Do Parents Value School Effectiveness?

Atila Abdulkadiroğlu  Parag A. Pathak  Jonathan Schellenberg
Duke University and NBER  MIT and NBER  UC Berkeley
Christopher R. Walters*
UC Berkeley and NBER

December 28, 2017

Abstract

School choice may lead to improvements in school productivity if parents’ choices reward effective schools and punish ineffective ones. This mechanism requires parents to choose schools based on causal effectiveness rather than peer characteristics. We study relationships among parent preferences, peer quality, and causal effects on outcomes for applicants to New York City’s centralized high school assignment mechanism. We use applicants’ rank-ordered choice lists to measure preferences and to construct selection-corrected estimates of treatment effects on test scores, high school graduation, college attendance, and college quality. Parents prefer schools that enroll high-achieving peers, and these schools generate larger improvements in short- and long-run student outcomes. Preferences are unrelated to school effectiveness and academic match quality after controlling for peer quality.

We gratefully acknowledge funding from the National Science Foundation and the W. T. Grant Foundation. Thanks also go to Nikhil Agarwal, Josh Angrist, Stephane Bonhomme, David Card, David Chan, Michael Dinerstein, Will Dobbie, James Heckman, Peter Hull, Pat Kline, Thibaut Lamadon, Magne Mogstad, Jack Mountjoy, Derek Neal, Ariel Pakes, Stephen Raudenbush, Jesse Rothstein, Alex Torgovitsky, Miguel Urquiola, Jeffrey Wooldridge, Danny Yagan, Seth Zimmerman, and seminar participants at the University of Chicago Committee on Education Workshop, the University of Chicago Harris School, UC Berkeley, the University of Chicago Econometrics Lunch, the University of Chicago Interactions Conference, UW Madison, the 2017 NBER Labor Studies Fall meetings, and the University of Virginia for suggestions and comments. We’re grateful to Eryn Heying for invaluable administrative support. Abdulkadiroğlu and Pathak are Scientific Advisory Board members of the Institute for Innovation in Public School Choice.
1 Introduction

Recent education reforms in the United States, including charter schools, school vouchers, and district-wide open enrollment plans, increase parents’ power to choose schools for their children. School choice allows households to avoid undesirable schools and forces schools to satisfy parents’ preferences or risk losing enrollment. Proponents of choice argue that this competitive pressure is likely to generate system-wide increases in school productivity and boost educational outcomes for students (Friedman, 1962; Chubb and Moe, 1990; Hoxby, 2003). By decentralizing school quality assessment and allowing parents to act on local information, school choice may provide better incentives for educational effectiveness than could be achieved by a centralized accountability system. Choice may also improve outcomes by allowing students to sort into schools that suit their particular educational needs, resulting in improved match quality (Hoxby, 2000). These arguments have motivated recent policy efforts to expand school choice (e.g., DeVos, 2017).

If choice is to improve educational effectiveness, parents’ choices must result in rewards for effective schools and sanctions for ineffective ones. Our use of the term “effective” follows Rothstein (2006): an effective school is one that generates causal improvements in student outcomes. Choice need not improve school effectiveness if it is not the basis for how parents choose between schools. For example, parents may value attributes such as facilities, convenience, student satisfaction, or peer composition in a manner that does not align with educational impacts (Hanushek, 1981; Jacob and Lefgren, 2007). Moreover, while models in which parents value schools according to their effectiveness are an important benchmark in the academic literature (e.g., Epple et al., 2004), it may be difficult for parents to separate a school’s effectiveness from the composition of its student body (Kane and Staiger, 2002). If parent choices reward schools that recruit higher-achieving students rather than schools that improve outcomes, school choice may increase resources devoted to screening and selection rather than better instruction (Ladd, 2002; MacLeod and Urquiola, 2015). Consistent with these possibilities, Rothstein (2006) shows that cross-district relationships among school choice, sorting patterns, and student outcomes fail to match the predictions of a model in which school effectiveness is the primary determinant of parent preferences.

This paper offers new evidence on the links between parent preferences, school effectiveness, and peer quality based on choice and outcome data for more than 250,000 applicants in New York City’s centralized high school assignment mechanism. Each year, thousands of New York City high school applicants rank-order schools, and the mechanism assigns students to schools using the deferred acceptance (DA) algorithm (Gale and Shapley, 1962; Abdulkadiroğlu et al., 2005). The DA mechanism is strategy-proof: truthfully ranking schools is a weakly dominant strategy for students (Dubins and Freedman, 1981; Roth, 1982). This fact motivates our assumption that applicants’ rankings measure their true preferences for schools.¹ We summarize these preferences by fitting discrete choice models to applicants’ rank-ordered preference lists.

¹As we discuss in Section 2, DA is strategy-proof when students are allowed to rank every school, but the New York City mechanism only allows applicants to rank 12 choices. Most students do not fill their preference lists, however, and truthful ranking is a dominant strategy in this situation (Haeringer and Klijn, 2009; Pathak and Sönmez, 2013). Fack et al. (2015) propose empirical approaches to measuring student preferences without requiring that truth-telling is the unique equilibrium.
We then combine the preference estimates with estimates of school treatment effects on test scores, high school graduation, college attendance, and college choice. Treatment effect estimates come from “value-added” regression models of the sort commonly used to measure causal effects of teachers and schools (Todd and Wolpin, 2003; Koedel et al., 2015). We generalize the conventional value-added approach to allow for match effects in academic outcomes and to relax the selection-on-observables assumption underlying standard models. Recent evidence suggests that value-added models controlling only for observables provide quantitatively useful but biased estimates of causal effects due to selection on unobservables (Rothstein, 2010, 2017; Chetty et al., 2014a; Angrist et al., 2017). We therefore use the rich information on preferences contained in students’ rank-ordered choice lists to correct our estimates for selection on unobservables. This selection correction is implemented by extending the classic multinomial logit control function estimator of Dubin and McFadden (1984) to a setting where rankings of multiple alternatives are known.

The final step of our analysis relates the choice model and treatment effect estimates to measure preferences for school effectiveness. The choice and outcome models we estimate allow preferences and causal effects to vary flexibly with student characteristics. Our specifications accommodate the possibility that schools are more effective for specific types of students and that applicants choose schools that are a good match for their student type. We compare the degree to which parent preferences are explained by overall school effectiveness, match quality, and peer quality, defined as the component of a school’s average outcome due to selection rather than effectiveness.

We find preferences are positively correlated with both peer quality and causal effects on student outcomes. More effective schools enroll higher-ability students, however, and preferences are unrelated to school effectiveness after controlling for peer quality. We also find little evidence of selection on match effects: on balance, parents do not prefer schools that are especially effective for their own children, and students do not enroll in schools that are a better-than-average match. These patterns are similar for short-run achievement test scores and longer-run postsecondary outcomes. Looking across demographic and baseline achievement groups, we find no evidence that any subgroup places positive weight on school effectiveness once we adjust for peer quality.

These findings do not indicate that parents choose schools irrationally; they may use peer characteristics to proxy for school effectiveness if the latter is difficult to observe, or value peer quality independently of impacts on academic outcomes. Regardless of the mechanism, however, our results imply that parents’ choices penalize schools that enroll low achievers rather than schools that offer poor instruction. As a result, school choice programs may generate stronger incentives for screening and selection than for improved academic quality. We provide suggestive evidence that schools have responded to these incentives by increasing screening since the introduction of centralized assignment in New York City.

Our analysis complements Rothstein’s (2006) indirect test with a direct assessment of the relationships among parent preferences, peer quality, and school effectiveness based on unusually rich choice and outcome data. The results also contribute to a large literature studying preferences for school quality (Black, 1999;...
Figlio and Lucas, 2004; Bayer et al., 2007; Hastings and Weinstein, 2008; Burgess et al., 2014; Imberman and Lovenheim, 2016). These studies show that housing prices and household choices respond to school performance levels, but they do not typically separate responses to causal school effectiveness and peer quality. Our findings are also relevant to theoretical and empirical research on the implications of school choice for sorting and stratification (Epple and Romano, 1998; Epple et al., 2004; Hsieh and Urquiola, 2006; Barseghyan et al., 2014; Altonji et al., 2015; Avery and Pathak, 2015; MacLeod and Urquiola, 2015; MacLeod et al., 2017). In addition, our results help to reconcile some surprising findings from recent studies of school choice. Cullen et al. (2006) find limited achievement effects of admission to preferred schools in Chicago, while Walters (forthcoming) documents that disadvantaged students in Boston are less likely to apply to charter schools than more advantaged students despite experiencing larger achievement benefits. Angrist et al. (2013) and Abdulkadiroğlu et al. (2017b) report on two settings where parents opt for schools that reduce student achievement. These patterns are consistent with our finding that school choices are not driven by school effectiveness.

Finally, our analysis adds to a recent series of studies leveraging preference data from centralized school assignment mechanisms to investigate school demand (Hastings et al., 2009; Harris and Larsen, 2014; Fack et al., 2015; Abdulkadiroğlu et al., 2017a; Glazerman and Dotter, 2016; Kapor et al., 2017; Agarwal and Somaini, forthcoming). Some of these studies analyze assignment mechanisms that provide incentives to strategically misreport preferences, while others measure academic quality using average test scores rather than distinguishing between peer quality and school effectiveness or looking at longer-run outcomes. We build on this previous work by using data from a strategy-proof mechanism to separately estimate preferences for peer quality and causal effects on multiple measures of academic success.

The rest of the paper is organized as follows. The next section describes school choice in New York City and the data used for our analysis. Section 3 develops a conceptual framework for analyzing school effectiveness and peer quality, and Section 4 details our empirical approach. Section 5 summarizes estimated distributions of student preferences and school treatment effects. Section 6 links preferences to peer quality and school effectiveness, and Section 7 discusses implications of these relationships. Section 8 concludes and offers some directions for future research.

2 Setting and Data

2.1 New York City High Schools

The New York City public school district annually enrolls roughly 90,000 ninth graders at more than 400 high schools. Rising ninth graders planning to attend New York City’s public high schools submit applications to the centralized assignment system. Before 2003 the district used an uncoordinated school assignment process in which students could receive offers from more than one school. Motivated in part
by insights derived from the theory of market design, in 2003 the city adopted a coordinated single-offer assignment mechanism based on the student-proposing deferred acceptance (DA) algorithm (Gale and Shapley, 1962; Abdulkadiroğlu et al., 2005, 2009). Abdulkadiroğlu et al. (2017a) show that introducing coordinated assignment reduced the share of administratively assigned students and likely improved average household welfare.

