Gains from Reassignment: Evidence from a Two-Sided Teacher Market *

Mariana Laverde
Boston College

Elton Mykerezi
University of Minnesota

Aaron Sojourner
W. E. Upjohn Institute

Aradhya Sood
University of Toronto

October 26, 2023

Abstract

Although the literature emphasizes the importance of efficiency based on agents' preferences, policymakers may hold different goals. For instance, school districts may focus on the mean student achievement and equity in student achievement across various groups. This paper studies both the potential efficiency and equity gains from within-district reassignment of teachers to classrooms using novel data that allows us to observe both decisions of teachers and school leaders in the internal-transfer process and test-score growth of students from observed assignments. We jointly model student achievement and teacher and school decisions to account for potential selection on test-score gains and to predict teacher effectiveness in unobserved matches. Teachers, but not schools, are averse to assignment based on the teachers' comparative advantage. However, estimates from counterfactual assignments of teachers to classrooms imply that under a constraint not to reduce any teacher's welfare, average student test scores could still rise by 7% of a standard deviation, most of the possible 8% gain absent the constraint. Although both high and low achievers experience average gains under this counterfactual, gains would be larger for high-achieving students.

*This paper benefited from discussions with Michael Dinerstein, Charlie Murry, Nikhil Agarwal, Minseon Park, Sukjoon Son, Yang Song, and Roman A. Zarate as well as the seminar participants at the the International Industrial Organization Conference, Southern Economic Association Conference, Society of Labor Economists Conference, and Western Economic Association Conference. Emails: laverdem@bc.com, myker001@umn.edu, sojourner@upjohn.org, and aradhya.sood@rotman.utoronto.ca.

1. Introduction

Mechanism design is frequently used to allocate limited resources, such as K-12 schools, college spots, doctors, human organs, and public housing. While the literature has emphasized the importance of efficiency based on agents' preferences, policymakers may want to balance agents' preferences with equity and various political economy issues. Like the burgeoning literature on outcome-based performance of assignment mechanisms Van Dijk (2019); Agarwal et al. (2020); Kapor et al. (2022), we study both the potential efficiency and equity gains in student achievement in the context of within-school district reassignment of teachers across schools and classrooms. How much can counterfactual teacher assignments under various constraints improve average student achievement and affect disparities across student racial and baseline-achievement groups? And how sharp are the efficiency-equity trade-offs?

The key challenge is measuring teacher effectiveness in unobserved assignments. Teacher effectiveness in observed matches might not represent counterfactual effectiveness elsewhere if teachers tend to apply to positions in which they have a comparative advantage or if schools tend to select teachers who do. Assuming exogenous mobility of teachers across schools is convenient and common but not necessarily credible. Instead, we jointly model student outcomes, teacher labor supply to schools, and school demand for teachers, allowing correlations between potential outcomes and decisions in the spirit of Roy (1951). Similarly to Agarwal et al.'s (2020) approach to the kidney-transplant market, we model student potential outcomes as a function of rich student, teacher, and match-specific observable and unobservable characteristics.

We leverage over 10 years of data from a large urban school district's internal transfer system (ITS), which governs the assignment of teachers to positions in the district. The set of open positions to which each teacher could apply, their choices of whether to apply, schools' choices to offer interviews and jobs to applicants, and applicants' decisions of whether to accept offers are observed. Knowing each teacher's and school's complete choice set and choices in each stage of the process (application, interview, offer, and acceptance) provides a rare opportunity to disentangle teacher preferences from the school hiring team's preferences.

We combine this data with longitudinal, individual-student-achievement data matched to the set of teachers in the same school, grade, year, and subject, modeling student achievement with a data structure similar to that of Hanushek and Rivkin (2010), Biasi (2021), and Biasi and Sarsons (2022).[1] The value-added model captures teacher effectiveness as the sum of a general-effectiveness component applying across all students, a rich set of interactions between observable teacher and student characteristics including race/ethnicity, sex, and prior achievement level (Condie et al., 2014; Delgado, 2022), and a set of teacher-by-school unobserved match components. Though many value-added models assume away the unobserved components, we follow Boyd et al. (2013), Mansfield (2015), Jackson (2013), and Aucejo et al. (2021) in including them.

We model teachers' and school hiring teams' decisions using a random-utility framework, capturing the potential correlation between student outcomes and these decisions via three channels. First, the student-achievement model includes interactions between observed teacher and student characteristics also present in the decision models. Second, school decisions may be correlated with each applicant's overall effectiveness. Last, teacher and school decisions depend on teacher-by-school idiosyncratic tastes, which may correlate with match effects in the student-outcome model.

