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

Robert Bifulco[1]
Syracuse University

Casey Cobb
University of Connecticut

Courtney Bell
Educational Testing Service

Abstract: In response to a landmark civil rights ruling, the state of Connecticut has adopted models of choice-based interdistrict desegregation that appear to satisfy current legal constraints. In this paper, we focus on Connecticut's interdistrict magnet schools, and estimate the effects these schools have had on student achievement. We use longitudinal data on individual student test performance and information from admissions lotteries to implement quasi-experimental, regression-based, and propensity score estimators. Preliminary analyses show that lottery based methods, propensity score methods, and regression analysis provide similar estimates of achievement effects of for the small set of schools for which all three methods can be implemented. We then proceed to use the latter two methods to estimate effects for all of the interdistrict magnet high schools and middle schools that serve students from Hartford, Waterbury and New Haven. Results indicate that, on average, interdistrict magnet high schools have positive effects on both math and reading achievement, and interdistrict magnet middle schools have positive effects on reading achievement. Extensions of our analysis indicate that interdistrict magnet high schools have positive effects particularly on the achievement of students in Hartford, New Haven and Waterbury and do so regardless of how much attending an interdistrict magnet high school reduces racial isolation. The positive effects of magnet middle schools appear to be limited to suburban students, except in those schools that are able to achieve substantial reductions in racial isolation for their central city students.

[1] Corresponding author: Center for Policy Research, 426 Eggers Hall, Syracuse, NY 13244, rbifulco@syr.edu, 315-443-3114

I. Introduction

More than 50 years after *Brown v. Board of Education*, many children attend schools with predominantly minority or poor student bodies. Racial and economic isolation are particularly marked in the Northeast and Midwest where geographically small school districts combined with high levels of residential segregation have slowed school desegregation efforts. In the year 2000, 51 percent of black students in the Northeast and 46 percent in the Midwest attended schools that were more than 90 percent non-white (Clotfelter, 2004).

Two U.S. Supreme Court rulings constrain efforts to reduce school segregation and the attendant racial and economic isolation of poor, minority students. In *Milliken v. Bradley* (1974), by a 5-4 vote, the Court ruled that, unless discriminatory actions on the part of district officials could be shown, federal courts could not order students to be assigned to schools across district lines for the purpose of desegregation. Because much of the segregation in Northeast and upper Midwest metropolitan areas is across districts rather than within single districts, this ruling has limited desegregation in those regions. More recently, and again by a 5-4 vote, the Court has limited the extent to which magnet school programs, which attempt to attract students from a broad geographic area to achieve integrated enrollments, can use the race of a student in making admissions decisions.[2] This decision may limit the ability to achieve desegregation goals through parental choice programs.

In this legal context, desegregation efforts in the state of Connecticut provide an important case study. The wealthiest state in the nation by several measures, Connecticut is home to cities with overwhelming minority populations and poverty rates among the highest in the nation. In a 1996 ruling, the Connecticut Supreme Court held that as a result of racial, ethnic

[2] *Parents Involved in Community Schools v. Seattle School District No. 1*, No. 05-908. Argued December 4, 2006—Decided June 28, 2007.

and economic isolation, Hartford public school students had been denied equal educational opportunity under the state constitution.[3] 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 of these programs are the Open Choice program, which allows transfers between urban and suburban districts in the Hartford, New Haven and Bridgeport areas, and an Interdistrict Magnet School program which has established more than 50 schools open to students from multiple districts.

Several features make these programs 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. Second, the programs are entirely voluntary; neither districts nor individual families are required to participate.[4] Third, although the extent to which the state has achieved racial integration goals in the Hartford area is monitored by the court, the race of individual students is not used in determining admission to the Open Choice program or any interdistrict magnet school. Thus, these programs offer models of choice-based interdistrict desegregation that appear to satisfy current legal constraints and hold the promise for racial integration.

In this paper, we focus on Connecticut's interdistrict magnet schools, and particularly on the effects these schools have had on student achievement. Below we summarize the evidence that Connecticut's interdistrict magnet schools do provide their students, especially those from Connecticut's larger central cities, more diverse peer environments than they would encounter in

[3] *Sheff v. O'Neill*, 238 Conn. 1, 678 A.2d 1267 (1996).
[4] 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). However, what the districts must do to meet this requirement is not specified and the extent of district participation in either the Open Choice or Interdistrict Magnet School programs has been left largely to local decision makers and varies widely across districts.

their home districts. One of the primary motives for reducing racial and economic isolation, however, is the belief that doing so will help to increase the academic achievement of those students who are provided the opportunity to attend more integrated schools. Thus, information on how interdistrict magnet schools effect student achievement is important for assessing the success of this approach to helping poor and minority students who are currently isolated in racially or economically homogenous schools.

The paper is organized as follows. In Section II, we describe the Interdistrict Magnet School program and summarize evidence indicating that interdistrict magnet schools do, on average, provide their students more diverse peer environments than they would encounter in their home districts. In Section III, we explain our approach to estimating the effects of attending an interdistrict magnet school on achievement, an approach which combines quasi-experimental methods that exploit lottery admissions for a small set of schools with multiple regression and propensity score methods that can be applied to a larger set of schools. Sections IV and V demonstrate that in this context, lottery based methods, propensity score methods, and regression analysis can provide similar estimates of magnet school effects on achievement. Section VI reports our best estimates of average achievement effects for the interdistrict magnet high schools and middle schools that serve students from Hartford, Waterbury and New Haven. Section VII explores whether the effects of magnet schools are different for city residents than for suburban residents, and among city residents, whether those who move to more diverse magnet schools benefit more. Section VIII offers conclusions and directions for further research.

II. Connecticut's Interdistrict Magnet School Program & Its Effect on Peer Environments

In Connecticut, an interdistrict magnet school is a publicly funded school operated by a local or regional school district, 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 based on 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, which are described in more detail below.

The state's Interdistrict Magnet School program has encouraged and supported the development of interdistrict magnet schools in several ways. If a magnet school is housed in a new building, the state provides 95 percent funding for the construction of the building. In order to receive the state construction money, the districts must submit a plan for approval by the State Department of Education and commit to the new school for at least twenty years. In evaluating and helping to develop plans, the State Department of Education looks for a governance structure composed of at least three districts and an educational program that will attract a diverse student body.

In the first few years following the *Sheff* decision, operational funding of interdistrict magnet schools was designed to encourage geographic diversity, with the hopes that such diversity would also result in racial and economic diversity. State funds for operating an interdistrict magnet school were based on per pupil payments determined as a percentage of the state's "foundation" level. If the districts participating in the magnet school sent no more than 30 percent of the students to the school, then the magnet school would receive 90 percent of the foundation level for each pupil from each such district. As the percentage from the sending district rose above the 30 percent threshold, the operational reimbursement would decrease correspondingly.

More recently, however, state operating funding has been provided on a flat per pupil basis. For interdistrict magnets operated by a local school district the amount of state operating funding that a student generates depends on whether the student lives inside or outside the district, and state operating funds are supplemented by the local district that operates the magnet. For magnet schools operated by Regional Educational Service Councils (RESCs), all students generate the same amount of state funding, and the local district where the student resides is expected to make payments to the RESC to supplement state funding. One financial incentive for maintaining a diverse student body has been maintained, namely if more than 75 percent of students come from a single district, schools are ineligible for magnet school funding.

The state has provided a further financial incentive for local districts to participate in magnet schools. Districts are permitted to count the students they are sending to the magnet school in their student counts for Education Cost Sharing purposes. That is, if a district sends children to an interdistrict magnet school, it receives the same amount in Education Cost Sharing funds from the state as it would if the students remained in the district's regular schools. In effect, interdistrict magnet school students generate state aid funding twice. The state also provides transportation funding for students who attend an interdistrict magnet school located outside the district in which they live.