Applicants report their preferences for schooling options to the assignment mechanism by submitting rank-ordered lists of up to 12 academic programs. An individual school may operate more than one program. To aid families in their decisionmaking the New York City Department of Education (DOE) distributes a directory that provides an overview of the high school admission process, key dates, and an information page for each high school. A school’s information page includes a brief statement of its mission, a list of offered programs, courses and extracurricular activities, pass rates on New York Regents standardized tests, and the school’s graduation rate (New York City Department of Education, 2003). DOE also issues annual schools reports that list basic demographics, teacher characteristics, school expenditures, and Regents performance levels. During the time period of our study (2003-2007) these reports did not include measures of test score growth, though such measures have been added more recently (New York City Department of Education, 2004, 2017).

Academic programs prioritize applicants in the centralized admission system using a mix of factors. Priorities depend on whether a program is classified as unscreened, screened, or an educational option program. Unscreened programs give priority to students based on residential zones and (in some cases) to those who attend an information session. Screened programs use these factors and may also assign priorities based on prior grades, standardized test scores, and attendance. Educational option programs use screened criteria for some of their seats and unscreened criteria for the rest. Random numbers are used to order applicants with equal priority. A small group of selective high schools, including New York City’s exam schools, admit students in a parallel system outside the main round of the assignment process (Abdulkadiroğlu et al., 2014).

The DA algorithm combines student preferences with program priorities to generate a single program assignment for each student. In the initial step of the algorithm, each student proposes to her first-choice program. Programs provisionally accept students in order of priority up to capacity and reject the rest. In subsequent rounds, each student rejected in the previous step proposes to her most-preferred program among those that have not previously rejected her, and programs reject provisionally accepted applicants in favor of new applicants with higher priority. This process iterates until all students are assigned to a program or all unassigned students have been rejected by every program they have ranked. During our study time period, students left unassigned in the main round participate in a supplementary DA round in which they rank up to 12 additional programs with available seats. Any remaining students are administratively assigned by the district. About 82 percent, 8 percent, and 10 percent of applicants are assigned in the main, supplementary, and administrative rounds, respectively (Abdulkadiroğlu et al., 2017a).
An attractive theoretical property of the DA mechanism is that it is strategy-proof: since high-priority students can displace those with lower priority in later rounds of the process, listing schools in order of true preferences is a dominant strategy in the mechanism’s canonical version. This property, however, requires students to have the option to rank all schools (Haeringer and Klijn, 2009; Pathak and Sönmez, 2013). As we show below, more than 70 percent of students rank fewer than 12 programs, meaning that truthful ranking of schools is a dominant strategy for the majority of applicants. The instructions provided with the New York City high school application also directly instruct students to rank schools in order of their true preferences (New York City Department of Education, 2003). In the analysis to follow, we therefore interpret students’ rank-ordered lists as truthful reports of their preferences. We also probe the robustness of our findings to violations of this assumption by reporting results based on students that rank fewer than 12 choices.2

2.2 Data and Samples

The data used here are extracted from a DOE administrative information system covering all students enrolled in New York City public schools between the 2003-2004 and 2012-2013 school years. These data include school enrollment, student demographics, home addresses, scores on New York Regents standardized tests, Preliminary SAT (PSAT) scores, and high school graduation records, along with preferences submitted to the centralized high school assignment mechanism. A supplemental file from the National Student Clearinghouse (NSC) reports college enrollment for students graduating from New York City high schools between 2009 and 2012. A unique student identifier links records across these files.

We analyze high school applications and outcomes for four cohorts of students enrolled in New York City public schools in eighth grade between 2003-2004 and 2006-2007. This set of students is used to construct several samples for statistical analysis. The choice sample, used to investigate preferences for schools, consists of all high school applicants with baseline (eighth grade) demographic, test score, and address information. Our analysis of school effectiveness uses subsamples of the choice sample corresponding to each outcome of interest. These outcome samples include students with observed outcomes, baseline scores, demographics, and addresses, enrolled for ninth grade at one of 316 schools with at least 50 students for each outcome. The outcome samples also exclude students enrolled at the nine selective high schools that do not admit students via the main DA mechanism. Appendix A and Appendix Table A1 provide further details on data sources and sample construction.

Key outcomes in our analysis include Regents math standardized test scores, PSAT scores, high school graduation, college attendance, and college quality. The high school graduation outcome equals one if a student graduates within five years of her projected high school entry date given her eighth grade cohort.

Along similar lines, Abdulkadiroğlu et al. (2017a) show that preference estimates using only the top ranked school, the top three schools, and all but the last ranked school are similar.
Likewise, college attendance equals one for students who enroll in any college (two or four year) within two years of projected on-time high school graduation. The college quality variable, derived from Internal Revenue Service tax record statistics reported by Chetty et al. (2017b), equals the mean 2014 income for children born between 1980 and 1982 who attended a student’s college. The mean income for the non-college population is assigned to students who do not enroll in a college. While this metric does not distinguish between student quality and causal college effectiveness, it provides an accurate measure of the selectivity of a student’s college. It has also been used elsewhere to assess effects of education programs on the intensive margin of college attendance (Chetty et al., 2011, 2014b). College attendance and quality are unavailable for the 2003-2004 cohort because the NSC data window does not allow us to determine whether students in this cohort were enrolled in college within two years of projected high school graduation.

Descriptive statistics for the choice and outcome samples appear in Table 1. These statistics show that New York City schools serve a disadvantaged urban population. Seventy-three percent of students are black or hispanic, and 65 percent are eligible for a subsidized lunch. Data from the 2011-2015 American Community Surveys shows that the average student in the choice sample lives in a census tract with a median household income of $50,136 in 2015 dollars. Observed characteristics are generally similar for students in the choice and outcome samples. The average PSAT score in New York City is 116, about one standard deviation below the US average (the PSAT is measured on a 240 point scale, normed to have a mean of 150 and a standard deviation of 30). The five-year high school graduation rate is 61 percent, and 48 percent of students attend some college within two years of graduation.

2.3 Choice Lists

New York City high school applicants tend to prefer schools near their homes, and most do not fill their choice lists. These facts are shown in Table 2, which summarizes rank-ordered preference lists in the choice sample. As shown in column (1), 93 percent of applicants submit a second choice, about half submit eight or more choices, and 28 percent submit the maximum 12 allowed choices. Column (2) shows that students prefer schools located in their home boroughs: 85 percent of first-choice schools are in the same borough as the student’s home address, and the fraction of other choices in the home borough are also high. Abdulkadiroğlu et al. (2017a) report that for 2003-04, 193 programs restricted eligibility to applicants who reside in the same borough. The preference analysis to follow, therefore, treats schools in a student’s home borough as her choice set and aggregates schools in other boroughs into a single outside option. Column (3), which reports average distances (measured as great-circle distance in miles) for each choice restricted to schools in the home borough, shows that students rank nearby schools higher within boroughs as well.

Applicants also prefer schools with strong academic performance. The last column of Table 2 reports the average Regents high school math score for schools at each position on the rank list. Regents scores are normalized to have mean zero and standard deviation one in the New York City population. To earn a
high school diploma in New York state, students must pass a Regents math exam. These results reveal that higher-ranked schools enroll students with better math scores. The average score at a first-choice school is 0.2 standard deviations ($\sigma$) above the city average, and average scores monotonically decline with rank. PSAT, graduation, college enrollment, and college quality indicators also decline with rank. Students and parents clearly prefer schools with high achievement levels. Our objective in the remainder of this paper is to decompose this pattern into components due to preferences for school effectiveness and peer quality.

3 Conceptual Framework

Consider a population of students indexed by $i$, each of whom attends one of $J$ schools. Let $Y_{ij}$ denote the potential value of some outcome of interest for student $i$ if she attends school $j$. The projection of $Y_{ij}$ on a vector of observed characteristics, $X_i$, is written:

$$Y_{ij} = \alpha_j + X_i' \beta_j + \epsilon_{ij},$$

where $E[\epsilon_{ij}] = E[X_i \epsilon_{ij}] = 0$ by definition of $\alpha_j$ and $\beta_j$. The coefficient vector $\beta_j$ measures the returns to observed student characteristics at school $j$, while $\epsilon_{ij}$ reflects variation in potential outcomes unexplained by these characteristics. We further normalize $E[X_i] = 0$, so $\alpha_j = E[Y_{ij}]$ is the population mean potential outcome at school $j$. The realized outcome for student $i$ is $Y_i = \sum_j 1\{S_i = j\} Y_{ij}$, where $S_i \in \{1...J\}$ denotes school attendance.

We decompose potential outcomes into components explained by student ability, school effectiveness, and idiosyncratic factors. Let $A_i \equiv (1/J) \sum_j Y_{ij}$ denote student $i$'s general ability, defined as the average of her potential outcomes across all schools. This variable describes how the student would perform at the average school. Adding and subtracting $A_i$ on the right-hand side of (1) yields:

$$Y_{ij} = \bar{\alpha} + X_i' \bar{\beta} + \check{\epsilon}_i + (\alpha_j - \bar{\alpha}) + X_i'(\beta_j - \bar{\beta}) + (\epsilon_{ij} - \bar{\epsilon}_i),$$

where $\bar{\alpha} = (1/J) \sum_j \alpha_j$, $\bar{\beta} = (1/J) \sum_j \beta_j$, and $\check{\epsilon}_i = (1/J) \sum_j \epsilon_{ij}$. Equation (2) shows that student $i$'s potential outcome at school $j$ is the sum of three terms: the student’s general ability, $A_i$; the school’s average treatment effect, $ATE_j$, defined as the causal effect of school $j$ relative to an average school for an average student; and a match effect, $M_{ij}$, which reflects student $i$'s idiosyncratic suitability for school $j$. Match effects may arise either because of an interaction between student $i$'s observed characteristics and the extra returns to characteristics at school $j$ (captured by $X_i'(\beta_j - \bar{\beta})$) or because of unobserved factors that make student $i$ more or less suitable for school $j$ (captured by $\epsilon_{ij} - \bar{\epsilon}_i$).

