Magnet Schools and Peers:
Effects on Mathematics Achievement

Dale Ballou

Vanderbilt University
November, 2007

I wish to acknowledge helpful discussions with Adam Gamoran, Julie Berry Cullen, Steve Rivkin, and Ellen Goldring, and the research assistance of Keke Liu. This research was supported by the Institute for Education Sciences.

Abstract

I estimate the impact of attending a magnet school on student achievement in mathematics in a moderately large Southern district. Admission to magnet schools is through lotteries. Actual attendance by lottery winners is of course voluntary. I use lottery outcomes as instruments to control for bias due to self-selection of enrollees. Because lottery winners would have attended zoned schools of varying quality in the absence of magnet schools, the response to treatment is necessarily heterogeneous. Even so, these instruments are capable of identifying the effect of treatment on the treated for all students who enter magnets through the lottery. I 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 furnishes a way of identifying peer effects free of bias caused by endogenous choice of peers. 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.

I. Introduction

Because parents seek out particular schools by residential decisions or special application (for example, to magnet schools and charter schools), the effect of the selected school on student learning is likely to be confounded with parental and family characteristics that influence both where students go to school and how much they learn. Fortunately, the administration of school choice programs frequently provides a way of disentangling the effect of the chosen school from the influence of factors that led to that choice. Many school choice programs are oversubscribed: the number of applicants exceeds the number of vacancies. Admissions are conducted by lottery. Differences between the students selected and the unsuccessful participants in the lottery therefore arise solely by chance. This makes unsuccessful participants a natural "control group" for purposes of measuring school effectiveness. As in a randomized experiment with treatment and control groups, the impact of the treatment (in this case, the difference in quality of education) can be ascertained by comparing achievement of successful applicants who enroll in magnet schools with the achievement of unsuccessful participants who enroll in zoned schools. Applications of this approach include analyses of private school voucher plans (Howell and Peterson, 2002), charter schools (Hoxby and Rockoff, 2005), intradistrict public school choice (Cullen, Jacob and Leavitt, 2003), and career magnet academies (Kemple and Snipes, 2000, Kemple and Scott-Clayton, 2004, Crain, Heebner, and Si, 1992, and Crain et al., 1999).

In this study I exploit randomization via admissions lotteries to examine the effect of magnet schools in a mid-sized Southern city. The district serves 70,000 students, of whom half are eligible for free or reduced price lunch. The district is racially mixed, serving 40% White students, 48% Black, and 8% Hispanic students during the 2003-04 school year. While the district operates magnet schools at all levels—elementary, middle school, and high school—end-of-year assessments have been most consistent in grades 3-8. As a result, I focus on academic outcomes for students who apply to one or more of the district's magnet middle schools.

Magnet schools, charter schools, and many other school choice plans serve a self-selected clientele. If analysis shows that these schools raise achievement relative to the alternative schools their students would have attended, it is reasonable to ask whether the effect is due to special features of the instructional program or whether it is merely the result of positive peer effects. A large literature has examined peer effects in education. However, the peers a student encounters at a school (like other features of that school) are endogenous to residential location decisions and other forms of school selection. Thus, estimated peer "effects" are apt to be confounded with the influence of unobservable family characteristics that underlie these decisions. Most of the literature on peer effects in education ignores this issue. Notable exceptions include two studies relying on longitudinal data to exploit within-school (and within-student) variation over time in peers to separate the influence of peers from other factors affecting the choice of school (Hanushek et al., 2001, 2002; Hoxby, 2000). In this paper I pursue a different strategy. Because the lottery that assigns students to a choice school also assigns them to the peers they will encounter at that school, admissions lotteries represent a source of exogenous variation in peer characteristics that can be exploited to identify peer effects independent of unobservable student (and family) characteristics. Unlike studies that rely on longitudinal data, I do not require the assumption that unobservable student influences are fixed to identify peer effects. This may be of particular benefit when studying achievement in the middle school years, when students are changing in many ways that influence both achievement and their susceptibility to peer influences.

Using lottery randomization to study the effects of school choice has some well-known limitations. Inferences are limited to the subpopulation of students that participate in lotteries. This excludes both the large number of students who do not express an interest in attending, as well as the smaller number admitted in other ways (e.g., through preferences shown siblings of students already enrolled or to students living in the neighborhood). Conclusions about peer effects must be similarly qualified. Estimates of the effect of peers on lottery participants are not necessarily the same as their effect on the average student.

The foregoing limitations affect the external validity of our findings: the extent to which these conclusions generalize beyond the students participating in admissions lotteries. There are several threats as well to internal validity. Despite the fact that lotteries mimic the experimental assignment of subjects into treatment and control groups, it does not immediately follow that estimated effects are uncontaminated by other influences. Given the importance of this question, I briefly survey these threats here, indicating the measures taken to meet these challenges. Several of these points are taken up at greater length below.

a. With random assignment, treatment and control groups differ only by chance, but this is reassuring only when the groups are sufficiently large that chance variation is inconsequential. When groups are small, treatment and control groups can be unbalanced with respect to important variables. This does not bias estimated treatment effects, but it does increase sampling error.

I take two steps to address this problem: (1) I combine several magnet schools into a single treatment (the composite control group is larger than the several control groups that arise when each school is considered a distinct treatment); (2) I include a number of student covariates in the model, improving the precision of the estimates and correcting for imbalance of treatment and control groups with respect to these characteristics.

b. Non-compliance with lottery outcomes (students accepted to magnet schools but enrolling elsewhere) creates the possibility of systematic, unobserved differences between magnet and non-magnet students. The usual solution is to use lottery assignment as an instrumental variable for attendance. However, with multiple instruments (at least one for each lottery entered) and heterogeneous responses to treatment, even random assignment may not provide valid instruments for the effect of attendance (Heckman, 1997). Accordingly, I test the exogeneity of the instruments using the omnibus overidentification test. In every instance I fail to reject the null hypothesis (by a wide margin), suggesting the instruments are valid.

c. Sample attrition can introduce systematic differences between treatment and control groups. There is a significant amount of atttrition in these data. As one might expect, it is greater among lottery losers than winners. However, bias from attrition appears to be slight, for two reasons. (1) Student covariates 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.

d. 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 are undiminished when teacher fixed effects are added to the model (indeed, they rise).

Peer effects can also be confounded with heterogeneous response to treatment, as described below. I 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, it suggests the contrary.

To conclude, the estimates of treatment and peer effects appear to possess internal validity, though there remain the limitations on external validity noted above. What, then, are the findings?

1. There are positive benefits on mathematics achievement from attending magnet schools in the district under study, though these are not uniform across grades or type of magnet school.

2. Some of the positive benefit appears to be attributable to peers. This is especially true of the district's selective, academic magnet, where the positive benefit falls to zero when the percentage of minority or low income peers is included in the model.

3. Peer effects have a substantial influence of their own on achievement. The effect of attending a school where 75 percent of the other students are poor and black compared to one in which 25 percent of the other students are poor and black is comparable to half a year or more normal growth in mathematics.

In Section II of this paper I describe the operation of the magnet school program in this district. Section III develops the model and describes in greater detail the threats to internal validity and the strategies for meeting these threats. Section IV presents the results. Discussion of their implications is taken up in Section V, along with concluding remarks.

II. Description of the Magnet Program

The district operates two types of magnet schools: selective academic magnets (with eligibility determined by grades and test scores) and non-academic magnets. At the middle school level there is one academic magnet serving grades 5-8. While there is a second academic magnet serving grades 7-12, the effectiveness of this school is not examined in this paper, which focuses on students in grades 5 and 6.

