ELSEVIER

Journal of Public Economics 89 (2005) 729–760

JOURNAL OF PUBLIC ECONOMICS

www.elsevier.com/locate/econbase

# The impact of school choice on student outcomes: an analysis of the Chicago Public Schools

Julie Berry Cullen[a,*], Brian A. Jacob[b], Steven D. Levitt[c]

[a] *Department of Economics, University of California, San Diego, 9500 Gilmano Dr., La Jolla, CA 92093-0508, USA*

[b] *John F. Kennedy School of Government, Harvard University, 79 JFK Street, Cambridge, MA 02138, USA*

[c] *Department of Economics, American Bar Foundation and University of Chicago, 1126 East 59th Street, Chicago, IL 60637, USA*

Available online 26 August 2004

**Abstract**

We explore the impact of school choice on student outcomes in the context of open enrollment within the Chicago Public Schools (CPS). Roughly half of the students opt out of their assigned high school to attend a different CPS school, and these students are much more likely than those who remain in their assigned schools to graduate. To determine the source of this apparent benefit, we compare outcomes across (i) similar students with differential access to schooling options and (ii) travelers and non-travelers within the same school. The results suggest that, other than for students who select career academies, the observed cross-sectional benefits are likely spurious.


*JEL classification:* I28; H72

*Keywords:* School choice; Student outcomes; Chicago Public Schools

## 1. Introduction

The current leading elementary and secondary education reform proposals involve introducing accountability through forms of market-based competition and expanded parental choice. Proponents claim that increased choice will force the current system of local monopolies to become more responsive and efficient. Critics worry that increased

* Corresponding author.

*E-mail addresses:* jbcullen@ucsd.edu (J.B. Cullen), brian_jacob@harvard.edu (B.A. Jacob), slevitt@midway.uchicago.edu (S.D. Levitt).

0047-2727/$ - see front matter © 2004 Published by Elsevier B.V.
doi:10.1016/j.jpubeco.2004.05.001

choice will isolate the most disadvantaged students in the worst schools and that parents may not be informed enough to make choices in the best interests of their children.[1] There is a growing body of theoretical and empirical work that presents a mixed picture of the net impact of various forms of choice.

This paper examines public school choice within the Chicago Public School system (hereafter abbreviated CPS). CPS has a number of features that make it an excellent case for analysis. First, parents have an enormous degree of flexibility in choosing which of the more than 60 CPS high schools their child will attend. For the cohorts we examine (those entering 9th grade in the Fall of 1993, 1994, or 1995), there were few restrictions on what school a child could attend within CPS. While a great deal of attention is presently focused on private school vouchers and charter schools, open enrollment is actually the most prevalent form of nontraditional parental choice. The type of intradistrict open enrollment that exists in CPS is available in more than one in every seven districts nationally, and in more than a third of large districts (NCES, 1996). Second, unlike most previous analyses of vouchers and magnet schools, which typically affect only a small percentage of students in the district, more than half of the students in CPS opt out of the neighborhood school to which they are assigned. Third, because minority and disadvantaged students are disproportionately represented in Chicago Public Schools, we are able to explore the impact of choice within an environment about which there is heightened concern.

Our empirical analysis is based on detailed student-level panel data for over 60,000 students who attend high school in the CPS. We find that students who opt out of their assigned neighborhood school are 7.6 percentage points more likely to graduate than peers who are observationally equivalent in 8th grade—off of a baseline graduation rate of 50%.[2] This increment to graduation is the same order of magnitude as the gap between students at Catholic and non-Catholic schools (Evans and Schwab, 1995).

There are several competing explanations for why students who opt out of their assigned school outperform those who stay. Higher graduation rates among those who opt out may be the result of these students attending better schools or finding a school that better matches their preferences. In either of these cases, the increased graduation rates represent true benefits of open enrollment. On the other hand, the higher graduation rates among those who opt out may instead reflect negative spillovers on those who are left behind or may arise spuriously if those who opt out are better on unobserved dimensions (e.g., student motivation, parental involvement).

We begin our attempt to disentangle these competing explanations by exploiting a unique survey administered to a subset of our sample in 8th grade. The survey asks students a wide range of questions about their expectations for the future, past educational record, and parental involvement. The responses capture a number of student qualities that are generally unobservable. Even after controlling for a rich set of covariates, these variables are strongly correlated with both the likelihood of graduation and with the

[1] See Chubb and Moe (1990) for a detailed discussion of the pros and cons of choice.

[2] We focus on graduation rates rather than test scores because of extreme sample selection in the latter (i.e., when a student drops out, we do not observe any further test scores). Even among the quartile of students with the highest test scores in 8th grade, graduation rates are less than 75%.

decision to opt out, suggesting that there are in fact systematic differences between students who opt out and those who do not.

We then compare how students with differential access to school choice fare in order to purge the analysis of the effects of unobserved student characteristics. If school choice has real effects on student outcomes, those children who live in areas with many nearby schools should on average derive the greatest benefit from the availability of school choice. Distance to nearby schools is shown to be a strong predictor of the likelihood a student will opt out of the assigned school.[3] Equally important, having controlled for other factors, distance is uncorrelated with the survey responses, providing a more direct and convincing test of the exogeneity assumption than is typically possible. Empirically, we find that easy access to a career academy is associated with substantial increases in graduation likelihood, implying that career academies confer real benefits.[4] In contrast, greater access to other types of schools—including high-achieving schools—is not associated with benefits on average, subject to the caveat that our estimates are relatively imprecise.

Examination of student outcomes *within* a given school lends further credence to the hypothesis that the gains associated with career academies are real, whereas gains for students who attend other types of schools are spurious. For most of the schools in CPS, the student body is a mix of individuals who were assigned to that school and students who elected to opt in. Presumably, these two groups experience similar peers, resources and teacher quality.[5] If students who opt in are similar to those assigned to the school on unobserved dimensions, then the outcomes of these two sets of students should be similar. At career academies, the difference in graduation rates between those who opt in and those who are assigned is not statistically significant. This suggests that it is something about the school itself (e.g., the teachers, curriculum, etc.) rather than the type of students who opt in that accounts for higher graduation rates.[6] In contrast, at high-achieving schools, those who opt in far outperform those assigned to the school, even after controlling for an extensive set of observable characteristics, including 8th grade test scores. This reinforces the idea that those who opt in to high-achieving schools are systematically better than observationally similar students who make other schooling choices and would outperform them regardless.

When evaluating these findings, it is important to bear in mind important weaknesses associated with the Chicago data that limit the scope of the analysis and the ability to

[3] Distance is especially important since no busing is provided. All students must make their own accommodations for getting to and from school.

[4] Career academies have three defining characteristics: (i) school-within-school programs; (ii) integrated academic/vocational curriculum; and (iii) employer partnerships. See Kemple and Snipes (2000) for a thorough description, as well as a discussion of their evolution.

[5] To account for the possibility that there may be different programs within a school or tracking, we also perform this analysis within ability quartiles at a particular school.

[6] Whatever it is about career academies that improves graduation rates is not captured by traditional measures of school quality, such as teacher–pupil ratios or teacher credentials, which are actually lower than at other schools. In their analysis of a random assignment experiment, Kemple and Rock (1996) attribute student performance gains partly to the more personalized student–teacher relationships that are facilitated by the school-within-a-school organization.

generalize our results. The most important limitation of our data set is that we only observe students in a regime where school choice has already been introduced. Consequently, we are not able to measure the impact of school choice on a variety of outcomes that are of interest, for example, does it keep high achieving students in the public schools, does it induce families with school-age children to live in Chicago, or does the competition induced by school choice improve overall school quality. Second, our primary analysis is limited to educational attainment in high school. We have no information about college attendance or labor market outcomes, and the drop out rate is so high that sample selection poses a substantial obstacle to convincingly analyzing test scores.

The remainder of this paper is structured as follows. Section 2 describes the implementation of the school choice program in the Chicago Public School system and establishes how our study relates to the existing literature on school choice. Section 3 introduces our data set and basic patterns of student participation in the CPS choice program. Section 4 estimates the cross-sectional relationship between opting out and graduation rates, while Section 5 explores potential explanations for the observed improved educational outcomes. Section 6 concludes.

## 2. Background on school choice

### *2.1. Previous literature*

Open enrollment is only one of a variety of initiatives, including private school vouchers, charter schools, and magnet schools, that fall under the umbrella of school choice. These programs share the common goal of breaking the traditional link between neighborhood and schooling. They have the potential to increase the degree of competition faced by local public schools and, thereby, to improve school performance.[7] From this perspective, school choice can improve all students' opportunities. However, another potential consequence of school choice is increased student sorting. If choice leads to a higher degree of sorting by ability and peer effects matter, then the distribution of educational benefits is likely to be quite unequal (Epple and Romano, 1998). The impact of school choice depends on which students take advantage of choice, the types of options that these students have, and what happens to the students who are left behind.[8]

Evidence on the impact of school choice on the students who participate is mixed. The Milwaukee voucher program is perhaps the most researched school choice program in the

[7] Hoxby (1994) demonstrates that increased Tiebout-style school choice is associated with more efficient public schools.

[8] There is a fairly large body of evidence about the types of students who take advantage of choice. A series of studies of different types of programs, including public and private alternatives, has found that children who participate are both more able and more advantaged (Coleman et al., 1993; Witte, 1993; Lankford et al., 1995; Buddin et al., 1998; Goldhaber et al., 1999). These studies suggest that choice schools increase social class and ability segregation across schools by siphoning off higher achieving and higher income students. Yet, other studies find evidence of participation that is more neutral with respect to background and even skewed toward minorities and low-income families (Duax, 1988; Lee et al., 1994; Schneider et al., 1996).

country. Beginning in the Fall of 1990, the Milwaukee public school district offered a limited number of low-income students vouchers to attend one of three private, nonsectarian schools in the district. Voucher recipients were selected from a pool of eligible applicants by a random lottery. Using a variety of methods to identify valid control groups, analyses of this program find anywhere from no achievement gains to large advantages (Witte et al., 1995; Green et al., 1997; Witte, 1997; Rouse, 1998).[9] Public school choice through the presence of magnet schools has been more consistently associated with improved outcomes for students (Blank, 1983; Metz, 1986; Crain et al., 1992; Gamoran, 1996), but the ability of these studies to account for student selection is often more restricted.

An aspect that limits what can be learned from studies of voucher and magnet school programs is that these programs, as they have been implemented, typically affect only a small percentage of students in the district. It is thus difficult to learn anything about either the degree of sorting or the gains and potential spillover effects associated with more systemic choice.[10] Open enrollment programs, in contrast, provide an opportunity to explore these questions. O'Brien and Murdoch (2000) examines a district-wide public school choice program in Texas, finding that academic performance in the choice district is higher than in similar non-choice districts in the state. They do not examine student-level achievement data or compare the achievement of those who do and do not opt out of their zone school.

One of the most interesting findings in our study involves the positive impact of career academies. Many prior studies have found benefits to career academies, but most have not been able to adequately deal with the problem of unobserved heterogeneity stemming from the selection by families and schools. A recent randomized experiment on career academies, however, provides a useful comparison for our results. Kemple and Rock (1996) compared students randomly selected to enroll in a career academy and those who applied and were not selected to enroll, following both groups through the end of their scheduled 12th grade year. They found that career academies improved outcomes for at-risk students, including reduced dropout rates, improved attendance, increased academic course-taking and the increased likelihood of earning enough credits to graduate on time. They also found low-risk students who attended career academies were more likely to graduate on time. While influencing attainment outcomes, career academies did not improve standardized math and reading achievement scores for any group.

[9] While in theory randomization provides an ideal context for the evaluation of school choice, in this case the control group was contaminated by selective attrition (Witte, 1997). Over half of the unsuccessful applicants never returned to the public schools and those who did return were from less educated, lower income families. More recent studies of publicly and privately funded voucher experiments in other cities that have been more carefully designed from the onset continue to present a wide range of estimated program effects (e.g., Peterson et al., 1998).