This decomposition allows us to interpret variation in observed outcomes across schools using three
terms. The average outcome at school $j$ is given by:

$$E[Y_i|S_i = j] = Q_j + ATE_j + E[M_{ij}|S_i = j].$$

(3)

Here $Q_j = E[A_i|S_i = j]$ is the average ability of students enrolled at school $j$, a variable we label “peer quality.” The quantity $E[M_{ij}|S_i = j]$ is the average suitability of $j$’s students for this particular school. In a Roy (1951)-style model in which students sort into schools on the basis of comparative advantage in the production of $Y_i$, we would expect this average match effect to be positive for all schools. Parents and students may also choose schools on the basis of peer quality $Q_j$, overall school effectiveness $ATE_j$, or the idiosyncratic match $M_{ij}$ for various outcomes.

## 4 Empirical Methods

The goal of our empirical analysis is to assess the roles of peer quality, school effectiveness, and academic match quality in applicant preferences. Our analysis proceeds in three steps. We first use rank-ordered choice lists to estimate preferences, thereby generating measures of each school’s popularity. Next, we estimate schools’ causal effects on test scores, high school graduation, college attendance, and college choice. Finally, we combine these two sets of estimates to characterize the relationships among school popularity, peer quality, and causal effectiveness.

### 4.1 Estimating Preferences

Let $U_{ij}$ denote student $i$’s utility from enrolling in school $j$, and let $\mathcal{J} = \{1...J\}$ represent the set of available schools. We abstract from the fact that students rank programs rather than schools by ignoring repeat occurrences of any individual school on a student’s choice list. $U_{ij}$ may therefore be interpreted as the indirect utility associated with student $i$’s favorite program at school $j$. The school ranked first on a student’s choice list is

$$R_{i1} = \arg \max_{j \in \mathcal{J}} U_{ij},$$

while subsequent ranks satisfy

$$R_{ik} = \arg \max_{j \in \mathcal{J} \setminus \{R_{im}, m < k\}} U_{ij}, \; k > 1.$$  

Student $i$’s rank-order list is then $R_i = (R_{i1}...R_{i\ell(i)})'$, where $\ell(i)$ is the length of the list submitted by this student.

We summarize these preference lists by fitting random utility models with parameters that vary according to observed student characteristics. Student $i$’s utility from enrolling in school $j$ is modeled as:

$$U_{ij} = \delta_c(X_{i,j}) - \tau_c(X_{i,j})D_{ij} + \eta_{ij},$$

(4)
where the function \( c(X_i) \) assigns students to covariate cells based on the variables in the vector \( X_i \), and \( D_{ij} \) records distance from student \( i \)'s home address to school \( j \). The parameter \( \delta_{cj} \) is the mean utility of school \( j \) for students in covariate cell \( c \), and \( \tau_c \) is a cell-specific distance parameter or "cost." We include distance in the model because a large body of evidence suggests it plays a central role in school choices (e.g., Hastings et al., 2009 and Abdulkadiroğlu et al., 2017a). We model unobserved tastes \( \eta_{ij} \) as following independent extreme value type I distributions conditional on \( X_i \) and \( D_i = (D_{i1} \ldots D_{iJ})' \). Equation (4) is therefore a rank-ordered multinomial logit model (Hausman and Ruud, 1987).

The logit model implies the conditional likelihood of the rank list \( R_i \) is:

\[
L(R_i|X_i, D_i) = \prod_{k=1}^{\ell(i)} \frac{\exp (\delta_{c(X_i)R_{ik}} - \tau_{c(X_i)}D_{iR_{ik}})}{\sum_{j \in J \setminus \{R_{im} : m < k\}} \exp (\delta_{c(X_i)j} - \tau_{c(X_i)}D_{ij})}.
\]

We allow flexible heterogeneity in tastes by estimating preference models separately for 360 covariate cells defined by the intersection of borough, sex, race (black, hispanic, or other), subsidized lunch status, above-median census tract income, and terciles of the mean of eighth grade math and reading scores. This specification follows several recent studies that flexibly parametrize preference heterogeneity in terms of observable characteristics (e.g., Hastings et al., 2017 and Langer, 2016). Students rarely rank schools outside their home boroughs, so covariate cells often include zero students ranking any given out-of-borough school. We therefore restrict the choice set \( J \) to schools located in the home borough and aggregate all other schools into an outside option with utility normalized to zero. Maximum likelihood estimation of the preference parameters produces a list of school mean utilities along with a distance coefficient for each covariate cell.

### 4.2 Estimating School Effectiveness

Our analysis of school effectiveness aims to recover the parameters of the potential outcome equations defined in Section 3. We take two approaches to estimating these parameters.

**Approach 1: Selection on observables**

The first set of estimates is based on the assumption:

\[
E[Y_{ij}|X_i, S_i] = \alpha_j + X_i' \beta_j, \quad j = 1 \ldots J.
\]

This restriction, often labeled "selection on observables,” requires school enrollment to be as good as random conditional on the covariate vector \( X_i \), which includes sex, race, subsidized lunch status, the log of median census tract income, and eighth grade math and reading scores. Assumption (5) implies that an ordinary least squares (OLS) regression of \( Y_i \) on school indicators interacted with \( X_i \) recovers unbiased estimates of \( \alpha_j \) and \( \beta_j \) for each school. This fully interacted specification is a multiple-treatment extension
of the Oaxaca-Blinder (1973) treatment effects estimator (Kline, 2011).\footnote{We also include main effects of borough so that the model includes the same variables used to define covariate cells in the preference estimates.} By allowing school effectiveness to vary with student characteristics, we generalize the constant effects “value-added” approach commonly used to estimate the contributions of teachers and schools to student achievement (Koedel et al., 2015).

The credibility of the selection on observables assumption underlying value-added estimators is a matter of continuing debate (Rothstein, 2010, 2017; Kane et al., 2013; Baicher-Hicks et al., 2014; Chetty et al., 2014a, 2016, 2017a; Guarino et al., 2015). Comparisons to results from admission lotteries indicate that school value-added models accurately predict the impacts of random assignment but are not perfectly unbiased (Deming, 2014; Angrist et al., 2016b, 2017). Selection on observables may also be more plausible for test scores than for longer-run outcomes, for which lagged measures of the dependent variable are not available (Chetty et al., 2014a). We therefore report OLS estimates as a benchmark and compare these to estimates from a more general strategy that relaxes assumption (5).

**Approach 2: Rank-ordered control functions**

Our second approach is motivated by the restriction:

\[
E[Y_{ij}|X_i, D_i, \eta_{i1}, \ldots, \eta_{iJ}, S_i] = \alpha_j + X_i' \beta_j + g_j(D_i, \eta_{i1}, \ldots, \eta_{iJ}), \quad j = 1, \ldots, J. \tag{6}
\]

This restriction implies that any omitted variable bias afflicting OLS value-added estimates is due either to spatial heterogeneity captured by distances to each school \((D_i)\) or to the preferences underlying the rank-ordered lists submitted to the assignment mechanism \((\eta_{ij})\). The function \(g_j(\cdot)\) allows potential outcomes to vary arbitrarily across students with different preferences over schools. Factors that lead students with the same observed characteristics, spatial locations, and preferences to ultimately enroll in different schools, such as school priorities, random rationing due to oversubscription, or noncompliance with the assignment mechanism, are presumed to be unrelated to potential outcomes.

Under assumption (6), comparisons of matched sets of students with the same covariates, values of distance, and rank-ordered choice lists recover causal effects of school attendance. This model is therefore similar to the “self-revelation” model proposed by Dale and Krueger (2002; 2014) in the context of postsecondary enrollment. Dale and Krueger assume that students reveal their unobserved “types” via the selectivity of their college application portfolios, so college enrollment is as good as random among students that apply to the same schools. Similarly, (6) implies that high school applicants reveal their types through the content of their rank-ordered preference lists.

Though intuitively appealing, full nonparametric matching on rank-ordered lists is not feasible in practice because few students share the exact same rankings. We therefore use the structure of the logit choice model in equation (4) to derive a parametric approximation to this matching procedure. Specifically, we
replace equation (6) with the assumption:

$$E[Y_{ij} | X_i, D_i, \eta_1, ..., \eta_J, S_i] = \alpha_j + X'_i \beta_j + D'_i \gamma + \sum_{k=1}^{J} \psi_k \times (\eta_{ik} - \mu_\eta) + \varphi \times (\eta_{ij} - \mu_\eta), \ j = 1...J,$$

where $\mu_\eta = E[\eta_{ij}]$ is Euler’s constant.\(^4\)

As in the multinomial logit selection model of Dubin and McFadden (1984), equation (7) imposes a linear relationship between potential outcomes and the unobserved logit errors. Functional form assumptions of this sort are common in multinomial selection models with many alternatives, where requirements for nonparametric identification are very stringent (Lee, 1983; Dahl, 2002; Heckman et al., 2008).\(^5\)

Equation (7) accommodates a variety of forms of selection on unobservables. The coefficient $\psi_k$ represents an effect of the preference for school $k$ common to all potential outcomes. This permits students with strong preferences for particular schools to have higher or lower general ability $A_i$. The parameter $\varphi$ captures an additional match effect of the preference for school $j$ on the potential outcome at this specific school. The model therefore allows for “essential” heterogeneity linking preferences to unobserved match effects in student outcomes (Heckman et al., 2006). A Roy (1951)-style model of selection on gains would imply $\varphi > 0$, but we do not impose this restriction.

By iterated expectations, equation (7) implies that mean observed outcomes at school $j$ are:

$$E[Y_i | X_i, D_i, R_i, S_i = j] = \alpha_j + X'_i \beta_j + D'_i \gamma + \sum_{k=1}^{J} \psi_k \lambda_k (X_i, D_i, R_i) + \varphi \lambda_j (X_i, D_i, R_i),$$

where $\lambda_k (X_i, D_i, R_i) = E[\eta_{ik} - \mu_\eta | X_i, D_i, R_i]$ gives the mean preference for school $k$ conditional on a student’s characteristics, spatial location, and preference list. The $\lambda_k (\cdot)$’s serve as “control functions” correcting for selection on unobservables (Heckman and Robb, 1985; Blundell and Matzkin, 2014; Wooldridge, 2015). As shown in Appendix B.1, these functions are generalizations of the formulas derived by Dubin and McFadden (1984), extended to account for the fact that we observe a list of several ranked alternatives rather than just the most preferred choice.