There are four ways to be admitted to a magnet: (1) lottery; (3) sibling preference; (3) geographic priority zones; (4) promotion from a feeder magnet. Students who have a sibling at a magnet school need not enter the lottery to enroll in that school. Geographic priority zones have been defined for some magnets. In principle, separate lotteries are to be held for students residing in one of these zones. In practice, the number of applicants from zones has not exceeded the number of reserved spaces, so that these lotteries were not held. There is one middle school magnet (West) to which students are automatically promoted if they attended a particular elementary magnet. (All school names are pseudonyms.) Students in the East middle school magnet who wish to attend the 7-12 academic magnet automatically transfer in grade 7 (or later) if they meet the admissions requirements.

Separate lotteries are held for each magnet school. Students can enter more than one lottery. Students who are not accepted outright on lottery day are placed on wait lists. Those admitted outright on lottery day must decide whether to accept any of the positions offered them. If they accept a position, they go to the bottom of the wait list for any other magnets. Students are accepted off the wait list as positions become open.

Middle school lotteries are held in the spring of the fourth grade for the following academic year. While lottery data have been furnished from spring 1997 on, achievement data are available only for school years 1998-99 through 2003-04. As it is desirable to control for fourth grade achievement, this limits our study to the five cohorts entering middle school between fall of 1999 and fall of 2003.

The number of magnet programs has increased over time.

Lottery Year	Magnet Schools	Grades Observed
1999	Academic, North, South	5-8
2000	Academic, North, South, East	5-8
2001	Academic, North, South, East	5-7
2002	Academic, North, South, East, West	5-6
2003	Academic, North, South, East, West, Central	5

The fifth non-academic magnet, Central, added in 2003, filled most of its places in that year through geographic priorities. As such, it contributes very few observations to treatment and control groups. Accordingly, for purposes of this study, Central is treated as a non-magnet school (which it was prior to the 2003-04 academic year).

Because the final year of achievement data is 2003-04, only the first two cohorts can be followed through all middle school grades. The limited amount of data for grades 7 and 8, plus the complications posed by a second academic magnet school conducting a lottery for 7th graders, prompted the decision to restrict attention to magnet school performance in grades 5 and 6. This restriction will be lifted in future work, when achievement data will be available to follow all cohorts to the end of 8th grade.

As shown in Table 1, there were approximately 5000 applications to middle school magnets this period.[1] Nearly half were to the academic magnet, where the probability of admission was just under 50 percent. I distinguish two categories of lottery winners: those admitted outright on lottery day, and those whose place on the wait list was reached by the start of the school year (delayed winners). Of the 1209 participants in the academic lottery not admitted by the start of 5th grade, all but 539 won admission to one of the other magnets, though many chose not to attend: only 287 enrolled as 5th graders in a non-academic magnet, compared to 834 who attended a non-magnet school. More than a quarter were not present for testing as 5th graders, in the great majority of cases because they were no longer enrolled in the district.[2]

Approximately 2600 students participated in one or more of the lotteries for the non-academic magnets. Numbers for the West magnet are low because the school was

[1] Students who reside in the district but do not attend a district elementary school are eligible to participate in middle school lotteries. These students are much less likely to enroll in the district as 5th graders if they lose than if they win the lotteries they enter. Because no subsequent achievement data are available on those who do not enroll, these different enrollment probabilities likely introduce systematic differences between the observable treatment and control groups. Following the practice of other researchers (e.g., Cullen, Jacob, and Leavitt), we omit from the sample all students who were not enrolled in the district as fourth graders. They are not counted in Table 1.

[2] 333 of the participants in the academic lottery never enrolled as 5th graders. An additional 103 students enrolled but were not tested. District records do not always make it clear when students leave the system. However, it appears that most of these students had left the system prior to testing (in spring), as 60 percent of those not tested were also not enrolled the following year.

organized as a magnet in 2002-03 and because most places have been taken by students promoted from a feeder magnet. In contrast to the academic magnet, most lottery participants win a place in the non-academic magnets, either outright or through delayed notice.[3] Because most participants apply to more than one school, few fail to secure a place in any. 388 of those who applied to one or more of the non-academic magnets were not admitted to any, but as some of these students also applied to the academic magnet, the number not admitted anywhere was still smaller, 199. The number enrolling in a non-magnet school is much larger, 846, so that many of those with an option to attend a magnet school chose not to. About 15 percent were not in the system as 5th graders.

Table 1 has several important implications for this study. 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. This defect is remedied by the use of instrumental variables to estimate the effect of treatment on those who enter magnet schools via the lottery.

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

[3] I count students as "delayed winners" based on their original position on the wait list: thus, a student is a delayed winner if his number is reached by the start of the school year. In practice, students who have already accepted a position at another magnet school are dropped to the bottom of the wait list maintained by the district and will not receive notice until everyone ahead of them has been offered a place. I rely on the original wait list because it is determined solely by lottery randomization and not by subsequent decisions by students and parents. The counts of outright and delayed winners in Table 1 therefore correspond to the variables used as instruments in the analyses reported below.

individually may be quite limited.[4] Accordingly, I 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.

Finally, 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. I return to this issue below.

Table 2 presents descriptive statistics on the characteristics of lottery participants. The biggest differences are between students applying to the academic magnet and applicants to other schools, who are much more likely to be low-income and black. As expected, applicants to the academic magnet also have significantly higher test scores as fourth graders. The average gain from one grade to the next is about 20 scale score points in mathematics. Thus, applicants to the academic magnet are on average a year ahead of applicants to the composite, non-academic magnet in mathematics.

Students who enroll in the academic magnet look much like the typical applicant. However, in three of the four non-academic magnets, the percentage of blacks is higher among enrollees than among participants generally. This probably reflects perceptions of

[4] Loosely, one might regard the students turned down by all magnets as the control group to whom we compare the lottery participants who are admitted to (and attend) magnet schools. However, the distinction between treatment and control is blurred by the use of instrumental variables to deal with the problem of non-compliance. The instrumental variables estimator replaces the binary treatment indicator with the predicted probability of receiving treatment. Lottery losers have zero probability, but winners' probabilities are less than one due to non-compliance. Because there is variation in these probabilities, some winners contribute less to the estimated treatment effect than others: implicitly, they are part treatment, part control. In principle, variation in these probabilities makes it possible to estimate a treatment effect even if there is no "pure" control group with zero chance of obtaining treatment, though this is hardly recommended.

the quality of neighborhood schools, leading fewer blacks to turn down places in magnets. The percentage of blacks among applicants who lose all the lotteries they enter is still higher (as is the share of low income students in the North and South lotteries). Because the likelihood of losing all lotteries is strongly related to the number of lotteries entered, this says more about application behavior than about the fairness of the lotteries, though the small number of students to lose all lotteries entered raises the probability that discrepancies of these magnitudes arise by chance.

Table 3 presents information on selected peer characteristics. These variables are calculated by averaging over fifth graders in schools attended. For students who switch schools in mid-year, they are weighted averages reflecting the proportion of the year spent in each school. Students attending the magnet school are designed as enrollees. Not only is the mean value of the peer characteristic likely to be different for enrollees than for participants generally, but the standard deviation is obviously much lower, given these students all attend the same school. (The standard deviation is not zero, as data are pooled across years, and there is variation from one fifth grade cohort to the next.) The much larger standard deviations for losing participants (and all participants, generally) underscores the role that lotteries play in determining peer characteristics in conjunction with residential decisions. Clearly, the attributes of one's peers can be very much affected by the outcome of the lottery.

There are some apparent discrepancies between Tables 2 and 3. For example, 65 percent of the lottery participants enrolling in South are black, yet the percentage of black peers at this school is 73. The explanation of the difference is the presence of peers who were admitted by non-lottery means (notably geographic priority zones). With the exception of the academic magnet, admission through this route tends to raise the share of students who are black, low-income, and special education and to lower mean fourth grade test scores.

