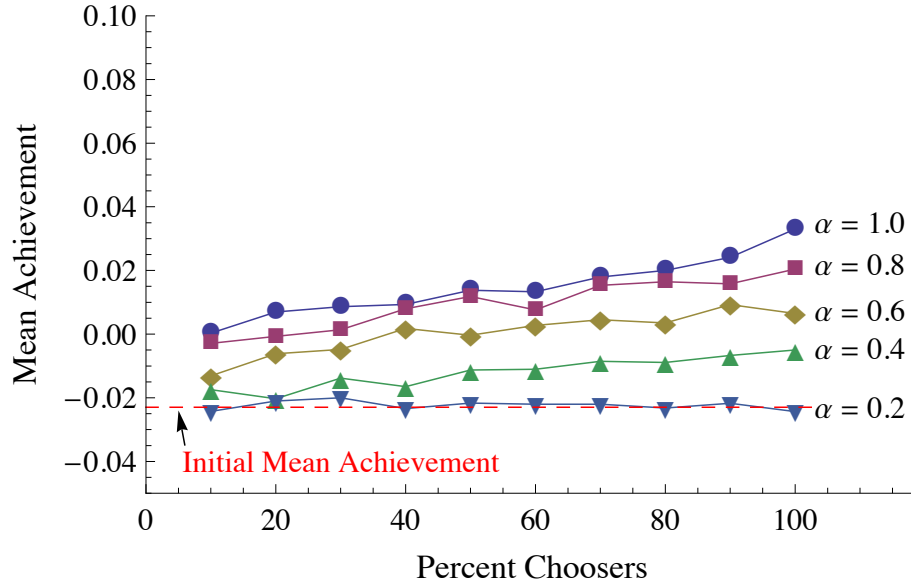


Figure 3: Treatment Effect vs. Percent Choosers

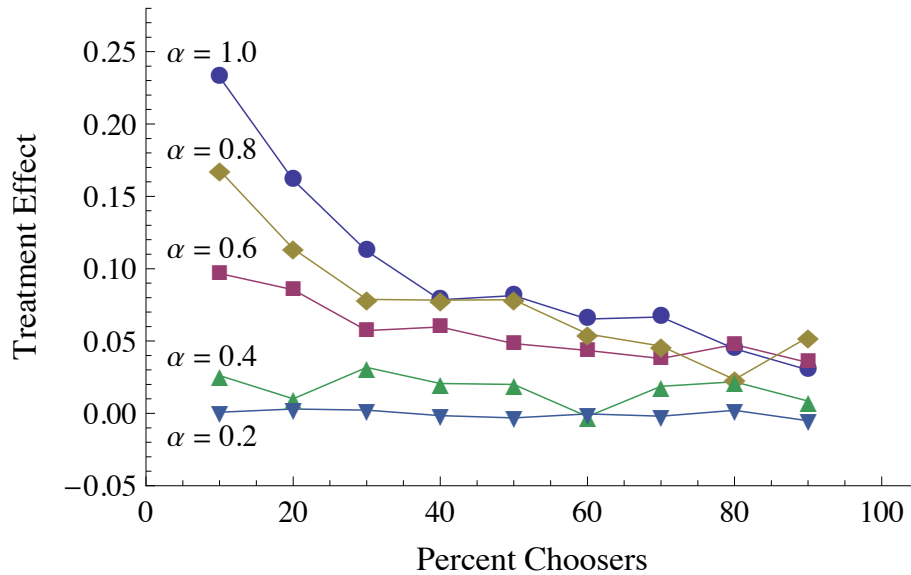
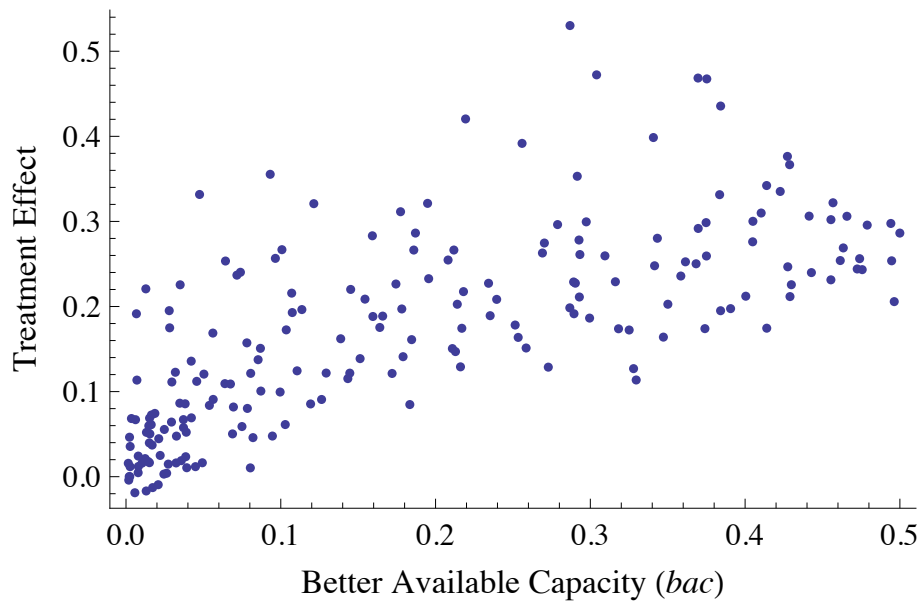


Figure 4: Mean Achievement vs. Better Available Capacity. Results from running the model 200 times, each time randomly assigning the percentage of students who choose, and the amount of initial excess capacity in the system. Each point represents a realization of the model. Results confirm that a higher level of better available capacity in a district *bac* correspond to larger differences in achievement between choosers and non-choosers. (*bac* is calculated by asking every student who is a chooser to calculate the quantity, *betterSpacesPerSchool*, the number of spaces per school available to them at a schools with a higher value-added. *bac* is the mean of *betterSpacesPerSchool* across all choosers in the initial time period.)



B.3 Additional Model Information

The model is comprised of two agents – students and schools – who operate on a landscape that represents the geography of a school district. The simulation begins in a state where all students attended their assigned neighborhood school, and the first time period of the simulation represents the first year when students can choose.

Each time period of the simulation proceeds as follows:

1. The model is populated with 5000 incoming students, and a fraction of them are randomly designated as “choosers”; the fraction is determined by the tunable parameter, *pctChoosers*
2. The “choosers” rank schools in accordance to their preferences, and in random order attempt to attend their top choice school; the remainder of the students attend their assigned neighborhood school.
3. If there are no available spaces at the student’s top choice, the student attempts to attend the next school on her ranked-list, and continues to try schools until she finds one with room. Regardless of availability, a student’s assigned neighborhood school must accept them.
4. Students updated their achievement level; the updated achievement depends both on the student’s individual-level attributes and the value-added of the school they attend.
5. Schools update their aggregate enrollment and achievement values; they also estimate the number of spaces available for new students next year.
6. Schools that do not meet a minimum threshold of enrollment are permanently closed.
7. Students completing their fourth year in a school, graduate from the system; a student stays at the same high school all four years.

Figure 5: Representative Initialization for Model. A circle represents a school. The circle's size is proportional to the school's enrollment, and its color indicates the value-added of the school (green = high, red = low). Schools are placed at the geographic location of their address. The small dots represent students, who are placed within their home census block and attend their assigned school.

