Educational Evaluation and Policy Analysis
December 2009, Vol. 31, No. 4, pp. 323–345
DOI: 10.3102/0162373709340917

Can Interdistrict Choice Boost Student Achievement? The Case of Connecticut's Interdistrict Magnet School Program

Robert Bifulco
Syracuse University

Casey D. Cobb
University of Connecticut

Courtney Bell
Educational Testing Service

Connecticut's interdistrict magnet schools offer a model of choice-based desegregation that appears to satisfy current legal constraints. This study presents evidence that interdistrict magnet schools have provided students from Connecticut's central cities access to less racially and economically isolated educational environments and estimates the impact of attending a magnet school on student achievement. To address potential selection biases, the analyses exploit the random assignment that results from lottery-based admissions for a small set of schools, as well as value-added and fixed-effect estimators that rely on pre–magnet school measures of student achievement to obtain effect estimates for a broader set of interdistrict magnet schools. Results indicate that attendance at an interdistrict magnet high school has positive effects on the math and reading achievement of central city students and that interdistrict magnet middle schools have positive effects on reading achievement.

Keywords: *school choice, magnet schools, desgregation programs, quasiexperimental methods*

School desegregation efforts constitute one of the most significant social policy endeavors of the last 50 years. Proponents of continued school desegregation efforts argue that racial and economic isolation undermines the achievement of minority and low-income students and that integration efforts can help to improve the educational outcomes of historically disadvantaged groups (Orfield, Frankenberg, & Garces, 2008). Others, however, have raised doubts about the potential benefits of desegregation programs, arguing that integrated school environments are difficult to achieve and maintain and that the evidence is inconclusive regarding the relationship between desegregated schools and academic achievement (Armor, 1995; Armor, Thernstrom, & Thernstrom, 2006).

Estimating the effects of school desegregation efforts on student achievement is difficult because students who participate in these programs are not randomly selected. Past evaluations of school desegregation programs have been criticized for not adequately addressing selection issues (Schofield, 1995). This shortcoming of the desegregation evaluation literature has allowed readers to draw markedly different conclusions. Consider two recent reviews: Linn and Welner (2007) conclude that "there is a relatively common finding that African American student achievement is enhanced by less segregated schooling" (p. 2), but Armor et al. (2006) conclude that "there is no evidence of a clear and consistent relationship between desegregation and academic achievement" (p. 4).

In this article, we examine Connecticut's interdistrict magnet schools and focus on estimating their effects on academic achievement. Connecticut's interdistrict magnet school program offers a model of choice-based desegregation that has allowed many students to attend less racially and economically isolated schools. It also appears to satisfy current legal constraints on desegregation programs. Thus, it provides a case study on the potential of desegregation efforts to help improve student achievement in the current legal environment.

To address the selection issues that have plagued past research, we use information from admission lotteries, as well as longitudinal data on individual students, to help isolate the effect of attending a magnet school on student achievement. Our analysis proceeds in three stages. First, following the example of several recent studies (Ballou, 2007; Betts, Rice, Zau, Tang, & Koedel, 2006; Cullen, Jacob, & Levitt, 2006; Howell & Peterson, 2002; Hoxby & Rockoff, 2005), we use data on admission lotteries from two magnet schools, to compare the achievement of applicants who were offered admission with the achievement of applicants who were denied admission. Because admission offers are determined solely by random lottery, comparison of these two groups should provide unbiased estimates of achievement effects. Next, we implement nonexperimental estimates of the effects of these two schools that rely on pretreatment measures of student achievement to control for differences between magnet and nonmagnet school students. Comparing these nonexperimental estimates to those obtained from the lottery analysis allows us to assess the extent of bias in the nonexperimental estimates. Finally, having demonstrated that our nonexperimental estimators do not suffer from substantial bias, we use those estimators to obtain impact estimates for a broader set of magnet schools.

For the two magnet schools for which data on admission lotteries are available, we find positive effects on mathematics and reading test scores and particularly large positive effects on reading. We also find that value-added and fixed-effect regressions that make use of pre–magnet school test scores are able to replicate closely the lottery-based estimates. Applying our nonexperimental estimators to the broader set of interdistrict magnet schools that serve Connecticut's central cities, we find that attendance at an interdistrict magnet high school has positive effects on the math and reading achievement of central city students and that interdistrict magnet middle schools have positive effects on reading achievement.

By providing defensible estimates of the impacts of a viable, choice-based interdistrict desegregation program, this study provides valuable information for assessing the potential that desegregation efforts have for improving student achievement. Nevertheless, there are three important caveats on the conclusions that can be drawn from our analysis. First, our lottery analysis, the validation of the nonexperimental estimates that we use, and those nonexperimental estimates themselves each rely on a number of untestable assumptions. Although we think that these assumptions are plausible, we try to highlight them in discussing our methods and results. Second, our analysis is unable to determine whether magnet schools improve student achievement because they offer less racially and economically isolated environments or because of other aspects of their educational programs. Finally, our analysis focuses on the short-term benefits that Connecticut's interdistrict magnets have for students who attend them. Additional information on longer-term benefits and the costs of the program are needed to draw policy conclusions.

Evidence on School Desegregation and Magnet Schools

Racially isolated schools and schools with high concentrations of poverty can undermine student achievement in a variety of ways. Several studies show that schools with concentrations of minority and low-income students have a difficult time attracting and retaining qualified teachers (Clotfelter, Ladd, & Vigdor, 2005; Freeman, Scafidi, & Sjoquist, 2005; Lankford, Loeb, &

Wyckoff, 2002; Loeb, Darling-Hammond, & Luczak, 2005). Historically, schools with high proportions of minorities have offered few opportunities for advanced placement and other academically challenging course work (Gamoran, 1992; Oakes, 1990). Also, because students from educationally disadvantaged backgrounds are more likely to encounter difficulties in school, such concentrations of students can influence teachers' expectations, student-based norms, forms of instruction, and levels of classroom disruption (Lavy, Passerman, & Schlosser, 2007).

Many believe that by providing access to more qualified teachers, more opportunities to take advanced course work, higher teacher expectations, and environments more conducive to learning, integrated schools can help poor and minority students improve academic achievement. Some also have argued that by organizing curriculum and instruction around a special theme, magnet schools can foster more student engagement and a stronger sense of membership and purpose and thereby help to improve student achievement (Gamoran, 1996).

Early studies of school desegregation efforts focused largely on the short-term effects of deliberately moving students to less racially segregated schools. Comprehensive reviews of this early research suggest that the impacts of desegregation on student achievement were mixed, with some evidence of small positive effects on the academic achievement of Black students (Cook, 1984; Schofield, 1995). Much of this literature is based on comparisons of students who attended desegregated schools with students who remained in segregated schools and has been criticized for failing to control adequately for unobserved differences between these two groups of students. Also, because desegregation efforts often were accompanied by conflict and resistance, estimates of the short-term effects might be contaminated by factors related to the desegregation process (Hanushek, Kain, & Rivkin, 2006).

Whereas early desegregation efforts often used mandatory student assignments to achieve integration targets, desegregation plans began to rely more heavily on magnet school programs during the 1970s and through the ensuing decades.[1] By offering parents a choice among several desegregated schools with different themes and sometimes enhanced resources, magnet school programs largely have avoided the resistance associated with other desegregation programs. Much of the early research on magnet schools focused on how effective they were in creating and maintaining integrated learning environments.[2] Until recently, however, "surprisingly little research has been done on the educational outcomes associated with magnet schools" (Yu, Taylor, Goldring, Smrekar, & Piche, 1997, p. 19).

In one of the few large-scale studies of magnet school effects, Gamoran (1996) uses a sample of city students from the National Education Longitudinal Study to estimate differences in 10th-grade achievement between students who attend comprehensive public high schools and students who attend magnet high schools. In analyses that control for differences in eighth-grade test scores and an extensive set of family background characteristics, Gamoran finds that magnet schools are more effective than regular schools at raising student proficiency in reading and social studies. Interestingly, estimates of magnet school benefits are virtually unchanged when controls are added for the student composition of the school and indicators of school environment, leaving questions about why magnet schools help improve student achievement.

More recently, a number of studies have tried to exploit admission lotteries to estimate the effects of magnet schools. These studies measure treatment effects by comparing the average outcomes of lottery winners who are offered admission to a magnet school with the average outcomes of students who apply but are denied admission because they lost the admission lottery. The studies most relevant here are those by Betts et al. (2006) and Ballou (2007). Betts et al. used lotteries to estimate the impact of magnet schools in San Diego and found that winning a magnet school lottery at the high school level increases mathematics achievement 2 and 3 years later by approximately 0.2 standard deviations. The researchers did not, however, find any statistically significant effects on reading achievement, nor did they find any effects for elementary or middle school magnets. Ballou examined four middle school magnets in a large Southern school district and found that attending a middle school magnet has positive impacts on mathematics achievement for fifth and sixth graders, although those impacts are uneven across schools.[3]

Although they are an excellent strategy for addressing potential biases due to self-selection, lottery-based analyses have important limitations. In any given choice program, some schools will not be oversubscribed or will not select students randomly, and those that are oversubscribed might not be representative of all schools. Moreover, admission lotteries are not typically held on a schoolwide basis; rather, admission lotteries are held for specific grades and, often, subgroups within grades. As a result, there are too few winners and losers in particular lotteries to gain the benefits of randomization. Conclusions from lottery studies are thus often limited to subgroups of schools and to types of students within schools, thereby undermining external validity. Because the mixed findings from existing studies of magnet schools suggest that magnet school effects vary, this limitation of the lottery approach is a serious concern.