[10] International policy experiments have involved more expansive systems of school choice. However, the lessons are somewhat limited by the lack of individual-level data. For example, New Zealand introduced unrestricted choice among all public and religious schools in 1991. Ladd and Fiske (2000) find that the gap between successful and unsuccessful schools has widened since then. However, because the researchers only have access to school-level data, whether this is due to real effects of choice or to changes in school composition cannot be determined.

## 10. Our framework

In this paper, we attempt to measure the impact of the CPS open enrollment program on student graduation rates and, more importantly, to identify the sources of any apparent benefits. Under the CPS open enrollment plan,[11] students are guaranteed a spot in a pre-assigned neighborhood school, known as the "attendance area" school, determined by a student's place of residence. At the time of our sample, students were also free to apply to other high schools anywhere in the CPS system.[12]

Because of data limitations, we cannot address a number of important questions related to whether school choice keeps students in the public schools, affects the residential location decisions of families, or leads to competition induced school quality improvements.[13] Instead, we take the residential decisions, private schooling decisions, and productivity of CPS schools as given, and holding these factors constant, we explore how giving public school students the choice to opt out of their neighborhood school to another public school within the district affects student outcomes. Our "partial equilibrium" approach induces two countervailing biases regarding the value of school choice to students. By ignoring the fact that private schools and moving to the suburbs are substitutes that parents could use to improve their child's education in the absence of school choice, we exaggerate the benefits of school choice. In other words, we assume that, absent school choice, children would stay in CPS, when they might (at some cost) choose to exit.[14] On the other hand, if school choice leads to increased competition and better schools (Hoxby, 1994), then our estimates of the impact of school choice understate the true benefits.

Although we do not develop a formal model of school choice, our empirical analysis is motivated by the following set of assumptions. Students differ along both observable

[11] CPS introduced public school choice as a part of efforts to achieve racial desegregation. In September 1980, the Chicago Board of Education signed a consent decree with the United States Department of Justice that required the Board to develop a comprehensive desegregation plan to alleviate the effects of historic segregation on Black and Hispanic students.

[12] More recently, some geographic restrictions on open enrollment have been imposed. Admissions criteria vary by school. Most hold lotteries if they are oversubscribed, others require written examinations or student interviews.

[13] Cullen et al. (2003) are able to analyze how access to school choice affects the decision to enroll in the Chicago Public Schools using randomization induced by lotteries in oversubscribed schools. That paper finds that winning an entrance lottery increases the likelihood that a student not currently enrolled in CPS will enroll the following year, but has a much smaller impact on the enrollment decisions of students already attending a CPS school.

[14] Given that many parents do opt for private schools (nearly 20% of school-aged children attend private schools in Chicago), this bias might be substantial. It is important to bear in mind, however, that we fail to find the expected benefits of school choice in our empirical analysis. This bias only serves to reinforce our pessimism about the benefits of school choice. For the reader's reference, private high schools are fairly evenly distributed across Chicago. The median public high school has three private high schools within 2.5 miles. About 10% of high schools have no private high school within that range, while 10% have as many as six or seven. The public and private schools are generally not close substitutes. One fact that makes this clear is the differences in peer characteristics. While the average fraction Black in public high schools is 62.8%, the average in private high schools in Chicago is 35.2%. Two-thirds of the private high schools have religious affiliations.

dimensions (e.g., past test scores, neighborhood of residence, race, gender, etc.) and unobservable dimensions (e.g., motivation to excel in school and tastes/skills that may interact with school characteristics in the production of educational outcomes). Schools vary according to teacher quality, the composition of peers, and idiosyncratic components that make a particular school more or less attractive to a given student (e.g., the presence of a specialized program). Students maximize utility, of which the educational outcome is just one component. Holding educational outcome constant, students prefer schools that are good matches for them on non-academic dimensions, and also presumably prefer schools that are closer to home in order to minimize travel costs.

Within such a framework, we would expect students who opt of out local schools to fare better academically for a number of reasons. First, holding student motivation and non-academic components of a student–school match constant, those who opt out of the neighborhood school pay a greater travel cost and thus will be willing to travel only if compensated by an academic benefit.[15] This academic gain may arise from better teachers, better peers, or a better match. Second, highly motivated students are presumably more likely to opt out of local schools and also likely to do better academically than observationally equivalent students with lower motivation. Thus, even if the link between opting out and attainment is not causal (in the sense that motivation drives both of them), opting out may still be correlated with higher attainment.

Whether the observed correlation between traveling and outcomes can be explained by unobserved qualities of the students who travel, by improved matches, or through differential peer and school quality has very different implications for the distributional impact of choice. If the link is purely spurious—driven by unobserved differences in motivation—then school choice provides no real benefit. If peer quality is important, then school choice will benefit those who opt out, but may also hurt those who remain behind. The net social gains are unambiguously positive only to the extent that the benefits to those who opt out arise from the facilitation of better matches between students and schools.

## 3. Data description and summary statistics

Over 400,000 students are enrolled in CPS in grades K-12. As in most urban districts, students in the system are disproportionately minority (more than two-thirds) and poverty rates are well above those for the nation as a whole. Given the high rates of disadvantage and poor overall performance relative to national norms, our analysis provides evidence on the net benefits of providing choice to students with otherwise limited opportunities.

In this study, we follow three cohorts of CPS students—those entering 9th grade for the first time in the Fall of 1993, 1994, and 1995—through their high school years. There are a total of 76,563 students in these cohorts.[16] Because we condition on 8th grade math and

[15] It is possible that opting out is associated with lower academic achievement if the primary reasons for attending schools other than neighborhood schools are not academically driven.

[16] This does not include students who attend schools that exclusively serve special needs populations or juvenile detention centers. Those schools operate outside the standard of open enrollment.

Table 1
Summary statistics

| | Full sample | By ability quartile (lowest to highest) | | | |
|---|---|---|---|---|---|
| | | I | II | III | IV |
| *Student characteristics* | | | | | |
| 8th Grade math test score | 7.643 | 6.204 | 7.163 | 7.876 | 9.264 |
| | (0.006) | (0.008) | (0.005) | (0.005) | (0.008) |
| 8th Grade reading test score | 7.492 | 5.242 | 6.870 | 8.043 | 9.719 |
| | (0.008) | (0.009) | (0.005) | (0.005) | (0.009) |
| Completes 10th grade at grade-level | 0.753 | 0.600 | 0.707 | 0.798 | 0.900 |
| | (0.002) | (0.004) | (0.004) | (0.003) | (0.002) |
| Completes 11th grade at grade-level | 0.599 | 0.404 | 0.529 | 0.647 | 0.806 |
| | (0.002) | (0.004) | (0.004) | (0.004) | (0.003) |
| Graduates in 4 years | 0.503 | 0.306 | 0.422 | 0.548 | 0.727 |
| | (0.002) | (0.004) | (0.004) | (0.004) | (0.004) |
| Male | 0.479 | 0.532 | 0.474 | 0.454 | 0.458 |
| | (0.002) | (0.004) | (0.004) | (0.004) | (0.004) |
| Black | 0.610 | 0.684 | 0.661 | 0.617 | 0.482 |
| | (0.002) | (0.004) | (0.004) | (0.004) | (0.004) |
| Hispanic | 0.273 | 0.265 | 0.269 | 0.282 | 0.275 |
| | (0.002) | (0.004) | (0.004) | (0.004) | (0.004) |
| Age in September of 9th grade | 14.558 | 14.763 | 14.598 | 14.491 | 14.387 |
| | (0.002) | (0.005) | (0.004) | (0.004) | (0.003) |
| Special education student in 8th grade | 0.102 | 0.271 | 0.091 | 0.035 | 0.014 |
| | (0.001) | (0.004) | (0.002) | (0.001) | (0.001) |
| Lives with one or both parents | 0.911 | 0.895 | 0.907 | 0.914 | 0.926 |
| | (0.001) | (0.003) | (0.002) | (0.002) | (0.002) |
| Moved tracts between 8th and 9th grades | 0.115 | 0.127 | 0.127 | 0.121 | 0.087 |
| | (0.001) | (0.003) | (0.003) | (0.003) | (0.002) |
| Social status index | −0.006 | −0.134 | −0.064 | 0.005 | 0.164 |
| (block level) | (0.003) | (0.005) | (0.005) | (0.005) | (0.006) |
| *Tract characteristics* | | | | | |
| Average 8th grade math and reading | 15.104 | 14.698 | 14.917 | 15.137 | 15.644 |
| combined score among peers in tract | (0.004) | (0.007) | (0.007) | (0.008) | (0.010) |
| Average 8th grade combined score among | 15.108 | 14.939 | 15.028 | 15.119 | 15.337 |
| peers in attendance area | (0.003) | (0.005) | (0.005) | (0.005) | (0.006) |
| Population | 4855 | 4706 | 4774 | 4870 | 5060 |
| | (12) | (23) | (23) | (23) | (23) |
| % School age | 0.236 | 0.252 | 0.244 | 0.235 | 0.216 |
| | (0.000) | (0.001) | (0.001) | (0.001) | (0.001) |
| % Hispanic | 0.212 | 0.215 | 0.208 | 0.212 | 0.216 |
| | (0.001) | (0.191) | (0.002) | (0.002) | (0.002) |
| % Black | 0.553 | 0.618 | 0.596 | 0.560 | 0.441 |
| | (0.002) | (0.004) | (0.004) | (0.004) | (0.004) |
| Median household income | 22,418 | 20,301 | 21,391 | 22,638 | 25,238 |
| | (37) | (73) | (73) | (72) | (72) |
| % Below poverty line | 0.278 | 0.330 | 0.301 | 0.270 | 0.214 |
| | (0.001) | (0.002) | (0.002) | (0.002) | (0.001) |
| Average years of education | 12.064 | 11.897 | 11.990 | 12.075 | 12.285 |
| | (0.003) | (0.006) | (0.006) | (0.006) | (0.007) |

Table 1 (*continued*)

| | Full sample | By ability quartile (lowest to highest) | | | |
|---|---|---|---|---|---|
| | | I | II | III | IV |
| *Tract characteristics* | | | | | |
| % Professionals | 0.169 | 0.155 | 0.163 | 0.169 | 0.188 |
| | (0.000) | (0.001) | (0.001) | (0.001) | (0.001) |
| % Female-headed households | 0.426 | 0.478 | 0.454 | 0.423 | 0.352 |
| | (0.001) | (0.002) | (0.002) | (0.002) | (0.002) |
| % Homeowners | 0.389 | 0.349 | 0.371 | 0.398 | 0.438 |
| | (0.001) | (0.002) | (0.002) | (0.002) | (0.002) |
| % Have lived in same house since 1985 | 0.568 | 0.566 | 0.567 | 0.572 | 0.567 |
| | (0.000) | (0.001) | (0.001) | (0.001) | (0.001) |
| Crime composite index | 0.018 | 0.168 | 0.096 | 0.014 | −0.201 |
| | (0.003) | (0.006) | (0.005) | (0.006) | (0.005) |
| Unemployment rate | 0.412 | 0.439 | 0.425 | 0.408 | 0.378 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| *Type of school attended* | | | | | |
| Assigned attendance area school | 0.466 | 0.634 | 0.537 | 0.439 | 0.260 |
| | (0.002) | (0.004) | (0.004) | (0.004) | (0.004) |
| Career academy | 0.156 | 0.088 | 0.163 | 0.226 | 0.146 |
| | (0.001) | (0.002) | (0.003) | (0.003) | (0.003) |
| High-achieving school | 0.161 | 0.020 | 0.042 | 0.106 | 0.467 |
| | (0.001) | (0.001) | (0.002) | (0.002) | (0.004) |
| Regular school | 0.218 | 0.259 | 0.258 | 0.229 | 0.128 |
| | (0.002) | (0.004) | (0.004) | (0.003) | (0.003) |
| Number of observations | 60,751 | 14,939 | 15,047 | 15,252 | 15,513 |

The first column corresponds to our full sample. The remaining four columns provide breakdowns of this sample by ability quartile, based on combined 8th grade reading and math scores. Standard errors are in parentheses.

reading achievement scores as proxies for baseline academic performance levels, students lacking such test scores are dropped from the sample (a total of 6446 or 8.4% of the full sample). Test scores are primarily missing because the student was absent the day the test was given (roughly 90% of the missing scores) or because the student was exempted from the test or the student transferred in from a different school district in 9th grade (roughly 10% of the missing scores).[17] In addition, 1089 students are dropped due to other missing data, primarily home address (which we use for the purpose of measuring distance to schools and linking to census-tract-level neighborhood characteristics). Finally, because we do not observe subsequent educational outcomes for roughly 8000 students who transfer out of CPS to another school system anytime during high school, we exclude these students from our individual-level regressions.[18] The students leaving CPS are primarily

[17] We have replicated the full set of analyses using imputed 8th grade test scores (and a dummy for whether 8th grade test scores are missing) to avoid having to drop these students. The results for the expanded sample are very similar to the results that are presented in the paper. Tables available from the authors upon request.

