Executive Summary: Public School Choice: Magnet Schools, Peer Effects, and Student Achievement

Goldring, Ellen

In this research we examine the performance of magnet schools, the oldest and most widely used form of intradistrict choice in the United States. We ask two questions: (1) do students attending magnet schools make greater academic gains (as measured by standardized achievement tests) than comparable students enrolled in conventional public schools; (2) do peer effects and other factors (such as teacher quality) contribute to school performance?

Because magnet schools are over-subscribed and vacant places are filled by lottery, the data meet the conditions for a quasi-randomized experiment: the unsuccessful magnet applicants constitute a natural control group against which we can compare the progress of students who attended magnets. This feature of our research design permits us to deal in a straightforward way with a self-selection bias that affects most studies of school choice: because parents exercising choice are widely believed to have stronger than average commitment to education, their children might well have made better than average progress even if they had remained in regular schools. Because the magnet lottery randomizes the assignment to magnet and regular programs, it removes this source of bias. Our investigation of peer effects also takes advantage of lottery randomization to ensure our estimates truly measure the influence of peers and not characteristics of students or their parents that have led them to seek particular peer groups.

Our research focuses on the Nashville, Tennessee metropolitan school district, a racially and economically diverse system in which ten percent of the students are served in magnet schools and closely related choice programs. Nashville offers an exceptionally good opportunity to study the impact of choice. Students apply to magnet schools via lottery with more applicants than spaces leading to randomization in our design. Students are tested annually in five different subjects. Scores of individual students are linked longitudinally over time, facilitating a value-added analysis of school quality.

We also exploit lottery outcomes to estimate the effect of peers on student achievement. Conditional on attendance zone, a magnet lottery determines whether a student is assigned to a school with the characteristics of the magnet school or one with the characteristics of their neighborhood school. Thus, within this group, there is randomized assignment with respect to the entire set of school characteristics, including peers. This provides an exogenous source of variation in peers that permits identification of peer effects free of bias caused by the fact that peers are typically self-selected through residential location decisions.

Results indicate that race and income of peers have a substantial impact on achievement: the estimated difference between a school where students are 75 percent black and one in which students are 25 percent black is more than half a year's normal growth in mathematics. Further analysis indicates that these peer characteristics are not proxies for other determinants of achievement, such as teacher quality or heterogeneity in the response to treatment.
This research has important implications. Although there have been previous studies of magnet schools, few have exploited lottery randomization to obtain estimates free of selection bias. Moreover, most previous research ignores peer influences as a channel through which magnet schools affect achievement. Although student assignment issues are among the most important decisions facing local districts, there is little empirical basis to guide decision makers. Our research establishes the feasibility of using lottery randomization to study both peer and magnet effects. There are a number of threats to validity in such a study, some of which we have already mentioned. Validity of the instruments must be established, given the near-certainty of response heterogeneity. Ostensible peer effects may in fact be the result of other factors merely correlated with peers (e.g., teacher quality). The peers that magnet school students would have had, were they not enrolled in a magnet school, are an unobserved counterfactual that must be estimated. Finally, as in any experiment, attrition of participants can introduce bias even when initial assignment to treatment and control groups is random. On all of these counts, the results of our analyses are reassuring.

Second, the results underscore that achievement can diverge dramatically within a relatively short time as a function of middle school peers. This may indicate a heightened sensitivity to peer influences just as students are reaching middle school. It is also possible that the lottery participants in this study are unusually susceptible to peer influences, compared to less motivated students who did not bother to enter the lottery.

Third, on average, attending a magnet school had a positive impact on mathematics achievement for the fifth and sixth graders in this study. Although needing to be cautious regarding the generalizability of these findings beyond the one district studied, we do note that the positive impact of magnet schools, coupled with the findings regarding peer effects, suggests that magnet schools may continue to be an important policy in achieving a mix of diverse peers that can help contribute to achievement gains.