This limitation of lottery-based analyses raises the question of whether more broadly applicable methods can provide similarly unbiased estimates. A fairly large literature has examined how well nonexperimental estimators can replicate estimates derived from randomized assignment. Much of this literature is based on the results of job training trials, and a major theme has been that the usefulness of nonexperimental estimators depends on context and, in particular, on the nature of the program selection process and on the availability of data related to program participation and outcomes (Cobb-Clark, & Crossley, 2003; Heckman, LaLonde, & Smith, 1999). Very few studies have assessed the ability of nonexperimental methods to replicate experimental estimates in the contexts of educational programs. Wilde and Hollister (2007) attempted to replicate the results of the Tennessee STAR class-size reduction experiments, using propensity score matching, and Agodini and Dynarski (2004) attempted to replicate the results of experimental evaluations of dropout prevention programs, also using propensity score matching. Both studies revealed that propensity score estimators are not able to replicate closely experimental estimates of program impacts. Neither study, however, was able to include controls for pretreatment measures of student achievement, which might considerably reduce bias in nonexperimental estimators.[4]

In this study, we demonstrate how lottery analyses like those used by Betts et al. (2006) and Ballou (2007) can buoy the results of nonexperimental analyses of the type used by Gamoran (1996). This approach allows us to provide defensible estimates of the impact of a legally feasible, choice-based desegregation program on short-term academic outcomes.

Connecticut's Interdistrict Magnet School Program

In a 1996 ruling, the Connecticut Supreme Court held that as a result of racial, ethnic, and economic isolation, Hartford public school students had been denied equal educational opportunity under the state constitution.[5] In response, the state has adopted a number of programs designed to provide students in the state's central cities opportunities to attend schools with students from suburban districts. The most significant response from the state has been its efforts to support the establishment and operation of interdistrict magnet schools.

What Is an Interdistrict Magnet School?

In Connecticut, an interdistrict magnet school is a publicly funded school operated by a local or regional school district, by a regional education service center, or by cooperative agreement involving two or more districts. Each magnet has an educational theme, and students choose to enroll on the basis of their interest in the school's theme. All students in the school districts participating in the magnet are eligible to attend; enrollment is by application only; and if a school is oversubscribed, admissions are made on the basis of lotteries (described in more detail below). The state has encouraged and supported the development of interdistrict magnet schools by providing funding for building construction and planning assistance and by allowing students in magnet schools to generate additional operating aid.[6]

At the beginning of the 2006–2007 school year, 54 interdistrict magnet schools serving 17,735 students were in operation. Six interdistrict magnet high schools were half-time programs (where students attend part of the school day at the magnet and part in their home school), and three interdistrict magnets were new in 2006–2007. Forty-one magnets served students who resided in Hartford, New Haven, or Waterbury, and in 2006–2007, students in these magnets accounted for 79.4% of all interdistrict magnet school students in the state.

Several features of Connecticut's interdistrict magnet schools make them important models to study in the current policy environment. First, the programs are designed to integrate students across district lines, which is crucial for achieving substantial amounts of racial integration in several regions of the United States. Second, participation is entirely voluntary; neither families nor districts are required to participate.[7] Third, although the court monitors the extent to which the state has achieved racial integration goals in the Hartford area, student race is not used in determining admission to any interdistrict magnet school. Thus, this program offers models of choice-based interdistrict desegregation that appear to satisfy current legal constraints.

Who Attends Interdistrict Magnet Schools?

Table 1 provides a profile of students in middle and high school magnets serving Hartford, New Haven, and Waterbury. About 63% of 10th-grade students and 58% of 8th-grade students in magnet schools that serve Hartford, New Haven, and Waterbury are from one of those central cities; the rest are from surrounding suburban towns. About 18% of all 10th-grade students who reside in a central city attend a magnet, and nearly 15% of all 8th-grade central city students do. Less than 4% of the students who reside in suburban areas attend magnets.[8]

Among students who reside in a central city, magnet school students are significantly more likely than nonmagnet school students to be Black and less likely to be Hispanic. Magnet students from the city are also more likely to be female, particularly among 10th graders, and have higher average test scores before entering middle school or high school. Among suburban students, the picture is different. Among 8th and 10th graders, magnet school students from the suburbs are substantially more likely than nonmagnet school students to be Black and less likely to be White. Magnet school students from the suburbs are also more likely to be eligible for free lunch and have lower average test scores. Differences in test scores between suburban magnet and nonmagnet school students are less marked at the middle school level. Interdistrict magnet schools, then, appear to be bringing together a substantial proportion of relatively high-achieving students from the central cities, with smaller proportions of relatively disadvantaged and lower-achieving students from the suburbs.

Do Interdistrict Magnet Schools Reduce Racial, Ethnic, and Economic Isolation?

Interdistrict magnet schools clearly provide at least some students of color from Connecticut's most isolated central cities the opportunity to join less-isolated learning environments. Figure 1 compares the percentage white and the percentage eligible for free or reduced-priced lunch in the average student of color's school among interdistrict magnet and non-magnet schools. These comparisons are at the high school level, although similar results are found in elementary and middle schools. Racial and ethnic isolation in Connecticut's central city districts is high. In Hartford and New Haven, the percentage White in the typical Black student's school is well below 10%, and the figures are similar for Hispanics. Waterbury has a larger population of White students. Students of color from these districts who attend magnet schools are, on average, in substantially more integrated environments than their counterparts in central city district schools. The percentage of free-lunch eligible in the interdistrict magnet schools attended by central city students of color is also much lower than that in the central city district schools.[9]

Although informative, the comparisons in Figure 1 do not tell us how magnet schools change the peer environments of their students, who may have been attending schools less isolated than the typical school in the district where they reside. Table 2 compares the peer environments of magnet school students with the peer environments of the schools they attended before enrolling in their current magnet school.[10]

TABLE 1
Magnet School Students, Compared to Nonmagnet School Students

	Urban students		Suburban students	
	Magnet	Nonmagnet	Magnet	Nonmagnet
Tenth graders				
Black	.533***	.465	.509***	.119
Hispanic	.299***	.392	.146**	.121
White	.150**	.127	.313***	.724
Free-lunch eligible	.684	.671	.343***	.193
Male	.429***	.506	.471**	.508
Grade 8 scores				
Mathematics	–.361***	–.735	–.157***	.186
Reading	–.308***	–.686	–.035*	.171
Grade 6 scores				
Mathematics	–.370***	–.699	–.218***	.151
Reading	–.393***	–.733	–.120***	.170
n	1,369	6,207	815	22,277
Eighth graders				
Black	.572***	.412	.356***	.118
Hispanic	.314***	.458	.123	.132
White	.104	.116	.493***	.706
Free-lunch eligible	.720***	.761	.301***	.239
Male	.482**	.515	.523	.515
Grade 6 scores				
Mathematics	–.392***	–.609	.104***	.193
Reading	–.343***	–.641	.180	.207
Grade 4 scores				
Mathematics	–.368***	–.576	.082**	.155
Reading	–.433***	–.659	.112**	.192
n	1,386	7,946	984	23,033

Note. Samples of urban students consist of students appearing in Connecticut State Department of Education test score files during 2005–2006 or 2006–2007 and residing in Hartford, New Haven, or Waterbury. Samples of suburban students consist of students appearing in the test score files during 2005–2006 or 2006–2007 and residing in a district in New Haven or Hartford county that participates in an interdistrict magnet school that serves Hartford, New Haven, or Waterbury. Figures reported are sample means. Test scores are standardized using year-specific means and standard deviations for the entire population. Test scores are missing for some students; as such, test score means are based on less than a full sample.
*$p < .10$. **$p < .05$. ***$p < .01$. Significance indicates difference between magnet and nonmagnet school students.

These comparisons confirm that, on average, interdistrict magnet schools reduce racial and economic isolation for the city students who attend them. Among magnet school students who reside in a central city, their current school has a significantly higher percentage of White students, a lower percentage of free-lunch eligible students, and higher average test scores. For the typical suburban magnet school student, his or her current magnet school has a higher percentage of minority and free-lunch-eligible students and, at the high school level, slightly lower average test scores.