At the beginning of the 2005-06 school year, 51 interdistrict magnet schools serving 15,884 students were in operation. Six of the interdistrict magnet high schools are half-time programs, where students attend part of the school day at the magnet and part in their home school, and four of the interdistrict magnets were new in 2005-06. Forty-one of the 51 magnets are located in the Hartford, New Haven or Waterbury areas, and in 2005-06, the magnets in these areas were serving 79.9 percent of all interdistrict magnet school students in the state.

Interdistrict magnet schools clearly provide students of color from Connecticut's most isolated central cities the opportunity to join less isolated learning environments. Panels A and B in Figure 1 compare the percent white in the average student of color's school in each city district's schools and in the interdistrict magnet schools that serve each of the city's residents. 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 very high. In Bridgeport and New Haven the percent white in the typical black student's school is less than 10 percent, and the figures are similar for Hispanics. In Hartford less than six percent of the students are white. Waterbury has a larger population of white students. The students of color from these districts who attend magnet schools are, on average, in substantially more integrated peer environments than their counterparts in central city district schools.

Panel C of Figure 1 compares the percent of students eligible for free and reduced-priced lunch in the average student of color's school in district schools and in the magnet schools that serve Connecticut's most isolated central cities. The percent free-lunch eligible in the interdistrict magnet schools attended by central city students of color is much lower than in the central city district schools these students would otherwise attend, suggesting magnets reduce economic isolation for their students.

The averages reported in Figure 1 hide substantial variation across magnet schools. Interdistrict magnet school programs that serve high school students from Hartford, New Haven or Waterbury range from 5 percent to 75 percent white. Across all 39 interdistrict magnet schools serving students from these cities during the 2005-06 school year, 16 were less than 20 percent white, 11 were between 20 and 30 percent white, and 12 were more than 40 percent white.

Although magnet schools are more diverse, on average, 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. Figure 2 shows that only small percentages of central city students of color attend diverse magnet schools. Less than 6 percent of black students and only 3 percent of Hispanic students residing in Hartford and New Haven attend an interdistrict magnet school with more than 25 percent white students.

Figure 3 compares the percent of students achieving proficiency in magnet schools to a weighted average of the percent of students achieving proficiency in the districts from which magnet school students are drawn.[5] These comparisons indicate that higher percentages of students in magnet schools are achieving proficiency than in the schools magnet school students might have otherwise attended. Of course, not all interdistrict magnet school students encounter an environment with higher levels of achievement as a result of attending a magnet school. In particular, achievement levels in most of the interdistrict magnet schools are not as high as in the highest achieving suburban districts. However, a relatively small percentage of interdistrict magnet school students are drawn from these high achieving suburban districts, and thus, most magnet school students encounter learning environments with higher levels of achievement than they would be exposed to in their home districts. Because they serve higher achieving students, magnets might be able to provide peer environments more conducive to academic success. However, we cannot tell whether differences in proficiency rates are due to more effective instruction in magnet schools, or because more motivated and able students, who would show relatively high levels of achievement choose to attend magnet schools. This question is the focus of this paper.

[5] Weights are based on interdistrict magnet school enrollment from the feeder district.

III. Empirical Strategy

The effects of a chosen school on its students are typically estimated by comparing the achievement of students who attend the school and students in other schools. Such estimates often confound differences in family and personal background between students with the effect 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 often travel longer distances and make other sacrifices to attend a magnet. Thus, we suspect that magnet school students might 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 be used to address unobserved variable bias resulting from self-selection. These studies measure treatment effects by comparing the average outcomes of lottery winners who enroll in a given type of school or program to the average outcomes of students who apply but are denied admission because they lost the lottery. Because lottery winners and losers are determined through a random process, we expect that if the lottery is large enough, the two groups will not differ significantly from each other on either observed or unobserved characteristics. Comparisons of average outcomes across the two groups will, then, be free of unobserved variable bias. This approach has been used to study voucher programs in Washington D.C., New York, and Dayton by Howell and Peterson (2002), intradistrict choice programs in Chicago by Cullen, Jacob and Leavitt (2003), charter schools in Chicago by Hoxby and Rockoff (2005), intradistict magnet schools in a large, Southern district by Ballou (2007), and a variety of choice programs in San Diego by Betts and colleagues (2006).

Although an excellent strategy for addressing potential biases due to self-selection, lottery based analyses do have important limitations. Threats to the internal validity in these studies will be discussed in the next section. Perhaps even more challenging is the fact that the lottery approach can only be applied in limited situations. In any given choice program, some schools either will not be oversubscribed or will not select students randomly, and those that are oversubscribed might not be representative of all schools.[6] What's more, admission lotteries are not typically held on a school wide basis. Rather admission lotteries are held for specific grades and often particular subgroups within grades. As a result, there are often too few winners and losers in particular lotteries to gain the benefits of randomization. Conclusions from lottery studies are often limited to subgroups of schools and types of students within schools, undermining external validity.

Alternative approaches to addressing bias due to self-selection use matching and/or statistical procedures to control for as many observable differences between treatment and comparison groups as possible. Regression analysis and propensity score matching are commonly used to achieve control for observable characteristics. The most convincing studies of this kind include pre-treatment test scores as control variables. Pre-treatment test scores can help control for many factors that influence student achievement and learning. When test scores are available from two or more pre-treatment periods, these methods can be used to determine if treatment group students make larger or smaller test score gains than students with similar pre-

[6] In an early national study, RPP (1997) surveyed charter schools on their admissions processes. Almost three-quarters indicated they were oversubscribed, but only 39 percent of those reported using a lottery to determine admissions. A study of private schools participating in the Milwaukee voucher program, which were required to select students by lottery, found that many were not using lotteries (People for the American Way website, http://www.pfaw.org/pfaw/general/default.aspx?oid=1486#, December 8, 2006).

treatment levels of and rates of growth in achievement. Because they do not use random assignment, estimates from these studies remain subject to potential biases from any unobserved differences between treatment and comparison group members who have similar levels of pre-treatment performance. They often have the advantage, however, of being applicable to a wider range of schools and students than lottery based studies.

Hoxby and Murarka (2008) suggest a way to combine the advantages of lottery-based analyses and comparison-with-controls analyses. 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 due to selection on unobservables are likely to influence estimates derived from propensity score or regression-based analyses. If the researcher is able to specify propensity score procedures or regression models that replicate the results of lottery-based procedures, then one can have more confidence that those same models and procedures 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 that influence selection into magnet schools for which lottery results can be obtained are similar to the factors that influence selection into other magnet schools. In our case, we can obtain effects estimates using lottery results for two schools--one that serves grades six through eight and another that serves grades six through twelve. The broader set of school that we examine using only propensity score matching and regression analysis are also limited to those serving either middle or high school grades. In addition, all of the schools in our analysis, including the two schools for which we use lottery results and the broader set of schools that we examine, serve students primarily from Hartford, New Haven or Waterbury and their surrounding inner ring suburbs. Thus, the schools for whom we have lottery results and the broader set of schools we examine draw students from similar types of districts and the students in these schools have chosen from a similar set of options. There are, then, reasons to think that the factors the influence selection into the schools for which we have lottery results are similar to those that influence selection into the other schools we examine. Therefore, the estimates we derive from propensity score matching and/or regression analyses are strengthened if confirmed by lottery-based analyses.

We begin our analyses here by providing that confirmation. After showing that lottery and comparison-with-controls methods can provide similar results for a small set of schools, we go on to develop the regression-based and propensity score results for a broader range of interdistrict magnet schools.

IV. Lottery-Based Analysis

For this study, we obtained the results of admission lotteries from six interdistrict magnet schools operated by the Capitol Region Education Council (CREC) – three elementary schools, two schools that begin in grade six, and a half-day high school. The analyses in this section focus on the two schools that begin in grade six. Our primary purpose is to determine if regression and propensity score methods that include pre-treatment measures of achievement can replicate estimates obtained from lottery-based analyses. Pre-treatment measures of achievement are not available for students in elementary school magnets, and thus we exclude the three elementary schools here. Also, we are unable to identify which students attend the half-day high school in the dataset we use for the regression and propensity score analyses, so that school is not included either. Both of the remaining schools are located in a first ring suburb close to the city of Hartford and serve the city of Hartford and four suburban districts.[7]

[7] 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.