III. Model and Analytical Strategy

To facilitate the exposition, assume (provisionally) the following: (1) A lottery randomly assigns students to a magnet school or a non-magnet school. There is full compliance with this assignment. (2) There is only one magnet school. (3) The lottery outcomes indicator is binary (win or lose). Achievement of student i in school j is represented as

$$Y_{ij} = \mu(X_i) + u_{ij} \ , \tag{1}$$

where X is a set of observed student characteristics. (I omit subscripts for year.) In particular, when student i would attend neighborhood school N if he does not enroll in magnet school M, his outcome is described either by

$$Y_{iM} = \mu(X_i) + u_{ijM} \tag{1a}$$

or

$$Y_{iN} = \mu(X_i) + u_{iN} \ . \tag{1b}$$

The treatment effect for student i is defined as $\delta_i = Y_{iM} - Y_{iN}$. Let d_i denote treatment, with $d_i = 1$ if he enrolls in magnet school M, 0 if he enrolls in regular school N. Then

$$Y_{ij} = \mu(X_i) + (u_{iM} - u_{iN})d_i + u_{iN} = \mu(X_i) + \delta_i d_i + u_{iN} \tag{2}$$

Note that the treatment effect is heterogeneous if for no other reason than because N varies across students.

While we might want to estimate the average of δ_i over all lottery participants (known as the Average Treatment Effect, ATE), given that the treatment effect is heterogeneous, the best we can do is to estimate the average δ_i over the students receiving treatment (known as the Effect of Treatment on the Treated, ETT). (However, under our provisional assumption that the treated are simply a random subset of participants, ATE and ETT coincide.) Define $\delta = E(u_{iM} - u_{iN} \mid d_i = 1)$, where the expectation is over the distribution of u_{iM} and u_{iN} among the treated participants. Define $\eta_i = (u_{iM} - u_{iN}) - E(u_{iM} - u_{iN}|d_i=1)$. Then

(3) $$Y_{ij} = \mu(X_i) + (u_{iM} - u_{iN})d_i + u_{iN} = \mu(X_i) + \delta d_i + \eta_i d_i + u_{iN}$$

where the error term has two components: a heterogeneous response to treatment ($=\eta_i d_i$) and the disturbance in the equation for Y_{iN}, u_{iN}.

Relax now the provisional assumptions, beginning with the assumption that the lottery outcome determines school assignment to allow for non-compliance. Self-selection of non-compliers has two implications. First, ETT no longer equals ATE. Second, d_i may be correlated with u_{iN}, making the treatment indicators endogenous. The conventional solution is to use lottery outcomes as instruments for d_i. However, this approach runs into new difficulties as soon as we relax the assumption that there is only a single magnet school. Suppose, rather, that there are multiple schools, each conducting its own lottery. Students can enter more than one lottery, though obviously they can attend only one magnet school. This means the lottery outcome indicator is multivalued. For concreteness, suppose there are two magnet schools so that the indicator is an ordered pair, with the first element a binary indicator of the outcome of the first lottery, and the second element a binary indicator of the outcome of the second lottery. The problem: this pair of outcomes may no longer constitute valid instruments for the endogenous treatment decision, inasmuch as the combination of two indicators may convey independent information about $\eta_i d_i$, the response heterogeneity term. Let η_{i1} denote response heterogeneity to the first magnet school. Compare $E[\eta_{i1} - E(\eta_{i1} \mid d_{i1} = 1) | d_{i1}=1, R_i = (1,0)]$ with $E[\eta_{i1} - E(\eta_{i1} \mid d_{i1} = 1) | d_{i1}=1, R_i=(1,1)]$. When $R_i = (1,0)$ student i doesn't have the option of attending school 2, but when $R_i = (1,1)$ he does. If the choice of school is based on private information about η_{i1}, this conditional expectation will be greater when student i might have selected school 2 than when school 2 is not in the choice set. Then $E[\eta_{i1} - E(\eta_{i1} \mid d_{i1} = 1) | d_{i1}=1, R_i = (1,0)] \neq E[\eta_{i1} - E(\eta_{i1} \mid d_{i1} = 1) | d_{i1}=1, R_i=(1,1)]$ and neither equals zero, as required if the instruments are to be valid.

The same argument arises when we relax the final provisional assumption and use more than one indicator to represent the outcome of a single lottery: e.g., outright winners and delayed winners. Many students who are not accepted outright may make other plans in the interim and decline a position when notice arrives late. Those who accept delayed offers may expect unusual benefits from attending.

The assumption that parents and students use private information about η_{i1} to choose a school is critical to this conclusion. There are plausible scenarios where this does not happen. For example, it could happen that while schools are recognized to be of different quality, everyone shares the same perception of these differences. Preferences for one magnet school over another may be based on factors unrelated to achievement (e.g., distance from home). If η_{ij} has no influence on decisions to accept a position in a magnet school, R_i is uninformative about η_{ij} and the conventional IV estimator works.

As a second example (in reality a special case of the first), suppose preferences over magnet schools are lexicographic: everyone prefers A to B, and no one with a choice of both selects the latter. Then $E[\eta_{i1} - E(\eta_{i1A} \mid d_{i1A} = 1) | d_{i1A}=1, R_i = (1,0)] = E[\eta_{i1A} -$

$E(\eta_{i1A} \mid d_{i1A} = 1) | d_{i1A}=1, R_i=(1,1)]$—availability of the second option is simply irrelevant and contains no information about response heterogeneity. This condition is very nearly met in our data, when the two treatments are defined as the academic magnet and a composite of the non-academic magnets. Very few students admitted to the former choose to attend the latter. Of 837 outright winners in the academic lottery, only 21 enrolled in a non-academic magnet. The ratio is not much higher among delayed winners: 22 of 712.

Further evidence supporting the validity of the instruments is presented below, in tests of overidentifying restrictions.

Peer Effects

We now make explicit the dependence of achievement on peer characteristics. Let P_j denote the value taken by one or more peer variables at school j.

$$(4) \qquad Y_{ij} = \mu(X_i) + \beta P_i + u_{ij} \ ,$$

from which we can obtain

$$\begin{aligned}(5) \qquad Y_{ij} &= \mu(X_i) + \delta d_i + \beta P_i + \eta_i d_i + u_{iN} \\ &= \mu(X_i) + \delta d_i + \beta[P_M d_i + P_N(1-d_i)] + \eta_i d_i + u_{iN}\end{aligned}$$

where η_i and u_{ij} are redefined to be net of the peer effects. While the peer variable is endogenous through its dependence on d_i, we can remedy this by using instruments to replace d_i with its prediction conditional on lottery outcomes, $\hat{d}_i$.

However, there is another sense in which peer characteristics are endogenous. While the lottery determines whether a student is assigned P_M or P_N, P_N itself depends on parental decisions, notably residential location. As such, it may be correlated with other determinants of achievement (families placing a strong value on education or with large incomes select homes in good school zones, etc.). Let $E(u_{iN}| P_N) = \gamma P_N$. We can control for this correlation by introducing P_N as an additional regressor, yielding

(6) $\quad Y_{ij} = \mu(X_i) + \delta d_i + \beta[P_M d_i + P_N(1-d_i)] + \gamma P_N + \eta_i d_i + v_{iN},$

where $v_{iN}, = u_{iN}- E(u_{iN}| P_N)$. Because γP_N picks up the correlation between γP_N and $u_{iN,}$ the estimate of β is not affected by this source of endogeneity bias. This may be somewhat easier to recognize in a re-arrangement of equation (6):

(7) $\quad Y_{ij} = \mu(X_i) + \delta d_i + \beta[P_M-P_N]d_i + (\beta+\gamma) P_N + \eta_i d_i + v_{iN,}$

where β is now seen to be the coefficient on the difference between magnet school peers and non-magnet school peers for students enrolling in the magnet school. Conditional on P_N, the lottery assigns students either to P_M or P_N. This means the instrument $[P_M-P_N]\hat{d}_i$ (or when equation (6) is estimated, $[P_M\hat{d}_i + P_N(1-\hat{d}_i)]$) is exogenous, provided P_N is also in the equation. Thus the final model contains two peer variables: one representing lottery-based exogenous variation in peers, the other representing residence-based, endogenous variation in peers. The causal effect of peers on achievement is represented by the coefficient on the former.