Empirical Strategy

The effects of a chosen school on its students are typically estimated by comparing the achievement of students who attend the school with that of students in other schools. Such estimates often confound the differences in family and personal background between students with the effects of the chosen school on learning. In the case of interdistrict magnets, students and parents who have selected magnet schools have made special efforts to seek out alternatives to their geographically assigned school, and they

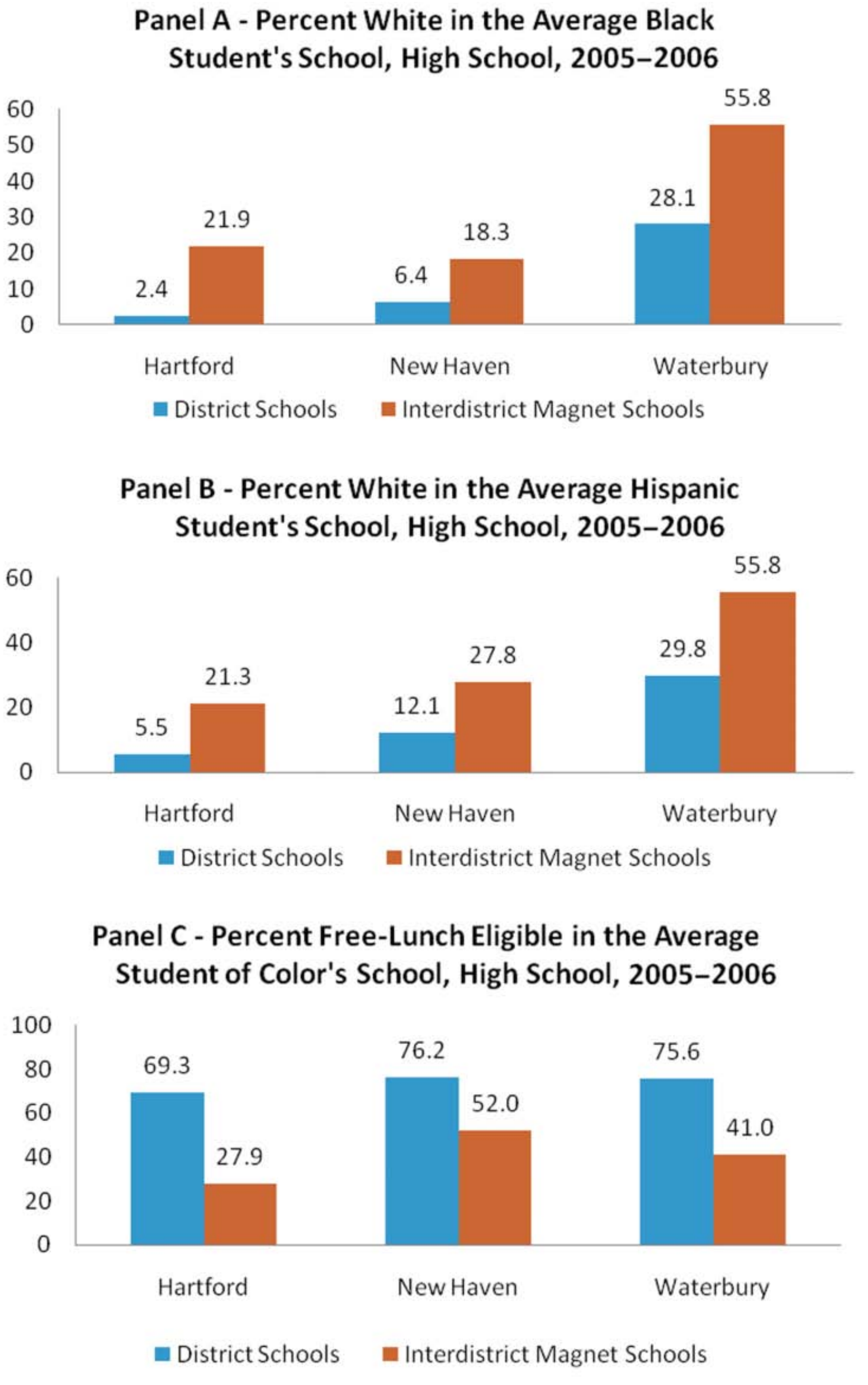

FIGURE 1. *Comparison of student composition in city and interdistrict magnet schools.*

TABLE 2

Change in Peer Environments for Magnet School Students

	Urban students		Suburban students	
	Previous school	Magnet school	Previous school	Magnet school
Tenth graders				
Black (%)	46.1	48.3**	34.4	49.0***
Hispanic (%)	37.2	25.3***	17.4	22.9***
White (%)	15.1	24.6***	45.2	25.6***
Free-lunch eligible (%)	72.0	59.9***	41.7	55.8***
Grade 8 scores (*Means*)				
Mathematics	−.549	−.330***	−.229	−.293***
Reading	−.618	−.322***	−.202	−.253**
n	970		626	
Eighth graders				
Black (%)	47.6	49.7**	28.2	41.8***
Hispanic (%)	38.3	23.7***	17.4	22.1***
White (%)	12.7	24.7***	50.6	33.0***
Free-lunch eligible (%)	71.7	55.8***	35.5	44.1***
Grade 4 scores (*Means*)				
Mathematics	−.553	−.255***	−.073	−.045
Reading	−.681	−.296***	−.049	−.053
n	874		706	

Note. Urban students include those in a magnet school serving students in Hartford, New Haven, or Waterbury during 2005–2006 or 2006–2007 and whom we can place in a nonmagnet school before their enrollment in their current magnet school. Suburban students consist of students appearing in the test score files during 2005–2006 or 2006–2007 who reside in a district in New Haven or Hartford county that participates in an interdistrict magnet school that serves Hartford, New Haven, or Waterbury and whom we can place in a nonmagnet school before their enrollment in their current magnet school.

*$p < .10$. **$p < .05$. ***$p < .01$. Significance indicates difference between previous school and magnet school.

often travel longer distances and make other sacrifices to attend a magnet. We already have seen in Table 1 that magnet school students differ on observable characteristics from nonmagnet school students. Magnet school students may also differ from other students with similar ethnic and socioeconomic backgrounds in terms of unobservables, such as motivation and parental support. Potential unobserved differences between interdistrict magnet school students and otherwise similar students make estimating magnet school effects difficult.

Recent studies of school choice programs demonstrate how admission lotteries can address unobserved variable bias resulting from self-selection. Because lottery winners and losers are determined through a random process, comparisons of average outcomes across the two groups will be free of selection bias. This approach has been used to study voucher programs in Washington, D.C., New York, and Dayton, Ohio (Howell & Peterson, 2002); intradistrict choice programs in Chicago (Cullen et al., 2006); charter schools in Chicago (Hoxby & Rockoff, 2005); intradistrict magnet schools in a large Southern district (Ballou, 2007); and a variety of choice programs in San Diego (Betts et al., 2006).

As already discussed, although the lottery is an excellent strategy for addressing potential biases attributed to self-selection, the external validity of lottery-based analyses is questionable. More broadly applicable approaches to addressing bias owing to self-selection use matching and/or statistical procedures to control for as many observable differences between treatment and comparison groups as possible. The most convincing studies of this kind include pretreatment test scores as matching variables, which can help to control for many factors that influence student achievement and learning. When test scores are available from two or more pretreatment periods, these methods can determine if treatment group students make larger or smaller test score gains, as compared to students with similar pretreatment levels of achievement. Because these studies do not use random assignment, their estimates remain subject to potential biases from any unobserved differences between treatment and comparison group members who have similar levels of pretreatment performance. Such studies often have the advantage, however, of being applicable to a wider range of schools and students than that of lottery-based studies.

Hoxby and Murarka (2008) suggest a way to combine the advantages of lottery-based analyses and nonexperimental methods that control for pretreatment test scores. Specifically, when lottery-based estimates of school effects are available for some students in some schools, these can be used to test the extent to which potential biases attributed to selection on unobservables are likely to influence estimates derived from nonexperimental analyses. If the researcher is able to specify matching procedures or statistical models that replicate the results of lottery-based procedures, then one can have more confidence that those same models and procedures, as applied to all students in the program, provide defensible estimates.

This strategy of leveraging the results from lottery analyses still requires untestable assumptions. One must assume that the factors influencing selection into the magnet schools for which lottery results can be obtained are similar to the factors influencing selection into other magnet schools. In our case, we can obtain effect estimates using lottery results for two schools—one that serves Grades 6 through 8 and one that serves Grades 6 through 12. The broader set of schools that we examine using only nonexperimental estimators are also limited to those serving either middle or high school grades. In addition, all the schools used in both types of analysis serve students primarily from Hartford, New Haven, or Waterbury and each city's surrounding inner ring of suburbs. Thus, we believe that the estimates that we derive using nonexperimental approaches are strengthened if confirmed by lottery-based analyses.

Lottery-Based Analysis

For this study, we obtained the results of admission lotteries from two interdistrict magnet schools operated by the Capitol Region Education Council. One of these schools serves Grades 6 through 8, and the other serves Grades 6 through 12. Both schools are located in a first-ring suburb close to the city of Hartford and serve the city of Hartford and four suburban districts.[11]

The admission policies for these two schools are identical. Each school allocates a predetermined number of seats for each district that it serves. Students apply in the spring of fifth grade for admission to sixth grade the following fall. When applications are received, siblings of students currently enrolled in the school are placed in the first seats allocated to their district. The remaining applicants are randomly assigned a number. Applicants from each district are then assigned to the remaining seats allocated to the district in order of the randomly assigned numbers. The students awarded seats through this process are contacted and offered admission, and the rest of the applicants from that district are placed on a waiting list, in order of their randomly assigned numbers. When a student turns down an admission offer, a seat in that district becomes available and is offered to the next applicant who is from the district and on the waiting list. Applications are accepted for only the sixth grade. If students leave the school after the start of sixth grade, those spots are filled with students from the original waiting list.