The admission policies for these two schools are identical. Each school allocates a pre-determined number of seats for each of the districts 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 number. 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 number. When a student from a specific district turns down an admission offer, a seat in that district becomes available and is offered to the next applicant from that district on the waiting list. Applicants are only accepted for sixth grade. If students leave the school after the start of sixth grade those spots are filled with individuals from the original waiting list.

For both schools, we have admissions data on applications submitted in 2003 and 2004. These data were matched to test score file records from 2001-02 through 2006-07 to provide measures of student achievement from the fall of fourth grade, the fall of sixth grade and the spring of eighth grade. These individual test score records were then matched over time. Thus, we have one post-treatment and two pre-treatment measures of student achievement.[8] Information on the students' age, gender, ethnicity, free-lunch status, and special education status is also available from the test score files.

[8] Prior to 2005-06, the Connecticut Mastery Tests (CMTs), 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 2005-06 tests were administered in the spring. All eighth grade test scores are from the spring of either 2005-06 or 2006-07.

Since admissions lotteries are district and year specific, we have a total of 22 potential lotteries.[9] 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 students from eight potential lotteries which did not have any losers. All of the applicants in these lotteries were eventually offered a seat in the school, and 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, so they did not participate in a true lottery. The two Hartford lotteries for the other school are also dropped for different reasons. First, attrition rates are high among participants in these lotteries—only 50 percent of these applicants have eighth grade tests scores, our outcome of interest. Second, unlike students from other districts, Hartford students have many different ways to opt out of the regular public schools—including other magnet schools, Open Choice, and charter schools. As a result, very few students who are lotteried out of the magnet school we are examining end up in Hartford public schools, which complicates interpretation of the magnet school effect we are trying to estimate.

Random assignment helps to ensure that lottery winners are similar to lottery losers on both 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 have substantial attrition from our samples. We are missing records from the test score files for any student who participated in a magnet school lottery but whom we could not match to a test score record either because they attended a school outside the Hartford metropolitan region, enrolled in a private school, or

[9] Five district specific lotteries in both 2003 and 2004 for both schools implies 5x2x2=20 lotteries. 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.

otherwise could not be located in the test score file. We observe eighth grade test scores, our outcomes of interest, for 67.4 percent of the lottery participants in our sample. Attrition rates are similar for lottery participants offered admission and those not offered admission--we observe eighth grade test scores for 70.0 percent of those offered admission and 66.0 percent of those never offered admission. Nonetheless, if the lottery losers who attrit from our sample differ in systematic ways from the lottery winners who attrit, then there might be important differences between lottery winners and lottery losers.

To test whether our lotteries produce balanced treatment and control groups, we examined differences between the non-attriting winners and losers of each lottery on a range of observable characteristics, including scores on math and reading tests administered at the beginning of the sixth and fourth grades. These tests revealed that in one of the smaller lotteries non-attriting winners had significantly lower test scores than non-attriting losers.[10] This lottery was dropped from our final sample. The remaining lotteries produced groups of non-attriting winners and losers that are similar on observed characteristics. Similarity on observable characteristics does not guarantee that systematic differences in attrition did not result in unobserved differences between winners and losers in these lotteries, but we have no strong reason to suspect such differences.

Our final sample includes applicants from 11 lotteries. For each lottery we can define those who are assigned low enough random numbers to be offered admission immediately, those who are not offered admission until other students have declined, and those who are never offered admission. We refer to these as on-time winners, delayed winners, and lotteried-out

[10] This lottery included 56 "winners" and only 11 students who were not offered admission. Of these 67 students, 25 are missing sixth grade test file records. For only five of these 25 attriters do we observe information from either the fourth or eighth grade test score files. Thus, we have no way to determine if this lottery produced an unbalanced set of winners or losers by chance, or whether differences between winners and losers result from differential attrition.

applicants, respectively. The 11 lotteries used in our analysis include a total of 866 applicants—229 on-time winners, 94 delayed winners, and 543 lotteried-out applicants. We can observe eighth grade test scores for 593 of these applicants, and among these, 173 are on-time winners, 53 are delayed winners, and 367 are lotteried-out.

To demonstrate that lottery winners and losers are balanced on observable characteristics Table 1 presents the results of a series of regressions. Each row in this table presents the regression of an observable characteristic on an indicator of whether or not the student won the lottery and a set of lottery dummy variables. In these regressions, on-time winners are counted as winners and delayed winners are excluded from the sample.[11] In all of the regressions except the first one, the coefficients on the lottery winner indicator is not statistically distinguishable from zero, which indicates that except for age there are not statistically significant differences between lottery winners and lottery losers. Given that t-tests from 12 separate regressions are reported in Table 1, it is not unreasonable to expect one significant result to emerge by chance. Most importantly, there are no significant differences between lottery winners and losers on pre-treatment measures of achievement.

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, where this variable can be defined to include on-time winners only, or both on-time winners and delayed winners; μ_L represents lottery specific fixed effects; and e_{iL} is a random error term. α can be estimated

[11] We also ran analogous regressions using all applicants that we observe in eighth grade and including on-time and delayed winners as lottery winners. The results of these regressions were quite similar to those reported in Table 1.

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 specific lottery, as random assignment helps to ensure, then the difference in mean eighth grade tests scores between the two groups is due solely 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 together the effects of magnet schools on the achievement of those who choose to enroll and the presumably zero effect on those who do not enroll. Thus, as an estimate of the magnet effect, α is biased toward zero. 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 in the case of many medical treatments where subjects' willingness and ability to comply with the treatment influences 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 solution to attenuation bias due to non-compliance 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:

$$\begin{aligned} &FirstStage: M_{iL} = \beta W_{iL} + \lambda_L + v_{iL} \\ &SecondStage: Y_{iL} = \gamma \hat{M}_{iL} + \theta_L + \omega_{iL} \end{aligned} \tag{2}$$

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, i.e., the effect of magnet schools on the students who attend them. There is some question about whether an indicator of on-time winning, with delayed winners excluded, or an indicator that includes on-time and delayed winners, is the most appropriate instrument for magnet school enrollment. Ballou (2007) argues that delayed winners who nonetheless 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.

If lotteries are truly random, we would not expect any significant differences between lottery winners and losers, and the simple regressions above 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, 2006). Second, in 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. Pre-treatment measures of achievement are especially useful for these purposes.

Table 2 presents estimated effects on eighth grade math and reading scores. The table includes: estimated effects of the intent to treat derived from regression (1); estimates of treatment on treated effects from (2) above; and estimates of the treatment on treated effects from versions of (2) that add individual level covariates including sixth grade math 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 including on-time and delayed winners.

The results are very 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 and including covariates in the two-stage least squares estimates increases precision substantially. The estimated magnet school effects are larger when covariates are included, especially for math where estimates are more than twice as large and become statistically significant when covariates are added. The difference between the models with and without covariates might seem to suggest that randomization is not working. However, these results are not necessarily inconsistent with the results in Table 1 which suggests that random assignment via the admission lotteries produced treatment and control groups with no significant differences. The estimated effects of the treatment on the treated obtained from the regressions without covariates are within the 95 percent confidence interval of the estimates obtained from the regressions that include covariates. Also, the results in Table 1 suggest that although the differences between lottery winners and lottery losers in sixth grade test scores are not statistically significant, they are also not zero. The differences between the estimates from models with and without covariates are roughly equal to the differences in sixth grade test scores between lottery winners and lottery losers presented in Table 1. Because the estimates that include covariates correct for these differences in pre-treatment test scores they are our preferred estimates.