Fourth, the research reported here, like other recent work on peer effects, has implications for the evaluation of teachers and schools on the basis of student test scores. Perhaps the most widely used of current methods for measuring educator value-added, the Educational Value-Added Assessment System at the SAS Institute (modeled on the Tennessee Value-Added Assessment System), includes no controls for student characteristics. With respect to the characteristics of individual students, this appears defensible when a long time-series of test scores for that student is available for analysis. However, such a time series is not generally an adequate substitute for peer characteristics. The larger the influence of peers, the more important it is to develop alternative assessment models that ensure teachers and schools are not held accountable for factors beyond their control.

Project Objective: (Section A)

Despite the overwhelming use of magnet schools as the most widespread form of public school choice, evidence is limited concerning the efficacy of magnet schools in promoting and sustaining student learning and achievement. The purpose of this paper is to study the educational consequences of enrolling in magnet schools. We also explore the importance of peer effects on student achievement.
Our research focuses on a racially and economically diverse metropolitan school district. Students apply to magnet schools via lottery. Students are tested annually and scores of individual students can be linked longitudinally over time. The study follows students two years after their participation in middle school magnet lotteries.

**Explanation of Progress**

We have completed an analysis of the impact of magnet schools on mathematics achievement in grades five and six. There were two types of magnet “treatment”—a selective, academic magnet school, and a composite of four non-selective magnets. Only students meeting specified standards were eligible for the former. Admission from this group was determined by lottery, as was admission to one or more of the non-academic magnets. Lotteries created a quasi-experiment in which students were randomly assigned to treatment (magnet) or control (non-magnet) groups.

Exploiting lottery-based randomization to obtain estimates of magnet and peer effects was not simple. There were several threats to validity. Much of the effort of the project went into identifying and resolving these threats.

First, there is a high degree of non-compliance: many students offered treatment (a place in a magnet school) do not accept it. This means that a simple comparison of the achievement of lottery winners and lottery losers (known as an intention-to-treat estimate) will be an attenuated measure of the treatment effect, as the effect (be it positive or negative) on winners who are actually treated will be diluted by the many winners who went untreated. A conventional solution is to use lottery outcomes as instrumental variables for treatment. However, despite randomization, lottery outcomes can fail to be valid instruments in conditions such as those here: students participate in multiple lotteries, and response to treatment is heterogeneous. Thus, we established the exogeneity of the instruments using the omnibus overidentification test. In every instance we failed to reject the null hypothesis (by a wide margin), indicating the instruments are valid.

Second, the number of students attending regular middle schools because they tried but failed to obtain a place in one of the non-academic magnets is small. This means the power to detect treatment effects for each of the non-academic magnets individually may be quite limited. Accordingly, we combine these four schools into a composite non-academic magnet treatment. Students are treated as applying to the composite if they apply to at least one of the four component schools. They are an outright winner to the composite if they win a place outright in at least one of the components; otherwise, they are a delayed winner in the composite if they are a delayed winner in at least one of the components.

Third, the large number of lottery participants who leave the district raises the possibility of significant bias if the attrition of lottery winners differs from lottery losers. We return to this point below, in the discussion of analytical strategies.

Because lotteries assign students to different sets of peers, they represent an exogenous source of variation in peer characteristics. When peers are determined by a family’s place of residence, investigation of peer influences using observational data can be badly biased by the fact that peers are essentially self-selected: what appears to be the
influence of peers might only be the influence of unobserved family variables correlated with choice of residence (and through it, school). Lottery-based variation in peers is a potentially valuable source of information on peer influences uncontaminated by this bias. However, to exploit this information, several challenges had to be met. First, as noted already, the validity of lottery outcomes as instruments had to be established. Second, the lottery assigns students randomly between magnet school peers and zoned school peers. The latter are still determined by residential location. Thus, it was necessary to condition on zoned school peers, meaning the model contains two peer variables: exogenous, lottery-determined peers (the coefficient on which represents the causal influence of peers on achievement; and an endogenous, residence-based peer variable (the coefficient on which is a function of self-selection and has no causal interpretation). Third, zoned school peers are not observed for students enrolled in magnet schools—they are a counterfactual that must be estimated (we principally use fourth grade schools for this purpose). Fourth, non-compliance means peers need not be those to which a student is assigned by the lottery. The lottery-based peer variable is therefore an instrument for actual peers. Fifth, while the estimated peer effect will be free of bias due to self-selection, it may still be the case that peers are confounded with other school-level variables with which they are correlated. One notable possibility is differences in teacher quality. Sixth, peers might also be confounded with heterogeneous responses to magnet school treatment—that is, the peers a student would have in his zoned school could be a signal of differences in students that cause some to flourish in magnet schools and others to struggle. We return to these final two issues below.