Sample and Data

For both schools, we collected admission data on applications submitted in 2003 and 2004. Staff at the Connecticut State Department of Education matched these data to test score file records from 2001–2002 through 2006–2007 to provide measures of student achievement from two pretreatment periods (the fall of fourth grade and the fall of sixth grade) and one posttreatment period (the spring of eighth grade).[12] Matches were based on name and date of birth; in some cases, the magnet school applicants could not be matched to a test score record, because the applicant attended a school outside the Hartford metropolitan region, enrolled in a private school, or could otherwise not be located in the test score file. We observe eighth-grade test scores (our outcomes of interest) for 67.4% of the lottery participants—including 70.0% of those offered admission and the 66.0% of those never offered admission. These individual test score records then were matched over time. In all the analyses that follow, test scores are converted to standard *z* scores using grade- and year-specific means and standard deviations. Information on the students' age, gender, ethnicity, free-lunch status, and special education status is also available from the test score files.

Because admission lotteries are district and year specific, we have a total of 22 potential lotteries.[13] In the analysis here, we drop applicants who did not participate in any of the lotteries because they had siblings enrolled in the school, and we drop students from the eight potential lotteries that did not have any losers. All the applicants in these latter lotteries eventually were offered a seat in the school; thus, these lotteries do not contribute randomly assigned comparison-group students. We also drop the remaining Hartford lotteries. All the applicants from Hartford to one of our schools were offered admission; as such, they did not participate in a true lottery. The two Hartford lotteries for the other school are also dropped but for a different reason. Unlike students from other districts, Hartford students have many ways to opt out of the regular public schools—including other magnet schools, an interdistrict open choice program schools, and charter schools. As a result, few students who are lotteried out of this magnet school end up in a Hartford public school, which complicates interpretation of the magnet school effect that we are trying to estimate.

We further restrict our sample of lottery participants to those students for whom we observe test scores in fourth and sixth grade. Students who apply to a magnet school from outside the public school system are less likely to enroll in public schools and to be observed in the post-treatment period, particularly if they are not offered admission to the magnet.[14] Thus, a control group of lottery losers observed in the post-treatment period does not necessarily provide appropriate matches for lottery winners who apply from outside the public school system. As Cullen et al. (2006) point out, excluding students not observed in public schools in the pre-treatment period does not invalidate the random assignment, because whether a student is observed pretreatment is determined before the lottery takes place. As in the case of both the Cullen et al. study of open enrollment and the Hoxby and Rockoff study of charter schools (2005), restricting the sample to those who are observed in public school during the pretreatment period is important for achieving balanced samples of lottery winners and lottery losers.

The sample that we used for this analysis includes 553 participants in 12 lotteries. Table 3 describes the students in this sample and compares them with nonmagnet school students in their previous schools. As compared to students in the schools from which they are drawn, the applicants to these two magnet schools are more likely to be Black or White and less likely to be Hispanic and free-lunch eligible, and they have higher average fourth-grade test scores.

Estimating Achievement Effects

Estimates of the effects of these magnet schools on achievement can be derived from this sample of lottery participants, using the following regression:

$$Y_{iL} = \alpha W_{iL} + \mu_L + e_{iL}, \tag{1}$$

where Y_{iL} is the eighth-grade test score of student i who participates in lottery L, W_{iL} is an indicator of whether student i won an admission offer through the lottery, μ_L represents lottery-specific fixed effects, and e_{iL} is a random error term. The parameter of interest, α, can be estimated using a fixed effect or least squares dummy variable estimator. This coefficient is a weighted average of the difference in mean eighth-grade test scores between the winners and losers of each lottery.

If there are indeed no systematic differences between lottery winners and losers in each lottery, as random assignment helps to ensure, then the difference in mean eighth-grade tests scores between the two groups is solely due to the lottery winners' enrollment in the interdistrict magnets. However, not all lottery winners accept their invitation to enroll. The estimates of α in Equation 1 average the effects of magnet schools

TABLE 3
Sample of Lottery Participants, Compared to Nonparticipants From the Same Districts

	Lottery sample	Nonmagnet sample
Black	.407**	.356
Hispanic	.109***	.212
White	.471***	.387
Free-lunch eligible	.235***	.394
Male	.495	.510
Grade 4 scores		
Mathematics	.088***	−.182
Reading	.208**	−.150
n	553	3,043

*$p < .10$. **$p < .05$. ***$p < .01$. Significance indicates difference between lottery sample and nonmagnet school students.

on the achievement of those who choose to enroll and the presumably zero effect on those who do not enroll. The estimates from this regression are sometimes referred to as the intention-to-treat effect (Ballou, 2007; Hoxby & Rockoff, 2005). Hoxby and Murarka (2008) argue that unlike the case of many medical treatments (where patients' willingness and ability to comply with the treatment influence its efficacy), the intention-to-treat effect has little relevance for evaluating the effect of choice schools. Those who choose not to accept admission are not receiving the treatment in any meaningful sense.

The standard approach to obtaining the effect of the treatment on the treated uses the indicator of winning a lottery as an instrument for an indicator of magnet school enrollment in a two-stage least squares or instrumental variables procedure (Ballou, 2007; Hoxby & Rockoff, 2005). The first- and second-stage equations in such a procedure are as follows:

$$\begin{aligned} &\textit{First stage: } M_{iL} = \beta W_{iL} + \lambda_L + v_{iL} \\ &\textit{Second stage: } Y_{iL} = \gamma \hat{M}_{iL} + \theta_L + \omega_{iL}, \end{aligned} \qquad (2)$$

where M_{iL} is an indicator that the student is enrolled in one of our two magnet schools during the eighth-grade test administration and $\hat{M}_{iL}$ is the predicted value of the magnet school indicator from the first-stage equation. The estimate of γ from this procedure can be interpreted as the effect of the treatment on the treated—that is, the effect of magnet schools on the students who attend them.

If lotteries are truly random, then we would not expect any significant differences between lottery winners and losers, and the simple regressions above would thus provide consistent estimates of the magnet school effect. Adding covariates to the above regressions is nonetheless desirable for two reasons. First, including covariates can significantly increase precision (Ballou, 2007; Betts et al., 2006). Second, for any finite sample, we do not expect differences between randomly assigned treatment and control groups to equal zero. Adding covariates can help to control for differences between treatment and controls that arise by chance. Pretreatment measures of achievement are especially useful for these purposes.

One issue in implementing these estimation procedures is that of defining a lottery winner. For each lottery, we know how many total seats were available to be filled, and we know the number of applicants who had siblings enrolled in the school. The difference between these two is the number of seats available to lottery participants. Each applicant can also be ordered by his or her randomly assigned lottery number. If the rank order of a lottery participant's randomly assigned number is less than or equal to the number of seats available to lottery participants, we labeled that participant an *on-time winner*. If on-time winners decline their admission offer or withdraw from the magnet after enrolling, applicants with the next-lowest lottery numbers are offered admission. So we also identified for each lottery the highest lottery number offered admission, and we counted all applicants with lottery numbers too high to be offered an on-time admission but low enough eventually to be offered admission, *delayed winners*. Of the 553 lottery participants in our sample, we have 142 on-time winners, 22 delayed winners, and 389 applicants who were never offered admission.

There is some question about whether an indicator of on-time winning, with delayed winners excluded from the sample, or an indicator that includes on-time and delayed winners is the most appropriate. Ballou (2007) argues that delayed winners who accept an invitation to enroll may expect especially large gains from attending a magnet school. If so, an indicator that includes delayed winners might not be a valid instrument. We have used both definitions of lottery winners in the estimations presented here, and it turns out that our results are not sensitive to this issue.

Attrition and Sample Balance

Random assignment helps to ensure that lottery winners are similar to lottery losers on observed and unobserved characteristics. However, randomization alone does not guarantee that our treatment and comparison groups have no significant differences. First, a few of the lotteries in these schools are small. When lotteries are small, large differences between lottery winners and losers can emerge by chance. Second, we are missing posttreatment test scores for any student who participated in a magnet school lottery but for whom we could not match to a test score record because she or he attended a school outside the Hartford metropolitan region, enrolled in a private school, or could otherwise not be located in the test score file. Once the sample is limited to students whom we observe in a public school during the pretreatment period, only 6.3% do not have a test score observed in the posttreatment period. Nonetheless, attrition of this kind is slightly higher among those not offered admission (8.2%) than among on-time winners (2.1%) and delayed winners (0.0%). If students who are not observed posttreatment are sufficiently different from those who are observed, this differential attrition could lead to nonrandom differences between the lottery winners and losers who are observed in the posttreatment period.

To demonstrate that lottery winners and losers are balanced on observable characteristics, Table 4 presents the results of a series of regressions. Each row presents a separate regression of an observable characteristic on an indicator of whether the student won the lottery or not and on a set of lottery dummy variables. The first three columns show the results of regressions run with all 553 lottery participants in the sample.[15] In each regression, the coefficient on the lottery winner indicator is not statistically distinguishable from zero. These results confirm that the initial lotteries were random.