[18] They are, however, included in the computation of average peer characteristics. They appear similar to those who remain with respect to 8th grade test scores.

moving out of the area, with fewer than 1 out of 5 leaving to attend private schools.[19] Thus, our final sample includes 60,623 students.

The primary data source for the analysis is administrative student record and test score files for students in the CPS. These files provide a complete record of each student's history in CPS, including which schools they attended each semester and records of any standardized test taken starting in 1st grade. These records include home and school addresses, which allow us to link census data to students, compute distances from homes to schools, and also to track any residential moves or school transfers. For a random subsample of students in the 1994 9th grade cohort, we also have survey data from their 8th grade that includes questions on family background as well as attitudes toward and experiences in school.

Summary statistics for the students in our sample are presented in the top panel of Table 1. The first column includes all students. CPS students on average test more than a grade level below the national average in both math and reading (the national mean score, measured in grade equivalents, is 8.8). More than 60% of incoming students are Black, and more than one-fourth are Hispanic.[20] Approximately 25% drop out prior to completing 10th grade, and only half graduate in 4 years. The remaining columns of the table categorize the students by ability quartile, based on combined math and reading ITBS test scores in the 8th grade. The lowest quartile students in 8th grade are on average performing near the 6th grade level in math, and read just above the 5th grade level. The top quartile of students is above national norms on both math and reading. Initial ability is a good predictor of educational attainment: less than one-third of students in the bottom quartile graduate in four years, compared to more than 70% of high-ability students.

The middle panel of Table 1 presents census-tract level characteristics based on student place of residence. Students in our sample are spread across 814 tracts. Higher ability students tend to live in census tracts with higher achieving peers, lower poverty rates, fewer female-headed households, greater home ownership, and lower crime rates.

During our sample period, there were 61 regular high schools in the CPS system.[21] Of these schools, 47 are attendance area schools. At the other 14 schools, students must apply to attend. Assignment to an attendance area school is done on the basis of a student's place of residence. The attendance area school is the default that a student will attend unless he or she submits an application to attend another school. A student's attendance area school is generally close to his or her house (on average within 2 miles), but is not necessarily the very

[19] The fact that four out of five students who leave the CPS do so because they leave the city (and often the state) suggests that the decision to leave the CPS is most often motivated by family location decisions that are less likely to bias our results. Sensitivity analyses confirm that our results are not affected by sample attrition. For example, we replicate our results using an expanded sample that includes all students who remained in the CPS through 10th grade, which allows us to include roughly 48% of the 8000 leavers. Using 10th grade completion as the dependent variable, we obtain results that are virtually identical for the restricted and expanded samples. Tables available from the authors upon request.

[20] In contrast to Census data, which treat Hispanic as an ethnicity, the Chicago schools treat Hispanic as a racial category, i.e., Black, White, and Hispanic are mutually exclusive categories. In the 1990 School District Data Book, 34% of Hispanic students in the public schools are self-classified as White, 2% as Black, and 1% as Native American or Asian. The remaining 64% report race as "Other."

[21] At the time of our sample, charter schools were not yet operational in Chicago.

Table 2
Characteristics of CPS high schools

| Characteristics of 9th graders | High-achieving schools | Career academies | Regular schools |
|---|---|---|---|
| Average 8th grade math score | 8.700 (0.193) | 7.437 (0.153) | 7.237 (0.054) |
| Average 8th grade reading score | 8.864 (0.232) | 7.290 (0.220) | 6.961 (0.069) |
| % Black | 0.543 (0.086) | 0.850 (0.074) | 0.606 (0.063) |
| % Hispanic | 0.207 (0.044) | 0.125 (0.062) | 0.300 (0.053) |
| % Lives with one or both parents | 0.929 (0.004) | 0.910 (0.008) | 0.900 (0.004) |
| Average social status (block level) | 0.283 (0.079) | − 0.175 (0.077) | − 0.057 (0.060) |
| Average distance traveled | 3.303 (0.405) | 2.208 (0.256) | 1.587 (0.135) |
| Number of schools | 12 | 10 | 39 |

The three columns present characteristics of the student bodies at the three types of high schools within CPS. High-achieving schools are the top quintile based on 8th grade achievement scores of the entering 9th grade class. "Regular" schools are the remaining schools that are neither high-achieving schools nor career academies. Standard errors are shown in parentheses.

closest school. Empirically, for 64% of the students in our sample, the attendance area school is the closest school, and for 93% of the students it is one of the three closest schools.

For the purposes of our analysis, we categorize schools into three mutually exclusive groups: career academies, "high-achieving" schools, and all other schools, which we refer to as "regular" schools. Career academies emphasize vocational skills and integrating school with work. High-achieving schools correspond to the top quintile of schools by average 8th grade test scores of the incoming students.[22] The results in the paper are not sensitive to variations in the fraction of schools classified as high-achieving, as long as the separate grouping of career academies is preserved.

The bottom panel of Table 1 presents attendance patterns by school type. Overall, less than half of all students actually enroll in their assigned attendance area school. Only 26% of students in the top quartile of ability attend their assigned schools. Even among the lowest ability group, more than one-third opt out of their local school. Thus, in contrast to previous studies of school choice in which take-up rates for students are quite low, in Chicago there is an enormous amount of student reallocation due to school choice.

There are stark differences in the patterns of school choice across ability quartiles. Those students in the bottom quartile who opt out are most likely to attend a regular school. In the middle quartiles, career academies are nearly as popular a choice. Among the top students, almost 50% opt out to attend high-achieving schools. As a consequence, open enrollment leads to substantial sorting by ability, relative to the initial attendance area assignments.

Table 2 further explores the characteristics of Chicago schools by type of school. Average 8th grade test scores are by definition highest at high-achieving schools. Living circumstances, as measured by either living with at least one parent or by a block-group level measure of social status (higher numbers imply better living conditions), are also

[22] The high-achieving schools include the three CPS schools that base admissions on test scores, most but not all of the non-career academy schools that are by application only, one career academy that is by application only, and a handful of attendance area schools, most of which are located in affluent areas.

better for students attending high-achieving schools. Students tend to travel further to attend these schools. Notably, career academies and regular schools appear similar on family background and peer quality measures, though career academies are disproportionately Black.

## 4. The academic benefit associated with opting out of the assigned school

In this section, we quantify the relationship between opting out and attainment.[23] We begin with models using the full sample and a single measure for opting out, and then investigate the extent to which the benefits of opting out vary along observable dimensions. We do not attempt to determine the underlying reasons for the higher relative outcomes for travelers in this section, instead deferring that discussion to the next section.

The basic estimates of the effect of opting out on student educational attainment are presented in Table 3. The specifications reported are from linear probability models; probits yield similar results evaluated at the mean. Standard errors, corrected for correlation within high schools, are in parentheses.

The top row of the table presents the coefficients associated with an indicator variable for whether a student opts out of his or her assigned school. Controlling for a wide range of other observables—most importantly the student's own test scores in 8th grade—opting out is associated with a 6.5 percentage point increase in completion of 10th grade, and roughly an 8 percentage point increase in completing 11th grade and graduating on time.[24] Each of these estimates is highly statistically significant. Compared to other educational interventions in the literature, opting out is associated with a large increase in years of schooling. These estimates imply that almost all of the effect of opting out occurs by 10th grade. This pattern of results is difficult to reconcile with plausible models of the benefits of school choice. Rather, we view this as the first of many indications that the observed correlation between opting out and attainment is largely spurious.

The bottom two rows of the table report coefficients on opting out from specifications with a restricted set of covariates. When we include only the neighborhood-level controls—excluding individual-level variables—the coefficient on opting out roughly doubles. Including all of the individual-level covariates except test scores also yields coefficients twice the size of the baseline specification. That result highlights the importance of the test-score variables, which single-handedly cut the magnitude of the

[23] Test score data are available, but extreme sample selection arises due to high dropout rates. In our initial analysis of test score data, we found the results to be quite sensitive to different sample selection corrections. Especially problematic is the fact that, controlling for prior test scores and other observable student characteristics, traveling is positively associated with remaining in school and, by default, taking the exams. Thus, students who take advantage of choice are more likely to take the exams than observably similar students who attend their assigned school. If students who learn less in school are those who are more likely not to take the exams, as seems likely, travelers will appear to learn less than they actually do relative to non-travelers due to differential test-taking.

[24] In our sample, 75.2% of students complete 10th grade, 59.9% complete 11th grade, and 50.3% graduate on time.

Table 3
Cross-sectional relationship between traveling and achievement

| Independent variable | Completes 10th grade | Completes 11th grade | Graduates high school |
|---|---|---|---|
| Opts out of assigned school | 0.065 (0.013) | 0.082 (0.011) | 0.076 (0.010) |
| 8th Grade math test score | 0.050 (0.003) | 0.071 (0.003) | 0.078 (0.003) |
| 8th Grade reading test score | 0.011 (0.002) | 0.016 (0.002) | 0.015 (0.002) |
| Male | − 0.036 (0.008) | − 0.065 (0.011) | − 0.085 (0.011) |
| Black | 0.037 (0.015) | 0.007 (0.018) | − 0.005 (0.018) |
| Hispanic | 0.046 (0.010) | 0.036 (0.012) | 0.030 (0.015) |
| Black male | − 0.061 (0.011) | − 0.077 (0.013) | − 0.071 (0.013) |
| Hispanic male | − 0.018 (0.008) | − 0.023 (0.013) | − 0.038 (0.013) |
| Age in September 9th grade | − 0.100 (0.005) | − 0.102 (0.005) | − 0.099 (0.005) |
| In special education in 8th grade | 0.032 (0.011) | 0.052 (0.011) | 0.067 (0.009) |
| Moved tracts between 8th and 9th | − 0.058 (0.006) | − 0.069 (0.007) | − 0.066 (0.007) |
| Lives with either or both parents | 0.023 (0.007) | 0.035 (0.007) | 0.041 (0.008) |
| Social status indicator (block level) | 0.015 (0.005) | 0.011 (0.004) | 0.013 (0.004) |
| $R^2$ | 0.113 | 0.147 | 0.159 |
| Number of Observations | 60,623 | 60,623 | 60,623 |
| Coefficient on opts out with only cohort controls | 0.124 (0.021) | 0.160 (0.026) | 0.158 (0.028) |
| Coefficient on opts out with all variables except 8th grade test scores | 0.107 (0.016) | 0.143 (0.017) | 0.142 (0.018) |

Each column corresponds to a separate ordinary least squares regression for the binary dependent variable indicated. The models are linear probability models. The sample is the subsample of students enrolled in 9th grade in the Fall of either 1993, 1994, or 1995 who took the 8th grade achievement exams in the prior year who remain within the Chicago Public School system through high school. In addition to the control variables shown, each regression also includes the average 8th grade math and reading combined score among peers in the student's residential tract and among peers assigned to the same attendance area school, as well as indicators for whether the attendance area school is a career academy or a high-achieving school. The specifications also include the following tract-level covariates: population, crime index, unemployment rate, mean years of education, median household income, and the percentage Black, Hispanic, homeowners, female-headed households, below the poverty line, school aged children, professionals and in same house since 1985. Standard errors, shown in parentheses, are corrected for correlation within 9th grade high school using White's generalized method.

opting out coefficients in half.[25] Although not shown in tabular form, we also find a strong positive relationship between the value of opting out and the gap between average peer ability in the assigned school and the school a student actually attends. However, even those students who opt to attend schools with lower-achieving peers than the school they were assigned to have positive (but statistically insignificant) effects of traveling.