However, there is still one more source of potential bias in this estimate. The peer effect is identified through the interaction of $[P_M-P_N]$ with d_i, meaning that it might be confounded with unobserved response heterogeneity (η_i), which is also interacted with d_i. For example, suppose that academics are taken more seriously in the magnet schools than in other middle schools and that instruction is more rigorous. Some students are readier to take advantage of this program (they have a high value of η_i), others are not. It is not implausible that P_N, as it is correlated with factors influencing residential location, might signal which students arrive at the magnet school readier to meet this challenge. If so, P_N

serves as a proxy for unobserved response heterogeneity, biasing in the estimate of β. Below I conduct some sensitivity tests exploring this possibility.

Variables

The student characteristics entering X are the variables displayed in Table 2: black, low income (as measured by eligibility for the free and reduced-price lunch program), special education, ESL, and grade 4 reading and mathematics achievement. To this group I add student gender.

The list of peer characteristics extends beyond school-level aggregates of these variables and includes the following: peer absenteeism, school size (enrollment), the proportion of the school year spent at the school by the average peer (a measure of potential disruption caused by intrayear mobility), and the number of disciplinary incidents occurring at the school relative to enrollment. Though point estimates of these effects were sometimes large, standard errors were still larger. None of them was statistically significant. The only peer characteristics that were regularly significant were black and low income, and I present only results for those. (Other results are available on request.)

As we have seen, multiple peer variables are required for each student: P_M, the value at each magnet school M to which a student applies, and P_N, the value at the regular middle school a student would attend if not enrolled in a magnet school. Calculation of P_M is straightforward (using students actually enrolled at M), but P_N is an unobserved counterfactual for students who attend a magnet school. I use each student's elementary school plus the variables in X to predict P_N.[5] These $\hat{P}_N$'s are used in equation (7) for all students to avoid introducing a difference between students whose P_N is observed (e.g., lottery losers) and those whose P_N is a counterfactual (lottery winners).

The main treatment indicators are dummy variables for enrollment in a magnet school. Students are regarded as enrolled in a magnet school if they finished the school year at the magnet. The relatively small number of students who started the year at a magnet before transferring out are treated as non-magnet students. I also explore two alternative measures of treatment: a cumulative exposure measure, equal to one in during the first year in a magnet and two during the second year, and a grade-specific enrollment indicator, allowing for different treatment effects in each grade.

The endogenous regressors (apart from $\hat{P}_N$, as noted above) are the treatment indicators and the interactions of those indicators with P_M and $\hat{P}_N$. For treatment measures of cumulative exposure, the instruments are outright winner and delayed winner multiplied by the maximum number of years a student could have spent at the magnet school (1 for 5th graders, 2 for 6th graders). Grade-specific treatment indicators are matched to grade-specific instruments (outright winner and delayed winner interacted with grade dummies).

The instrument for peer variables is constructed as $\sum P_j \hat{d}_j + \hat{P}_N(1\text{-}\sum \hat{d}_j)$,

---

[5] The sample used for this prediction comprised lottery participants who lost all the lotteries they entered and therefore attended non-magnet, neighborhood schools. This is the arguably the best comparison group for estimating where lottery winners at the same elementary school would have gone had they not won, but the sample is small: in many elementary schools, there are only a few such losers. To test the robustness of my results, I explore alternatives to these estimates of P_N below. Because students at a given elementary school often attend a variety of middle schools, the fit of the first-stage regression in the 2SLS estimator might be improved if P_N were predicted from a sample that included all students not attending a magnet school the following year, whether they were lottery participants or not. I have experimented with such a prediction of P_N, but the conjecture is not borne out: the fit of the first-stage regression is actually poorer, and the 2SLS standard errors slightly higher.

where $\hat{d}_j$ is an estimate of the probability that a student attends magnet school j. Predictors in this equation are the lottery outcome indicators and student demographic variables. $\hat{d}_j$ is identically zero for students who were not winners in lottery j. The composite peer characteristic, P_j for j = the composite magnet, was constructed as a weighted average of values at the individual magnet schools, with weights based on planned enrollment of lottery students (for this purpose, I used the number of outright winners).

To control for differences in test forms across grades and year, all models also include a full set of grade by year interactions. Finally, because lotteries randomize only among the participants in a given lottery, indicators are required of the lotteries a student entered.[6]

IV. Results

Baseline estimates appear in Table 4. These models do not contain peer characteristics. The three models correspond to the three ways of measuring magnet treatment described above. When the treatment variable is a binary indicator (Model 1), point estimates are positive but not statistically significant. When the treatment variable measures cumulative exposure (Model 2), both the academic magnet and the cumulative magnet have positive and statistically significant effects on mathematics achievement. The effect in the academic magnet is about one tenth of a normal year's growth. In the

[6] These variables capture any selection effects related to the choice of lotteries. Given that I combine multiple schools (and lotteries) into a single composite, lottery participation indicators must be defined for each combination of lotteries a student can enter. For example, students entering three lotteries have greater chances of winning the composite lottery than students entering only one or two. Winners in this group therefore need to be compared with losers in the same group, as there arise non-random differences between winners and losers across groups reflecting the factors that determine the number of lotteries entered.

composite magnet it is nearer three-tenths. The temptation to compare these estimates to one another should be resisted. The treatment effect is defined relative to expected outcomes in the schools that magnet enrollees would have attended, had they not gone to the magnet school. The mix of alternative schools differs for students at the academic and non-academic magnets, so no inferences can be drawn from these results about the comparative effectiveness of the academic and non-academic magnets.

Model 3, which allows for grade specific treatment effects, sheds further light on the effectiveness of magnet schools. The pattern is quite different between the academic and composite magnets. Students at the academic magnet benefit in grade 5, but the gains are largely given back in grade 6. By contrast, all the gains in the composite magnet occur in grade 6. An explanation for this difference is still to be uncovered.

The other coefficients in Table 4 pertain to student characteristics. As one would expect, prior scores in mathematics (and to a lesser extent, in reading) are strong predictors of later performance. Black students and low-income students score significantly lower, but students learning English as a second language do somewhat better in mathematics than the average student, controlling for the other demographic variables.

These models are overidentified. For every treatment variable there are two excluded instruments, based on the two lottery outcome indicators (outright win and delayed win). Results of omnibus overidentification tests support the exogeneity of the instruments in all three models.

Table 5 presents estimates of models that include peer characteristics. The introduction of peer variables serves two purposes. First, it explores the extent to which the positive magnet effects reported in Table 4 are attributable to more favorable peers. Second, by distinguishing exogenous, lottery-based variation in peers from endogenous, residence-based variation in peers, we obtain estimates of peer effects free of bias resulting from the self-selection of peers. Because the Model 3 estimates in Table 4 showed that magnet effects are neither constant across grades nor additively cumulative, I use the more flexible Model 3 specification to explore peer effects. In addition, as noted earlier, results are shown only for two characteristics: percent black and percent low income.

The results are dramatic. Including either of these peer characteristics completely overturns the positive grade 5 effect in the academic magnet (Models 1 and 2). The grade 6 gains in the non-academic composite survive. The implication is that students who lost in the lottery for the academic magnet but attend a middle school whose racial and SES composition is the same as the academic magnet do just as well as lottery winners. The same is not true, however, of the losers in the composite lottery.

In addition, the estimates of the peer effects are quite large. Reducing the percentage of black peers from 75 percent to 25 percent yields an estimated gain of 12 points on the mathematics test—about six-tenths of a year's normal growth. This point estimate is somewhat imprecise, but even if the effect were only half this large, it would still be extremely important.

Finally, the coefficients on the residence-based peer variable are much smaller and statistically insignificant, suggesting that residential location (as expressed in P_N) is not strongly correlated with other influences on achievement, at least in these data.[7]

[7] As all of the students in this sample are lottery participants, their families may not be typical residents of their school attendance zones.