The last three columns in Table 4 show the results of regressions including only those lottery participants that we observe in the posttreatment period. These results indicate whether differential attrition created any observable differences between lottery winners and losers. In all these regressions except the first, the coefficients on the lottery winner indicator is not statistically distinguishable from zero, which indicates that except for age, there are no statistically significant differences between the lottery winners and losers whom we observe posttreatment. Given that t tests from 12 separate regressions are reported in Table 1, it is not unreasonable to expect one significant result at the $p < .10$ level to emerge by chance. Most important, the differences on pretreatment measures of achievement between lottery winners and losers are substantively small and statistically insignificant. These results suggest that neither small sample sizes nor differential attrition has created a substantial imbalance between the treatment and control groups.

Results of the Lottery Analysis

Table 5 presents estimated effects on eighth-grade mathematics and reading scores. The table includes the following: the estimated effects of the intent to treat derived from Equation 1; the estimates of treatment on treated effects from Equation 2; and the estimates of the treatment on treated effects from versions of Equation 2 that add individual-level covariates, including fourth- and sixth-grade mathematics and reading test scores. The first three columns present results from regressions that define a lottery winner as an on-time winner. These regressions exclude delayed winners from the sample. The last three columns present the results from regressions that define lottery winners as on-time and delayed winners.

The results are similar regardless of how we define lottery winners. As expected, the point estimates of the treatment-on-treated effect are larger than the estimates of the intent-to-treat effect; furthermore, including covariates in the two-stage least squares estimates substantially increases precision. The results indicate that

TABLE 4
Testing the Balance of Lottery Samples

Dependent Variable	All lottery participants (n = 553) Coeff.	*SE*	*p*	Participants observed in eighth grade (n = 517) Coeff.	*SE*	*p*
Age (in years)	.025	.042	.552	.066*	.037	.074
Black	−.047	.040	.243	.000	.041	.301
Hispanic	.017	.028	.545	.023	.028	.389
White	−.066	.042	.114	.059	.043	.170
Asian	−.026	.017	.110	−.028	.017	.103
Free-lunch eligible	.004	.040	.912	.014	.040	.730
Special education	.007	.021	.889	.006	.022	.798
Male	−.050	.048	.297	−.065	.048	.179
Grade 6 scores						
Mathematics	.011	.079	.889	.006	.080	.943
Reading	.046	.083	.582	.047	.083	.576
Grade 4 scores						
Mathematics	.011	.083	.894	.014	.085	.870
Reading	.034	.087	.696	.038	.088	.665

Note. Coefficient, standard error, and p value reported for indicator of whether the student was a lottery winner or not—including on-time and delayed winners. Each row represents a separate regression; all regressions include lottery-fixed effects. Test scores are standardized using year-specific means and standard deviations for the entire population.
*$p < .10$.

TABLE 5
Lottery-Based Estimates of the Effect of Interdistrict Magnet Schools on Achievement

Grade 8	On-time lottery winners ITT	TOT	TOT-WC	On-time + delayed lottery winners ITT	TOT	TOT-WC
Mathematics	.110	.142	.139***	.109	.139	.138***
	(.080)	(.103)	(.054)	(.076)	(.097)	(.050)
R^2	.088	.083	.767	.084	.079	.772
n		492			514	
Reading	.243***	.312***	.283***	.252***	.318***	.278***
	(.093)	(.120)	(.070)	(.088)	(.112)	(.064)
R^2	.072	.055	.703	.077	.062	.709
n		494			516	

Note. Each set of results are from separate regressions. Dependent variables include test scores standardized using year-specific mean and standard deviation for the population. Results in column labeled *ITT* (intent to treat) are ordinary least squares regressions of test score on indicator of whether student won the admission lottery or not. Results in columns labeled *TOT* (treatment on treated) are two stage least squares estimates using an indicator of students who won lottery as instrument for enrollment in an interdistrict magnet school during eighth grade. The covariates included in the models presented in columns labeled TOT-WC include student's age, gender, ethnicity, free-lunch eligibility in Grade 4, special education status in Grade 4, and Grade 4 and Grade 6 mathematics and reading scores. In the first three columns, only on-time lottery winners are counted as lottery winners; that is, delayed winners are excluded from the sample. In the last three columns, delayed winners are included and counted as lottery winners. All regressions include lottery fixed effects. Standard errors robust to clustering with in schools are in parentheses.
*$p < .10$. **$p < .05$. ***$p < .01$.

these two interdistrict magnet schools have had positive effects on student achievement. The dependent variable in these regressions are test scores that have been standardized using the year-specific mean and standard deviation for the population. Thus, the estimates of the treatment-on-treated effect from the models that include covariates indicate that the reading test scores of students in these magnet schools are nearly 0.28 standard deviations higher and their mathematics scores are nearly 0.14 standard deviations higher than what they would be if those students had attended other schools.

Do Nonexperimental Methods Replicate Lottery-Based Estimates?

To answer this question, we assembled a data set consisting of students who reside in the suburban districts served by the two interdistrict magnet schools used in the lottery-based analysis and who appear in the 2006 or 2007 eighth-grade test score files maintained by the state. Each of these student records was matched to sixth- and fourth-grade test score records for the same student. Connecticut only recently has begun phasing in student identification numbers to facilitate the tracking of students across test score files from different years; so, many of the longitudinal matches were made using name, date of birth, and other identifying information in the test score files. The state department of education was able to match successfully 85% of these student records to a sixth-grade test score record and 75% to both a sixth-grade and a fourth-grade test score record. We use the sample of students with test scores for fourth, sixth, and eighth grade to estimate two regression models.

The first regression, the value-added regression, can be formulated as follows:

$$Y_{i8} = \alpha M_{i8} + \mathbf{X}_i \mathrm{B} + \mu_t + v_i, \tag{3}$$

where Y_{i8} is student i's eighth-grade test score, M_{i8} is a binary variable indicating whether the student was enrolled in one of the two magnet schools used in our lottery analysis at the time of the eighth-grade test administration, $\mathbf{X}_i$ is a vector of individual-level covariates, μ_t is a year fixed effect, and v_i is a random error term. Covariates include age, gender, ethnicity, special education status in fourth grade, free-lunch eligibility in fourth grade, and pretreatment mathematics and reading test scores from sixth grade and fourth grade.

The second model, the fixed-effect regression, can be formulated as follows:

$$Y_{it} = \alpha' M_{it} + \omega_i + \mu'_t + v'_i, \tag{4}$$

where Y_{it} is student i's test score in year t, M_{it} is an binary variable equal to 1 if the student is in eighth grade and in a magnet school in year t (0, otherwise), ω_i is an individual student fixed effect, μ'_t is a year fixed effect, and v'_i is a random error term. Estimates of α' that control for the individual fixed effects can be obtained by differencing all variables from the individual student mean, and employing what is sometimes called the within estimator (Baltagi, 1995). In this case, the estimator effectively controls for time-invariant unobserved differences between magnet school students and nonmagnet school students.[16]

Both estimators in Equations 3 and 4 will provide unbiased estimates of the effect of attending a magnet school (α and α') as long as the error terms v_i and v'_i are uncorrelated with magnet school enrollment during eighth grade. For the value-added estimator, Equation 3, including pretreatment test scores in the equation helps to control for any differences between magnet school enrollees and other students that might cause magnet school enrollment to be correlated with the error term. For the fixed-effect estimator, Equation 4, any time-invariant effects of student characteristics on achievement are absorbed in the individual-based fixed effect, thus eliminating many potential sources of selection bias. If, however, magnet school enrollees would have had test score trajectories different from those of other students in the absence of these two magnet schools, then the estimated effects from both these equations will be biased (Bifulco & Ladd, 2006; Rouse, 1998). By comparing the effect estimates obtained from these regression models to the presumably unbiased estimates obtained from the preceding lottery analysis, we can assess whether unobserved differences in test score trajectories are a source of substantial bias in this context.[17]

Table 6 presents the nonexperimental estimates of the effect of these two interdistrict magnet schools. The first column presents results from the value-added regressions

TABLE 6

Comparison of Nonexperimental Estimates With Lottery-Based Estimates

Grade 8	Value-added regression	Fixed-effect regression	Lottery-based estimate
Mathematics	.144*	.130**	.138***
	(.074)	(.052)	(.050)
R^2	.811	.897	.772
n	4,026	12,018	514
Reading	.340***	.306***	.278***
	(.019)	(.035)	(.064)
R^2	.731	.879	.709
n	4,024	11,982	516

Note. Dependent variables are test scores standardized using the grade- and year-specific mean and standard deviation for the population. Valued-added regressions include age, gender, ethnicity, free-lunch eligibility, special education status, year fixed effect, and fourth- and sixth-grade mathematics and reading test scores, as well as a magnet enrollment indicator. The coefficient on the magnet school enrollment indicator is reported. The fixed-effect regression includes magnet school indicator, year fixed effects, and controls for individual fixed effects. Lottery-based estimates are taken from last column of Table 5. The figures in parentheses are standard errors, adjusted for clustering at the school level.
*$p < .10$. **$p < .05$. ***$p < .01$.