Table 3 also presents the results from other variables of interest that are included in the regressions. Controlling for other factors, males are less likely to graduate than females, especially Black males who are 15 percentage points less likely to graduate than otherwise similar Black females. There are relatively small differences in graduation rates across races.

[25] In our specifications in Table 3, 8th grade test scores are constrained to affect attainment in a linear fashion. Given the importance of these variables, we have thoroughly investigated the relationship between past test scores and future attainment. Including four-part splines in both 8th grade math and reading scores to allow for non-linearity in the impact of lagged test scores has little impact. Adding 7th grade test scores, either linearly or in a four-part spline, also has little impact on the coefficient for opting out.

Table 4
The value of opting out by ability and type of school attended

| Independent variable | Dependent variable = graduates high school | | | |
|---|---|---|---|---|
| | Student ability quartile (lowest to highest) | | | |
| | I | II | III | IV |
| Opts out | 0.062 (0.011) | 0.078 (0.015) | 0.070 (0.015) | 0.081 (0.015) |
| Opts out to career academy | 0.092 (0.026) | 0.105 (0.025) | 0.059 (0.027) | 0.031 (0.027) |
| Opts out to high-achieving school | 0.154 (0.051) | 0.145 (0.034) | 0.152 (0.026) | 0.122 (0.015) |
| Opts out to regular school | 0.046 (0.011) | 0.052 (0.015) | 0.042 (0.013) | 0.018 (0.021) |
| Number of observations | 14,898 | 15,017 | 15,221 | 15,487 |

The top row reports the coefficients on opting out from ordinary least squares regressions predicting whether a student graduates. The next three rows present the coefficients on binary indicators for whether a students opts out by type of school, where these variables are included jointly. Each of the four columns presents results from regressions on subsamples based on quartiles of the overall student ability distribution, where ability is measured by combined 8th grade math and reading achievement test scores. Although only the coefficients on the opting out variables are shown, the full set of covariates described in Table 3 is included. Standard errors (in parentheses) are corrected for correlation within 9th grade high school using White's generalized method.

There is some evidence of positive effects of living with at least one parent and living in a neighborhood with higher socio-economic status. Moving between census tracts in the year prior to 9th grade is associated with lower graduation rates. Students assigned to special education in 8th grade, controlling for other factors including 8th grade achievement, are more likely to graduate.

Table 4 presents the impact of opting out by ability quartile and by type of school attended. The value of opting out is highest among top quartile students (8.1 percentage points) and lowest among bottom quartile students (6.2 percentage points). Across all ability groups, opting out to a high-achieving school is associated with the greatest increases—always over 10 percentage points. Career academies are associated with greater attainment gains than regular schools across the board.

## 5. Disentangling the reasons why students who opt out perform better

Having established a correlation between opting out of the assigned school and higher educational attainment, we turn our attention to trying to understand the source of those gains. We focus on three main hypotheses as to why those who opt out might do better: (1) they are better along unobservable dimensions like motivation or parental involvement, (2) they are on average traveling to "better" schools, or (3) by opting out they find a school which better matches their educational needs for idiosyncratic reasons.

Our approach in this section is as follows. We first provide an intuitive discussion of the identification issues involved in studying school choice (interested readers are directed to the appendix for a formal empirical model). This framework highlights the fact that OLS estimates may not be closely related to the true causal effects that are relevant from a policy perspective, but that a valid instrumental variable would allow us to estimate something much closer to the causal effects. We then demonstrate empirically that the concerns about OLS are justified: students who opt out systematically differ from other

students along a number of important (and typically unobservable) dimensions likely to influence graduation, even after controlling for the extensive set of variables in our baseline specification. We then show that our proposed instrument (distance from a student's home to schooling alternatives) appears uncorrelated with these same student characteristics, but is a good predictor of what school a student attends. Thus, we present reduced-form estimates of the link between distance to schooling and outcomes. Finally, we complement these results with an analysis of whether the performance of students who opt *into* a school systematically differs from the performance of students who are assigned to that same school and, therefore, experience similar peers and school resources. If travel still appears to benefit students, then it is likely due to unobserved heterogeneity or match quality, rather than shared school effects.

### *5.1. Identification of the causal effect of choice*

We assume that outcomes for students who are similar according to observable dimensions will differ systematically due to unobserved student qualities, the quality of the specific match between the student and the school, and the shared quality of the school attended. Students who opt out potentially benefit from school choice in two ways: (1) they are able to attend a school with a better match to their particular needs, and (2) they may attend schools that are higher quality than their assigned school on average. Students who stay in their assigned schools are potentially affected by choice both through competitive pressures on the school and because the composition of the student body at their school changes. If average peer quality is an important determinant of student outcomes, then the impact of choice on those who stay is likely to be negative since the most able students travel. Alternatively, this effect could be positive if the change in student composition leads to a reallocation of resources that benefits the remaining students.

Because we do not observe the counterfactual outcomes of students absent school choice, we cannot directly estimate the effects described in the preceding paragraph. The OLS estimates in the previous section are based on a comparison of students who do and do not choose to opt out, rather than from a comparison of the same students' outcomes in two different regimes. More specifically, the OLS coefficients are affected by three separate biases relative to the true impact of choice for students who travel. The first bias is that those who opt out may systematically differ from those who stay behind on unobservable individual characteristics. The empirical evidence presented below strongly suggests that this bias will overstate the true benefits to school choice. The second bias is that those who elect to opt out may systematically differ in the quality of their match at the school initially assigned. This is likely to bias the OLS coefficient downward: bad initial matches make it more likely a student will choose to opt out—i.e., students who choose to stay are likely to be particularly well suited for their assigned school. The third bias is very different in nature from the first two. Since the OLS coefficient reflects the difference between the outcomes of those who opt out and those who stay, if those who remain behind are hurt by choice through spillovers, then the OLS estimate of the benefits of school choice to travelers is biased upwards. The opposite is true if those who remain behind benefit from the fact that other students opt out.

In order to estimate something closer to a causal effect than the OLS estimates, one would like to identify an instrumental variable $z$, which predicts whether a given student opts out, but is otherwise uncorrelated with student outcomes. In our case, distance from a student's residence to the nearest career academy, high-achieving school, or regular school (excluding the student's attendance area school) will serve as the instrument. If our distance measure is uncorrelated with unobserved student ability and initial match quality, the reduced form estimate (i.e., educational attainment on the left-hand side and the instrumental variable and observable characteristics **X** on the right-hand side) tells us whether students who live close to different types of schools benefit on average. It reflects both the direct benefit to travelers of attending alternative schools under choice as well as spillovers to those who remain at their assigned school. Note that, without further assumptions, we cannot separately identify benefits to those who opt out versus those who stay. Nonetheless, the reduced-form estimates are much more informative regarding the causal effects of the policy than are the OLS estimates because they do not include the effects of unobserved student ability.

Typically in instrumental variables settings a further assumption is made, namely, that those who do not receive "treatment" (in our case opting out) are unaffected by the fact that others receive "treatment." In that case, 2SLS allows one to identify the causal effect on those who opt out. In the current setting, however, this assumption is highly suspect. To the extent that school quality is affected by the composition of the students, a school choice regime in which over half of all students opt out is likely to lead to important changes in school quality for those who remain behind. As a consequence, we do not present 2SLS estimates, although we do briefly discuss them.

### *5.2. Do students who opt out systematically differ on "unobservable" dimensions?*

If students who opt out are systematically different than other students on unobservable dimensions, then as noted above, the OLS estimates of opting out are unlikely to have a causal interpretation. Of course, it is not possible to directly test for unobservables. In our data, however, we do have access to two reasonable sets of proxies for unobserved student characteristics. The first set of variables is lagged test score measures and lagged school and residential moves. Although these variables are available to us, we have not included them in the baseline specification because these variables are missing for many students. More interesting and unique, however, is that for a subset of our sample we have access to a survey conducted in 8th grade. Students were asked detailed questions about their attitudes and background, including expectations for the future, degree of preparedness for high school, past grades, and parental education and involvement. These are precisely the sort of typically unobservable characteristics that we are concerned might be biasing our OLS estimates of opting out.

The concern about unobservables appears justified: the 8th grade survey responses are excellent predictors of eventual graduation (even after controlling for the rich array of covariates that are available to us) and the survey responses are systematically related to the decision to opt out. With respect to predicting graduation, students who

self-report ever having failed a grade are 18.0 percentage points less likely to graduate, controlling for the full set of covariates included in our OLS specifications. Similarly strong effects in the opposite direction are associated with higher self-reported grades (an increase of 16.5 percentage points) and expecting to graduate (an increase of 10.2 percentage points). Those who feel well-prepared for high school, have high expectations for the future more generally, or have spoken with their parents about what high school to attend are 1.2, 4.6, and 2.8 percentage points more likely to graduate, respectively.

Table 5 demonstrates that these survey responses are also systematically related to the decision to opt out. We run regressions identical to those presented earlier, but with the survey responses, rather than graduation, as the dependent variables. The first four columns of Table 5 present the results. Each row of the table corresponds to a different dependent variable. Columns 1–3 report the coefficients on indicator variables for opting out by type of school. Column 4 shows the results of an *F*-test of the joint significance of the opting-out variables. The estimates provide striking evidence that students who opt out are systematically different from those who remain in their assigned schools on every single dimension tested. Conditional on the full set of control variables in the baseline specification (including 8th grade test scores), those who opt out did better on 6th and 7th grade tests and were less likely to have switched schools or moved residences in junior high. Those who opt out self-report that they are better prepared for high school, have higher expectations for graduation and the future more generally, have better junior-high grades, are less likely to have failed a grade or been suspended, and are absent fewer days. The degree of positive selection of these forms is greatest at high-achieving schools. Also, students attending high achieving schools (but not other students who opt out) have better educated parents and greater parental involvement in school activities.

### *5.3. Using distance to alternative schools as an instrument for opting out*

An ideal instrument for opting out would have a large impact on a student's decision to opt out, but would be uncorrelated with a student's unobserved ability and idiosyncratic match quality with the assigned school. The variables we propose are distances to the nearest schools of various kinds, e.g., the closest high-achieving school, career academy, etc. Fig. 1 displays a map of Chicago with the location of each public high school, the category to which it is assigned, and median incomes by neighborhood. Schools are not located uniformly across the city. In some areas, a number of schools are clumped close together, whereas other schools are relatively geographically isolated. This uneven distribution of schools is critical to our identification strategy.