Note that equation (8) includes main effects of distance to each school; we do not impose an exclusion restriction for distance. Identification of the selection parameters $\psi_k$ and $\varphi$ comes from variation in preference rankings for students who enroll at the same school conditional on covariates and distance. Intuitively, if students who rank school $j$ highly do better than expected given their observed characteristics at all schools, we will infer that $\psi_j > 0$. If these students do better than expected at school $j$ but not elsewhere, we will infer that $\varphi > 0$.

We use the choice model parameters to build first-step estimates of the control functions, then estimate equation (8) in a second-step OLS regression of $Y_i$ on school indicators and their interactions with $X_i$,

\(^4\)The means of both $X_i$ and $D_i$ are normalized to zero to maintain the interpretation that $\alpha_j = E[Y_{ij}]$.

\(^5\)As discussed in Section 6, we also estimate an alternative model that includes fixed effects for first choice schools.
controlling for \(D_i\) and the estimated \(\lambda(k)\) functions.\(^6\) We adjust inference for estimation error in the control functions via a two-step extension of the score bootstrap procedure of Kline and Santos (2012). As detailed in Appendix B.2, the score bootstrap avoids the need to recalculate the first-step logit estimates or the inverse variance matrix of the second-step regressors in the bootstrap iterations.

The joint distribution of peer quality and school effectiveness

Estimates of equations (5) and (7) may be used to calculate each school’s peer quality. A student’s predicted ability in the value-added model is

\[
\hat{A}_i = \frac{1}{J} \sum_{j=1}^{J} [\hat{\alpha}_j + X'_i \hat{\beta}_j],
\]

where \(\hat{\alpha}_j\) and \(\hat{\beta}_j\) are OLS value-added coefficients. Predicted ability in the control function model adds estimates of the distance and control function terms in equation (8). Estimated peer quality at school \(j\) is then

\[
\hat{Q}_j = \frac{1}{1 \{S_i = j\}} \sum_{i \{S_i = j\}} \frac{\hat{A}_i}{1 \{S_i = j\}},
\]

the average predicted ability of enrolled students.

The end result of our school quality estimation procedure is a vector of estimates for each school, \(\hat{\theta}_j = (\hat{\alpha}_j, \hat{\beta}_j', \hat{Q}_j)\). The vector of parameters for the control function model also includes an estimate of the selection coefficient for school \(j\), \(\hat{\psi}_j\). These estimates are unbiased but noisy measures of the underlying school-specific parameters \(\theta_j\). We investigate the distribution of \(\theta_j\) using the following hierarchical model:

\[
\hat{\theta}_j | \theta_j \sim N(\theta_j, \Omega_j),
\]

\[
\theta_j \sim N(\mu_\theta, \Sigma_\theta).
\]

Here \(\Omega_j\) is the sampling variance of the estimator \(\hat{\theta}_j\), while \(\mu_\theta\) and \(\Sigma_\theta\) govern the distribution of latent parameters across schools. In a hierarchical Bayesian framework \(\mu_\theta\) and \(\Sigma_\theta\) are hyperparameters describing a prior distribution for \(\theta_j\). We estimate these hyperparameters by maximum likelihood applied to model (10), approximating \(\Omega_j\) with an estimate of the asymptotic variance of \(\hat{\theta}_j\).\(^7\) The resulting estimates of \(\mu_\theta\) and \(\Sigma_\theta\) characterize the joint distribution of peer quality and school treatment effect parameters, purged of the estimation error in \(\hat{\theta}_j\).

This hierarchical model can also be used to improve estimates of parameters for individual schools. An empirical Bayes (EB) posterior mean for \(\theta_j\) is given by

\[
\theta_j^* = \left(\hat{\Omega}_j^{-1} + \hat{\Sigma}_\theta^{-1}\right)^{-1} \left(\hat{\Omega}_j^{-1} \hat{\theta}_j + \hat{\Sigma}_\theta^{-1} \hat{\mu}_\theta\right),
\]

where \(\hat{\Omega}_j\), \(\hat{\mu}_\theta\) and \(\hat{\Sigma}_\theta\) are estimates of \(\Omega_j\), \(\mu_\theta\) and \(\Sigma_\theta\). Relative to the unbiased but noisy estimate \(\hat{\theta}_j\), this EB shrinkage estimator uses the prior distribution to reduce sampling variance at the cost of increased

\(^6\)The choice model uses only preferences over schools in students’ home boroughs, so \(\lambda(k)\) is undefined for students outside school \(k\)’s borough. We therefore include dummies for missing values and code the control functions to zero for these students. We similarly code \(D_{k,i}\) to zero for students outside of school \(k\)’s borough and include borough indicators so that the distance coefficients are estimated using only within-borough variation. Our key results are not sensitive to dropping students attending out-of-borough schools from the sample.

\(^7\)The peer quality estimates \(\hat{Q}_j\) are typically very precise, so we treat peer quality as known rather than estimated when fitting the hierarchical model.
bias, yielding a minimum mean squared error (MSE) prediction of \( \theta_j \) (Robbins, 1956; Morris, 1983). This approach parallels recent work applying shrinkage methods to estimate causal effects of teachers, schools, neighborhoods, and hospitals (Chetty et al., 2014a; Hull, 2016; Angrist et al., 2017; Chetty and Hendren, 2017; Finkelstein et al., 2017). Appendix B.3 further describes our EB estimation strategy. In addition to reducing MSE, empirical Bayes shrinkage eliminates attenuation bias that would arise in models using elements of \( \hat{\theta}_j \) as regressors (Jacob and Lefgren, 2008). We exploit this property by regressing estimates of school popularity on EB posterior means in the final step of our empirical analysis.

### 4.3 Linking Preferences to School Effectiveness

We relate preferences to peer quality and causal effects with regressions of the form:

\[
\hat{\delta}_{cj} = \kappa_c + \rho_1 Q^*_j + \rho_2 ATE^*_j + \rho_3 M^*_cj + \xi_{cj},
\]

where \( \hat{\delta}_{cj} \) is an estimate of the mean utility of school \( j \) for students in covariate cell \( c \), \( \kappa_c \) is a cell fixed effect, and \( Q^*_j \) and \( ATE^*_j \) are EB posterior mean predictions of peer quality and average treatment effects. The variable \( M^*_cj \) is an EB prediction of the mean match effect of school \( j \) for students in cell \( c \). Observations in equation (11) are weighted by the inverse sampling variance of \( \hat{\delta}_{cj} \). We use the variance estimator proposed by Cameron et al. (2011) to double-cluster inference by cell and school. Two-way clustering accounts for correlated estimation errors in \( \hat{\delta}_{cj} \) across schools within a cell as well as unobserved determinants of popularity common to a given school across cells.

We estimate equation (11) separately for Regents test scores, PSAT scores, high school graduation, college attendance, and college quality. The parameters \( \rho_1 \), \( \rho_2 \), and \( \rho_3 \) measure how preferences relate to peer quality, overall school effectiveness, and match quality.

### 5 Parameter Estimates

#### 5.1 Preference Parameters

Table 3 summarizes the distribution of household preference parameters across the 316 high schools and 360 covariate cells in the choice sample. The first row reports estimated standard deviations of the mean utility \( \delta_{cj} \) across schools and cells, while the second row displays the mean and standard deviation of the cell-specific distance cost \( \tau_c \). School mean utilities are deviated from cell averages to account for differences in the reference category across boroughs, and calculations are weighted by cell size. We adjust these standard deviations for sampling error in the estimated preference parameters by subtracting the average squared standard error from the sample variance of mean utilities.

Consistent with the descriptive statistics in Table 1, the preference estimates indicate that households
dislike more distant schools. The mean distance cost is 0.33. This implies that increasing the distance to a
particular school by one mile reduces the odds that a household prefers this school to another in the same
borough by 33 percent. The standard deviation of the distance cost across covariate cells is 0.12. While
there is significant heterogeneity in distastes for distance, all of the estimated distance costs are positive,
suggesting that all subgroups prefer schools closer to home.

The estimates in Table 3 reveal significant heterogeneity in tastes for schools both within and between
subgroups. The within-cell standard deviation of school mean utilities, which measures the variation in \( \delta_{cj} \)
across schools \( j \) for a fixed cell \( c \), equals 1.12. This is equivalent to roughly 3.4 \((1.12/0.33)\) miles of distance,
implying that households are willing to travel substantial distances to attend more popular schools. The
between-cell standard deviation, which measures variation in \( \delta_{cj} \) across \( c \) for a fixed \( j \), is 0.50, equivalent to
about 1.5 \((0.50/0.33)\) miles of distance. The larger within-cell standard deviation indicates that students
in different subgroups tend to prefer the same schools.

5.2 School Effectiveness and Peer Quality

Our estimates of school treatment effects imply substantial variation in both causal effects and sorting
across schools. Table 4 reports estimated means and standard deviations of peer quality \( Q_j \), average
treatment effects \( ATE_j \), and slope coefficients \( \beta_j \). We normalize the means of \( Q_j \) and \( ATE_j \) to zero and
quantify the variation in these parameters relative to the average school. As shown in column (2), the
value-added model produces standard deviations of \( Q_j \) and \( ATE_j \) for Regents math scores equal to 0.29\( \sigma \).
This is somewhat larger than corresponding estimates of variation in school value-added from previous
studies (usually around 0.15 – 0.2\( \sigma \); see, e.g., Angrist et al., 2017). One possible reason for this difference
is that most students in our sample attend high school for two years before taking Regents math exams,
while previous studies look at impacts after one year.

As shown in columns (3) and (4) of Table 4, the control function model attributes some of the variation
in Regents math value-added parameters to selection bias. Adding controls for unobserved preferences and
distance increases the estimated standard deviation of \( Q_j \) to 0.31\( \sigma \) and reduces the estimated standard
deviation of \( ATE_j \) to 0.23\( \sigma \). Figure 1, which compares value-added and control function estimates for
all five outcomes, demonstrates that this pattern holds for other outcomes as well: adjusting for selection
on unobservables compresses the estimated distributions of treatment effects. This compression is more
severe for high school graduation, college attendance, and college quality than for Regents math and PSAT
scores. Our findings are therefore consistent with previous evidence that bias in OLS value-added models
is more important for longer-run and non-test score outcomes (see, e.g., Chetty et al., 2014b).