The results in Table 2, then, indicate that these two interdistrict magnet schools have had positive effects on student achievement. The estimated effects on reading are larger than the estimated effects on math. To give an idea of the magnitude of these effects, the standard deviation in both eighth grade reading and math scores is approximately 45, and thus the estimates of the treatment on treated effect from the models that include covariates indicate that reading test scores of students in these magnet schools are between 0.26 and 0.29 standard deviations higher and their math scores are approximately 0.14 standard deviations higher than they would be if those students had attended other schools.

V. Do Comparison-with-Controls Methods Replicate Lottery Based Estimates?

To answer this question we assembled a dataset 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 were matched back to sixth and fourth grade test score records for the same student. Connecticut has only recently 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 successfully match 85 percent of these student records to a sixth grade test score record and 75 percent to both a sixth grade and a fourth grade test score record.

We begin by using these data to estimate the following regression model:

$$Y_{i8} = \alpha \mathrm{M}_{i8} + X_i \mathrm{B} + \mu_t + v_i \tag{3}$$

where Y_{i8} is student i's eighth grade test score; M_{i8} is an binary variable indicating whether or not 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; 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, free-lunch eligibility, and pre-treatment math and reading test scores. A model that includes sixth grade test scores controls for test score levels at the beginning of middle school, and a model that includes both sixth and fourth grade test scores controls for test score gains made between grade four and six as well as test score levels.

Although it does not necessarily control for potential unobserved differences between treatment and comparison group students, regression analysis can be an effective method to control for observable characteristics. When pre-treatment measures of achievement are available, such controls for observables can provide plausible effect estimates. If the effects of observed covariates on the outcome of interest are not linear, however, a linear regression like that specified in (3) may not provide effective controls for differences between treatment and comparison group members. This problem can lead to large biases if treatment and comparison groups have very different covariate distributions (Stuart, 2007). Also, if treatment effects vary across individuals, then regression returns a weighted average of the treatment effects on different individuals where the weights are designed to minimize variance in coefficient estimates. This weighted average need not correspond to the average effect of the treatment on the treated, which is usually the parameter of policy interest (Cobb-Clark & Crossley, 2003).

Propensity score matching refers to a set of techniques for controlling for observed characteristics that avoids these potential problems with regression analyses. Propensity score methods begin by estimating a probability model to predict the likelihood that a student will select into a treatment. Here we estimate a logit model with enrollment in one of our two magnet schools in eighth grade as the dependent variable, and all the variables included as covariates in our regression analyses as independent variables. Estimates of this model can be used to compute a predicted likelihood of selecting one of our magnet schools, what is referred to as a propensity score.

To ensure full conditioning on observable characteristics, specification of the probability model used to estimate the propensity score deserves careful attention. Dehejia and Wahba (2002) recommend an iterative procedure for determining a specification that balances the covariates across treatment and comparison group members. We use that procedure here.[12] We estimated propensity scores using a model that, in addition to other covariates, includes sixth grade test scores only, and another that makes use of fourth grade test scores as well.

The next step in propensity score matching is to limit the sample to the area of common support. In this case, we limit the set of potential comparison group students to those whose propensity scores are at least as high as the treatment group student with the lowest propensity score, and similarly limit our treatment group to students whose propensity score is at least as low as the comparison group student with the highest propensity score. This step ensures that the treatment and comparison students used to estimate treatment effects have similar distributions on the covariates.[13]

Propensity scores can be used to compute estimates of treatment effects in a number of ways. Here we use two. The simplest approach is the nearest neighbor method, in which each student in our treatment group is matched to the one comparison group member with the closest propensity score, and the effect estimate is calculated as the difference in average eighth grade test scores between the treatment group and the matched comparisons.[14] The second method uses a kernel density based estimator. In this procedure the eighth grade test score of each treatment group student is compared to a weighted average of all the comparison group

[12] The resulting specifications of the logit models, along with results of estimating the models, are available upon request.

[13] For the estimations presented in Table 3, this step resulting in dropping 0.9 percent of the comparison groups students and none of the interdistrict magnet school students.

[14] Matching can be done with or without replacement. Matching without replacement maximizes the number of comparison group members selected, but can lead to large differences between particular treatment group members and their matched comparisons. Results of matching without replacement can also be sensitive to the order in which treatment units are matched. Here we use matching with replacement.

members, where the weights are determined by a measure of the distance between the comparison group student's propensity score and the treatment group member's propensity score. The average of these differences across all treatment group members provides an estimate of the average effect of treatment on the treated. The nearest neighbor method ensures that treatment group members are compared only to the most similar comparisons available, and thereby minimize potential biases. In contrast the kernel density estimator makes maximal use of all the available information, and can substantially increase precision.

Table 3 presents the results of our comparison-with-controls estimates of the effect of these two interdistrict magnet schools. The first column presents results from OLS estimates of equation (3). The next two columns present estimates based on propensity score matching. The second column presents estimates based on nearest neighbor comparisons and the third column presents results from kernel density matching. Estimates in the upper panel are derived using only sixth grade test scores as pre-treatment achievement measures, and estimates in the lower panel are derived from models that include fourth grade test scores as well as sixth grade test scores. To facilitate comparison, we have also included lottery-based estimates of the average treatment-on-treated effect taken from the third and sixth columns of Table 2.

The results from the comparison-with-controls methods are similar whether test scores from two pre-treatment periods or only one pre-treatment period are used. Results are also similar across the three comparison-with-controls methods. In this case at least, regression based estimates do not appear to be subject to the potential problems advocates of propensity score matching emphasize. As expected, the kernel density matching provides much more precise estimates than nearest neighbor matching. As with the lottery estimates, the estimated effects of attending one of these magnet schools are positive and statistically significant for both math and reading, and the estimated effects on reading are approximately twice as large as the estimated effects on math.

For both math and reading, the point estimates from the comparison-with-controls methods are up to 25 percent larger than the estimates from the lottery-based analysis presented in the next to last column.[15] The differences, however, are substantially small. For reading, the comparison-with-controls methods imply an effect size of about 0.33 standard deviations compared to an effect size between 0.26 and 0.29 standard deviations implied by the lottery-based analysis. For math, the comparison-with-controls estimates in the lower panel of Table 4 imply an effect size of approximately 0.17 standard deviations compared with an effect size of about 0.14 standard deviations from the lottery analysis. Such small differences are unlikely to influence policy conclusions. Also, all of the comparison-with-controls estimates are comfortably within the 95 percent confidence interval for the corresponding lottery-based estimates, indicating that the differences between estimates from the comparison-in-control methods and the lottery-based analyses are not statistically significant.

VI. The Average Effects of Interdistrict Magnet Schools

Having shown that regression analysis and propensity score methods that use pre-treatment 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 3 suggest that students' who self-select into the two magnet schools we have examined so far do not differ from comparison group students picked out by regression and propensity score methods in any unobserved ways that substantially bias estimated effects on

[15] Reading estimates from the comparison-with-controls methods are closer to 15 percent higher when compared with lottery-based estimates that include on-time and delayed winners, i.e. when compared to estimates in the last column.

achievement. These results do not, however, guarantee that students who select into other magnet schools do not differ in unobserved ways from observationally similar students, or that selection-on-unobservables will not bias effect estimates for those other magnets. Nonetheless comparison-with-controls methods that include pre-treatment test scores represent the best available methods for estimating the effects of larger sets of interdistrict magnet schools, and given the confirmation of these methods provided in Table 3, 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 another set of 6 interdistrict middle schools.[16] These schools include all of 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.[17] We focus on estimating the effects of the interdistrict magnet high schools on 10th grade reading and math Connecticut Academic Performance Test (CAPT) exams, and the effects of the interdistrict magnet middle schools on 8th grade math and reading CMTs. The CAPT is the high school statewide testing programs administered by the state.