(Equation 3), and the second column presents the results from the fixed-effect regressions (Equation 4). To facilitate comparison, we have included lottery-based estimates of the average treatment-on-treated effect taken from the last column of Table 5.

The results from the nonexperimental methods are quite similar to one another and to the lottery-based estimates. For mathematics, the value-added estimate is 4% larger, and the fixed-effect estimate 6% smaller, than the lottery-based estimate. For reading, the point estimate from the value-added regression is 22% larger, and the point estimate from the fixed-effect regression 10% larger, than the lottery-based estimates. These differences are all substantively small. For the largest of the differences—that between the value-added and lottery-based estimates of the effect on reading—the value-added estimates imply an effect size of 0.34 standard deviations, compared to an effect size about 0.28 standard deviations as implied by the lottery-based analysis. Such small differences are unlikely to influence policy conclusions. Also, all the nonexperimental estimates are comfortably within the 95% confidence interval for the corresponding lottery-based estimates, indicating that the differences between estimates from the nonexperimental estimators and the lottery-based analyses are not statistically significant.

The Average Effects of Interdistrict Magnet Schools

Having shown that nonexperimental methods that use pretreatment measures of achievement can provide results similar to those derived from lottery-based analyses, we now proceed to use them to estimate average achievement effects for larger sets of interdistrict magnet schools. These estimates should be viewed with some caution. The estimates in Table 6 suggest that students who self-select into the two magnet schools examined so far do not have unobserved achievement trajectories that differ from those of nonmagnet students in ways that substantially bias effect estimates obtained from value-added and fixed-effect estimators. These results do not, however, guarantee that selection on unobservables will not bias effect estimates obtained for other magnets. For instance, estimates that examine effects on 10th-grade test scores might be subject to sources of bias not prevalent in estimated effects on eighth-grade test scores. Specifically, if the high school dropout rate among low-achieving students is different in nonmagnet than in magnet high schools, then nonexperimental methods might provide biased estimates of the magnet school effect. This issue is less likely to influence estimates on eight-grade test scores. Nonetheless, regression models that control for pretreatment test scores represent the best available methods for estimating the effects of larger sets of interdistrict magnet schools, and given the support for these methods provided in Table 6, the estimates presented in this section are plausible.

We develop estimates of average achievement effects for a set of 12 interdistrict magnet high schools and a set of 7 interdistrict middle schools.[18] These schools include all the full-day interdistrict magnet high schools and all but two of the interdistrict magnet middle schools that serve students from Hartford, New Haven, or Waterbury.[19] We focus on estimating the effects of the interdistrict magnet high schools on 10th-grade mathematics and reading Connecticut Academic Performance Test scores and the effects of the interdistrict magnet middle schools on 8th-grade mathematics and reading Connecticut Mastery Tests. The Connecticut Academic Performance Test is the high school statewide testing program.

To construct our student sample for the analysis of the 12 interdistrict magnet high schools, we asked officials at the state department of education to extract 2005–2006 and 2006–2007 10th-grade Connecticut Academic Performance Test records for all the students attending either one of those interdistrict magnets or a high school in a district that sends at least 10 students each year to one of those interdistrict magnets. We then asked the state officials to match those student records to records from earlier eighth- and sixth-grade test score files. Our sample for the middle school analysis was constructed in an analogous manner. A sample of 1,731 magnet high school students and 11,091 students from feeder districts were extracted from the 10th-grade test score files. State officials were able to match 74.6% of these students to eighth-grade test score records and 60.4% to sixth-grade test score records. Of the 1,188 magnet school students and 11,231 students from feeder districts extracted from the eighth-grade test score files, state officials were able to match 80.1% to sixth-grade test score records and 63.6% to fourth-grade test score records.[20] Table 7 presents summary statistics on the sample of 10th-grade students whom we were able to match to an eighth- and sixth-grade record and the sample of eighth-grade students whom we were able to match to a sixth- and fourth-grade record.

We use Equation 3 and Equation 4 exactly as described above to compute value-added and fixed-effect estimates of effects on mathematics and reading achievement. Table 8 presents the results. The top panel of the table presents results for the entire set of 7 magnet middle schools and 12 magnet high schools serving Hartford, New Haven, and Waterbury. Separate value-added and fixed-effect estimates were computed using the sample of students residing in one of the three central cities and the sample of students from the surrounding suburbs. Estimates of the effects of magnet middle schools on eighth-grade tests scores and the effects of magnet high schools on 10th-grade test scores are presented separately.

In general, the estimated effects of magnet schools are positive, and the fixed-effect estimates are somewhat smaller than the value-added estimates. In no case, however, do the fixed-effect estimates indicate substantively or statistically different conclusions. The eighth-grade results indicate that magnet middle schools have had similar effects on the mathematics achievement of suburban and central city students. These effect estimates are positive but statistically insignificant for suburban students and only marginally significant for city students. The estimated effects of magnet middle schools on reading are larger than those for mathematics, and they are statistically significant. The positive effects of magnet middle schools on reading scores are larger for students from the suburbs, but they are statistically significant for students from the central cities as well. The estimates imply that 3 years of exposure to a magnet school in the middle school years increases reading achievement between 0.093 and 0.152 standard deviation for city students and between 0.219 and 0.265 standard deviations for suburban students.

The high school results indicate that, on average, interdistrict magnet schools have had positive and statistically significant effects on the 10th-grade mathematics and reading achievement of central city students. The estimated effects on mathematics range from 0.108 to 0.135 standard deviations and, on reading, range from 0.110 to 0.153. These represent the effects of 2 years of exposure. If we assume similar effects over the second half of these students' high school careers, these estimates imply effect sizes of between 0.22 and 0.27 for mathematics and between 0.22 and 0.30 for reading. The estimated effects of magnet high schools on the 10th-grade achievement of suburban students are somewhat smaller than those for central city students, and they are not statistically significant.

The sample of magnet middle schools used to generate the estimated effects on eighth-grade test scores presented in the top panel of

TABLE 7
Treatment and Comparison Group Samples

	Central city students		Suburban students	
	Magnet	Nonmagnet	Magnet	Nonmagnet
Tenth graders				
n	700	2,151	373	4,525
Black	.520	.497	.450***	.231
Hispanic	.329**	.379	.121***	.190
White	.130	.110	.408***	.550
Asian	.017	.011	.016	.027
Free-lunch eligible	.673***	.731	.305*	.356
Special education	.069***	.102	.064**	.099
Male	.403*	.440	.428*	.476
Age	16.0***	16.1	15.9	15.9
	(.515)	(.592)	(.432)	(.452)
Grade 8 scores				
Mathematics	−.337***	−.599	−.068**	−.167
	(.767)	(.828)	(.799)	(.937)
Reading	−.283***	−.538	.049***	−.115
	(.776)	(.829)	(.841)	(.923)
Grade 6 scores				
Mathematics	−.399***	−.629	−.142	−.219
	(.843)	(.929)	(.834)	(.972)
Reading	−.448***	−.705	.003***	−.187
	(.857)	(.881)	(.870)	(.989)
Eighth graders				
n	376	2,770	473	4,275
Black	.378	.371	.277***	.198
Hispanic	.439	.463	.082***	.231
White	.176	.149	.611***	.528
Asian	.005	.014	.030	.042
Free-lunch eligible	.601***	.744	.203***	.354
Special education	.051***	.105	.055*	.080
Male	.441	.472	.491	.529
Age	14.0***	14.2	13.8***	13.9
	(.510)	(.626)	(.388)	(.435)
Grade 6 scores				
Mathematics	−.113***	−.461	.231***	.041
	(.816)	(.842)	(.889)	(1.024)
Reading	−.093***	−.531	.322***	.027
	(.800)	(.809)	(.901)	(.968)
Grade 4 scores				
Mathematics	−.202***	−.552	.225***	−.050
	(.851)	(.847)	(.908)	(1.041)
Reading	−.224***	−.625	.289***	−.008
	(.867)	(.834)	(.926)	(1.027)

Note. Means (with standard deviations in parentheses). Test scores are z scores computed using the year-specific mean and standard deviation for entire population of students.
*$p < .10$. **$p < .05$. ***$p < .01$. Significance indicates difference between magnet and nonmagnet school students.