Model 3 incorporates both percent black and percent low income. The coefficient on percent black remains large and significant, while the effect of low income peers falls nearly to zero. Some of this may be due to collinearity of the two variables. Standard errors also rise.

The estimated peer effects are so large, the question arises: are they capturing something other than the effect of peers? I remarked above that the lottery-based peer variable, which is interacted with treatment in the model, could be confounded with unobserved heterogeneity in the treatment response. A direct test of this hypothesis is difficult to mount. However, we can test it indirectly by considering whether the estimated peer effect is diminished when controls are introduced that capture observable heterogeneity in treatment response. If the estimated peer effect is in fact a proxy for the way different individuals respond to treatment, then controlling for race, income, and other student-level characteristics, all of which are correlated with characteristics of peers, should reduce the magnitude of the ostensible "peer effect." Thus, I re-estimate Models 1 and 2, including interactions of the treatment indicators with all of the student-level variables in the model, including grade 4 achievement. This does not produce the anticipated change. Point estimates of the peer effects are larger than before.

In addition, it is somewhat difficult to see why an association between peer characteristics and response heterogeneity would produce strongly negative coefficients on these peer variables. Recall the story sketched above of a student who was not prepared for the more rigorous program offered in the magnet school, by virtue of the quality of the elementary school attended. Coincidentally (but not causally), peers in this elementary school were largely poor and minority. Had this student continued in his neighborhood middle school, P_N (percent minority or percent low income) would have been high, implying a strongly negative value of $P_M - P_N$. If this is the typical case, then such strongly negative values of P_M-P_N will be associated with poorer test performance in grades 5 and 6, imparting a positive bias to the coefficient on P_M-P_N. But the estimated coefficients in Table 5 are strongly negative, so this cannot be the typical case. Positive response heterogeneity must be associated with *less* (not more) favorable values of P_N if the peer effects reported in Table 5 are an artifact of the correlation between peers and response heterogeneity. While it is not impossible to construct scenarios that do the job ("students who were misfits in their elementary schools are overjoyed to be accepted into a magnet school and redouble their efforts to take advantage of this opportunity to escape their neighborhoods"), such stories are not particularly compelling, given the immaturity of these students and the greater plausibility of scenarios with the opposite implication.

If the estimated peer effects are not picking up response heterogeneity, perhaps they are proxies for other characteristics of the schools that magnet losers must attend. A likely candidate is the quality of the teaching staff. If (as often alleged) less effective teachers tend to be assigned to schools with high percentages of poor minority students, peer effects are confounded with teacher quality. To test this hypothesis, I re-estimate Models 1 and 2, introducing teacher fixed effects into the model. This greatly reduces the variation in the data available to estimate peer effects. Average differences in peer characteristics between teachers are absorbed in the estimated teacher effects. Only the within-teacher variation remains. Fortunately for our purposes, these years were a transitional period in which the district moved from a court-ordered desegregation plan to a neighborhood-based school system. The resultant churning of school assignments, both for teachers and students, has contributed to variation in peer characteristics over the course of teacher's career. Indeed, in the sample used to estimate the achievement model, just over one-half of the estimated variation in neighborhood peers ($\hat{P}_N$) is within-teacher as opposed to between teachers.

Results are presented in columns 6 and 7 of Table 5. As expected, standard errors are much higher, given the amount of data required to estimate the teacher effects. However, there is no indication at all that the large peer effects estimated earlier were the result of peers acting as proxies for teachers. Controlling for teacher quality, the estimated peer effects are even larger than before.

I have also examined the robustness of these results to alternative ways of dealing with the fact that $\hat{P}_N$ is endogenous and, in the case of magnet enrollees, a counterfactual. As noted in footnote five, the samples used to predict P_N are small, comprising lottery participants who lost all the lotteries they entered. Because students at a given elementary school often attend a variety of middle schools, there could be considerable sampling error in estimates based on small numbers of lottery losers. The fit of the first-stage regression in the 2SLS estimator might be improved if P_N were predicted from a sample that included all students not attending a magnet school the following year, whether they were lottery participants or not. I have experimented with such a prediction of P_N, but the conjecture is not borne out: the fit of the first-stage regression is actually poorer, and the 2SLS standard errors slightly higher. The estimated coefficient on percent black remained strongly negative (-25.1), but the coefficient on percent low income fell to -8.2 and was no longer statistically significant.

Because P_N is based on residential location, the instrument for the peer variable, $\sum P_j \hat{d}_j + \hat{P}_N (1\text{-}\sum \hat{d}_j)$, is correlated with all the factors that influence this decision. It was to ensure the exogeneity of the instrument that $\hat{P}_N$ was included in the model. An alternative securing the same result uses an indicator of the elementary school attended in fourth grade in place of $\hat{P}_N$. This was the key regressor in the equation that generated $\hat{P}_N$. Using dummy variables for elementary schools is a more general specification in that it captures other residence-related influences on achievement in addition to peers. Because attendance zones changed in the district over this period, altering the relationship between residence and school, the school dummies are interacted with indicators for years. The findings reported above are robust to this alternative specification. The coefficient on percent black remains strongly negative -28.0 (with a standard error of 8.6), as does the coefficient on percent low income (-20.2, with a standard error of 8.8).

Finally, in the event there remain misgivings about the instrumental variables estimator, I report results of a model that uses lottery outcomes rather than treatment to explain achievement (known in the literature as "intention to treat" equation). The lottery outcomes are outright win and delayed win, for both the academic and composite lotteries, each interacted with grade level. For peer characteristics, the intention to treat model employs the value at the academic magnet for those students entering only the academic lottery, multiplied by a dummy variable equal to one if the student was either an outright or a delayed winner. The regressor for students entering only the composite lottery was constructed analogously. For students who participated in both lotteries, I exploit the near-lexicographic preferences and assign them the peers of the academic magnet if they won that lottery, otherwise the peers of the composite magnet if they were winners in that lottery. Students who lost the lotteries they entered were assigned $\hat{P}_N$.

The results are reported in Table 6. As expected, the intention to treat estimates are attenuated with respect to the earlier instrumental variables estimates, inasmuch as non-compliance with lottery assignment dilutes the effect. However, the earlier patterns are still apparent. There is a positive effect in fifth grader for the academic magnet and a positive sixth grade effect for the composite magnet. Including peer characteristics erases this effect in the academic magnet, but not in the composite. The peer variables continue to have large negative effects on achievement.

Attrition

Attrition rates by lottery participants are presented in Table 6. There are two indicators of attrition: (1) there is no record of a student enrolling in a year subsequent to the lottery; (2) there are no spring test results for a student. The second certainly does not mean a student has left the system, as there are other reasons that no test scores may be reported. However, for our purposes, a student without test scores poses the same problem as a student who has left the system: neither contributes to the estimation of magnet school effects.

Attrition is ignorable if there are no systematic differences between winners and losers who leave the system. The evidence suggests otherwise. To begin, losers are significantly less likely to remain in the district than winners. The discrepancy is particularly pronounced between fouth and fifth grades, after lottery results have been made known. Losers in the composite lottery have been about 50 percent more likely to leave the system than winners. In the academic lottery the gap is still more pronounced.

More than a fifth of lottery losers do not return for fifth grade, compared to one-eighth of the winners. It appears that winning the lottery makes the difference between leaving the system and staying for some families. As these families may be more concerned than the average about the quality of their children's schools, or at least have above average means to act on these concerns, attrition may well introduce systematic differences between treatment and control groups related to achievement.

Table 6 also provides a breakdown on the percentage of students without test scores in mathematics. Evidence of systematic differences between winners and losers is considerably weaker. Losers are more likely to be missing test scores as fifth graders than winners of the academic lottery, but the reverse is true of the composite lottery, and in the other grades the comparison is a wash. Accordingly, for the rest of this analysis I will use the term attrition to refer to students who do not return to the system in the fall.