To construct our student sample for the analysis of the twelve interdistrict magnet high schools, we asked officials at the state department of education to extract from the 2005-06 and 2006-07 tenth grade CAPT records for all of the students attending either one of those interdistrict magnets or a high school in a district that sends students to one of those interdistrict magnets. We then asked the state officials to match those student records to records from earlier eighth grade and sixth grade test score files. Our sample for the middle school analysis was

[16] High schools here are schools that serve grades 9-12, and middle schools are schools that begin in grade 6 or 7. Four of the six "middle schools" end in grade 8, but two serve high school grades as well.
[17] Two interdistrict magnet schools that serve students from New Haven start in Grade 5 and are not included in this analysis.

constructed in an analogous manner. A sample of 1,730 magnet high school students and 13,507 students from feeder districts were extracted from the tenth grade test score files. State officials were able to match 71.5 percent of these students to eighth grade test score records and 58.7 percent to sixth grade test score records. Longitudinal matching rates were higher for 8^{th} graders. Of the sample of 1,248 magnet school students and 18,563 students from feeder districts extracted from the eighth grade test score files, state officials were able to match 80.4 percent to sixth grade test score records and 65.9 percent to fourth grade test score records. Table 4 presents summary statistics on the sample of tenth grade students who we were able to match to an eighth grade test score record and the sample of eighth grade students who we were able to match to a sixth grade test score.

We use exactly the same procedures as described in the previous section to compute OLS, nearest neighbor, and kernel density matching estimates of effects on math and reading achievement. The results are presented in Table 5. The top panel presents OLS, nearest neighbor, and kernel density estimates of the average effects of interdistrict magnet high schools on tenth grade math and reading achievement. The first three columns present results derived from models that use only eighth grade test scores, as well as other individual level covariates, and the last three columns, include estimates from models that use measures of achievement from two pre-treatment periods. The bottom panel shows that same set of estimates of the average effects of interdistrict middle schools on eighth grade math and reading achievement.

The high school results indicate that, on average, interdistrict magnet schools have had positive and statistically significant effects on both tenth grade math and tenth grade reading achievement. The lone exception is the nearest neighbor estimates of the effects on grade 10 math based on propensity scores derived from models that include both sixth and eighth grade test scores as controls. The estimated effects of attending a magnet school are small, but not insignificant from a policy perspective. The estimated effects on reading are between 0.08 and 0.12 standard deviations, and on math are between 0.05 and 0.11 standard deviations. These represent the effects of two 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 0.16 and 0.24 for reading and between 0.10 and 0.22 for math.

The estimated effects for middle schools are weaker for math, but stronger for reading. The results indicate that interdistrict middle school magnets, on average, have small, and with the exception of the kernel density estimates, statistically insignificant effects on eighth grade math scores. The estimated effects on eighth grade reading, in contrast, are positive, statistically significant, and quite large. The estimates in the bottom row imply effect sizes between 0.17 and 0.21 standard deviations over the seventh and eighth grade years.

VII. Extensions

The estimates in Table 5 represent the average effects of interdistrict magnet schools on student achievement. We are also interested to know whether the effects of magnet schools vary across students and schools. Connecticut's interdistrict magnet school program is motivated in large part by a court order requiring state efforts to reduce the isolation and raise the educational achievement of central city minority students. Thus, the effects of interdistrict magnet schools specifically on students from Connecticut's racially isolated central cities are of interest. To explore this question we split both our high school and middle school samples into students who reside in Hartford, New Haven or Waterbury (city students) and students who reside in other districts served by interdistrict magnet schools (suburban students). We then calculated our regression and propensity score estimates of magnet school effects separately using these two subsamples. The results are presented in Table 6.

The first two columns of the top panel of Table 6 present the estimated effects of attending a magnet high school on grade 10 test scores obtained from OLS regressions. These estimates indicate that interdistrict magnet high schools have had significant, positive effects on both the math and reading achievement of their city students. The coefficient estimates imply that attending an interdistrict magnet high school increases the tenth grade math achievement of central city students by 0.12 standard deviations and the tenth grade reading achievement of those students by 0.15 standard deviations. The estimated effects of interdistrict magnet high schools on their suburban students are smaller than the estimated effects on central city students and are not significantly different than zero. The differences between the estimated effects on city students and on suburban students, however, are not statistically significant. The results from the propensity score methods, presented in columns 3-6, largely confirm the OLS results with one exception—the nearest neighbor estimates indicate effects on grade 10 math are larger for suburban students than for city students. The nearest neighbor estimates, however, are relatively imprecise, and thus we put more stock in the kernel density estimates which are very similar to the OLS estimates.

The story is different for middle school magnets. The estimates in the bottom panel of Table 6 indicate that interdistrict magnet middle schools have small, statistically insignificant effects on city students and larger, statistically significant impacts on suburban students. What's more, the estimated effects on grade 8 reading for suburban students are statistically different than the estimated effects on grade 8 reading for city students. These results suggest that the positive, average effects of interdistrict magnet middle schools are driven primarily by positive effects on suburban students, and that interdistrict magnet middle schools are not doing as much as interdistrict magnet high schools to increase the achievement of students from Connecticut's central cities.

One of the animating ideas of interdistrict magnet schools is that they can help inner city minority students improve their academic achievement by providing access to less racially isolated learning environments. As we have seen magnet schools do on average provide their students access to less isolated schools. However, there is substantial variation across interdistrict magnet schools in the extent to which they achieve this goal. We are interested in whether or not those magnet schools that do more to reduce racial isolation for their students have different effects on the achievement of central city students than magnet schools that have been less successful in reducing racial isolation.

To explore this question, for each magnet school student residing in a central city, we identified the differences between the percent white in the magnet school they attend and the percent white in their home district—which we label Δ%white. This difference is intended as a measure of the extent to which the magnet school reduces racial isolation for the student.[18] We then estimated two regressions, using our sample of city students, to determine whether the effect of attending a magnet school varied with this measure. First, we add to the regressions presented earlier (Equation 3) an interaction between the magnet school indicator and our measure of the extent to which attending the magnet school reduces racial isolation. Second, we replace the single magnet school indicator in the regressions presented earlier with separate indicators of whether or not the student attends a magnet school that reduces racial isolation by less than 10

[18] A better measure would be the difference in the racial composition of the magnet school and the school the student would otherwise attend, but we don't have that information, and so the difference between the magnet school and the home district is our best approximation.

percentage points, between 10 and 30 percentage points, and more than 30 percentage points. The results of these regressions are presented in Table 7.

For high school magnets, the coefficients on the magnet*Δ%white interaction are positive. The fact that these coefficients are not statistically distinguishable from zero, however, indicates that the effect of magnet high schools does not vary significantly with the extent to which attending a magnet reduces racial isolation. The results for the second specification indicate that interdistrict magnet high schools have significant, positive effects regardless of how much they reduce racial isolation. The estimated effects are largest when the magnet school reduces racial isolation by more than 40 points. That particular estimate, however, is based on only one magnet school, and thus it is difficult to attribute this large positive effect to racial composition rather than to other unique aspects of the school. In general, then, the effect of interdistrict magnet high schools do not appear to depend on the extent to which they reduce racial isolation.

Among middle schools, the relationship between the magnet school effect and the extent to which attending a magnet reduces racial isolation is stronger. The coefficients on the magnet*Δ%white interaction are larger for middle school magnets than for high school magnets, and is statistically significant in the Grade 8 math regression. The second specifications indicate that the effect of attending a magnet middle school is negative when racial isolation is reduced by less than 10 percentage points and significantly positive only when racial isolation is reduced by more than 40 points. Only one magnet middle school reduces racial isolation of its city students by less than 10 percentage points and only two reduce it by more than 40 percentage points. Thus, we cannot draw strong conclusions from these regression results. Nonetheless, the bottom panel of Table 7 suggests that the effect of magnet middle schools on central city students might depend on whether or not they are able to reduce racial isolation.