The bottom rows of Table 4 show evidence of substantial treatment effect heterogeneity across students.
For example, the standard deviation of the slope coefficient on a black indicator equals 0.12\( \sigma \) in the control
function model. This implies that holding the average treatment effect \( ATE_j \) fixed, a one standard deviation
improvement in a school’s match quality for black students boosts scores for these students by about a tenth of a standard deviation relative to whites. We also find significant variation in slope coefficients for gender (0.06σ), hispanic (0.11σ), subsidized lunch status (0.05σ), the log of median census tract income (0.05σ), and eighth grade math and reading scores (0.11σ and 0.05σ). The final row of column (3) reports a control function estimate of ϕ, the parameter capturing matching between unobserved preferences and Regents scores. This estimate indicates a positive relationship between preferences and the unobserved component of student-specific test score gains, but the magnitude of the coefficient is very small.8

Our estimates imply that high-ability students tend to enroll in more effective schools. Table 5 reports correlations between \( Q_j \) and school treatment effect parameters based on control function estimates for Regents math scores. Corresponding value-added estimates appear in Appendix Table A2. The estimated correlation between peer quality and average treatment effects is 0.59. This may reflect either positive peer effects or higher-achieving students’ tendency to enroll in schools with better inputs. Our finding that schools with high-ability peers are more effective contrasts with recent studies of exam schools in New York City and Boston, which show limited treatment effects for highly selective public schools (Abdulkadiroğlu et al., 2014; Dobbie and Fryer, 2014). Within the broader New York public high school system, we find a strong positive association between school effectiveness and average student ability.

Table 5 also reports estimated correlations of \( Q_j \) and \( ATE_j \) with the elements of the slope coefficient vector \( β_j \). Schools with larger average treatment effects tend to be especially good for girls: the correlation between \( ATE_j \) and the female slope coefficient is positive and statistically significant. This is consistent with evidence from Deming et al. (2014) showing that girls’ outcomes are more responsive to school value-added. We estimate a very high positive correlation between black and hispanic coefficients, suggesting that match effects tend to be similar for these two groups.

The slope coefficient on eighth grade reading scores is negatively correlated with peer quality and the average treatment effect. Both of these estimated correlations are below -0.4 and statistically significant. In other words, schools that enroll higher-ability students and produce larger achievement gains are especially effective at teaching low-achievers. In contrast to our estimate of the parameter ϕ, this suggests negative selection on the observed component of match effects in student achievement. Section 6 presents a more systematic investigation of this pattern by documenting the net relationship between preferences and treatment effects combining all student characteristics.

Patterns of estimates for PSAT scores, high school graduation, college attendance, and college quality are generally similar to results for Regents math scores. Appendix Tables A3-A6 present estimated distributions of peer quality and school effectiveness for these longer-run outcomes. For all five outcomes, we find substantial variation in peer quality and average treatment effects, a strong positive correlation between these variables, and significant effect heterogeneity with respect to student characteristics. Overall,

---

8The average predicted value of \( (η_{ij} - μ_η) \) for a student’s enrolled school in our sample is 2.0. Our estimate of ϕ therefore implies that unobserved match effects increase average test scores by about one percent of a standard deviation \((0.006σ \times 2.0 = 0.012σ)\).
causal effects for the longer-run outcomes are highly correlated with effects on Regents math scores. This is evident in Figure 2, which plots EB posterior mean predictions of average treatment effects on Regents scores against corresponding predictions for the other four outcomes. These results are consistent with recent evidence that short-run test score impacts reliably predict effects on longer-run outcomes (Chetty et al., 2011; Dynarski et al., 2013; Angrist et al., 2016a).

5.3 Decomposition of School Average Outcomes

We summarize the joint distribution of peer quality and school effectiveness by implementing the decomposition introduced in Section 3. Table 6 uses the control function estimates to decompose variation in school averages for each outcome into components explained by peer quality, school effectiveness, average match effects, and covariances of these components.

Consistent with the estimates in Table 4, both peer quality and school effectiveness play roles in generating variation in school average outcomes, but peer quality is generally more important. Peer quality explains 47 percent of the variance in average Regents scores (0.093/0.191), while average treatment effects explain 28 percent (0.054/0.191). The explanatory power of peer quality for other outcomes ranges from 49 percent (PSAT scores) to 83 percent (high school graduation), while the importance of average treatment effects ranges from 10 percent (PSAT scores) to 19 percent (log college quality).

Despite the significant variation in slope coefficients documented in Table 4, match effects are unimportant in explaining dispersion in school average outcomes. The variance of match effects accounts for only five percent of the variation in average Regents scores, and corresponding estimates for the other outcomes are also small. Although school treatment effects vary substantially across subgroups, there is not much sorting of students to schools on this basis, so the existence of potential match effects is of little consequence for realized variation in outcomes across schools.

The final three rows of Table 6 quantify the contributions of covariances among peer quality, treatment effects, and match effects. As a result of the positive relationship between peer quality and school effectiveness, the covariance between $Q_j$ and $ATE_j$ substantially increases cross-school dispersion in mean outcomes. The covariances between match effects and the other variance components are negative. This indicates that students at highly effective schools and schools with higher-ability students are less appropriately matched on the heterogeneous component of treatment effects, slightly reducing variation in school average outcomes.
6 Preferences, Peer Quality, and School Effectiveness

6.1 Productivity vs. Peers

The last step of our analysis compares the relative strength of peer quality and school effectiveness as predictors of parent preferences. Table 7 reports estimates of equation (11) for Regents math scores, first including $Q_j^*$ and $ATE_j^*$ one at a time and then including both variables simultaneously. Mean utilities, peer quality, and treatment effects are scaled in standard deviations of their respective school-level distributions, so the estimates can be interpreted as the standard deviation change in mean utility associated with a one standard deviation increase in $Q_j$ or $ATE_j$.

Bivariate regressions show that school popularity is positively correlated with both peer quality and school effectiveness. Results based on the OLS value-added model, reported in columns (1) and (2), imply that a one standard deviation increase in $Q_j$ is associated with a 0.42 standard deviation increase in mean utility, while a one standard deviation increase in $ATE_j$ is associated with a 0.24 standard deviation increase in mean utility. The latter result contrasts with studies reporting no average test score impact of attending preferred schools (Cullen et al., 2006; Hastings et al., 2009). These studies rely on admission lotteries that shift relatively small numbers of students across a limited range of schools. Our results show that looking across all high schools in New York City, more popular schools tend to be more effective on average.

While preferences are positively correlated with school effectiveness, however, this relationship is entirely explained by peer quality. Column (3) shows that when both variables are included together, the coefficient on peer quality is essentially unchanged, while the coefficient on the average treatment effect is rendered small and statistically insignificant. The $ATE_j$ coefficient also remains precise: we can rule out increases in mean utility on the order of 0.06 standard deviations associated with a one standard deviation change in school value-added at conventional significance levels. The control function estimates in columns (5)-(7) are similar to the value-added estimates; in fact, the control function results show a small, marginally statistically significant negative association between school effectiveness and popularity after controlling for peer quality.

Columns (4) and (8) of Table 7 explore the role of treatment effect heterogeneity by adding posterior mean predictions of match quality to equation (11), also scaled in standard deviation units of the distribution of match effects across schools and cells. The match coefficient is negative for both the value-added and control function models, and the control function estimate is statistically significant. This reflects the negative correlation between baseline test score slope coefficients and peer quality reported in Table 5: schools that are especially effective for low-achieving students tend to be more popular among high-achievers and therefore enroll more of these students despite their lower match quality. This is consistent with recent studies of selection into early-childhood programs and charter schools, which also find negative selection on test score match effects (Cornelissen et al., 2016; Kline and Walters, 2016; Walters, forthcoming).
Figure 3 presents a graphical summary of the links among preferences, peer quality, and treatment effects by plotting bivariate and multivariate relationships between mean utility (averaged across covariate cells) and posterior predictions of $Q_j$ and $ATE_j$ from the control function model. Panel A shows strong positive bivariate correlations for both variables. Panel B plots mean utilities against residuals from a regression of $Q_j^*$ on $ATE_j^*$ (left-hand panel) and residuals from a regression of $ATE_j^*$ on $Q_j^*$ (right-hand panel). Adjusting for school effectiveness has little effect on the relationship between preferences and peer quality. In contrast, partialing out peer quality eliminates the positive association between popularity and effectiveness.

6.2 Preferences and Effects on Longer-run Outcomes

Parents may care about treatment effects on outcomes other than short-run standardized test scores. We explore this by estimating equation (11) for PSAT scores, high school graduation, college attendance, and log college quality.

Results for these outcomes are similar to the findings for Regents math scores: preferences are positively correlated with average treatment effects in a bivariate sense but are uncorrelated with treatment effects conditional on peer quality. Table 8 reports results based on control function estimates of treatment effects. The magnitudes of all treatment effect coefficients are small, and the overall pattern of results suggests no systematic relationship between preferences and school effectiveness conditional on peer composition. We find a modest positive relationship between preferences and match effects for log college quality, but corresponding estimates for PSAT scores, high school graduation, and college attendance are small and statistically insignificant. This pattern contrasts with results for the Norwegian higher education system, reported by Kirkeboen et al. (2016), which show sorting into fields of study based on heterogeneous earnings gains. Unlike Norwegian college students, New York City’s high school students do not prefer schools with higher academic match quality.

6.3 Heterogeneity in Preferences for Peer and School Quality

Previous evidence suggests that parents of higher-income, higher-achieving students place more weight on academic performance levels when choosing schools (Hastings et al., 2009). This pattern may reflect either greater responsiveness to peer quality or more sensitivity to causal school effectiveness. If parents of high-achievers value school effectiveness, choice may indirectly create incentives for schools to improve because better instruction will attract high-ability students, raising peer quality and therefore demand from other households. In Table 9 we investigate this issue by estimating equation (11) separately by sex, race, subsidized lunch status, and baseline test score category.