In evaluating whether distance to schools is a reasonable choice of instrument, we first address the question of whether distance is actually an important determinant of opting out, turning later to the issue of whether distance is likely to be uncorrelated with student unobservables. Because students must provide their own transportation to school, distance is likely to be an important factor in choosing schools. The first-stage results presented in Table 6 confirm this conjecture. The dependent variable in each

Table 5
The relationship between student characteristics and distance to alternative schools

| Dependent variable | Opts out by type of school | | | | In(distance to the nearest school by type) | | | |
|---|---|---|---|---|---|---|---|---|
| | Career academy | High-achieving | Regular school | *F*-test of joint significance | Career academy | High-achieving | Regular school | *F*-test of joint significance |
| *Prior achievement scores and mobility* | | | | | | | | |
| 7th Grade math score | 0.169** | 0.339** | 0.030** | $F=180.61$ | −0.008 | −0.002 | 0.006 | $F=0.34$ |
| [$\mu=6.809$; $N=57{,}056$] | (0.013) | (0.015) | (0.009) | [$p=0.000$] | (0.010) | (0.010) | (0.013) | [$p=0.795$] |
| 7th Grade reading score | 0.268** | 0.514** | 0.055** | $F=203.83$ | −0.007 | −0.007 | 0.007 | $F=0.30$ |
| [$\mu=6.654$; $N=57{,}355$] | (0.016) | (0.024) | (0.013) | [$p=0.000$] | (0.011) | (0.012) | (0.014) | [$p=0.829$] |
| 6th Grade math score | 0.127** | 0.258** | 0.009 | $F=90.00$ | −0.004 | −0.015 | 0.004 | $F=0.47$ |
| [$\mu=5.987$; $N=55{,}453$] | (0.013) | (0.017) | (0.009) | [$p=0.000$] | (0.011) | (0.013) | (0.014) | [$p=0.706$] |
| 6th Grade reading score | 0.119** | 0.387** | 0.026** | $F=124.94$ | −0.008 | 0.007 | 0.011 | $F=0.44$ |
| [$\mu=5.554$; $N=55{,}663$] | (0.013) | (0.016) | (0.011) | [$p=0.000$] | (0.012) | (0.013) | (0.014) | [$p=0.728$] |
| Changes schools between Fall of 7th grade and Spring of 8th grade [$\mu=0.132$; $N=57{,}820$] | −0.050** | −0.060** | −0.028** | $F=38.17$ | −0.002 | −0.001 | −0.003 | $F=0.16$ |
| | (0.006) | (0.006) | (0.005) | [$p=0.000$] | (0.006) | (0.007) | (0.006) | [$p=0.924$] |
| Moved residential tracts between Fall of 7th and Spring of 8th grade [$\mu=0.285$; $N=57{,}720$] | −0.017** | −0.029** | 0.002 | $F=7.38$ | −0.005 | −0.005 | −0.008 | $F=0.92$ |
| | (0.007) | (0.007) | (0.006) | [$p=0.000$] | (0.006) | (0.006) | (0.008) | [$p=0.430$] |
| *Survey responses (1994 Cohort Only)* | | | | | | | | |
| I feel well-prepared for high school | 0.061** | 0.056** | 0.041** | $F=6.38$ | 0.005 | −0.009 | 0.009 | $F=0.34$ |
| [$\mu=0.366$; $N=7522$] | (0.016) | (0.018) | (0.015) | [$p=0.000$] | (0.012) | (0.011) | (0.015) | [$p=0.796$] |
| I am likely to graduate high school | 0.074** | 0.068** | 0.050** | $F=14.34$ | −0.002 | −0.012 | −0.013 | $F=1.41$ |
| [$\mu=0.811$; $N=7563$] | (0.012) | (0.013) | (0.014) | [$p=0.000$] | (0.008) | (0.008) | (0.012) | [$p=0.240$] |
| Positive expectations for the future (1–5 scale) [$\mu=4.136$; $N=7626$] | 0.087** | 0.106** | 0.021 | $F=5.58$ | 0.006 | −0.029 | 0.010 | $F=0.83$ |
| | (0.026) | (0.032) | (0.027) | [$p=0.001$] | (0.020) | (0.019) | (0.024) | [$p=0.479$] |
| Self-reported grades A's or B's | 0.118** | 0.158** | 0.027 | $F=24.95$ | 0.016 | 0.023** | −0.024 | $F=1.56$ |
| [$\mu=0.592$; $N=7037$] | (0.019) | (0.018) | (0.019) | [$p=0.000$] | (0.019) | (0.013) | (0.020) | [$p=0.201$] |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Ever failed a course [$\mu = 0.564$; $N = 8205$] | −0.104** (0.019) | −0.172** (0.021) | −0.053** (0.016) | $F = 22.93$ [$p = 0.000$] | −0.025 (0.019) | 0.008 (0.016) | 0.009 (0.019) | $F = 0.82$ [$p = 0.483$] |
| Spoke several times with parents about which high school to attend [$\mu = 0.669$; $N = 7778$] | 0.069** (0.017) | 0.077** (0.017) | 0.047** (0.015) | $F = 9.45$ [$p = 0.000$] | −0.008 (0.012) | 0.002 (0.010) | −0.004 (0.016) | $F = 0.16$ [$p = 0.926$] |
| Mother or father attended some college [$\mu = 0.556$; $N = 6262$] | −0.005 (0.019) | 0.081** (0.019) | −0.005 (0.016) | $F = 7.47$ [$p = 0.000$] | 0.018 (0.015) | −0.004 (0.013) | −0.002 (0.019) | $F = 0.58$ [$p = 0.632$] |
| I participate in groups outside of school (e.g., church youth groups) [$\mu = 0.338$; $N = 8705$] | 0.017 (0.015) | 0.057** (0.019) | 0.017 (0.013) | $F = 3.17$ [$p = 0.025$] | 0.015 (0.011) | −0.012 (0.010) | −0.011 (0.014) | $F = 1.22$ [$p = 0.304$] |
| Number of days absent this year [$\mu = 6.071$; $N = 7986$] | −1.205** (0.218) | −1.864** (0.248) | −0.492** (0.179) | $F = 21.74$ [$p = 0.000$] | 0.133 (0.150) | −0.052 (0.150) | 0.223 (0.207) | $F = 0.59$ [$p = 0.623$] |
| I was suspended from school this year [$\mu = 0.291$; $N = 8215$] | −0.048** (0.017) | −0.072** (0.016) | −0.017 (0.016) | $F = 7.23$ [$p = 0.000$] | 0.023** (0.012) | 0.003 (0.012) | −0.004 (0.018) | $F = 1.23$ [$p = 0.299$] |
| An unmarried sister got pregnant in the last 2 years [$\mu = 0.111$; $N = 7662$] | −0.021** (0.011) | −0.038** (0.011) | −0.002 (0.010) | $F = 4.44$ [$p = 0.005$] | 0.012 (0.008) | −0.012 (0.008) | −0.010 (0.010) | $F = 2.39$ [$p = 0.070$] |
| My parents have attended a school meeting this year [$\mu = 0.481$; $N = 7296$] | 0.014 (0.020) | 0.082** (0.020) | −0.026 (0.017) | $F = 9.69$ [$p = 0.000$] | −0.013 (0.018) | 0.017 (0.014) | −0.026 (0.020) | $F = 1.15$ [$p = 0.331$] |
| My parents have volunteered at school this year [$\mu = 0.324$; $N = 7083$] | −0.015 (0.018) | 0.081** (0.019) | 0.027 (0.016) | $F = 8.08$ [$p = 0.000$] | 0.003 (0.014) | 0.018* (0.011) | −0.028 (0.018) | $F = 1.64$ [$p = 0.181$] |

All specifications include the full set of covariates described in Table 3. Each row presents the results for regressions where the variable indicated is the dependent variable. The left panel adds indicators for the type of school to which the student opts out to the control set, while the right panel adds the three distance measures shown, as well as log distance from the attendance area school. The survey variables are based on responses to an 8th grade survey administered to a random third of the 1994 cohort. *F*-statistics (*p*-values are shown in square brackets) from tests of the joint significance of the three opting out variables or the three distance variables are shown in the last column in each panel. Standard errors (in parentheses) are corrected for correlation within eighth grade school using White's generalized method.

* Significant at the 10% level.

** Significant at the 5% level.

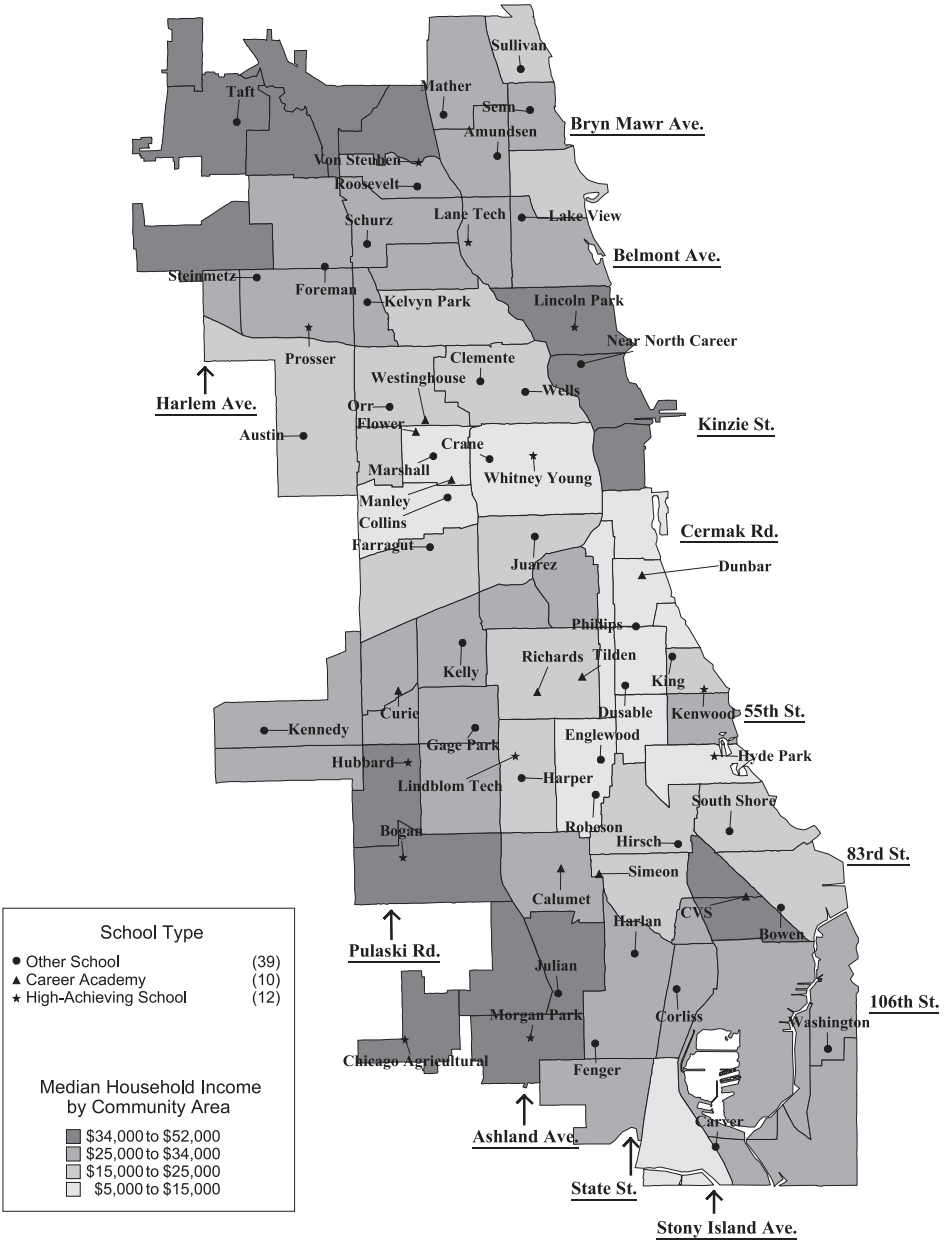


Fig. 1. Geographic distribution of CPS high schools.

case is whether a student opted out. In all specifications, we estimate linear probability models with the full set of covariates utilized in the earlier tables, adding three separate distance measures corresponding to the log distance in miles from the