VIII. Conclusions

Its 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, minority students in central city schools. Although, interdistrict magnets provide only a small fraction of students in Connecticut's central cities access to diverse schools, they do provide less racially isolated and higher achieving environments than most of their students would otherwise encounter. The key question we have tried to address here is whether interdistrict magnet school students also have higher levels of achievement than 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 School program, regression and propensity score methods that make use of pre-treatment 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 interdistrict magnet high schools have positive effects on both math and reading achievement, and interdistrict magnet middle schools have positive effects on reading achievement. That interdistrict magnets 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.

The extensions of our analysis indicate that interdistrict magnet high schools have positive effects particularly on the achievement of students in Hartford, New Haven and Waterbury who would otherwise be attending very racially and economically isolated schools. We find that the effects of interdistrict magnet high schools on central city students are positive regardless of how much attending an interdistrict magnet high school reduces racial isolation. Although this finding does not diminish the value of interdistrict magnet schools to their students, it does suggest that their positive effects on achievement are due to something other than their effect on racial isolation.

The story is somewhat different for magnet middle schools. The positive effects of magnet middle schools appear to be limited to suburban students, except in those schools that are able to achieve substantial reductions in racial isolation for their central city students. Whether those interdistrict magnet middle schools that have positive effects on the achievement of central city students do so because they reduce racial isolation or because of other positive aspects of those schools is impossible to determine from such a small sample of schools.

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 non-pecuniary 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 costs, we believe the results here show that Connecticut's interdistrict magnet schools provide a promising model for other states concerned with the effects of racial and economic isolation—a model that deserves more attention.

REFERENCES

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

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

Clotfelter, Charles. (2004). *After Brown: The rise and retreat of school desegregation*. Princeton, NJ: Princeton University Press.

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

Cullen, J.B., Jacob, B.A. & Levitt, S. (2003). The effect of school choice on student outcomes: Evidence from randomized lotteries. National Bureau of Economic Research Working Paper 10113. Cambridge MA: NBER.

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 Earlbaum & Associates.

Hoxby, C.M. & Rockoff, J. (2005). The Impact of Charter Schools on Student Achievement. Unpublished paper.

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

RPP International. (1997). *A national study of charter schools: Second-year report*. Washington, DC: U.S. Department of Education. Office of Educational Research and Improvement.

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

Table 1: Comparison of Lotteried-in and Lotteried-Out Students

	On-Time Winner	Constant	Observations	R-squared
Dependent Variable				
Age in Years	0.080*	13.837*	540	0.031
	(0.036)	(0.039)		
Female	0.085	0.523*	540	0.027
	(0.047)	(0.051)		
Black	-0.015	0.417*	540	0.218
	(0.040)	(0.043)		
Hispanic	0.020	0.359*	540	0.153
	(0.044)	(0.047)		
White	0.021	0.176*	540	0.052
	(0.029)	(0.031)		
Asian	-0.026	0.046*	540	0.013
	(0.019)	(0.021)		
free lunch eligible	-0.003	0.260*	540	0.024
	(0.038)	(0.041)		
Special Education	0.027	0.031	540	0.041
	(0.023)	(0.025)		
6th Grade Math Score	-3.142	244.866*	496	0.086
	(3.901)	(4.173)		
6th Grade Reading Score	-2.514	243.210*	497	0.048
	(4.016)	(4.299)		
4th Grade Math Score	-0.583	241.159*	458	0.062
	(4.260)	(4.621)		
4th Grade Reading Score	-0.611	241.030*	458	0.068
	(4.256)	(4.620)		

Samples include students with either an eighth grade math or eight grade reading test score. All regressions include a lottery fixed effect. Standard errors in parentheses.* significant at 0.05 level.

Table 2: Lottery Based Estimates of the Effect of Interdistrict Magnet Schools on Achievement

	On-Time Lottery Winners			*On-Time+Delayed Lottery Winners*		
	Intent to Treat	Treatment on Treated	Treatment-on-Treated (w covariates)	Intent to Treat	Treatment on Treated	Treatment-on-Treated (w covariates)
Grade 8 Math	2.284	2.735	6.242*	1.837	2.245	6.442*
	(3.445)	(4.124)	(2.219)	(3.158)	(3.859)	(2.049)
N	537	537	492	590	590	541
R-squared	0.075	0.076	0.770	0.068	0.068	0.769
Grade 8 Reading	7.599*	9.075*	11.959*	7.722*	9.414*	12.961**
	(3.718)	(4.444)	(2.756)	(3.430)	(4.187)	(2.563)
N	538	538	493	591	591	542
R-squared	0.070	0.069	0.696	0.067	0.065	0.697

Each set of results are from separate regessions. Results in column labeled Intent to Treat are OLS regressions of test score on indicator of whether or not student won the admission lottery. Results in Treatment on Treated columns are IV estimates using indicator of students who won lottery as instrument for enrollment in interdistrict magnet school during eighth grade. Covariates included in models presented in colums 3 and 6 include student's age, gender, ethnicity, free-lunch eligibility in grade 4, special education status in grade 4, grade 6 math and reading scores. In the first three columns only on-time lottery winners are counted as lottery winners and 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 in parentheses. * indicates statistically significant at 0.05 level.

Table 3: Regression, Propensity Score and Lottery Based-Estimates of the Average Effect of Treatment of Treated

	Using Only Grade 6 Test Score with Other Covariates				
	OLS	Nearest Neighbor	Kernel Density Matching	Lottery Estimates	
Grade 8 Math	7.701*	10.074*	7.866*	6.242*	6.442*
	(3.553)	(3.534)	(1.597)	(2.219)	(2.049)
Grade 8 Reading	15.277*	15.655*	15.164*	11.959*	12.961**
	(0.921)	(3.166)	(1.458)	(2.756)	(2.563)
	Using Grade 4 and Grade 6 Test Scores with Other Covariates				
	OLS	Nearest Neighbor	Kernel Density Matching	Lottery Estimates	
Grade 8 Math	7.482*	6.959*	7.475*	6.242*	6.442*
	(3.440)	(3.503)	(1.587)	(2.219)	(2.049)
Grade 8 Reading	15.090*	14.925*	14.991*	11.959*	12.961**
	(0.910)	(3.413)	(1.708)	(2.756)	(2.563)

OLS regressions include age, gender, ethnicity, free-lunch eligibility, special education status, and year as well as pre-treatment test scores and magnet enrollment indicator. The same variables, plus some higher order terms and interactions are used to estimate propensity scores used in nearest neighbor and kernel density estimators. The figures in parentheses are standard errors. OLS standard errors are adjusted for clustering at the school level. Nearest neighbor and kernel density standard errors are bootstrapped. * indicates statistically significant at 0.05 level.

Table 4: Treatment and Comparison Group Samples

	10th Grade Magent Students	10th Grade Students from Feeder Districts	8th Grade Magent Students	8th Grade Students from Feeder Districts
N	1083	8062	1060	14873
Age	15.4	15.4	13.4	13.5
	(0.494)	(0.508)	(0.473)	(0.554)
Male	0.394	0.472	0.498	0.499
	(0.489)	(0.499)	(0.500)	(0.500)
Black	0.516	0.287	0.384	0.263
	(0.500)	(0.452)	(0.487)	(0.440)
Hispanic	0.267	0.246	0.256	0.274
	(0.443)	(0.431)	(0.436)	(0.446)
White	0.194	0.444	0.340	0.436
	(0.396)	(0.497)	(0.474)	(0.496)
Asian	0.020	0.021	0.019	0.026
	(0.141)	(0.142)	(0.136)	(0.159)
Free lunch eligible	0.591	0.466	0.408	0.445
	(0.492)	(0.499)	(0.492)	(0.497)
Special Education	0.065	0.115	0.058	0.110
	(0.246)	(0.319)	(0.234)	(0.313)
8th Grade Math Score	234.0	232.1		
	(35.7)	(42.3)		
8th Grade Reading Score	241.5	237.5		
	(41.8)	(47.7)		
6th Grade Math Score	239.4	238.3	248.3	239.7
	(37.1)	(43.5)	(41.4)	(48.0)
6th Grade Reading Score	235.8	234.5	245.1	233.1
	(40.4)	(46.2)	(41.7)	(47.2)
4th Grade Math Score			246.1	237.6
			(42.4)	(45.8)
4th Grade Reading Score			243.6	235.4
			(40.6)	(43.9)

Means with standard deviations in parentheses.