Student characteristics have been included in the model in part to defend against attrition-induced bias. To the extent that attrition-related achievement differences are a linear function of student demographic variables and fourth grader test scores, these controls restore the equivalence of treatment and control groups. But attrition could be a function of unobservable variables, and the differences between winners and losers may not be simple linear functions of student characteristics even when the latter are observed. For example, if high achieving students are more likely to leave the system when put into a school with uncongenial peers, then differential attrition could be contributing to the very large estimated peer effects by disproportionately removing high achievers from schools serving predominantly minority and low income populations.

To investigate these matters further, I estimate a model of attrition between fourth and fifth grades among participants in the academic lottery. Even though the regressors are mostly limited to variables already in the achievement model, estimating the model gives us a sense of the relationship of attrition to achievement. If attrition turns out not to be closely related to observed variables that predict achievement, it becomes less plausible that unobservables with comparable predictive power are driving the decision. In addition, I explore the possibility just raised of selective attrition among high performing students faced with the prospect of attending schools that serve predominantly low income, minority peers. Four variables in the model are functions of peer characteristics. The first characterizes the peers a student will have if he loses the lottery (P_N in the above notation). Second is the interaction of P_N with lottery outcome. I use the indicator for whether a student was an outright winner, as application to private schools generally must be made in the spring, whereas delayed notification of acceptance to the academic magnet often is not sent out until summer or just prior to the start of the school year. One would expect the coefficient on the interaction of P_N with winning to offset the effect of the stand-alone P_N: peers in the neighborhood school shouldn't matter to a student who has won the lottery. I also interact P_N with a student's fourth grade mathematics test score (P_NMath) as well as introducing a three-way interaction of P_NMath with lottery outcome. A positive coefficient on P_NMath would confirm that such students are more likely to leave the system, in which case one would expect a negative coefficient on the three-way interaction of this variable with winning. Two models were estimated, one in which the peer characteristic was percent black and the other in which it was percent low income.

Results are displayed in Table 7. In both equations the point estimate on the peer characteristic is positive, as expected. Winning the lottery tends to offset this influence, also as expected. However, the interaction with fourth grade mathematics achievement works opposite to expectations: higher performing students are less likely to exit when their neighborhood school peers are unfavorable, unless (remarkably) they win the lottery.

To facilitate the interpretation of these results, I have calculated attrition rates for students whose fourth grade mathematics score was one standard deviation above or below the sample mean, under the following scenarios: neighborhood peers unfavorable to achievement (75 percent black or low income) and the student loses the lottery; unfavorable peers but the student wins the lottery; peers that are favorable to achievement (only 25 percent black or low income) and the student loses the lottery; the same peers, but the student wins the lottery.

Results are presented in Table 8. Overall, low performing students are more likely to leave the system than high performers if their peers are unfavorable to their own achievement (columns 1 and 4). The reverse is true when peers are favorable—then high achievers are more likely to exit (columns 2 and 5). This pattern works against finding strong peer effects. However, the critical issue is whether lottery winners differ from lottery losers in this regard. The differences turn out to be quite modest. Losers are more likely to leave across the board, but the differential is greater for low than high achievers when peers are unfavorable. This pattern is reversed, with a greater spread in attrition rates among high than low achievers, when peers are only 25 percent poor or minority.

These patterns run counter to those hypothesized above, suggesting that differential attrition is not behind the large peer effects estimated in the achievement equation.

All of the coefficients in Table 7 are estimated quite imprecisely. This is in part a consequence of the numerous interactions in the model. However, the linear combinations of these coefficients reported in Table 8 also have large standard errors and overlapping confidence intervals, so that we cannot reject the hypothesis that there is no difference between high and low achieving students, whatever their peers and lottery outcomes. And in fact, when the attrition model is re-estimated using a different indicator of lottery outcome (combined outright and delayed win), different patterns are revealed.

Finally, to assess the net effect of attrition on the estimates reported in Tables 4 and 5, I estimate a more general model that includes participants in the composite lottery as well as the academic lottery, with separate equations for attrition after grade four and grade five.[8] I then re-estimate the achievement model, weighting sample observations by the inverse of the predicted probability that a student remains in the system, so that the reweighted sample resembles what it would have been had no attrition taken place. The results are virtually identical to Models 1 and 2 of Table 5. It makes no difference whether lottery outcomes are measured using the outright win indicator or the indicator combining outright and delayed winners.

To conclude, although there is substantial attrition from the district, with higher rates among lottery losers than winners, I can find no evidence that attrition has caused the remaining sample of lottery losers to differ systematically from the sample of lottery

[8] The equation for attrition after grade 5 includes fifth grade test performance as an additional explanatory variable.

winners with respect to observed factors influencing achievement. While lottery losers are more likely to seek alternative schooling options (not a surprise), the tendency to do so cuts across prior achievement levels and peer types, leaving estimates of magnet and peer effects unaffected.

V. Discussion and Conclusions

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. These findings contrast with two recent investigations of school choice. Betts et al. (2006) found no positive effects of San Diego magnet schools on mathematics achievement at these grade levels. Using a research strategy very similar to that employed here, Cullen, Jacob and Leavitt (2003) found no academic gains for winners of lotteries to oversubscribed high schools in Chicago's school choice plan. Given the diversity of magnet programs and the districts in which they are located, generalization from any of these results is hazardous. Still, given the latest negative findings, the results here remind us that in at least some times and places, students benefit from enrolling in magnet schools.

The peer effects found here echo those in some of the recent literature, except that they are substantially larger. In a study using longitudinal data on Texas schools, using a combination of student and grade-within-school effects to isolate the influence of peers, Hanushek, Kain, and Rivkin (2002) found that reducing percent black by 50 percentage points in grades five through seven would produce an increase in seventh grade achievement of about six-tenths of a year's growth. This is identical to the estimate obtained in this study, except for two things. First, the HKR finding pertained to black students. Among white and Hispanic students the response to a decrease in the share of black schoolmates was only about half as great. Second, the HKR estimate aggregates a sequence of three one-year gains to arrive at the final figure for seventh grade achievement. The gain reported in this paper is an average over fifth and sixth graders and has therefore taken half as a long to produce as the postulated gains in Texas. Taking these two factors into account, the peer effects reported here would appear to be three to four times the magnitude of comparably-defined gains found in that research.

Also using Texas data, Hoxby (2000) found that a 10 percentage point increase in the share of blacks caused a drop of .28 points on the Texas Assessment of Academic Skills for blacks, but a decline of only .03 points among whites. Given that a normal year's growth is in the range of 2 – 2.5 points, her estimate for blacks is greater than the effect reported here, though the impact on whites is much smaller. However, like the HKR study, Hoxby relies on variation in student cohorts to identify peer effects. The estimated impact in grade five therefore represents the cumulative effect of peers in a cohort and has been several years in the making. Again, the peer effects reported in this study emerge much faster, within one to two years of the middle school lottery.

The findings reported here indicate that achievement can diverge dramatically within a relatively short time as a function of middle school peers. This may reflect 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.

Contrary to the findings of Hoxby (2000) and Hanushek et al. (2001), I found no effect of peers' prior academic achievement. This was unexpected, but it should be remembered that the variable in question is a school average. Average achievement may simply not be particularly relevant by middle school, when departmentalized instruction and ability grouping become prevalent (as is the case in this district). Although the Hoxby and Hanushek et al. studies included some of the same grades as this paper, they were in elementary schools, increasing the likelihood that the mix of students in any given class mirrored the school as a whole.

I found no influence of student absenteeism, disciplinary incidents, or within-year mobility on achievement. This is interesting in light of the common lament in the literature, that researchers must rely on student race and income as proxies for the behavioral variables they wished they had, and suggests that the relevant behavioral variables may not be easy to identify or measure. The middle school students in this district seem to be fairly impervious to misbehavior and poor attendance by their peers. The channels by which peer effects are transmitted would appear to be more elusive: the communication of peer approval and disapproval, establishing norms inimical to academic achievement.