Table 6
The impact of distance on the likelihood of opting out

| Independent variable | Student ability quartile (lowest to highest) | | | |
|---|---|---|---|---|
| | I | II | III | IV |
| *Dependent variable = opts out* | | | | |
| In(miles from closest career academy) | − 0.025 (0.014) | − 0.063 (0.013) | − 0.091 (0.011) | − 0.057 (0.010) |
| In(miles from closest high-achieving school) | − 0.003 (0.014) | − 0.007 (0.013) | − 0.016 (0.014) | − 0.057 (0.014) |
| In(miles from closest regular school) | − 0.105 (0.019) | − 0.074 (0.018) | − 0.065 (0.015) | − 0.055 (0.014) |
| *F*-test of joint significance | 11.42 | 13.89 | 25.90 | 18.43 |
| | [*p* = 0.000] | [*p* = 0.000] | [*p* = 0.000] | [*p* = 0.000] |
| Pr(opts out) | 0.366 | 0.463 | 0.561 | 0.740 |
| *Dependent variable = opts out to career academy* | | | | |
| In(miles from closest career academy) | − 0.074 (0.008) | − 0.117 (0.010) | − 0.155 (0.010) | − 0.098 (0.012) |
| In(miles from closest high-achieving school) | 0.006 (0.006) | 0.009 (0.008) | 0.002 (0.009) | 0.007 (0.007) |
| In(miles from closest regular school) | 0.008 (0.007) | 0.013 (0.009) | 0.010 (0.011) | 0.014 (0.010) |
| *F*-test of joint significance | 30.38 | 45.10 | 79.16 | 24.66 |
| | [*p* = 0.000] | [*p* = 0.000] | [*p* = 0.000] | [*p* = 0.000] |
| Pr(opts to career academy) | 0.088 | 0.163 | 0.226 | 0.146 |
| *Dependent variable = opts out to high-achieving school* | | | | |
| In(miles from closest career academy) | 0.005 (0.004) | − 0.001(0.005) | − 0.007 (0.007) | − 0.009 (0.014) |
| In(miles from closest high-achieving school) | − 0.001 (0.003) | − 0.015 (0.006) | − 0.029 (0.007) | − 0.082 (0.014) |
| In(miles from closest regular school) | − 0.000 (0.003) | 0.007 (0.005) | 0.002 (0.008) | − 0.018 (0.014) |
| *F*-test of joint significance | 0.89 | 3.29 | 5.46 | 12.13 |
| | [*p* = 0.446] | [*p* = 0.021] | [*p* = 0.001] | [*p* = 0.000] |
| Pr(opts to high-achieving school) | 0.020 | 0.042 | 0.106 | 0.467 |
| *Dependent variable = opts out to regular school* | | | | |
| In(miles from closest career academy) | 0.043 (0.012) | 0.056 (0.012) | 0.070 (0.011) | 0.050 (0.009) |
| In(miles from closest high-achieving school) | − 0.008 (0.014) | − 0.000 (0.012) | 0.011 (0.012) | 0.017 (0.008) |
| In(miles from closest regular school) | − 0.113(0.018) | − 0.093 (0.017) | − 0.076 (0.014) | − 0.051 (0.009) |
| *F*-test of joint significance | 18.78 | 20.27 | 25.06 | 21.41 |
| | [*p* = 0.000] | [*p* = 0.000] | [*p* = 0.000] | [*p* = 0.000] |
| Pr(opts to regular school) | 0.259 | 0.258 | 0.229 | 0.128 |
| Number of observations | 14,898 | 15,017 | 15,221 | 15,487 |

The dependent variable in the specifications in the top panel is a binary indicator for whether a student opts out of the assigned school. The dependent variables in the following panels are binary indicators for whether a students opts out by type of school. Each of the four columns presents results from regressions on subsamples based on quartiles of the overall student ability distribution. Student ability is measured by combined 8th grade math and reading achievement test scores. Although only the coefficients on the distance variables are shown, the full set of covariates described in Table 3 is included as well as log miles to the attendance area school. The distance variables are defined to be the logarithm of distance in miles from the student's home to the closest school of the type indicated. Standard errors (in parentheses) are corrected for correlation within eighth grade school using White's generalized method. We report *F*-statistics (*p*-values in square brackets) from tests of the joint significance of the three distance measure. Unconditional rates of opting out are included in the bottom row of each panel.

student's home to the closest career academy, the closest high-achieving school, and the closest school that does not fall into either of these categories.[26] The impact of these distance variables on opting out varies substantially across the ability distribution, so we present separate estimates by 8th grade test score quartile.

The top panel of Table 6 shows results when the dependent variable is an indicator for opting out to any school other than the assigned school. The coefficient on log distance is negative for every school type in every quartile, consistent with the hypothesis that greater distances are associated with a lower likelihood of opting out. An *F*-test of the joint significance of the distance measures is highly statistically significant in each of the four quartiles. The relative magnitudes of the coefficients are sensible. Distance to the closest career academy has the greatest impact on students in the middle two quartiles—precisely the students who are most likely to opt out to such schools. Distance to high-achieving schools has almost no impact on students in the bottom two quartiles, a moderate impact on those in the third quartile, and a large impact on those in the top quartile. These results parallel the frequency with which students at different points of the ability distribution attend high-achieving schools. Similarly, distance to regular schools has the greatest impact at the bottom of the ability distribution.

Panels 2–4 of Table 6 explore the relationship between distances to schools and the decision to opt out to a particular type of school. In panel 2, for instance, distance to the closest career academy is shown to be a strong predictor of the likelihood of opting out to a career academy, with *t*-statistics around 10. Panel 3 presents similar results for high-achieving schools. Panel 4, which corresponds to opting out to regular schools, also shows a big influence of distance to regular schools on opting out to such schools. Interestingly, career academies appear to act as good substitutes for regular schools, i.e., the presence of a nearby career academy greatly reduces the probability that a student will opt out to a regular school.[27]

Given the findings of Table 6, it is clear that distances influence the choices that students make concerning which school to attend. If it is also the case that distances are uncorrelated with unobserved student characteristics, then these distances can be used to generate reduced form estimates of the effect of school choice that are purged of contamination from that type of unobservable. Before controlling for observables, distance is likely to be correlated with student characteristics. A student's place of residence is not randomly assigned, but rather, chosen by parents, presumably with school options as one component in the optimization. School locations are not randomly assigned either. Career academies are concentrated in poorer areas of the cities with higher concentrations of Blacks. High-achieving schools are disproportionately located in affluent neighbor-

[26] Also included in these regressions is the log distance to the assigned attendance area school. We do not present the coefficient on that variable since it cannot be interpreted as an indicator of access to choice in the same way as the other distance measures. As would be expected, distance to one's own attendance area school is highly positively correlated with the likelihood of opting out.

[27] Although not shown in tabular form, we have also replicated the results in Table 6 replacing the dependent variable with an indicator variable for whether the student actually opts out to the *nearest* career academy, the *nearest* high-achieving school, or the *nearest* other school. The coefficients on the distance variables are similar in magnitude in these specifications, as one would hope.

Table 7
Reduced form estimates of the impact of distance on the probability of graduating

| Dependent variable = binary indicator for whether student graduates | | | | |
|---|---|---|---|---|
| Independent variable | Student ability quartile (lowest to highest) | | | |
| | I | II | III | IV |
| ln(miles from closest career academy) | − 0.019 (0.014) | − 0.028 (0.012) | − 0.027 (0.012) | − 0.015 (0.010) |
| ln(miles from closest high-achieving school) | − 0.014 (0.011) | − 0.016 (0.011) | − 0.008 (0.010) | 0.006 (0.009) |
| ln(miles from closest regular school) | 0.005 (0.012) | − 0.005 (0.016) | 0.016 (0.011) | − 0.004 (0.008) |

The dependent variable is a binary indicator for whether a student graduates. The four columns present results from regressions on subsamples based on quartiles of the overall student ability distribution. Student ability is measured by combined 8th grade math and reading achievement test scores. Although only the coefficients on the distance variables are shown, log distance to the attendance area school and the full set of covariates described in Table 3 are included. The distance measures are the logarithm of distance in miles from the student's home to the closest school of the type indicated. Standard errors (in parentheses) are corrected for correlation within 9th grade high school using White's generalized method.

hoods.[28] We attempt to minimize concerns about endogenous location by controlling for an unusually rich set of baseline covariates. Worth highlighting among this set are the controls for whether a student moved residences in the summer prior to beginning high school and average student test scores in 8th grade for both the census tract of residence and the attendance area.

Most importantly, we can directly test whether the distance measures are correlated with the proxies for unobservable characteristics that were so strongly correlated with the decision to opt out. While by no means definitive, the existence of these survey data provides us a much more direct and convincing test of the underlying exogeneity assumption than is typically available. Results are presented in columns 5–8 of Table 5. For none of the variables that proxy for unobservables can we reject at the 0.05 level the null hypothesis that the coefficients on the distance measures are jointly equal to zero. Only 3 of the individual coefficients out of 57 are statistically significant at the 0.10 level. These findings are extremely encouraging regarding the exogeneity of the distance measures, although it is of course possible that these measures may still be correlated with other unobservables we cannot measure.[29]

The reduced-form results relating an indicator for whether a student graduates to these distance measures are presented in the top panel of Table 7. Results are presented for each ability quartile separately. In addition to the distance measures reported in the table, these regressions include the full set of covariates. Estimation is done with linear probability models, correcting the standard errors for correlation within 9th grade high schools.

[28] Although it also appears that the Chicago School Board has made a conscious effort to locate attractive schools in areas with high proportions of poor and minority students as part of the desegregation project. For instance, Whitney Young and Lindblom Tech—two of the three highest achieving schools in the city—are located in relatively poor neighborhoods.

[29] Although not shown here due to space constraints, estimating the results separately for each ability quartile yields comparable results. The distance variables are not correlated with prior achievement, mobility or survey measures for any of the four ability quartiles. Tables available from the authors upon request.

If shorter travel distances to alternative schools benefit students on average, then the coefficients on the three distance measures should be negative in Table 7 (i.e., longer travel distances reduce graduation rates). The strongest results emerge for career academies. Within each quartile of the ability distribution, being farther away from the nearest career academy reduces the likelihood of graduation. The largest effects are in the middle of the ability distribution, consistent with the earlier finding that these students are most likely to opt out to career academies and that a change in distance has the largest impact on their probability of opting out. Based on the estimates in Table 7, a student between the 25th and 50th percentiles of the ability distribution who lives 2 miles closer to a career academy (roughly one standard deviation) will be 2.1 percentage points–about 5%—more likely to graduate high school than a comparable peer.[30]

There is some weak evidence that proximity to high-achieving schools may provide benefits to students in the lower two quartiles. Given that only a few percent of these students opt out to such schools and distance to a high-achieving school had virtually no impact on the probability of opting out for these quartiles, it is implausible that the pattern reflects direct benefits to travelers. One explanation for any apparent benefits to low-ability students of having high-achieving schools nearby is that the presence of high ability students imposes a negative externality on low-achieving students, such as through a less appropriate curriculum.

Surprisingly, there is no apparent benefit to students in the highest ability quartile—the students most likely to opt out to attend such schools. This is in stark contrast to the OLS estimates, which implied that going to a high-achieving school had an enormous impact on graduation. If one assumes that there are no spillovers to those who remain behind, the 2SLS coefficient on opting out to a high-achieving school in the top quartile implied by the reduced form estimate is $-0.071$ (standard error = 0.080). Despite the imprecision of the estimates, we can reject the null hypothesis that the OLS estimate (0.122 with a standard error of 0.011) and the 2SLS estimate are the same at the 0.05 level. The gap between the reduced form and OLS results is consistent with students who choose to attend high-achieving schools being systematically different from observationally equivalent students who do not select (or are not selected by) such schools.

The lack of benefit on average to being near a high-achieving school for top quartile students cannot be easily explained by the fact that the students who leave

[30] Although not shown in the table, the 2SLS coefficient on opting out to a career academy ranges between 0.155 to 0.259 across the four quartiles, with all four coefficients statistically significant at the 0.10 level and three of the four significant at the 0.05 level. If one accepts the assumption of no spillovers to those who remain in their assigned school, the 2SLS coefficient can be interpreted as the causal effect on those who opt out. Note that these 2SLS coefficients are larger than the OLS estimates of opting out to a career academy in Table 4. This suggests either that (1) there is negative selection into career academies according to unobservable characteristics—unlikely given the survey results in Table 5, (2) match quality at the attendance area schools for students who opt out to career academies is especially low (quite likely given that career academies tend to attract students who have struggled in the conventional system—see Kemple and Snipes, 2000), or (3) students who remain behind experience positive externalities when other students elect to opt out to career academies (also plausible if their departure either led to competition-induced improvement or increased resources devoted to other students by, for example, decreasing class sizes).

make the high-achieving students who stay behind worse off through reduced peer quality. When we replicate specifications identical to those in Table 7, but with average peer 8th grade test score added to the regression, the coefficients on the distance variables are virtually unchanged.[31] Moreover, students in the lower half of the ability distribution appear if anything to benefit on average from closer proximity to high-achieving schools, so it is hard to make a compelling case that the result is being driven by an omitted variable that is correlated with both proximity to high-achieving schools and low graduation rates.