Table 5: Estimates of the Average Effect of Treatment of Treated for High School and Middle School Interdistrict Magnet Schools

	High School Magnets					
	Using Only Grade 8 Test Scores			*Using Grade 6 & 8 Test Scores*		
	OLS	Nearest Neighbor	Kernel Density Matching	OLS	Nearest Neighbor	Kernel Density Matching
Grade 10 Math	4.601*	4.804*	3.601*	3.908*	2.229	3.476*
	(2.105)	(1.850)	(0.828)	(2.018)	(1.610)	(0.704)
Grade 10 Reading	4.705*	3.650*	4.153*	4.547*	5.555*	4.680*
	(2.053)	(1.738)	(0.823)	(2.066)	(1.871)	(0.841)
	Middle School Magnets					
	Using Only Grade 6 Test Scores			*Using Grade 4 & 6 Test Scores*		
	OLS	Nearest Neighbor	Kernel Density Matching	OLS	Nearest Neighbor	Kernel Density Matching
Grade 8 Math	2.902	2.209	4.230*	3.063	1.639	3.979*
	(2.827)	(2.230)	(0.890)	(2.848)	(2.200)	(0.873)
Grade 8 Reading	8.328*	8.000*	9.316*	8.045*	7.575*	8.759*
	(2.476)	(1.935)	(1.159)	(2.519)	(2.404)	(0.855)

OLS regressions include age, gender, ethnicity, free-lunch eligibility, special educaiton status, and year as well as pre-treatment test scores and magnet enrollment indicator. The same variables, plus some higher order terms and interactions are used to estimate propensity scores used in nearest neighbor and kernel density estimators. The figures in parentheses are standard errors. OLS standard errors are adjusted for clustering at the school level. Nearest neigbor and kernel density standard errors are bootstrapped. * indicates statistically significant at 0.05 level.

Table 6: Estimated Magnet School Effects, By Student's Residence

	High School Magnets					
	OLS		Nearest Neighbor		Kernel Density Matching	
	City Students	Suburban Students	City Students	Suburban Students	City Students	Suburban Students
Grade 10 Math	5.339*	3.208	2.230	5.171	5.237*	2.846
	(2.222)	(2.470)	(2.543)	(3.032)	(0.913)	(1.543)
Grade 10 Reading	6.629*	3.643	3.562	4.527	6.426*	4.020*
	(2.058)	(2.819)	(2.548)	(3.202)	(1.132)	(2.124)

	Middle School Magnets					
	OLS		Nearest Neighbor		Kernel Density Matching	
	City Students	Suburban Students	City Students	Suburban Students	City Students	Suburban Students
Grade 8 Math	1.523	5.846*	-1.335	3.846	0.831	6.657*
	(2.986)	(2.544)	(2.426)	(3.486)	(1.209)	(1.038)
Grade 8 Reading	3.077	12.000*	0.466	10.455*	2.844*	12.629*
	(2.580)	(2.291)	(2.492)	(2.915)	(1.230)	(1.173)

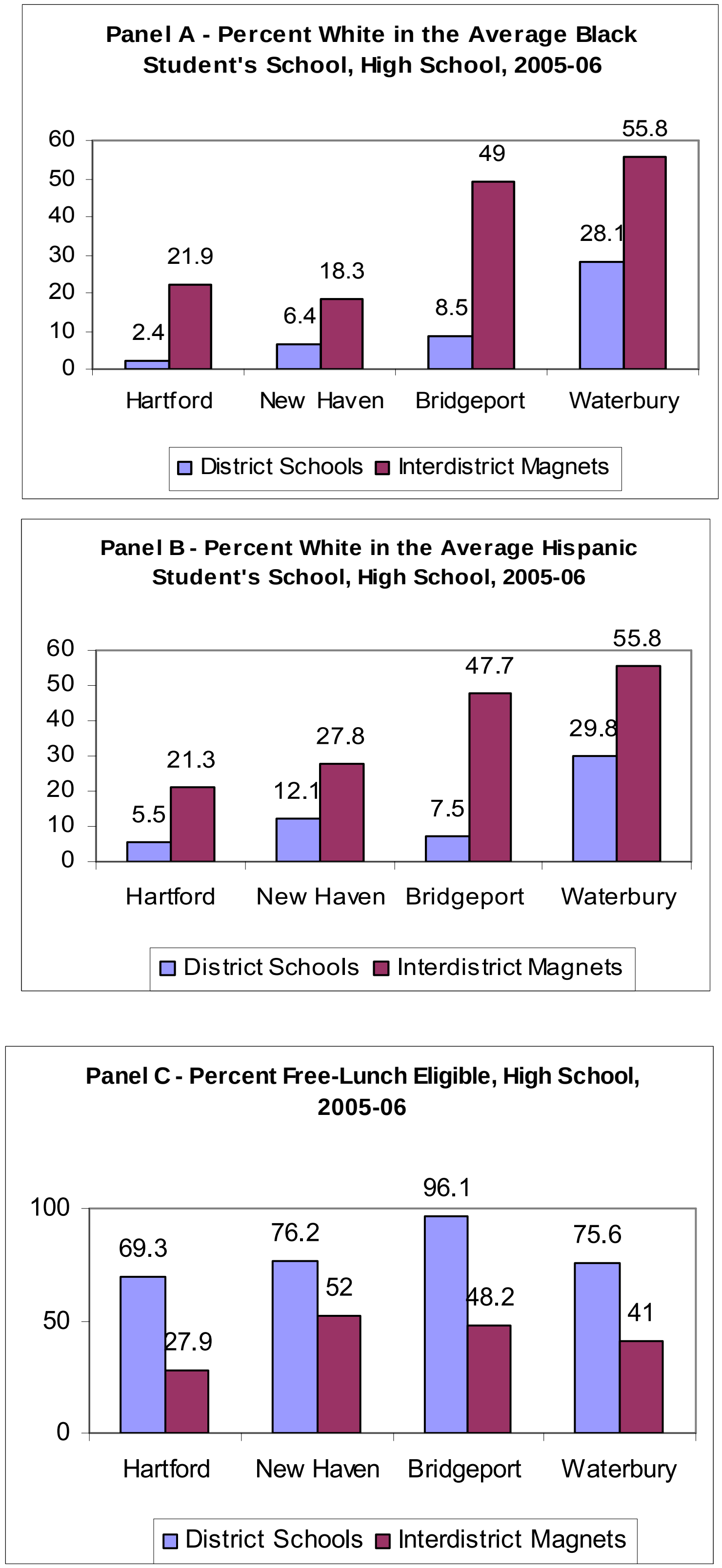
OLS regressions include age, gender, ethnicity, free-lunch eligibility, special education status, year and pre-treatment math and reading test scores from two previous periods. The same variables, plus some higher order terms and interactions are used to estimate propensity scores used in nearest neighbor and kernel density estimators. The figures in parentheses are standard errors. OLS standard errors are adjusted for clustering at the school level. Nearest neighbor and kernel density standard errors are bootstrapped. * indicates statistically significant at 0.05 level.