Model and Analytical Strategy

We modeled mathematics achievement among middle school students as a function of student characteristics (race, low income, special ed, ESL, gender, 4th grade achievement), peer characteristics (percent black, percent low income, other), and binary indicators of whether the student attended a magnet school. There were two indicators, one for the academic magnet school, the other a composite of the four non-selective magnet schools. Lottery outcomes were used as instruments for whether students enrolled in a magnet school. The instrument for peers was constructed as \( \sum P_j \hat{d}_j + \hat{P}_N (1-\Sigma \hat{d}_j) \), where \( P_j \) are the peers a student will have if he attends magnet school \( j \), \( \hat{d}_j \) is the predicted probability he attends magnet school \( j \) (based on lottery outcomes), and \( \hat{P}_N \) represents the middle school peers a student is predicted to have if he does not attend a magnet school (predicted on basis of elementary school attended and student demographic variables).

The achievement equation was

\[
Y_{ij} = \mu(X_i) + \sum \lambda L_{ij} + \delta_i d_{ij} + 8[\Sigma P_j d_{ij} + P_N(1-\Sigma d_{ij})] + \gamma \hat{P}_N + \eta_{ij} d_{ij} + \nu_{IN}
\]

where \( Y_{ij} \) is mathematics achievement by student \( i \) in school \( j \) (time subscripts are omitted but understood), \( X_i \) are student covariates, \( L_{ij} \) is a lottery participation indicator for student \( i \) and lottery \( j \), \( \lambda \) are lottery selection effects (representing differences in the pool of applicants to different lotteries), \( \delta_i \) is the causal effect of magnet school \( j \), relative to the expected outcome in non-magnet schools for the students at \( j \), \( \Sigma P_j d_{ij} + P_N(1-\Sigma d_{ij}) \) is a
characteristic of peers at the school actually attended by student $i$, $\beta$ is the causal effect of peers on achievement, $\gamma$ is the information contained in $\hat{P}_N$ about achievement (through residential location), $\eta_{ij}$ is unobserved response heterogeneity (how student $i$ responds to school $j$), and $v_{iN}$ is an idiosyncratic error term.

We conducted a variety of specification tests and checks on robustness to address concerns described above. Although admissions lotteries create exogenous variation in peer characteristics, estimated peer effects may represent the contribution of unobserved factors correlated with peers. One plausible candidate is teacher quality. However, estimated peer effects were undiminished when teacher fixed effects were added to the model.

Peer effects can also be confounded with heterogeneous response to treatment. (The source of the identification problem is the fact that both $\Sigma P_{jdij} + P_N(1-\Sigma dij$, the peer variable, and $\eta_{ijdij}$, heterogeneous response, involve interactions with $d_{ij}$.) We test for this by allowing the response to treatment to vary with observed heterogeneity, interacting treatment indicators with student characteristics. Estimated peer effects are undiminished. While this does not conclusively rule out the possibility that peer characteristics proxy for unobservable response heterogeneity, the fact that observable heterogeneity is unrelated to peer effects makes it less plausible that the latter are proxies for unobserved heterogeneity.

Finally, sample attrition can introduce systematic differences between treatment and control groups. There is a significant amount of attrition in these data. Of the 2315 students who applied to the selective magnet during this sample period, 436 had left the system or for other reasons were not tested the following year as fifth graders. As one might expect, attrition is greater among lottery losers than winners. However, bias from attrition appears to be slight, for two reasons. (1) The student covariates in the model include prior test scores, removing the source of greatest potential bias. (2) Losers leaving the system do not appear to differ systematically from winners who leave, at least with respect to observable variables that predict achievement. As before, this analysis cannot rule out conclusively that there is no attrition based on unobserved characteristics that influence achievement. However, on the basis of all the factors we can observe, attrition does not seem to have introduced differences between treatment and control groups systematically related to achievement.