Table 8 include the two schools used in the lottery analyses presented above. The middle panel presents estimates that do not include those two schools. The estimated effect for these five magnet middle schools for which lottery data were not available to us are smaller than the

TABLE 8
Estimated Magnet School Treatment on Treated Effects, by Students' Residence

	Value-added estimates		Fixed-effect estimates	
	Central city students	Suburban students	Central city students	Suburban students
Grade 8				
Mathematics	.126**	.104	.082*	.095
	(.058)	(.077)	(.049)	(.067)
n	3,062	4,690	9,186	14,070
Reading	.152***	.265***	.093***	.219***
	(.050)	(.048)	(.019)	(.051)
n	3,063	4,693	9,189	14,079
Grade 10				
Mathematics	.135***	.085*	.108***	.061*
	(.044)	(.047)	(.034)	(.036)
n	2,709	4,740	8,127	14,220
Reading	.153***	.082	.110**	.030
	(.042)	(.055)	(.042)	(.040)
n	2,725	4,759	8,175	14,277
Lottery schools excluded				
Grade 8				
Mathematics	.077	.103**	.038	.057
	(.051)	(.052)	(.033)	(.048)
n	2,989	2,935	8,967	8,805
Reading	.123**	.147***	.062	.095*
	(.056)	(.055)	(.037)	(.049)
n	2,989	2,936	8,967	8,808
Hartford-area schools only				
Grade 8				
Mathematics	.199**	.124	.148**	.107
	(.082)	(.079)	(.075)	(.077)
n	1,690	4,568	5,070	13,704
Reading	.237***	.301***	.147***	.249***
	(.038)	(.043)	(.053)	(.060)
n	1,697	4,572	5,091	13,716
Grade 10				
Mathematics	.277***	.165***	.255***	.126*
	(.045)	(.049)	(.045)	(.069)
n	1,035	1,770	3,105	5,310
Reading	.228***	.193***	.155*	.134***
	(.070)	(.065)	(.094)	(.049)
n	1,050	1,779	3,150	5,337

Note. Dependent variables are test scores standardized using the grade- and year-specific mean and standard deviation for the population. Valued-added regressions include age, gender, ethnicity, free-lunch eligibility, special education status, year fixed effect, and fourth- and sixth-grade mathematics and reading test scores, as well as magnet enrollment indicator. The coefficient on the magnet school enrollment indicator is reported. The fixed-effect regression includes magnet school indicator, year fixed effects, and controls for individual fixed effects. The figures in parentheses are standard errors, adjusted for clustering within schools.
$^{*}p < .10$. $^{**}p < .05$. $^{***}p < .01$.

estimates presented in the top panel of Table 8 and substantially smaller than the effect estimates for the two schools for which we do have lottery data (reported in Table 6). These results highlight the limitation of lottery-based studies that examine only a small number of schools. The effectiveness of any type of school—in this case, interdistrict magnet schools—is likely to vary across schools of that type, and there is little reason to believe that the effects of those schools for which admission lottery data are most readily available will be representative of the broader set of schools.

As discussed above, the value-added and fixed-effect estimators used here are subject to potential selections bias, particularly if the unobserved test score trajectories of magnet school students are different from those of nonmagnet school students who have similar pre–magnet school achievement levels. The fact that the value-added and fixed-effect estimators provide effect estimates similar to lottery-based estimates for the two schools for which we have lottery data is somewhat reassuring in that magnet and nonmagnet school students with similar pre–magnet school test scores do not have different unobserved test score trajectories. Both the schools included in the lottery analysis, however, are located in the Hartford area; thus, one might doubt that the validation of the value-added and fixed-effect estimators provided in Table 6 are relevant for magnet schools in the Waterbury and New Haven areas.

The bottom panel of Table 8 presents results for Hartford-area schools only. This sample includes five magnet middle schools and five magnet high schools. The estimated effects for this sample of Hartford-area magnet schools are generally larger than those in the top panel of Table 8. It could be that magnet school students in the Waterbury and New Haven areas have negative test score trajectories relative to those of nonmagnet students with the similar pre–magnet school test scores, which would make the effect estimates in the top panel of Table 8 biased downward. Another explanation for these results is that magnet schools in the Hartford area are more effective for the students they serve, relative to the districts from which they draw their students, than are the magnet schools in Waterbury and New Haven area. In any case, the conclusion that magnet schools have positive and statistically significant effects is clearly robust to limiting the sample to students from Hartford-area schools.

Conclusion

Reliance on voluntary choice to promote integration makes Connecticut's interdistrict magnet school program an interesting model for reducing racial and economic isolation and improving educational outcomes for poor and minority students in central city schools. Although interdistrict magnets provide only a small fraction of students in Connecticut's central cities with access to diverse schools, the schools do provide less racially isolated and higher-achieving environments than what most of their students would otherwise encounter. The key question that we have tried to address here is whether interdistrict magnet school students also have higher levels of achievement than what they would in the absence of the interdistrict magnet schools.

Several recent studies have advanced methods that use admission lotteries as a way to eliminate selection bias from evaluations of school choice programs. The analyses above show that in the context of Connecticut's interdistrict magnet schools, regression analyses that make use of pretreatment test score measures can provide estimates similar to those obtained from lottery-based analyses.

Our best estimates of the effects of interdistrict magnet schools on student achievement indicate that attendance at an interdistrict magnet high school has positive effects on the mathematics and reading achievement of central city students and that interdistrict magnet middle schools have positive effects on reading achievement. That interdistrict magnet schools, on average, succeed in providing their students more integrated, higher-achieving peer environments and that they also, on average, have positive effects on achievement suggests that they represent a promising model for helping to address the ills of racial and economic isolation.

We should note a few caveats on our findings. First, the value-added and fixed-effect regression analysis on which our effect estimates are based is subject to potential biases. Most notable, if magnet and nonmagnet school students with similar pre–magnet school test scores have different unobserved test score trajectories, our value-added and fixed-effect estimates will be biased. The fact that these value-added and fixed-effect estimators were able to replicate closely the results of lottery-based analyses provides some assurance that our effect estimates are not substantially biased. That an estimator can provide unbiased effect estimates for two particular schools, however, does not guarantee that it will provide unbiased estimates when applied to different schools. Estimates of impacts on 10th-grade test scores, for instance, might be subject to different sources of selection bias than estimates of impacts on eighth-grade test scores.

Second, our results do not tell us which aspects of interdistrict magnet schools benefit students. The number of magnet schools in our sample is too small to determine whether less racially and economically isolated environments, the organization of instruction around a particular theme, or other aspects of magnet schools are most closely associated with positive achievement effects.

Finally, additional studies are needed before any policy conclusions can be drawn about Connecticut's interdistrict magnet school program. In particular, the costs, as well as the benefits, of interdistrict magnet schools must be considered. Not only do interdistrict magnet schools create the pecuniary costs of providing new school buildings and transporting students over longer distances, but they may also generate nonpecuniary costs in the form of more disadvantaged peer environments and lower levels of achievement for central city students left behind in their neighborhood schools. Despite these questions about methods, mechanisms, and costs, we believe that the results here show that Connecticut's interdistrict magnet schools provide a promising model for other states that are concerned with the effects of racial and economic isolation—a model that warrants further examination.

Notes

[1]Magnet schools were approved first by federal courts as a means of satisfying desegregation court orders in 1975; by 1991–1992, more than 1.2 million students were attending magnet schools in 230 districts (Yu, Taylor, Goldring, Smrekar, & Piche, 1997).

[2]See Hawley and Smylie (1986) and Rossell (1988) for contrasting assessments from this early literature on how effective voluntary magnet school programs are at promoting integration relative to mandatory assignment policies.

[3]An ongoing evaluation by MDRC uses lotteries to isolate the impact of career academies established in nine high schools (Kemple & Willner, 2008). Although these academies can be classified as magnet programs, three things make them less relevant for our purposes: First, they are operated as programs within a school rather than as stand-alone schools; second, the focus on vocational training is not typical of magnet schools in most desegregation programs; and, third, desegregation is not a primary purpose of the career academies.

[4]Agodini and Dynarski (2004) did have pretreatment measures of their outcomes—namely, dropping out, educational aspirations, self-esteem, and absenteeism. However, pretreatment measures of these outcomes might not be as predictive of posttreatment outcomes as in the case of academic achievement measures.

[5]*Sheff v O'Neill*, 238 Conn. 1, 678 A.2d 1267 (1996).

[6]In the first few years following the *Sheff* decision, the operational funding of magnet schools was designed to encourage geographic diversity, with the hopes that such diversity would result in racial and economic diversity. More recently, however, state operating funding has been provided on a flat, per-pupil basis, and schools have to maintain a specified level of diversity to qualify for this aid.

[7]Connecticut education law does require districts to "provide educational opportunities for its students to interact with students and teachers from other racial, ethnic and economic backgrounds" (Public Act 97-290 §1).

[8]The sample of suburban nonmagnet school students used to compute the figures in Table 1 are limited to those suburban districts in the Hartford and New Haven counties that participate in an interdistrict magnet school that serves central city students.

[9]Although magnet schools are, on average, more diverse than central city district schools, they provide access to less-isolated learning environments for only a small percentage of students of color in the state's central cities. This limits the overall effect of the program on racial, ethnic, and economic isolation. Less than 6% of Black students and only 3% of Hispanic students residing in Hartford and New Haven attend an interdistrict magnet school with more than 25% White students.

[10]These comparisons are based on students from the Hartford, New Haven, and Waterbury areas whom we observed in 8th or 10th grade during the 2005–2006 and 2006–2007 school year and whom we also observed in a nonmagnet school sometime earlier.

[11]One of the schools serves four suburban districts west of the Connecticut River, and the other serves four districts east of the Connecticut River. Hartford is the only district served by both schools.

[12]Before 2005–2006, the Connecticut Mastery Tests, which are part of Connecticut's statewide testing program, were administered in the fall—early in the school year and only in Grades 4, 6, and 8. So, applicants in 2003 did not take statewide tests in seventh grade, and none of the applicants in our sample have fifth-grade test scores. Beginning in 2005–2006, tests were administered in the spring. All eighth-grade test scores are from the spring of 2005–2006 or 2006–2007. We count tests in the fall of sixth grade as pretreatment measures.