We find that no subgroup of households responds to causal school effectiveness. Consistent with previous work, we find larger coefficients on peer quality among non-minority students, richer students (those
ineligible for subsidized lunches), and students with high baseline achievement. We do not interpret this as direct evidence of stronger preferences for peer ability among higher-ability students; since students are more likely to enroll at schools they rank highly, any group component to preferences will lead to a positive association between students’ rankings and the enrollment share of others in the same group.\footnote{This is a version of the “reflection problem” that plagues econometric investigations of peer effects (Manski, 1993).} The key pattern in Table 9 is that, among schools with similar peer quality, no group prefers schools with greater causal impacts on academic achievement.

**6.4 Alternative Specifications**

We investigate the robustness of our key results by estimating a variety of alternative specifications, reported in Appendix Tables A7-A9. To assess the sensitivity of our estimates to reasonable changes in our measure of school popularity, Appendix Table A7 displays results from models replacing $\hat{\delta}_{cj}$ in equation (11) with the log share of students in a cell ranking a school first or minus the log sum of ranks in the cell (treating unranked schools as tied). These alternative measures of demand produce very similar results to the rank-ordered logit results in Table 7.

Estimates based on students’ submitted rankings may not accurately describe demand if students strategically misreport their preferences in response to the 12-choice constraint on list length. As noted in Section 2, truthful reporting is a dominant strategy for the 72 percent of students that list fewer than 12 choices. Appendix Table A8 reports results based on rank-ordered logit models estimated in the subsample of unconstrained students. Results here are again similar to the full sample estimates, suggesting that strategic misreporting is not an important concern in our setting.

Equation (8) parameterizes the relationship between potential outcomes and preference rankings through the control functions $\lambda_k(\cdot)$. Columns (1)-(4) of Appendix Table A9 present an alternative parameterization that replaces the control functions with fixed effects for first choice schools. This approach ignores information on lower-ranked schools but more closely parallels the application portfolio matching approach in Dale and Krueger (2002; 2014). As a second alternative specification, columns (5)-(8) report estimates from a control function model that drops the distance control variables from equation (8). This model relies on an exclusion restriction for distance, a common identification strategy in the literature on educational choice (Card, 1995; Neal, 1997; Booker et al., 2011; Walters, forthcoming; Mountjoy, 2017). These alternative approaches to estimating school effectiveness produce no meaningful changes in the results.

**7 Discussion**

The findings reported here inform models of school choice commonly considered in the literature. Theoretical analyses often assume parents know students’ potential achievement outcomes and choose between schools on this basis. For example, Epple et al. (2004) and Epple and Romano (2008) study models in
which parents value academic achievement and consumption of other goods, and care about peer quality only insofar as it produces higher achievement through peer effects. Hoxby (2000) argues that school choice may increase achievement by allowing students to sort on match quality. Such models imply that demand should be positively correlated with both average treatment effects and match effects conditional on peer quality, a prediction that is inconsistent with the pattern in Table 7.

Parents may choose between schools based on test score levels rather than treatment effects. Cullen et al. (2006) suggest confusion between levels and gains may explain limited effects of admission to preferred schools in Chicago. Since our setting has substantial variation in both levels and value-added, we can more thoroughly investigate this model of parent decision-making. If parents choose between schools based on average outcomes, increases in these outcomes due to selection and causal effectiveness should produce equal effects on popularity. In contrast, we find that demand only responds to the component of average outcomes that is due to enrollment of higher-ability students. That is, we can reject the view that parental demand is driven by performance levels: demand places no weight on the part of performance levels explained by value-added but significant weight on the part explained by peer quality.

It is important to note that our findings do not imply parents are uninterested in school effectiveness. Without direct information about treatment effects, for example, parents may use peer characteristics as a proxy for school quality, as in MacLeod and Urquiola (2015). In view of the positive correlation between peer quality and school effectiveness, this is a reasonable strategy for parents that cannot observe treatment effects and wish to choose effective schools. Effectiveness varies widely conditional on peer quality, however, so parents make substantial sacrifices in academic quality by not ranking schools based on effectiveness. Table 10 compares Regents math effects for observed preference rankings vs. hypothetical rankings in which parents order schools according to their effectiveness. The average treatment effect of first-choice schools would improve from $0.07\sigma$ to $0.43\sigma$ if parents ranked schools based on effectiveness, and the average match effect would increase from $-0.04\sigma$ to $0.16\sigma$. This implies that the average student loses more than half a standard deviation in math achievement by enrolling in her first-choice school rather than the most effective option.

The statistics in Table 10 suggest that if information frictions prevent parents from ranking schools based on effectiveness, providing information about school effectiveness could alter school choices considerably. These changes may be particularly valuable for disadvantaged students. As shown in Appendix Table A10, gaps in effectiveness between observed first-choice schools and achievement-maximizing choices are larger for students with lower baseline achievement. This is driven by the stronger relationship between peer quality and preferences for more-advantaged parents documented in Table 9. These results suggest reducing information barriers could lead to differential increases in school quality for disadvantaged students and reduce inequality in student achievement. On the other hand, the patterns documented here may also reflect parents’ valuation of school amenities other than academic effectiveness rather than a lack of information about treatment effects.
Regardless of why parents respond to peer quality rather than school effectiveness, our results have important implications for the incentive effects of school choice programs. Since parents only respond to the component of school average outcomes that can be predicted by the ability of enrolled students, our estimates imply a school wishing to boost its popularity must recruit better students; improving outcomes by increasing causal effectiveness for a fixed set of students will have no impact on parent demand. Our results therefore suggest that choice may create incentives for schools to invest in screening and selection.

The evolution of admissions criteria used at New York City’s high schools is consistent with the implication that schools have an increased incentive to screen applicants due to parents’ demand for high-ability peers. After the first year of the new assignment mechanism, several school programs eliminated all lottery-based admissions procedures and became entirely screened. In the 2003-04 high school brochure, 36.8 percent of programs are screened, and this fraction jumps to 40.3 percent two years later. The Beacon High School in Manhattan, for example, switched from a school where half of the seats were assigned via random lottery in 2003-04 to a screened school the following year, where admissions is based on test performance, an interview and a portfolio of essays. Leo Goldstein High School for Sciences in Brooklyn underwent a similar transition. Both high schools frequent lists of New York City’s best public high schools (Linge and Tanzer, 2016). Compared to the first years of the new system, there has also been growth in the number of limited unscreened programs, which use a lottery but also give priority to students who attend an open house or high school fair. Compared to unscreened programs, prioritizing applicants who attend an information session provides an ordeal that favors applicants with time and resources thus resulting in positive selection (Disare, 2017). The number of limited unscreened programs nearly doubled from 106 to 210 from 2005 to 2012 (Nathanson et al., 2013).

8 Conclusion

A central motivation for school choice programs is that parents’ choices generate demand-side pressure for improved school productivity. We investigate this possibility by comparing estimates of school popularity and treatment effects based on rank-ordered preference data for applicants to public high schools in New York City. Parents prefer schools that enroll higher-achieving peers. Conditional on peer quality, however, parents’ choices are unrelated to causal school effectiveness. Moreover, no subgroup of parents systematically responds to causal school effectiveness. We also find no relationship between preferences for schools and estimated match quality. This indicates that choice does not lead students to sort into schools on the basis of comparative advantage in academic achievement.

This pattern of findings has important implications for the expected effects of school choice programs. Our results on match quality suggest choice is unlikely to increase allocative efficiency. Our findings regarding peer quality and average treatment effects suggest choice may create incentives for increased screening rather than academic effectiveness. If parents respond to peer quality but not causal effects,
school’s easiest path to boosting its popularity is to improve the average ability of its student population. Since peer quality is a fixed resource, this creates the potential for socially costly zero-sum competition as schools invest in mechanisms to attract the best students. MacLeod and Urquiola (2015) argue that restricting a school’s ability to select pupils may promote efficiency when student choices are based on school reputation. The impact of school choice on effort devoted to screening is an important empirical question for future research.

While we have shown that parents do not choose schools based on causal effects for a variety of educational outcomes, we cannot rule out the possibility that preferences are determined by effects on unmeasured outcomes. Parents may be sensitive to school safety or other non-academic amenities, for example. Our analysis also does not address why parents put more weight on peer quality than on treatment effects. If parents rely on student composition as a proxy for effectiveness, coupling school choice with credible information on causal effects may strengthen incentives for improved productivity and weaken the association between preferences and peer ability. Distinguishing between true tastes for peer quality and information frictions is another challenge for future work.
Figure 1: Comparison of value-added and control function estimates of school average treatment effects

Notes: This figure plots school average treatment effect (ATE) estimates from value-added models against corresponding estimates from models including control functions that adjust for selection on unobservables. Value-added estimates come from regressions of outcomes on school indicators interacted with gender, race, subsidized lunch status, the log of census tract median income, and eighth grade math and reading scores. Control function models add distance to school and predicted unobserved tastes from the choice model. Points in the figure are empirical Bayes posterior means from models fit to the distribution of school-specific estimates. Dashed lines show the 45-degree line.
Figure 2: Relationships between effects on test scores and effects on long run outcomes

A. Regents math scores and PSAT scores
B. Regents math scores and high school graduation
C. Regents math scores and college attendance
D. Regents math scores and log college quality