Summary of Findings:
On average, attending a magnet school had a positive impact on mathematics achievement for the fifth and sixth graders in this study. Gains were uneven. In the academic magnet, they amounted to one-sixth of a year’s normal growth in grade 5, but were largely surrendered in grade six. In the non-academic magnets, gains were not apparent until grade six, but then they were much larger, amounting to half a year’s growth. Results indicate that race and income of peers have a substantial impact on achievement: the estimated difference between a school where students are 75 percent black and one in which students are 25 percent black is more than half a year’s normal growth in mathematics. As noted above, these peer characteristics are not proxies for other determinants of achievement, such as teacher quality or heterogeneity in the response to treatment.
Work planned but not undertaken (e.g., design changes) or completed should be explained.

The original plan called for collecting data on students’ neighborhoods (e.g., crime rates) and adding neighborhood variables to the achievement model. This was not done, for two reasons. First, there would be considerable measurement error in neighborhood variables. We lacked data on students’ exact place of residence, meaning that neighborhood would have to be defined broadly, using attendance zones for fourth grade. This would introduce measurement error to the extent that neighborhoods vary within a single zone. For students attending magnet schools as fourth graders we lack even this much information about their neighborhood. The same problem arises with respect to the measurement of counterfactual peer characteristics, but there it affects only the instrumental variables, which remain valid despite it. Here the regressors themselves would be measured with error.

Second (and weighing more heavily in the decision), neighborhood characteristics are like peer characteristics in that they represent influences of a student’s environment on academic achievement. Among the peer characteristics available for this research were attendance, numbers of disciplinary incidents, student intra-year mobility, and prior (fourth grade) achievement. By comparison with neighborhood characteristics, where uncertainty about the neighborhood in which a student resides is a source of imprecision, these variables are measured with much less error. However, none of them was found to have a significant effect on achievement. This did not bode well for the neighborhood variables and suggested that the effort spent collecting such data would not be repaid.

The original proposal indicated that we would investigate the contribution of teacher attributes to magnet school effectiveness. Such attributes would include certification status, experience, and teachers’ scores on the Tennessee Value-Added Assessment System (TVAAS) scores prior to joining the magnet school faculty. In the end we took another approach to this issue by controlling for all fixed attributes of teachers in the form of teacher fixed effects. We deemed this approach superior for three reasons. (1) Teacher attributes would be measured with error. For example, teacher experience represented only within-district experience. (2) The research literature shows that exceedingly little of the variation in teacher ability is explained by attributes such as licensure and advanced degrees. Teacher fixed effects, by contrast, capture these as well as otherwise unmeasured differences in teacher quality. (3) TVAAS scores were available only as school level averages of teacher scores prior to coming to the school. They would be based only on a portion of the teachers working at the school, as teachers with no prior experience in the district would not contribute to the average. This would result in measurement error, with different schools subject to greater or lesser amounts of error depending on the make-up of their staffs. More important, it became apparent that the most pressing question with respect to teachers was whether the very large peer effects we were estimating were in fact proxies for differences in teacher quality. For an investigation of this question, TVAAS scores were of no value, as TVAAS does not control for peer characteristics when estimating teacher effects. Thus, an examination of TVAAS scores cannot resolve the question of whether estimated peer effects identify influences of peers or teachers.

The project as originally proposed included focused interviews with principals of the district’s middle schools with regard to discretionary recourses and their allocation. We learned through interviews that principals have limited discretionary resources. Most
resources, such as number teachers and number of library books are allocated by district formula based on school size. Discretionary resources by and large come from PTA fundraising and other non-school activities, such as athletics and coke machines.