Distance to regular schools is not statistically significant for any of the quartiles and flips signs erratically. While it is difficult to draw any strong conclusions from this, it casts some doubt on the validity of the systematic benefits that were seen in the cross-section.

### *5.4. Additional evidence on why opting out improves attainment from within-school estimates*

Of the 61 schools in the CPS, 47 are attendance area schools, meaning that some students are assigned to them. Other students, however, are allowed to opt into these schools. As a consequence, for this subset of schools, we are able to identify the impact of opting in to the school controlling for school-fixed effects.[32] Within a school, students are likely to experience similar average peer quality, teacher quality, resource levels, and curriculum. Thus, if any of those factors are important in determining why those who opt out are more likely to graduate, adding school-fixed effects should diminish the estimated impact on graduation rates. On the other hand, if adding school dummies has little impact on the travel coefficient, then other explanations such as unobserved motivation or better match quality for those who travel must explain the cross-sectional relationship.

The first column of Table 8 simply reports baseline OLS estimates of the impact of opting out to the three types of schools, similar to the estimates in Table 4. The second column of the table reports the results from the exact same regression, but restricts the sample to students who attend one of the 47 attendance area schools, since this is the sample that actually provides identifying variation to the within-school estimates. The results obtained in the subsample are similar to those of the overall sample.

Column 3 adds school-fixed effects to the regression on this subsample. For career academies, the coefficient shrinks by 80% and is no longer statistically significant. Thus, the graduation benefits of attending career academies seem to accrue to all students—those who attend via choice and those who attend by default because they are assigned to the school. These findings are consistent with the idea that there is something about the teachers, programs, or curricula at such schools that keeps students in school.

[31] This approach to estimating peer effects is admittedly crude, but would be expected to overstate the role of mean peer quality (1993). Lefgren (2001) uses the extent of tracking within a school to identify peer effects in CPS. He finds little or no evidence for peer effects operating through mean peer quality.

[32] For the 14 schools in which enrollment is by application only, all students who attend must by definition opt in, so there is no within-school variation in the opting out variable.

Table 8
The impact of opting out controlling for school fixed effects

| Independent variable | Dependent variable = binary indicatory for whether graduates | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Opts out to career academy | 0.078 (0.022) | 0.106 (0.047) | 0.020 (0.021) | 0.024 (0.021) |
| Opts out to high-achieving school | 0.146 (0.016) | 0.127 (0.026) | 0.093 (0.024) | 0.080 (0.024) |
| Opts out to regular school | 0.043 (0.010) | 0.045 (0.011) | 0.032 (0.006) | 0.030 (0.007) |
| Attendance area school sample | No | Yes | Yes | Yes |
| School-cohort fixed effects | No | No | Yes | Yes |
| School-cohort fixed effects × ability quartile | No | No | No | Yes |
| Number of Observations | 60,623 | 46,513 | 46,513 | 46,513 |

Each column presents estimates from a separate specification. In each case, the dependent variable is a binary indicator for whether the student graduates and the full set of control variables from Table 3 is included. Only the coefficients (and standard errors) on the three indicators for opting out of the assigned school are shown. The first column presents our baseline estimates for the full sample of students described in Table 3. Because the impact of opting out is identified only from schools that have both assigned students and students who opt in once school fixed effects are included, the second column shows the baseline estimates for this sample of schools for reference. The last two columns add school-cohort and school-cohort-ability quartile fixed effects, respectively, to the baseline model. Robust standard errors are shown in parentheses.

For high-achieving and regular schools, the results are very different. Adding school-fixed effects has little impact on the measured benefit of opting out. This suggests that the higher graduation rates among students opting out to those types of schools is unlikely to be due to peers, resources, teachers, or curricula at those schools. Instead, consistent with the earlier findings, the apparent benefit to students opting out to attend such schools is likely to be spuriously driven by unobserved characteristics of these students.[33]

One critique of the above interpretation is that many attendance area schools feature magnet programs. One might expect that travelers disproportionately take advantage of such programs. If that were the case, a single fixed-effect per school may not capture the range of experiences of students at the school. The inclusion of school and ability quartile interactions to more closely capture shared school experiences in column 4, however, reduces the estimates only slightly.

## 6. Conclusion

The Chicago Public Schools have implemented a system of school choice in which more than half of all students elect to opt out of their local assigned school to attend another public school in Chicago. Although students who opt out are more likely to graduate than observationally equivalent students who attend their neighborhood school, our evidence suggests that the leading explanation for this correlation is that

[33] If student–school match quality were responsible, then the reduced form results would have been expected to show an average benefit.

those who opt out are superior along unobservable dimensions such as their motivation level and parental involvement. The notable exception to this finding is career academies, which appear to provide real improvement in graduation rates. This finding is consistent with recent experimental evidence on the positive impacts of career academies among high-risk youth.

Perhaps the greatest puzzle raised by our analysis is why so many students choose to opt out of their assigned school, despite the fact that we are able to identify little tangible academic benefit of doing so. One possible explanation is that there are non-academic reasons for opting out. Most students who leave their local school attend a school with more academically talented peers who come from neighborhoods with higher socio-economic characteristics. It is possible that such schools have lower levels of violence, for instance.[34] Alternatively, it may simply be the case that students and their parents do not accurately gauge the education production function. Students and parents have little information about educational inputs (e.g., the raw ability of students attending a particular school), but good information about educational outputs (e.g., average test scores and graduation rates). Given this limited information, it may be rational to opt out of local schools to attend schools with higher levels of outputs, even though in reality, the latter set of schools on average confer no additional benefits to those who attend.

Despite these puzzles, the findings presented have important policy implications. With the exception of career academies, we find that systemic choice within a public school district does not seem to benefit those who participate. This casts doubt on the power of this form of choice to improve educational outcomes, or to serve as an efficient form of discipline for low quality neighborhood schools.[35] On the other hand, given that attending career academies leads to higher attainment, focusing greater attention on these schools could improve understanding of the aspects of the learning environment that matter for at-risk students.

## Acknowledgements

We would like to thank John Easton of the Consortium on Chicago School Research and the Chicago School Board for providing access to the Chicago Public Schools data. The National Science Foundation and Sloan Foundation provided financial support. We would also like to gratefully acknowledge Michael Greenstone, Jinyong Hahn, Caroline Hoxby, Lawrence Katz, Alan Krueger, Susanna Loeb, Charles Manski, Bruce Meyer, Kevin Murphy, Steve Rivkin, Jeff Smith, Chris Taber, three anonymous referees, the editor Lans Bovenberg, and seminar participants too numerous to mention for helpful comments.

[34] Nonetheless, if this were the case, one might predict that students attending such schools would be more likely to graduate. We do not find this to be true.

[35] We cannot rule out that involving the private sector or nontraditional public schools, such as charter schools, would lead to a more favorable picture of the role of choice in school reform.

## Appendix A. Data appendix

| Data | Source | Construction |
|---|---|---|
| Student outcomes | CPS Board | Standardized test scores are taken directly from student test and transcript files provided by the Board. Information on graduation is from administrative records provided by the Board. We classify a student as having graduated if the student is not actively enrolled in the CPS in the Fall following their fourth year of high school and the reason the student left the CPS is marked as graduation. |
| Attendance area school | CPS Board | Board staff matched students to attendance area schools using students' Fall 9th grade addresses and current (Fall 1999) attendance area boundaries provided by the CPS. A few 1999 attendance area schools did not exist as options for our 1993–1995 cohorts. In these cases, we assume that the attendance area school is the closest neighborhood school that is attended by the majority of affected students in our cohorts. |
| Distance variables | CPS Board | Student and school addresses were geocoded using MapInfo software and used to calculate home-to-school distances. |
| Student demographics | CPS Board | Student demographic variables (race, gender, age) come directly from student records provided by the Board. Family composition is based on information indicating the child's guardian in the Fall of 9th grade. Special education status covers a variety of disabilities ranging from mild learning disabilities to severe physical handicaps. |
| Student SES | 1990 Census | We created a measure of social status using 1990 Census measures at the student block group level. The index combines information on the percent of employed residents who are professionals and the mean education level of residents in the student's block group in the Fall of 9th grade. |
| Neighborhood characteristics | 1990 Census and CPS Board and CCSR | Basic information on the student's census tract, such as median household income and percent below the poverty line, comes from the 1990 Census. The crime composite is an index created by factor analysis using official crime statistics for 1994 provided by the Consortium on Chicago School Research (CCSR). The index was created at the block group level. The variable used in this analysis is a tract-level average (for the student's tract in the Fall of 9th grade), weighted by the total population in each block group. |
| Survey variables | CPS Board | A survey, "Charting Reform: The Student Speaks," was administered to a random third of all 8th graders in the Spring of 1994 (our 1994 cohort). Students were asked a variety of questions about family background, past performance, and expectations for the future. |

## Appendix B. Technical appendix: treatment effects framework

Consider the following linear probability model:

$$y_{is} = \mathbf{X}_i\beta + \alpha_i + \theta_{is} + \gamma_s + \nu_{is}, \quad (1)$$

where $y$ is an indicator for whether or not the student graduates, $\mathbf{X}$ includes background characteristics, and $i$ and $s$ indicate student and school, respectively. We assume that outcomes for students who are similar according to observable dimensions will differ systematically due to unobserved student qualities ($\alpha_i$), the quality of the specific match between the student and the school ($\theta_{is}$), and the shared quality of the school attended ($\gamma_s$). The final term, $v_{is}$, is a zero mean random component that is independent of the other variables. It is useful to think of the problem in a causal effects framework with two policy regimes: the "choice" regime (C) which we observe in the data and a counterfactual "no choice" regime (N) that we do not observe, but would like to observe in ideal data. Within the "choice" regime, students can be classified into one of two categories: those who "opt out" of their local school in the choice regime ($t_i = 1$; travelers) and those who attend the attendance area school ($t_i = 0$; stayers).

Define the causal effect of introducing school choice on students who opt out as follows:

$$\begin{aligned}\delta = E[y_{ic}^{C} \mid \mathbf{X}_i, t_i = 1] - E[y_{ia}^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 1] = (E[\theta_{ic}^{C} \mid \mathbf{X}_i, t_i = 1] - E[\theta_{ia}^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 1]) \\ + (E[\gamma_c^{C} \mid \mathbf{X}_i, t_i = 1] - E[\gamma_a^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 1]), = \delta_\theta + \delta_\gamma \qquad (2)\end{aligned}$$

where $a$ indicates the assigned school and $c$ indicates the chosen school. The treatment effect for travelers can be further decomposed into:

$$\begin{aligned}\delta = (E[\theta_{ic}^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 1] - E[\theta_{ia}^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 1]) + (E[\gamma_c^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 1] \\ -E[\gamma_a^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 1]) + (E[\theta_{ia}^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 1] - E[\theta_{ia}^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 1]) \\ +(E[\gamma_a^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 1] - E[\gamma_a^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 1]) = (\delta_{\theta'} + \delta_{\gamma'}) + (\delta_{\theta''} + \delta_{\gamma''}) \\ = \delta_{\mathrm{alt}} + \delta_{\mathrm{ge}} \qquad (3)\end{aligned}$$

Outcomes for travelers change both because of the change in match and school qualities due to the choice of an alternative school under choice, and because the qualities at the assigned school are affected by choice. The first set of effects comprise an alternative treatment effect (which we denote $\delta_{\mathrm{alt}}$) that measures how much better travelers do merely because they have chosen an alternative school under the choice regime. The second set ($\delta_{\mathrm{ge}}$) captures any general equilibrium effects on match and school quality.