Notes: This figure plots estimates of causal effects on Regents math scores against estimates of effects on longer-run outcomes. Treatment effects are empirical Bayes posterior mean estimates of school average treatment effects from control function models. Panel A plots the relationship between Regents math effects and effects on PSAT scores. Panels B, C, and D show corresponding results for high school graduation, college attendance, and log college quality.
Notes: This figure plots school mean utility estimates against estimates of peer quality and Regents math average treatment effects. Mean utilities are school average residuals from a regression of school-by-covariate cell mean utility estimates on cell indicators. Peer quality is defined as the average predicted Regents math score for enrolled students. Regents math effects are empirical Bayes posterior mean estimates of school average treatment effects from control function models. The left plot in Panel A displays the bivariate relationship between mean utility and per quality, while the right plot shows the bivariate relationship between mean utility and Regents math effects. The left plot in Panel B displays the relationship between mean utility and residuals from a regression of peer quality on Regents math effects, while the right plot shows the relationship between mean utility and residuals from a regression of Regents math effects on peer quality. Dashed lines are ordinary least squares regression lines.
<table>
<thead>
<tr>
<th></th>
<th>Choice sample</th>
<th>Regents math</th>
<th>PSAT</th>
<th>HS graduation</th>
<th>College</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>Female</td>
<td>0.497</td>
<td>0.518</td>
<td>0.532</td>
<td>0.500</td>
<td>0.500</td>
</tr>
<tr>
<td>Black</td>
<td>0.353</td>
<td>0.377</td>
<td>0.359</td>
<td>0.376</td>
<td>0.372</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.381</td>
<td>0.388</td>
<td>0.384</td>
<td>0.399</td>
<td>0.403</td>
</tr>
<tr>
<td>Subsidized lunch</td>
<td>0.654</td>
<td>0.674</td>
<td>0.667</td>
<td>0.680</td>
<td>0.700</td>
</tr>
<tr>
<td>Census tract median income</td>
<td>$50,136</td>
<td>$50,004</td>
<td>$49,993</td>
<td>$49,318</td>
<td>$49,243</td>
</tr>
<tr>
<td>Bronx</td>
<td>0.231</td>
<td>0.221</td>
<td>0.226</td>
<td>0.236</td>
<td>0.239</td>
</tr>
<tr>
<td>Brooklyn</td>
<td>0.327</td>
<td>0.317</td>
<td>0.335</td>
<td>0.339</td>
<td>0.333</td>
</tr>
<tr>
<td>Manhattan</td>
<td>0.118</td>
<td>0.118</td>
<td>0.119</td>
<td>0.116</td>
<td>0.116</td>
</tr>
<tr>
<td>Queens</td>
<td>0.259</td>
<td>0.281</td>
<td>0.255</td>
<td>0.250</td>
<td>0.253</td>
</tr>
<tr>
<td>Staten Island</td>
<td>0.065</td>
<td>0.063</td>
<td>0.064</td>
<td>0.059</td>
<td>0.059</td>
</tr>
<tr>
<td>Regents math score</td>
<td>0.000</td>
<td>-0.068</td>
<td>0.044</td>
<td>-0.068</td>
<td>-0.044</td>
</tr>
<tr>
<td>PSAT score</td>
<td>120</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>115</td>
</tr>
<tr>
<td>High school graduation</td>
<td>0.587</td>
<td>0.763</td>
<td>0.789</td>
<td>0.610</td>
<td>0.624</td>
</tr>
<tr>
<td>Attended college</td>
<td>0.463</td>
<td>0.588</td>
<td>0.616</td>
<td>0.478</td>
<td>0.478</td>
</tr>
<tr>
<td>College quality</td>
<td>$31,974</td>
<td>$33,934</td>
<td>$35,010</td>
<td>$31,454</td>
<td>$31,454</td>
</tr>
</tbody>
</table>

Notes: This table shows descriptive statistics for applicants to New York City public high schools between the 2003-2004 and 2006-2007 school years. Column (1) reports average characteristics and outcomes for all applicants with complete information on preferences, demographics, and eighth-grade test scores. Columns (2)-(5) display characteristics for the Regents math, PSAT, high school graduation, and college outcome samples. Outcome samples are restricted to students with data on the relevant outcome, enrolled in for ninth grade at schools with at least 50 students for each outcome. Regents math scores are normalized to mean zero and standard deviation one in the choice sample. High school graduation equals one for students who graduate from a New York City high school within five years of the end of their eighth grade year. College attendance equals one for students enrolled in any college within two years of projected high school graduation. College quality is the mean 2014 income for individuals in the 1980-1982 birth cohorts who attended a student's college. This variable equals the mean income in the non-college population for students who did not attend college. The college outcome sample excludes students in the 2003-2004 cohort. Census tract median income is median household income measured in 2015 dollars using data from the 2011-2015 American Community Surveys. Regents math, PSAT, graduation, and college outcome statistics exclude students with missing values.
### Table 2. Correlates of preference rankings for New York City high schools

<table>
<thead>
<tr>
<th>Choice</th>
<th>Fraction reporting</th>
<th>Same borough</th>
<th>Distance</th>
<th>Regents math score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choice 1</td>
<td>1.000</td>
<td>0.849</td>
<td>2.71</td>
<td>0.200</td>
</tr>
<tr>
<td>Choice 2</td>
<td>0.929</td>
<td>0.844</td>
<td>2.94</td>
<td>0.149</td>
</tr>
<tr>
<td>Choice 3</td>
<td>0.885</td>
<td>0.839</td>
<td>3.04</td>
<td>0.116</td>
</tr>
<tr>
<td>Choice 4</td>
<td>0.825</td>
<td>0.828</td>
<td>3.12</td>
<td>0.085</td>
</tr>
<tr>
<td>Choice 5</td>
<td>0.754</td>
<td>0.816</td>
<td>3.18</td>
<td>0.057</td>
</tr>
<tr>
<td>Choice 6</td>
<td>0.676</td>
<td>0.803</td>
<td>3.23</td>
<td>0.030</td>
</tr>
<tr>
<td>Choice 7</td>
<td>0.594</td>
<td>0.791</td>
<td>3.28</td>
<td>0.009</td>
</tr>
<tr>
<td>Choice 8</td>
<td>0.523</td>
<td>0.780</td>
<td>3.29</td>
<td>-0.013</td>
</tr>
<tr>
<td>Choice 9</td>
<td>0.458</td>
<td>0.775</td>
<td>3.31</td>
<td>-0.031</td>
</tr>
<tr>
<td>Choice 10</td>
<td>0.402</td>
<td>0.773</td>
<td>3.32</td>
<td>-0.051</td>
</tr>
<tr>
<td>Choice 11</td>
<td>0.345</td>
<td>0.774</td>
<td>3.26</td>
<td>-0.071</td>
</tr>
<tr>
<td>Choice 12</td>
<td>0.278</td>
<td>0.787</td>
<td>3.04</td>
<td>-0.107</td>
</tr>
</tbody>
</table>

Notes: This table reports average characteristics of New York City high schools by student preference rank. Column (1) displays fractions of student applications listing each choice. Column (2) reports the fraction of listed schools located in the same borough as a student's home address. Column (3) reports the mean distance between a student's home address and each ranked school, measured in miles. This column excludes schools outside the home borough. Column (4) shows average Regents math scores in standard deviation units relative to the New York City average.
<table>
<thead>
<tr>
<th></th>
<th>Mean (1)</th>
<th>Within cells (2)</th>
<th>Between cells (3)</th>
<th>Total (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>School mean utility</td>
<td>-</td>
<td>1.117</td>
<td>0.500</td>
<td>1.223</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.045)</td>
<td>(0.003)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>Distance cost</td>
<td>0.330</td>
<td>-</td>
<td>0.120</td>
<td>0.120</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
</tbody>
</table>

Number of students 270157
Number of schools 316
Number of covariate cells 360

Notes: This table summarizes variation in school value-added and utility parameters across schools and covariate cells. Utility estimates come from rank-ordered logit models fit to student preference rankings. These models include school indicators and distance to school and are estimated separately in covariate cells defined by borough, gender, race, subsidized lunch status, an indicator for above or below the median of census tract median income, and tercile of the average of eighth grade math and reading scores. Column (1) shows the mean of the distance coefficient across cells weighted by cell size. Column (2) shows the standard deviation of school mean utilities across schools within a cell, and column (3) shows the standard deviation of a given school's mean utility across cells. School mean utilities are deviated from cell averages to account for differences in the reference category across cells. Estimated standard deviations are adjusted for sampling error by subtracting the average squared standard error of the parameter estimates from the total variance.
**Table 4. Distributions of peer quality and treatment effect parameters for Regents math scores**

<table>
<thead>
<tr>
<th></th>
<th>Value-added model</th>
<th>Control function model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (1)</td>
<td>Std. dev. (2)</td>
</tr>
<tr>
<td>Peer quality</td>
<td>0</td>
<td>0.288</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>(0.012)</td>
</tr>
<tr>
<td>ATE</td>
<td>0</td>
<td>0.290</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Female</td>
<td>-0.048</td>
<td>0.062</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Black</td>
<td>-0.112</td>
<td>0.130</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.097</td>
<td>0.114</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Subsidized lunch</td>
<td>0.001</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Log census tract median income</td>
<td>0.020</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Eighth grade math score</td>
<td>0.622</td>
<td>0.105</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Eighth grade reading score</td>
<td>0.159</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Preference coefficient ($\psi_j$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Match coefficient ($\varphi$)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table reports estimated means and standard deviations of peer quality and school treatment effect parameters for Regents math scores. Peer quality is a school's average predicted test score given the characteristics of its students. The ATE is a school's average treatment effect, and other treatment effect parameters are school-specific interactions with student characteristics. Estimates come from maximum likelihood models fit to school-specific regression coefficients. Columns (1) and (2) report estimates from an OLS regression that includes interactions of school indicators with sex, race, subsidized lunch, the log of the median income in a student's census tract, and eighth grade reading and math scores. This model also includes main effects of borough. Columns (3) and (4) show estimates from a control function model that adds distance to each school and predicted unobserved preferences from the choice model. Control functions and distance variables are set to zero for out-of-borough schools and indicators for missing values are included.
Table 5. Correlations of peer quality and treatment effect parameters for Regents math scores