Table 7: Variation in Estimated Magnet School Effect on City Students, by Difference in Percent White Between Magnet and Home District

	High Schools			
	Grade 10 Math		Grade 10 Reading	
	(1)	(2)	(1)	(2)
Magnet	4.907*		5.315*	
	(2.099)		(2.168)	
Magnet*Δ%White	0.049		0.103	
	(0.102)		(0.123)	
Magnet (Δ%White<10)		4.781*		5.552*
		(1.890)		(1.889)
Magnet (Δ%White 10-30)		5.868		6.988*
		(3.229)		(2.991)
Magnet (Δ%White>40)		11.094*		17.398*
		(1.420)		(1.377)
	Middle Schools			
	Grade 8 Math		Grade 8 Reading	
	(1)	(2)	(1)	(2)
Magnet	-7.948		-2.381	
	(4.951)		(4.998)	
Magnet*Δ%White	0.396*		0.230	
	(0.147)		(0.153)	
Magnet (Δ%White<10)		-10.207*		-6.040*
		(2.539)		(1.321)
Magnet (Δ%White 10-30)		2.765		4.248
		(2.318)		(2.398)
Magnet (Δ%White>30)		6.874*		6.665*
		(3.296)		(2.040)

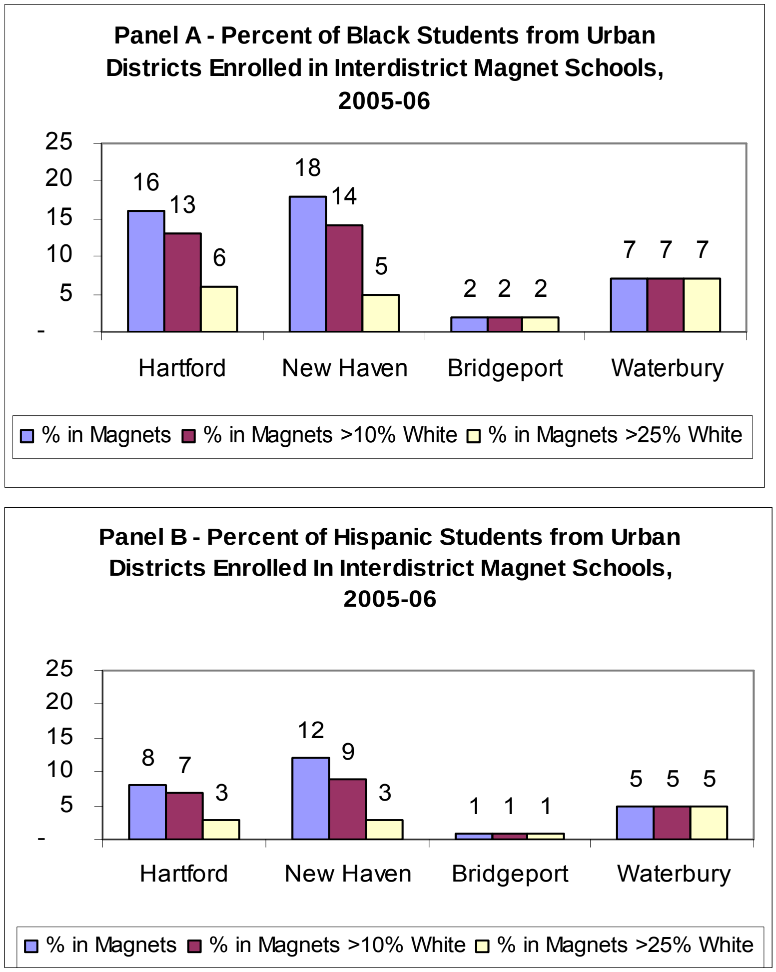
Δ%White is the difference in percent white students between the magnet school attended by the student and the the student's home district. Each column presents results from separate OLS regressions. All regressions are estimate using sample of students who reside in Hartford, New Haven or Waterbury, and include controls for age, gender, ethnicity, free-lunch eligibility, special educaiton status, year and pre-treatment math and reading test scores from two previous periods. The figures in parentheses are standard errors adjusted for clustering at the school level. * indicates statistically significant at 0.05 level.

Figure 1: Comparison of Student Composition in City and Interdistrict Magnet Schools



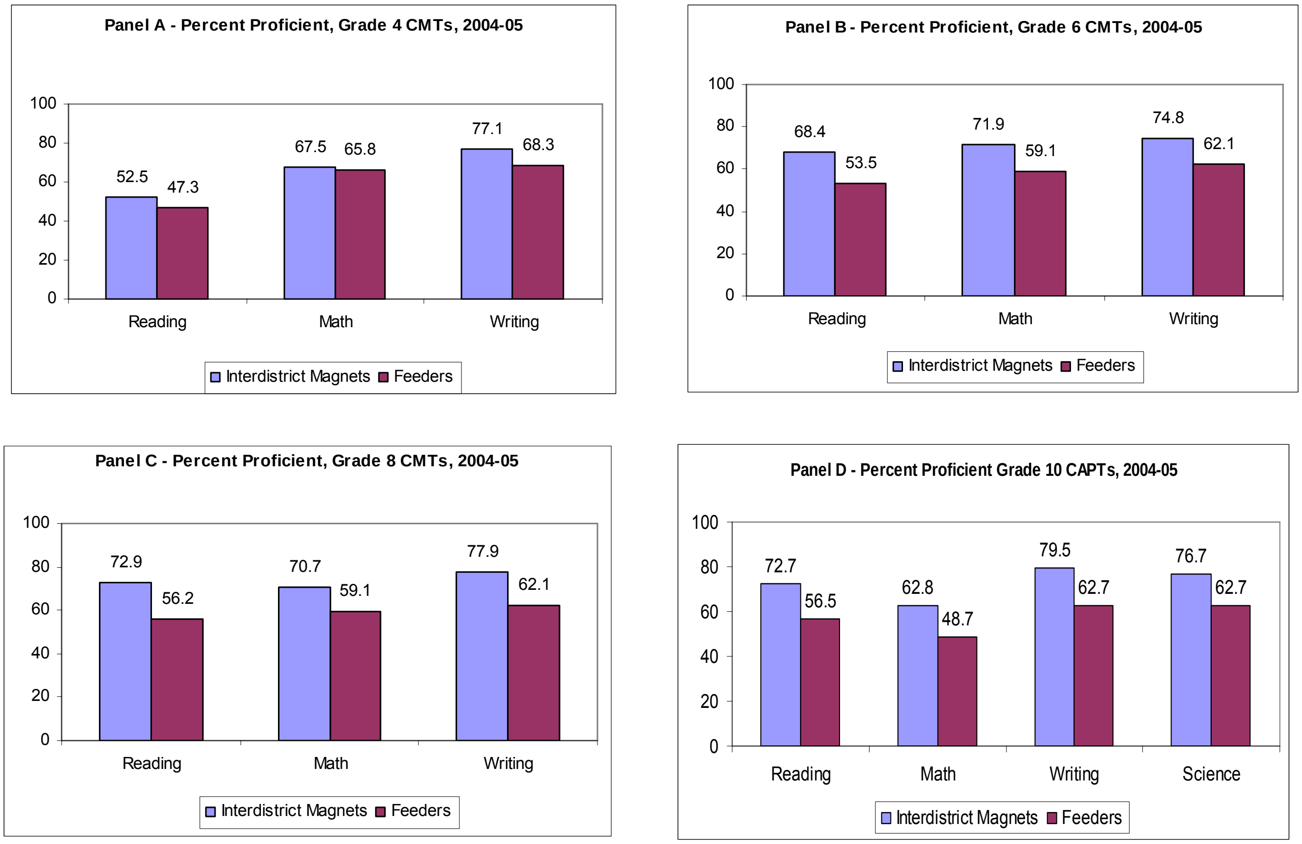
Source: Authors' computations using data provide by the Connecticut State Department of Education

Figure 2: Participation in Diverse Interdistrict Magnet Schools



Source: Authors' computations using data provide by the Connecticut State Department of Education

Figure 3: Achievement Levels in Interdistrict Magnet Schools and Their Feeder Districts



Source: Authors' computations using data from the 2004-05 CMT and CAPT reports available at www.cmtreports.com and www.captreports.com