Define the causal effect of introducing school choice on students who stay as:

$$\begin{aligned}\lambda = E[y_{ia}^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 0] - E[y_{ia}^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 0] = (E[\theta_{ia}^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 0] - E[\theta_{ia}^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 0]) \\ + (E[\gamma_a^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 0] - E[\gamma_a^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 0]) = \lambda_\theta + \lambda_\gamma, \qquad (4)\end{aligned}$$

The OLS estimate of the relationship between traveling and attainment under choice measures:

$$\begin{aligned}\delta^{\mathrm{OLS}} = E[y_{ic}^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 1] - E[y_{ia}^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 0] = \delta + (E[\alpha_i \mid \mathbf{X}_i, t_i = 1] \\ -E[\alpha_i \mid \mathbf{X}_i, t_i = 0]) + (E[\theta_{ia}^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 1] - E[\theta_{ia}^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 0]) \\ +(E[\gamma_a^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 1] - E[\gamma_a^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 0]) = \delta + \mu_\alpha + \mu_\theta + \mu_\gamma \qquad (5)\end{aligned}$$

This formulation treats stayers as the comparison group, but their outcomes under choice differ from what the outcomes for travelers would be in the absence of choice for three reasons. First, stayers are likely to have lower unobserved ability than stayers. Second, the initial match quality for stayers at the assigned school is likely to be better than for those who opt out, and is also likely affected by the existence of choice. The reverse is true of shared school quality.

The fact that the last two biases arise from both initial differences between travelers and stayers and from general equilibrium or spillover effects is more easily seen in the reformulation below:

$$\begin{aligned}\delta^{\mathrm{OLS}} &= \delta + \mu_{\alpha} + (E[\theta_{ia}^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 1] - E[\theta_{ia}^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 0]) + (E[\theta_{ia}^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 0] \\ &\quad - E[\theta_{ia}^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 0]) + (E[\gamma_a^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 1] - E[\gamma_a^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 0]) \\ &\quad + E[\gamma_a^{\mathrm{N}} \mid \mathbf{X}_i, t_i = 0] - E[\gamma_a^{\mathrm{C}} \mid \mathbf{X}_i, t_i = 0] = \delta + \mu_{\alpha} + \mu_{\theta'} + \mu_{\gamma'} - \lambda \end{aligned} \tag{6}$$

If we assume that average shared school quality does not vary across travelers and non-travelers in the absence of choice conditional on $\mathbf{X}$,[36] then $\mu_{\gamma'} = 0$ and we are left with the three biases described intuitively in the text.

In order to estimate something closer to a causal effect, one would like to identify an instrumental variable $z$, which predicts whether a given student opts out, but is otherwise uncorrelated with student outcomes. Take the simplest case in which there is one dichotomous instrumental variable that affects whether a student opts out, but is uncorrelated with unobserved student ability and the error term in Eq. (1). A reduced form regression with educational attainment on the left-hand side and the instrumental variable and observable characteristics $\mathbf{X}$ on the right-hand side yields a coefficient on the instrument of:

$$\begin{aligned}\pi &= E[y_{ic}^{\mathrm{C}} \mid \mathbf{X}_i, z_i = 1] - E[y_{ic}^{\mathrm{C}} \mid \mathbf{X}_i, z_i = 0] = E[\theta_{ic}^{\mathrm{C}} + \gamma_c^{\mathrm{C}} \mid \mathbf{X}_i, z_i = 1] \\ &\quad - E[\theta_{ic}^{\mathrm{C}} + \gamma_c^{\mathrm{C}} \mid \mathbf{X}_i, z_i = 0] \end{aligned} \tag{7}$$

If we assume that the instrument is also uncorrelated with what match quality and shared school quality would be in the absence of choice (which is reasonable since the average ability of students assigned to the same attendance area school is included in $\mathbf{X}$),[37] this can be expressed as:

$$\begin{aligned}\pi &= E[\theta_{ic}^{\mathrm{C}} + \gamma_c^{\mathrm{C}} - \theta_{ia}^{\mathrm{N}} - \gamma_a^{\mathrm{N}} \mid \mathbf{X}_i, z_i = 1] - E[\theta_{ic}^{\mathrm{C}} + \gamma_c^{\mathrm{C}} - \theta_{ia}^{\mathrm{N}} - \gamma_a^{\mathrm{N}} \mid \mathbf{X}_i, z_i = 0] \\ &= (P_1\delta^1 - P_0\delta^0) + ((1 - P_1)\lambda^1 - (1 - P_0)\lambda^0) \end{aligned} \tag{8}$$

where $P_1$ and $P_0$ indicate the fractions of students who travel when the instrument is equal to 1 and to 0, respectively. The superscripts signify the causal treatment effects for the populations according to the value of the instrument.

[36] More formally, the assumption is that $E[\gamma_s^{\mathrm{N}}|\mathbf{X}_i, t_i = 1] = E[\gamma_s^{\mathrm{N}}|\mathbf{X}_i, t_i = 0]$. The mean ability of students assigned to the same school is among the variables included in our control set, which can be thought of as a proxy for shared school quality under the counterfactual of no choice.

[37] This means $E[\theta_{ia}^{\mathrm{N}}|\mathbf{X}_i, z_i = 1] = E[\theta_{ia}^{\mathrm{N}}|\mathbf{X}_i, z_i = 0]$ and $E[\gamma_a^{\mathrm{N}}|\mathbf{X}_i, z_i = 1] = E[\gamma_a^{\mathrm{N}}|\mathbf{X}_i, z_i = 0]$. We are assuming that initial match and school quality would not vary across assigned schools with the same types of peers assigned to them in the absence of school choice.

It is easier to interpret how spillovers affect this estimate by recasting the reduced form using the following relationships. The first term, after adding and subtracting two terms, becomes:

$$\begin{aligned} &E[\theta_{ic}^{C} + \gamma_{c}^{C} - \theta_{ia}^{N} - \gamma_{a}^{N} \mid \mathbf{X}_i, z_i = 1] \\ &\quad = E[\theta_{ic}^{C} + \gamma_{c}^{C} - \theta_{ia}^{C} - \gamma_{a}^{C} + \theta_{ia}^{C} + \gamma_{a}^{C} - \theta_{ia}^{N} - \gamma_{a}^{N} \mid \mathbf{X}_i, z_i = 1] \\ &\quad = P_1(\delta_{\text{alt}}^{1}) + E[\theta_{ia}^{C} + \gamma_{a}^{C} - \theta_{ia}^{N} - \gamma_{a}^{N} \mid \mathbf{X}_i, z_i = 1] \end{aligned} \tag{9}$$

The second term can be recast in the same way:

$$E[\theta_{ic}^{C} + \gamma_{c}^{C} - \theta_{ia}^{N} - \gamma_{a}^{N} \mid \mathbf{X}_i, z_i = 0] = P_0(\delta_{\text{alt}}^{0}) + E[\theta_{ia}^{C} + \gamma_{a}^{C} - \theta_{ia}^{N} - \gamma_{a}^{N} \mid \mathbf{X}_i, z_i = 0] \tag{10}$$

Assuming that the average impact of opting out of the assigned school under the choice regime does not vary with the value of the instrument, the coefficient can then be expressed as:

$$\pi = (P_1 - P_0)(\delta_{\text{alt}}) + (E[\theta_{ia}^{C} + \gamma_{a}^{C} \mid \mathbf{X}_i, z_i = 1]) - (E[\theta_{ia}^{C} + \gamma_{a}^{C} \mid \mathbf{X}_i, z_i = 0]) \tag{11}$$

The terms reflecting match and school quality in the no choice regime ($\theta_{ia}^{N}$ and $\gamma_{a}^{N}$) drop out since we have assumed these do not vary by the value of our instrument. The first term is proportional to the change in outcomes for travelers because they choose an alternative school, and the second term captures spillovers to the quality of neighborhood schools due to differential degrees of choice. Negative spillovers will lower the reduced form estimate, while positive spillovers will raise it. In the absence of spillovers, one could obtain the traditional indirect least squares estimate by dividing by $(P_1 - P_0)$, which is simply the first-stage estimate of the effect of the instrument on opting out.

## References


Blank, R.K., 1983. Survey of Magnet Schools: Analyzing a Model of Quality Integrated Education. ABT Associates for the Department of Education, Washington, DC.

Buddin, R.J., Cordes, J.J., et al., 1998. School choice in California: who chooses private schools. Journal of Urban Economics 44, 110–134.

Chubb, J.E., Moe, T.M., 1990. Politics, Markets and America's Schools. Brookings Institution, Washington, DC.

Coleman, J., Schiller, K., et al., 1993. Parent choice and inequality. In: Schneider, B., Coleman, J. (Eds.), Parents, Their Children and Schools. Westview Press, Boulder, CO, pp. 147–182.

Crain, R.L., Heebner, A.L., et al., 1992. The Effectiveness of New York City's Career Magnet Schools: An Evaluation of Ninth-Grade Performance Using an Experimental Design. National Center for Research in Vocational Education, Berkeley, CA.

Cullen, J.B., Jacob, B.A., Levitt, S., 2003. The effect of school choice on student outcomes: evidence from randomized lotteries. National Bureau of Economics Research Working Paper 10113, Cambridge, MA.

Duax, T., 1988. The Impact of Magnet Schools on a Predominantly Black Community. University of Wisconsin, Madison.

Epple, D., Romano, R.E., 1998. Competition between private and public schools, vouchers and peer-group effects. American Economic Review 88 (1), 33–62.

Evans, W., Schwab, R., 1995. Finishing high school and starting college: do catholic schools make a difference? Quarterly Journal of Economics 110 (4), 941–974.

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

Goldhaber, D.D., Brewer, D.J., et al., 1999. Testing for sample selection in the Milwaukee school choice experiment. Economics of Education Review 18 (2), 259–267.

Green, J.P., Peterson, P.E., et al., 1997. The Effectiveness of School Choice: The Milwaukee Experiment. Program on Education Policy and Governance, Harvard University, Cambridge, MA.

Hoxby, C.M., 1994. Does Competition Among Public Schools Benefit Students and Taxpayers? National Bureau of Economic Research, Cambridge, MA.

Kemple, J.J., Rock, J.L., 1996. Career Academies: Early Implementation Lessons from a 10-Site Evaluation. Manpower Demonstration Research, New York.

Kemple, J.J., Snipes, J.C., 2000. Career Academies' Impact on Students' Engagement and Performance in High School. Manpower Demonstration Research, New York.

Ladd, H.F., Fiske, E.B., 2000. When Schools Compete: A Cautionary Tale. Brookings Institution Press, Washington, DC.

Lankford, R.H., Lee, E.S., et al., 1995. An analysis of elementary and secondary school choice. Journal of Urban Economics 38, 236–251.

Lee, V.E., Croninger, R.G., et al., 1994. Parental choice of schools and social stratification. Educational Evaluation and Policy Analysis 16, 434–457.

Lefgren, L., 2001. Are Education Peer Effects Real? Unpublished manuscript. University of Chicago Graduate School of Business.

Metz, M.H., 1986. Different by Design: The Context and Character of Three Magnet Schools. Routledge Press, New York.

National Center for Education Statistics (1996). Public School Choice Programs 1993–1994: Availability and Student Participation (NCES 97-909).

O'Brien, D.M., Murdoch, J.C., 2000. School choice in a large Texas school district. American Education Finance Association Annual Meeting, Austin, Texas.

Peterson, P.E., Myers, D., et al., 1998. An Evaluation of the New York City School Choice Scholarship Program: The First Year. Mathematica Policy Research and the Harvard Program on Education Policy and Governance, Cambridge, MA.

Rouse, C.E., 1998. Private school vouchers and student achievement: an evaluation of the milwaukee parental choice program. Quarterly Journal of Economics 113 (2), 553–602.

Schneider, B., Schiller, K.S., et al., 1996. Public school choice: some evidence from the national education longitudinal study of 1988. Educational Evaluation and Policy Analysis 18 (1), 19–29.

Witte, J.F., 1993. The Milwaukee Private-School parental choice program. In: Rasell, E., Rothstein, R. (Eds.), School Choice: Examining the Evidence. Economic Policy Institute, Washington, DC, pp. 69–110.

Witte, J.F., 1997. Achievement Effects of the Milwaukee Voucher Program. American Economics Association Annual Meeting, New Orleans.

Witte, J.F., Sterr, T.D., et al., 1995. Fifth-Year Report: Milwaukee Parental Choice Program. University of Wisconsin, Madison, WI.