<table>
<thead>
<tr>
<th>Peer quality quality</th>
<th>ATE</th>
<th>Female</th>
<th>Black</th>
<th>Hispanic</th>
<th>Sub. lunch</th>
<th>Log tract inc.</th>
<th>Math score</th>
<th>Reading score</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
</tr>
<tr>
<td>ATE</td>
<td>0.588</td>
<td>(0.052)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.078</td>
<td>0.299</td>
<td>(0.078)</td>
<td>(0.101)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.006</td>
<td>0.107</td>
<td>-0.177</td>
<td>(0.077)</td>
<td>(0.106)</td>
<td>(0.142)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.013</td>
<td>0.115</td>
<td>-0.235</td>
<td>0.922</td>
<td>(0.080)</td>
<td>(0.112)</td>
<td>(0.150)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Subsidized lunch</td>
<td>0.045</td>
<td>-0.168</td>
<td>0.066</td>
<td>-0.038</td>
<td>0.004</td>
<td>(0.086)</td>
<td>(0.117)</td>
<td>(0.140)</td>
</tr>
<tr>
<td>Log census tract income</td>
<td>0.035</td>
<td>0.068</td>
<td>-0.010</td>
<td>-0.239</td>
<td>-0.045</td>
<td>-0.280</td>
<td>(0.099)</td>
<td>(0.134)</td>
</tr>
<tr>
<td>Eighth grade math score</td>
<td>-0.075</td>
<td>0.037</td>
<td>-0.074</td>
<td>-0.005</td>
<td>-0.007</td>
<td>0.060</td>
<td>0.027</td>
<td>(0.064)</td>
</tr>
<tr>
<td>Eighth grade reading score</td>
<td>-0.418</td>
<td>-0.452</td>
<td>-0.193</td>
<td>-0.090</td>
<td>-0.078</td>
<td>0.004</td>
<td>0.086</td>
<td>0.256</td>
</tr>
<tr>
<td>Preference coefficient ((\theta_j))</td>
<td>0.429</td>
<td>0.247</td>
<td>0.212</td>
<td>-0.083</td>
<td>-0.058</td>
<td>-0.127</td>
<td>0.316</td>
<td>-0.241</td>
</tr>
</tbody>
</table>

Notes: This table reports estimated correlations between peer quality and school treatment effect parameters for Regents math scores. The ATE is a school's average treatment effect, and other treatment effect parameters are school-specific interactions with student characteristics. Estimates come from maximum likelihood models fit to school-specific regression coefficients from a control function model controlling for observed characteristics, distance to school and unobserved tastes from the choice model.
<table>
<thead>
<tr>
<th></th>
<th>Regents math</th>
<th>PSAT score/10</th>
<th>High school graduation</th>
<th>College attendance</th>
<th>Log college quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total variance of average outcome</td>
<td>0.191</td>
<td>1.586</td>
<td>0.012</td>
<td>0.016</td>
<td>0.021</td>
</tr>
<tr>
<td>Variance of peer quality</td>
<td>0.093</td>
<td>0.781</td>
<td>0.010</td>
<td>0.010</td>
<td>0.009</td>
</tr>
<tr>
<td>Variance of ATE</td>
<td>0.054</td>
<td>0.160</td>
<td>0.002</td>
<td>0.003</td>
<td>0.004</td>
</tr>
<tr>
<td>Variance of match</td>
<td>0.008</td>
<td>0.027</td>
<td>0.002</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td>2Cov(peer quality, ATE)</td>
<td>0.081</td>
<td>0.745</td>
<td>0.005</td>
<td>0.008</td>
<td>0.011</td>
</tr>
<tr>
<td>2Cov(peer quality, match)</td>
<td>-0.023</td>
<td>-0.061</td>
<td>-0.003</td>
<td>-0.003</td>
<td>-0.002</td>
</tr>
<tr>
<td>2Cov(ATE, match)</td>
<td>-0.022</td>
<td>-0.068</td>
<td>-0.004</td>
<td>-0.005</td>
<td>-0.003</td>
</tr>
</tbody>
</table>

Notes: This table decomposes variation in average outcomes across schools into components explained by student characteristics, school average treatment effects (ATE), and the match between student characteristics and school effects. Estimates come from control function models adjusting for selection on unobservables. Column (1) shows results for Regents math scores in standard deviation units, column (2) reports estimates for PSAT scores, column (3) displays estimates for high school graduation, column (4) reports results for college attendance, and column (5) shows results for log college quality. The first row reports the total variance of average outcomes across schools. The second row reports the variance of peer quality, defined as the average predicted outcome as a function of student characteristics and unobserved tastes. The third row reports the variance of ATE, and the fourth row displays the variance of the match effect. The remaining rows show covariances of these components.
Table 7. Preferences for peer quality and Regents math effects

<table>
<thead>
<tr>
<th></th>
<th>Value-added models</th>
<th>Control function models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Peer quality</td>
<td>0.416</td>
<td>0.438</td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>ATE</td>
<td>0.244</td>
<td>-0.033</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>Match effect</td>
<td>-0.072</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>21684</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table reports estimates from regressions of school popularity on peer quality and school effectiveness. School popularity is measured as the estimated mean utility for each school and covariate cell in the choice model from Table 4. Covariate cells are defined by borough, gender, race, subsidized lunch status, an indicator for students above the median of census tract median income, and tercile of the average of eighth grade math and reading scores. Peer quality is constructed as the average predicted Regents math score for enrolled students. Treatment effect estimates are empirical Bayes posterior mean predictions of Regents math effects. Mean utilities, peer quality, and treatment effects are scaled in standard deviation units. Columns (1)-(4) report results from value-added models, while columns (5)-(8) report results from control function models. All regressions include cell indicators and weight by the inverse of the squared standard error of the mean utility estimates. Standard errors are double-clustered by school and covariate cell.
Table 8. Preferences for peer quality and school effectiveness by outcome

<table>
<thead>
<tr>
<th></th>
<th>PSAT score</th>
<th>High school graduation</th>
<th>College attendance</th>
<th>Log college quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Peer quality</td>
<td>0.467</td>
<td>0.430</td>
<td>0.235</td>
<td>0.322</td>
</tr>
<tr>
<td></td>
<td>(0.070)</td>
<td>(0.070)</td>
<td>(0.054)</td>
<td>(0.065)</td>
</tr>
<tr>
<td>ATE</td>
<td>0.325</td>
<td>-0.092</td>
<td>0.103</td>
<td>-0.174</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.074)</td>
<td>(0.045)</td>
<td>(0.054)</td>
</tr>
<tr>
<td>Match effect</td>
<td>-0.049</td>
<td>-0.065</td>
<td>-0.017</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.044)</td>
<td>(0.050)</td>
<td>(0.061)</td>
</tr>
</tbody>
</table>

N 21684

Notes: This table reports estimates from regressions of school popularity on peer quality and school effectiveness separately by outcome. School popularity is measured as the estimated mean utility for each school and covariate cell in the choice model from Table 4. Covariate cells are defined by borough, gender, race, subsidized lunch status, an indicator for students above the median of census tract median income, and tercile of the average of eighth grade math and reading scores. Peer quality is constructed as the average predicted outcome for enrolled students. Treatment effect estimates are empirical Bayes posterior mean predictions from control function models. Mean utilities, peer quality, and treatment effects are scaled in standard deviation units. All regressions include cell indicators and weight by the inverse of the squared standard error of the mean utility estimates. Standard errors are double-clustered by school and covariate cell.
Table 9. Heterogeneity in preferences for peer quality and Regents math effects

<table>
<thead>
<tr>
<th></th>
<th>Male (1)</th>
<th>Female (2)</th>
<th>Black (3)</th>
<th>Hispanic (4)</th>
<th>Other (5)</th>
<th>Eligible (6)</th>
<th>Ineligible (7)</th>
<th>Eligible (8)</th>
<th>Ineligible (9)</th>
<th>Lowest (10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer quality</td>
<td>0.432</td>
<td>0.441</td>
<td>0.396</td>
<td>0.370</td>
<td>0.705</td>
<td>0.410</td>
<td>0.501</td>
<td>0.251</td>
<td>0.395</td>
<td>0.686</td>
</tr>
<tr>
<td>ATE</td>
<td>-0.075</td>
<td>-0.021</td>
<td>-0.047</td>
<td>-0.011</td>
<td>-0.192</td>
<td>-0.036</td>
<td>-0.076</td>
<td>-0.015</td>
<td>-0.029</td>
<td>-0.117</td>
</tr>
<tr>
<td>Match effect</td>
<td>-0.177</td>
<td>-0.169</td>
<td>-0.200</td>
<td>-0.144</td>
<td>-0.149</td>
<td>-0.180</td>
<td>-0.155</td>
<td>-0.166</td>
<td>-0.169</td>
<td>-0.125</td>
</tr>
<tr>
<td>N</td>
<td>10795</td>
<td>10889</td>
<td>7467</td>
<td>7433</td>
<td>6784</td>
<td>11043</td>
<td>10641</td>
<td>7264</td>
<td>7286</td>
<td>7134</td>
</tr>
</tbody>
</table>

Notes: This table reports estimates from regressions of school popularity on peer quality and school effectiveness separately by student subgroup. School popularity is measured as the estimated mean utility for each school and covariate cell in the choice model from Table 4. Peer quality is constructed as the average predicted Regents math score for enrolled students. Treatment effect estimates are empirical Bayes posterior mean predictions of Regents math effects from control function models. Mean utilities, peer quality, and treatment effects are scaled in standard deviation units. Peer quality is constructed as the average predicted Regents math score for enrolled students. All regressions include cell indicators and weight by the inverse of the squared standard error of the mean utility estimates. Standard errors are double-clustered by school and covariate cell.
### Table 10. Potential achievement gains from ranking schools by effectiveness

<table>
<thead>
<tr>
<th>Observed rankings</th>
<th>Rankings based on effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer quality</td>
<td>Peer quality</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Choice 1</td>
<td>0.112</td>
</tr>
<tr>
<td>Choice 2</td>
<td>0.057</td>
</tr>
<tr>
<td>Choice 3</td>
<td>0.021</td>
</tr>
<tr>
<td>Choice 4</td>
<td>-0.013</td>
</tr>
<tr>
<td>Choice 5</td>
<td>-0.046</td>
</tr>
<tr>
<td>Choice 6</td>
<td>-0.074</td>
</tr>
<tr>
<td>Choice 7</td>
<td>-0.097</td>
</tr>
<tr>
<td>Choice 8</td>
<td>-0.114</td>
</tr>
<tr>
<td>Choice 9</td>
<td>-0.127</td>
</tr>
<tr>
<td>Choice 10</td>
<td>-0.139</td>
</tr>
<tr>
<td>Choice 11</td>
<td>-0.146</td>
</tr>
<tr>
<td>Choice 12</td>
<td>-0.156</td>
</tr>
</tbody>
</table>

Notes: This table summarizes Regents math score gains that parents could achieve by ranking schools based on effectiveness. Columns (1)-(3) report average peer quality, average treatment effects, and average match quality for students' observed preference rankings. Columns (4)-(6) display corresponding statistics for hypothetical rankings that list schools in order of their treatment effects. Treatment effect estimates come from control function models. All calculations are restricted to ranked schools within the home borough.
References