[13]Five district-specific lotteries in 2003 and 2004 for both schools imply 20 lotteries ($5 \times 2 \times 2 = 20$). However, for one of the districts served by one of these interdistrict magnets, seats are allocated by the middle school to which the student would be assigned; so, there are two separate lotteries each year for that district.

[14]Only 24% of lottery participants who apply from outside the public school system and are not offered admission have test scores observed in eighth grade, compared to 44% of lottery winners who apply from outside the public school system.

[15]In these regressions, lottery winners are defined as on-time or delayed winners. We also ran analogous regressions dropping delayed winners from the sample, as well as regressions where the student's assigned lottery number was substituted for the binary indicator of winning or losing the lottery. The results of these regressions were similar to those reported in Table 1.

[16]Because all the covariates in our data are constant over time or because changes in the variable are influenced by school policy, their effects cannot be distinguished from both the treatment and student fixed effects, and thus, they are not included here.

[17]We also implemented propensity score estimators with the same covariates used in the value-added regression analysis (Equation 3). We used nearest-neighbor, caliper-matching, and kernel density estimators. The literature on propensity score matching indicates that if the effects of observed covariates on the outcome variable are nonlinear and the treatment and comparison groups have different covariate distributions, ordinary least squares regression can produce biased estimates of program impacts (Stuart, 2007). In this case, however, there is substantial overlap in the covariate distributions; thus, the propensity score estimators all provided estimates similar to those obtained from the value-added regression. Because they are so similar to the regression results, we do not report the propensity score estimates here.

[18]High schools here are schools that serve Grades 9 through 12, and middle schools are schools that begin in Grade 6 or 7. Four of the seven middle schools end in Grade 8, but three serve high school grades as well.

[19]Two interdistrict magnet middle schools that serve students from New Haven start in Grade 5 and, thus, are not included in this analysis.

[20]Once the matching was completed and to ensure that students in the treatment group had received full exposure to their magnet school's educational program, we dropped any magnet school student whom we could identify as entering his or her current magnet school sometime after the entry grade. However, because we do not observe students during the entry grade for high schools, we were able to identify these cases for the middle school magnets only.

References

Agodini, R., & Dynarski, M. (2004). Are experiments the only option? A look at dropout prevention programs. *Review of Economics and Statistics, 86,* 180–194.

Armor, D. J. (1995). *Forced justice: School desegregation and the law*. Oxford, UK: Oxford University Press.

Armor, D. J., Thernstrom, A., & Thernstrom, S. (2006, August). Social science brief to the Supreme Court of the United States in support of petitioners in *Parents Involved v. Seattle School District No. 1* and *Meredith v. Jefferson County Board of Education*. http://www.thernstrom.com/pdf/Amicus_Brief.pdf, accessed July 13, 2009.

Ballou, D. (2007). *Magnet schools and peers: Effects on student achievement*. Unpublished paper.

Baltagi, B. H. (1995). *Econometric analysis of panel data.* New York: Wiley.

Betts, J., Rice, L., Zau, A., Tang, E., & Koedel, C. (2006). *Does school choice work? Effects on student integration and academic achievement*. San Francisco: Public Policy Institute of California.

Bifulco, R., & Ladd, H. F. (2006). The impacts of charter schools on student achievement: Evidence from North Carolina. *Education Finance and Policy, 1,* 50–90.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. (2005). Who teaches whom? Race and the distribution of novice teachers. *Economics of Education Review, 24,* 377–392.

Cobb-Clark, D. A., & Crossley. T. (2003). Econometrics for evaluations: An introduction to recent developments. *The Economic Record, 79,* 491–511.

Cook. T. (1984). What have Black children gained academically from school integration? Examination of meta-analytic evidence. In T. Cook et al. (Eds.), *School desegregation and Black achievement* (pp. 7–42). Washington, DC: National Institute of Education.

Cullen, J. B., Jacob, B. A., & Levitt, S. (2006). The effect of school choice on student outcomes: Evidence from randomized lotteries. *Econometrica, 74,* 1191–1230.

Freeman, C., Scafidi, B., & Sjoquist, D. L. (2005). Racial segregation in Georgia public schools, 1994–2001: Trends, causes and impact on teacher quality. In J. Boger, C. Edley, & G. Orfield (Eds.), *School resegregation: Must the South turn back?* (pp. 148–163). Chapel Hill: University of North Carolina Press.

Gamoran, A. (1992). Access to excellence: Assignment to honors English classes in the transition from middle to high school. *Educational Evaluation and Policy Analysis, 14,* 185–204.

Gamoran, A. (1996). Student achievement in public magnet, public comprehensive, and private city high schools. *Educational Evaluation and Policy Analysis, 18,* 1–18.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2006). *New evidence about* Brown v. Board of Education*: The complex effects of school racial composition on achievement*. Unpublished manuscript, University of Texas, Dallas.

Hawley, W., & Smylie, M. A. (1986). The contribution of school desegregation to academic achievement and racial integration. In P. Katz & D. Taylor (Eds.), *Eliminating racism: Means and controversies* (pp. 281–297). New York: Pergamon Press.

Heckman, J. J., LaLonde, R. J., & Smith, J. A. (1999). The economics and econometrics of active labor market programs. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 3, pp. 1865–2097). New York: Elsevier.

Howell, W. G., & Peterson, P. E. (2002). *The education gap: Vouchers and urban schools*. Washington, DC: Brookings Institution Press.

Hoxby, C. M., & Murarka, S. (2008). Methods of assessing the achievement of students in charter schools. In M. Berends, M. G. Springer, & H. J. Walberg (Eds.), *Charter school outcomes* (pp. 7–38). New York: Lawrence Erlbaum.

Hoxby, C. M., & Rockoff, J. (2005). *The impact of charter schools on student achievement*. Unpublished paper.

Kemple, J. J., & Willner, C. J. (2008). *Long-term impacts on labor market outcomes, educational attainment, and transitions to adulthood*. New York: MDRC.

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Education Evaluation and Policy Analysis, 24,* 37–62.

Lavy, V., Passerman, D. M., & Schlosser, A. (2007). *Inside the black box of ability peer effects: Evidence from variation in high and low achievers in the classroom*. Unpublished paper.

Linn, R. L., & Welner, K. G. (2007). *Race-conscious policies for assigning students to schools: Social science research and the Supreme Court cases*. Washington, DC: National Academy of Education.

Loeb, S., Darling-Hammond, L., & Luczak, J. (2005). How teaching conditions predict teacher turnover in California schools. *Peabody Journal of Education, 80,* 44–70.

Oakes, J. (1990). Opportunities, achievement, and choice: Women and minority students in science and mathematics. *Review of Research in Education, 16,* 153–222.

Orfield, G., Frankenberg, E., & Garces, L. M. (2008). Statement of American social scientists of research on school desegregation to the U. S. Supreme Court in *Parents v. Seattle School District* and *Meredith v. Jefferson County*. *Urban Review, 40,* 96–136.

Rossell, C. H. (1988). How effective are voluntary plans with magnet schools? *Educational Evaluation and Policy Analysis, 10,* 325–342.

Rouse, C. E. (1998). Private school vouchers and student achievement: An evaluation of the Milwaukee parental choice program. *The Quarterly Journal of Economics, 113,* 553–602.

Schofield, J. W. (1995). Review of research on school desegregation's impact on elementary and secondary school students. In J. A. Banks & C. A. McGee Banks (Eds.), *Handbook of research on multicultural education* (pp. 597–617). New York: Macmillan.

Stuart, E. (2007). Estimating causal effects using school-level data sets. *Educational Researcher, 36,* 187–198.

Wilde, E. T., & Hollister, R. (2007). How close is close enough? Evaluating propensity score matching using data from a class size reduction experiment. *Journal of Policy Analysis and Management, 26,* 455–477.

Yu, C. M., Taylor, W. T., Goldring, E., Smrekar, C., & Piche, D. (1997). *Do magnet schools serve children in need?* Washington, DC: Citizens' Commission on Civil Rights.

Authors

ROBERT BIFULCO is an associate professor of public administration at the Maxwell School of

Citizenship and Public Affairs, where he is associated with the Center for Policy Research, Syracuse University, 426 Eggers Hall, Syracuse, NY 13244; rbifulco@syr.edu. He has expertise in the evaluation of educational policies and has conducted research on charter schools, the relationship between school choice and student segregation, school peer effects, performanced-based accountability, the distribution of educational resources, and whole school reform.

CASEY D. COBB is an associate professor of educational leadership in the Neag School of Education and the director of the Center for Education Policy Analysis, University of Connecticut, Gentry Building, Storrs, CT 06269; casey.cobb@uconn.edu. His areas of expertise include the analysis of policies on school choice, desegregation, and education reform.

COURTNEY BELL is an associate research scientist in the Teaching and Learning Research Center at the Educational Testing Service, 10545 Borgman Avenue, Huntington Woods, MI 48070; cbell@ets.org. Her research focuses on the intersections of policy and practice in parental choice, teaching policy, teacher learning, and the measurement of teaching.

Manuscript received August 20, 2008
Final revision received May 15, 2009
Accepted May 27, 2009