Finally, the research reported here, like other recent work on peer effects, has implications for the use of student test scores to evaluate teachers and schools. 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 attributes 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 models of assessment that ensure teachers and schools are not held accountable for factors beyond their control.

References

Betts, Julian, Lorien Rice, Andrew Zau, Emily Tang, & Cory Koedel. 2006. Does School Choice Work? Effects on Student Integration and Academic Achievement. Public Policy Institute of California.

Crain, Robert, Amy Heebner & Yiu-Pong Si. 1992. The Effectiveness of New York City's Career Magnet Schools: An Evaluation of Ninth Grade Performance Using an Experimental Design. Berkeley, CA: National Center for Research in Vocational Education.

Crain, Robert L., Anna Allen, Robert Thaler, Deborah Sullivan, Gail L. Zellman, Judith Warren Little, Denise D. Quigley. 1999. The Effects of Academic Career Magnet Education on High Schools and Their Graduates. Berkeley, CA: National Center for Research in Vocational Education.

Cullen, Julie Berry, Brian A. Jacob and Steven Leavitt. 2003. The Effect of School Choice on Student Outcomes: Evidence from Randomized Lotteries. National Bureau of Economic Research Working Paper 10113. Cambridge MA: NBER.

Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin. 2001. Does Peer Ability Affect Student Achievement? National Bureau of Economic Research Working Paper 8502. Cambridge MA: NBER.

Hanushek, Eric A., John F. Kain and Steven G. Rivkin. 2002. New Evidence About Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement. National Bureau of Economic Research Working Paper 8741. Cambridge MA: NBER.

Heckman, James. 1997. Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations. Journal of Human Resources, 32(3), 441-462.

Howell, William G. and Paul E. Peterson. 2002. The Education Gap: Vouchers and Urban Schools. Washington DC: Brookings Institution Press.

Hoxby, Caroline. 2000. Peer Effects in the Classroom: Learning from Gender and Race Variation. National Bureau of Economic Research Working Paper 7867. Cambridge MA: NBER.

Hoxby, Caroline, and Jonah Rockoff. 2005. The Impact of Charter Schools on Student Achievement. Unpublished.

Kemple, James, J. and Jason C. Snipes. 2000. Career Academis: Impacts on Students' Engagement and Performance in High School. Manpower Demonstration Research Corporation.

Kemple, James J. and Judith Scott-Clayton. 2004. Career Academies: Impacts on Labor Market Outcomes and Educational Attainment. Manpower Demonstration Research Corporation.

TABLE 1: MAGNET SCHOOL LOTTERIES AND ENROLLMENT

		Non-Academic:				
	Academic	North	South	East	West	Composite Non-Academic
Lottery Participants[a]	2315	1277	1320	1395	292	2594
Winners						
Outright	883	520	723	496	52	1450[c]
Delayed	223	341	314	385	49	756
Losers						
This Lottery	1209	416	283	514	191	388
All Lotteries[b]	539	69	49	46	17	199
Grade 5 Enrollment[b]						
This Magnet	758	246	397	388	30	1061
Other Magnets	287	435	330	416	144	339
Non-Magnets	834	416	419	415	91	846
Left System/Untested						
5th Grade	436	183	175	170	26	346
6th Grade	184	135	163	131	14	266

a. Counts only students enrolled in the district as 4th graders, when lottery was conducted.
b. Counts only students tested in mathematics as 5th graders.
c. Students admitted outright to at least one non-academic magnet.

TABLE 2: STUDENT CHARACTERISTICS IN GRADE 5

Magnet School	Black Pct	Low Income Pct	Special Educ. Pct	ESL Pct	4th Grade Math Mean	4th Grade Math SD	4th Grade Reading Mean	4th Grade Reading SD
Academic								
Academic								
Participants	23	15	12	9	666	28	684	29
Enrolling Participants	21	12	14	10	668	27	685	28
Losing Participants[a]	20	15	10	7	666	28	683	29
Non-Academic								
North								
Participants	46	30	10	7	646	34	664	65
Enrolling Participants	51	32	7	3	646	29	664	32
Losing Participants[a]	55	41	8	6	639	41	661	37
South								
Participants	54	32	9	5	637	33	655	34
Enrolling Participants	65	36	8	4	630	30	649	33
Losing Participants[a]	74	47	11	2	625	38	640	31
East								
Participants	44	23	11	11	651	34	665	34
Enrolling Participants	49	21	10	12	649	32	662	32
Losing Participants[a]	47	22	12	14	651	32	665	29
West								
Participants	93	40	7	3	615	29	636	27
Enrolling Participants	75	56	6	6	628	39	647	38
Losing Participants[a]	61	43	9	9	640	39	656	37
Composite Non-Academic								
Participants	48	28	10	8	645	34	661	34
Enrolling Participants	57	30	8	7	640	32	657	33
Losing Participants[a]	57	35	9	9	640	36	658	33

Characteristics of students enrolled in the district as 4th and 5th graders and tested in mathematics both years.

a. Students who lost all lotteries they entered.

TABLE 3: PEER CHARACTERISTICS, GRADE 5

Magnet Lottery Participants	Percent Black		Percent Low Income		Percent Special Ed.		Percent ESL		Mean 4th Grade Math		Mean 4th Grade Reading	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Academic												
All	36	18	30	19	15	5	10	6	641	33	657	35
Enrollees	21	4	12	3	13	2	10	3	667	8	684	8
Losers[a]	20	40	15	35	10	30	7	26	619	40	633	42
North												
All	50	20	40	19	13	6	8	6	634	24	648	25
Enrollees	53	7	34	6	6	2	4	3	643	5	660	4
Losers[a]	59	22	53	20	17	6	7	6	619	15	632	16
South												
All	56	21	44	17	14	5	7	7	627	26	642	27
Enrollees	73	5	51	7	13	3	3	2	620	3	637	4
Losers[a]	59	22	49	18	15	5	9	8	615	59	628	61
East												
All	50	21	39	17	14	5	9	6	635	25	649	26
Enrollees	61	3	44	8	13	3	10	2	633	17	645	18
Losers[a]	51	22	43	18	16	5	12	9	624	14	683	15
West												
All	59	23	46	17	14	5	8	7	630	24	645	26
Enrollees	96	0	54	0	12	0	3	0	616	0	630	0
Losers[a]	61	21	61	18	17	5	9	8	618	23	632	24
Non-Academic Composite												
All	52	21	41	18	14	5	8	7	631	25	646	26
Enrollees	65	11	45	10	11	4	6	4	630	14	645	15
Losers[a]	56	22	49	20	16	5	10	8	619	31	633	33

Peers of students enrolled in the district as 4th and 5th graders and tested in mathematics both years.
a. Students who lost all lotteries they entered.