Although the original proposal focused on the impact of magnet schools and peers on academic achievement, we noted that we could also investigate the impact of magnet schools on behaviors, such as absenteeism and the number of disciplinary incidents in which students were involved. We have in fact conducted such analyses; however, our work to this point should be regarded as preliminary. By and large estimates are in the direction expected: attending a magnet school tends to improve attendance and reduce the number of disciplinary incidents in which students are involved. Most of these estimates do not reach the threshold of statistical significance. Because we will be obtaining additional data that will increase the number of observations for this study by approximately 50 percent (see below), we are putting off until that time further work on this question.

Section C: (other information –response to required questions).

1. Utilizing your evaluation results, draw conclusions about the success of the project and its impact. Describe any unanticipated outcomes or benefits from your project and any barriers that you may have encountered.

   Our project makes two contributions: substantive and methodological. In terms of the focus on magnet schools, despite the overwhelming use of magnet schools as the most widespread form of public school choice, there is limited evidence about the efficacy of magnet schools in promoting and sustaining student learning and achievement. Our finding of a significant magnet school effect in 5th and 6th grades makes a strong contribution because of its internal validity, given the quasi-experimental nature of the research design.

   We also find large peer influences on achievement. Given that our population comprises middle school students, this is not unexpected, though our estimates are larger than those reported in the recent literature. One unanticipated outcome: peer influences do not appear to operate through such channels as student absenteeism, disruptive behavior, or even prior achievement. This suggests that the relevant behavioral variables may not be easy to identify or measure using administrative data. The channels by which peer effects are transmitted would appear to be more elusive: ways of communicating peer approval and disapproval, enforcing norms that either support or undercut academic achievement.

   The methodological contributions are of two kinds. First, we identify the principal threats to validity facing such research. Second, the specification of the model and the subsequent tests and checks on robustness that we conduct to meet these threats illustrate what other researchers will need to do when attempting to exploit such data.

2. What would you recommend as advice to other educators that are interested in your project? How did your original ideas change as a result of conducting the project?

   Generalization on the basis of a single study of one district is hazardous. However, the main lessons of this study are the following.
1. Peer effects are very important at the middle school level. Educators need to develop strategies and techniques for countering the negative influences arising from high concentration of minority and low income students.

2. Despite the obvious reasons to exploit admissions lotteries in studying magnet and peer effects (the quasi-experimental nature of the research design), there are many pitfalls in such work. Researchers should read the narrative paper in the appendix to this report.

3. If applicable, describe your plans for continuing the project (sustainability; capacity building) and/or disseminating the project results.

Continuation of the research.

The data originally provided for this research covered academic years 1998-99 through 2003-04. Because 1998-99 test scores were used as controls for prior achievement, the first lottery cohort we could study entered fifth grade in 1999-2000. The fifth and final cohort entered middle school in 2003-04. As a result, we have many more observations on students in the first two years of middle school (grades 5 and 6) than in the final two years. This was our principal reason for restricting analysis to fifth and sixth graders.

We now have an agreement with the district to update these files through the 2006-07 school year. This allows us to follow the five lottery cohorts in this study all the way through middle school. We will therefore extend the analysis of achievement to include effects in grades 7 and 8. (Preliminary analysis, using the smaller data set available to this point, indicates that there may be substantial cumulative effects from magnet school attendance.)

In addition, we have found that district administrative records permit us to place middle school students in a particular classroom in each period of the day. Thus, we know not only what class a student was taking from which teacher, but also who the student’s classmates were. This allows us to undertake a much more fine-grained analysis of peer effects than we have been able to do using peer variables constructed at the school level. Peer characteristics that were unimportant at the school level (e.g., prior achievement) may be of greater consequence when measured at the classroom level.

Dissemination. The results of this project have been presented at the American Educational Research Association, The Association for Public Policy Analysis and Management, the IES annual conferences, and an invited address at The University of Connecticut conference on magnet schools. We will be meeting with the Director of Research and Evaluation from the district where we conducted this study to discuss our results and will continue to prepare them for publication in refereed journals.

4. Report on any statutory reporting requirements for this grant program. N/A