TABLE 4: MAGNET SCHOOL IMPACT ON MATHEMATICS ACHIEVEMENT

Independent Variables	Model 1	Model 2	Model 3
Magnet Treatments:			
Academic	2.60		
	(1.54)		
Composite	3.29		
	(3.59)		
Cumulative Academic		1.91	
		(1.05)	
Cumulative Composite		6.05	
		(2.30)	
Academic			
Academic Grade 6			
Composite			
Composite Grade 6			
Student Characteristics			
Grade 4 math	0.47	0.47	
	(0.01)	(0.01)	
Grade 4 reading	0.19	0.19	
	(0.01)	(0.01)	
Black	-6.62	-7.16	
	(0.88)	(0.87)	
Special Education	1.66	1.88	
	(1.08)	(1.08)	
Low Income	-3.90	-3.88	
	(0.86)	(0.86)	
ESL	3.87	3.94	
	(1.28)	(1.28)	
Female	-0.90	-0.98	
	(0.66)	(0.67)	
Overidentification Test p-value	0.55	0.76	
No. obs.	5637	5609	

TABLE 5: MAGNET AND PEER EFFECTS ON MATHEMATICS ACHIEVEMENT

Independent Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Magnet Treatments:							
Academic	-2.53	-0.37	-2.07	Magnet Treatments Interacted with Student Characteristics (28 Coefficients)		-14.45	-11.28
	(2.41)	(3.08)	(3.05)			(5.14)	(5.52)
Academic Grade 6	-2.67	-2.68	-2.96			-2.47	-4.30
	(2.76)	(2.77)	(2.77)			(6.17)	(6.13)
Composite	0.22	-0.93	0.77			14.17	22.72
	(3.60)	(3.62)	(3.88)			(26.11)	(25.56)
Composite Grade 6	10.36	9.28	9.55			46.07	50.43
	(4.04)	(4.06)	(4.10)			(40.02)	(39.49)
Peer Effects:							
(lottery-based)							
Percent Black	-23.97		-25.37	-30.57		-54.44	
	(6.94)		(8.95)	(8.44)		(18.83)	
Percent Low Income		-12.43	2.27		-13.72		-27.20
		(6.89)	(8.98)		(7.30)		(14.74)
(residence-based)							
Percent Black	1.65		-10.67	2.76		2.90	
	(2.62)		(4.01)	(2.79)		(2.61)	
Percent Low Income		-4.74	-10.67		-5.45		-2.93
		(2.79)	(4.01)		(2.89)		(2.68)
Student Characteristics							
Grade 4 math	0.47	0.47	-5.56	0.48	0.47	0.44	0.44
	(0.01)	(0.01)	(0.93)	(0.03)	(0.03)	(0.01)	(0.01)
Grade 4 reading	0.19	0.18	0.47	0.13	0.12	0.17	0.17
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.01)	(0.01)
Black	-5.14	-6.18	1.56	-3.07	-5.41	-5.03	-4.97
	(0.93)	(0.91)	(1.08)	(2.46)	2.19	(0.91)	(0.88)
Special Education	1.55	1.79	-3.00	0.91	-1.41	3.69	3.78
	(1.08)	(1.08)	(0.92)	(3.02)	(2.98)	(1.16)	(1.15)
Low Income	-3.14	-2.77	3.90	-2.13	-2.26	-2.67	-2.44
	(0.89)	(0.91)	(1.29)	(2.52)	(2.54)	(0.96)	(0.98)
ESL	3.98	3.95	-0.97	0.07	-1.09	4.20	4.80
	(1.29)	(1.28)	(0.67)	(3.74)	(3.72)	(1.43)	(1.40)
Female	-0.92	-1.04	8.95	-5.20	-5.17	-1.08	-1.18
	(0.67)	(0.67)	(3.77)	(1.77)	(1.75)	(0.68)	(0.16)
Teacher Fixed Effects	No	No	No	No	No	Yes	Yes
Overidentification Test							
p-value	0.72	0.73	0.65	0.30	0.39	----	----
No. obs.	5618	5618	5618	5618	5618	5618	5618

TABLE 6: INTENTION TO TREAT ESTIMATES

Independent Variables:	(1)	(2)	(3)
Academic Lottery			
Outright winner	2.82	-0.69	-0.01
	(1.17)	(1.69)	(1.94)
Outright winner x Grade 6	-3.84	-4.66	-4.41
	(1.71)	(1.72)	(1.72)
Delayed winner	4.97	1.34	2.16
	(1.96)	(2.31)	(2.49)
Delayed winner x Grade 6	-5.43	0.20	-0.80
	(2.85)	(1.49)	(1.45)
Composite Lottery			
Outright winner	-0.79	0.26	-0.91
	(1.45)	(1.67)	(1.62)
Outright winner x Grade 6	4.45	4.32	3.65
	(1.64)	(1.64)	(1.65)
Delayed winner	-0.92	3.11	2.34
	(1.62)	(1.96)	(1.96)
Delayed winner x Grade 6	3.27	-6.02	-5.73
	(1.95)	(2.85)	(2.85)
Peer Variables			
Lottery-based			
Percent black		-13.08	
		(4.30)	
Percent low income			-9.36
			(4.84)
Residence-based			
Percent black		-2.05	
		(2.04)	
Percent low income			-6.24
			(2.16)
No. obs.	5637	5618	5618

TABLE 7: ATTRITION RATES, BY LOTTERY OUTCOME

		Lotteries			
		Academic		Composite	
		Winners	Losers	Winners	Losers
Will leave system after grade					
	4	12.8%	21.3%	8.4%	12.4%
	5	8.0%	14.2%	11.7%	8.7%
	6	8.9%	10.9%	10.1%	4.2%
	7	6.2%	9.2%	12.0%	15.8%
No spring math test in grade					
	5	4.7%	7.6%	7.8%	2.8%
	6	4.4%	4.7%	5.8%	3.3%
	7	3.8%	3.5%	3.4%	3.5%
	8	2.0%	3.7%	2.8%	2.7%

TABLE 8: EFFECT OF PEERS & LOTTERY OUTCOME ON ATTRITION, ACADEMIC MAGNET

	Model 1		Model 2	
Independent Variables:	Coeff.	s.e.	Coeff.	s.e.
Outright winner	-4.003	(4.805)	-1.206	(4.673)
Percent black	3.599	(5.460)		
Percent black x winner	-1.619	(9.770)		
Percent black x Grade 4 math	-0.005	(0.008)		
Percent black x Grade 4 math x winner	0.002	(0.015)		
Percent low income			2.051	(5.657)
Percent low income x winner			-8.875	(10.001)
Percent low income x Grade 4 math			-0.003	(0.008)
Percent low income x Grade 4 math x winner			0.013	(0.015)
Grade 4 math score	0.001	(0.004)	0.001	(0.004)
Grade 4 math x winner	0.002	(0.007)	-0.002	(0.007)
Grade 4 reading score	0.001	(0.002)	0.001	(0.002)
Grade 4 reading x winner	0.004	(0.003)	0.004	(0.003)
Black	-0.502	(0.130)	-0.440	(0.125)
Black x winner	0.154	(0.237)	0.118	(0.229)
Low income	-0.121	(0.130)	-0.071	(0.131)
Low income x winner	-0.024	(0.252)	0.006	(0.253)
Special ed	-0.268	(0.171)	-0.249	(0.170)
Special ed x winner	0.489	(0.268)	0.486	(0.268)
ESL	0.041	(0.129)	0.034	(0.128)
ESL x winner	-0.114	(0.221)	-0.111	(0.221)
Female	0.023	(0.082)	0.021	(0.082)
Female x winner	0.101	(0.144)	0.078	(0.145)
No. obs.	2286		2286	

Model includes dummy variables for each year and their interactions with winner.

TABLE 9: ATTRITION PROBABILITIES AS FUNCTION OF PRIOR ACHIEVEMENT, PEER CHARACTERISTICS AND ACADEMIC LOTTERY OUTCOME (OUTRIGHT WIN)

	Percent Black			Percent Low Income		
Prior Achievement	High	Low	Difference	High	Low	Difference
Lottery Losers						
High	14.6%	13.0%	1.6%	11.6%	15.4%	-3.8%
Low	15.8%	11.2%	4.6%	12.4%	14.1%	-1.7%
Difference	-1.2%	1.8%	-3.0%	-0.8%	1.3%	-2.1%
Lottery Winners						
High	10.5%	9.3%	1.3%	7.8%	10.7%	-2.9%
Low	11.5%	7.9%	3.6%	8.4%	9.7%	-1.3%
Difference	-0.9%	1.4%	-2.3%	-0.6%	1.0%	-1.6%
Difference, Losers - Winners						
High	4.1%	3.7%		3.8%	4.7%	
Low	4.3%	3.3%		4.0%	4.4%	