Identification relies on assuming (1) conditionally independent assignment of teachers to students within, but not across, schools and (2) two shifters that separate supply and demand from outcomes.[2] We use driving time from a teacher's home to each school as the supply shifter. Driving time is correlated with teacher supply decisions and assumed to be independent of student potential outcomes and school demand.[3] For job

[1]Students and teachers are matched at the year-school-grade-subject, rather than classroom, level. Our value-added estimates based on this linking strongly correlate with teacher value-added estimates that the district created based on linking at the classroom level. Section 5.1 provides details.

[2]A valid demand shifter is an instrument that shifts demand (relevance) but not student outcomes, except through the induced demand shift (exclusions).

[3]Intuitively, to uncover the direction and magnitude of selection by teachers, the method contrasts the effectiveness of teachers living closer to their school, whose choices are assumed to be driven more by proximity rather than selection on effectiveness, against the effectiveness of teachers living farther from their school, whose choices are assumed to be more driven by selection on effectiveness. Finding the former less effective than the latter would be interpreted as positive selection.

applications, we use a school's share of same-race teacher peers as the demand shifter. School hiring teams are more likely to offer interviews and jobs to teachers when the school already has a higher share of teachers of the same race. The analysis assumes this variable does not affect student outcomes, after conditioning on the observable characteristics in the outcomes model. This approach follows applications in Agarwal et al. (2020), Geweke et al. (2003), Hull (2018), Kapor et al. (2022), and Van Dijk (2019) and ideas developed in Lewbel (2007), Heckman and Vytlacil (2005), and Heckman (1990). The distribution of the unobserved school-teacher match component of effectiveness is identified from within-teacher, across-school variation in effectiveness.

We estimate this three-equation model using Bayesian inference and a Gibb's sampler as in Geweke et al. (2003), Agarwal et al. (2020), and Kapor et al. (2022). Unlike maximum likelihood, this does not require numerical integration over a high-dimensional space nor maximization of a potentially nonconcave function. Our simulation recovers the joint posterior distribution of model parameters and latent variables. The Bayesian approach allows us to use the nested structure of the data to model teacher overall effectiveness and the unobserved teacher-school match effects using hierarchical priors. Hierarchical modeling allows information on student test scores to be shared among teachers. Teacher effectiveness and match effects are "shrunk" toward the distributions derived by the sampling model in proportion to the signal-to-noise ratio, analogously to Bayesian shrinkage of parameters in OLS estimation.

The model estimates that the standard deviation of teacher general effectiveness is 0.08, with match on observed teacher-student characteristics about an order of magnitude less important. Teachers tend to prefer schools with a higher share of high-income students, as in Boyd et al. (2011, 2013) and are averse to the schools where they would be most effective. Schools' demand decisions are uncorrelated with teacher education and experience, but schools do value teachers' general effectiveness.

Crucially, we find that the observed equilibrium assignment results in similar average student-achievement scores and share of proficient students as assigning teachers to positions at random would, though the observed assignment favors white students and high-achieving students relative to random assignment. We examine different counterfactual assignments of teachers across positions. A policy maker aiming to maximize average student-achievement scores who could compel any assignment could increase the average score by 8% of a standard deviation over scores in the observed equilibrium assignment. Gains would come mainly from matching more effective teachers to larger classrooms, rather than from leveraging comparative advantage.

While this reassignment would raise average achievement, more-advantaged students would benefit more, presenting an efficiency-equity trade-off. More specifically, although all groups experience increases in average test scores under this counterfactual, larger gains for higher-achieving students and white students lead to widening achievement gaps. Some of the implied assignments may make teachers worse off than under the status quo and so may not be feasible. For this reason, we consider a counterfactual reassignment that maximizes average scores conditional on not reducing assigned teachers' modeled welfare relative to the status quo. Average student gains in this constrained counterfactual are close to those under the unconstrained assignment—average achievement rises 7% of a standard deviation—indicating that teacher preferences are not the main barrier to achieving gains.

Teachers create enormous social value, and making the best use of their talents might generate large social returns (Rockoff, 2004; Chetty et al., 2014). Effectiveness, or the causal effect of a teacher on student test-score growth, varies substantially across teachers, and different teachers may have a different comparative advantage with various student types (Delgado, 2022; Aucejo et al., 2022; Bates et al., 2022; Biasi and Sarsons, 2022). Is the observed assignment of teachers to classrooms optimal for student learning? The challenge of optimally matching talent to roles is a very general organizational problem. It has been studied in education (Boyd et al., 2013; Bates et al., 2022), other parts of the public sector (Ba et al., 2021; Fenizia, 2022; Bergeron et al., 2022), the private sector (Cowgill et al., 2021), and in general terms (Prendergast and Topel, 1996).

Our paper is close to Bates et al. (2022), which also studies within-district teacher assignments using a model of teacher and principal preferences and student outcomes.

The two papers complement each other, serving as independent investigations of similar questions in different contexts. Both papers (1) study potential gains from the reallocation of teachers across schools within a district, (2) estimate their models with data from a district's internal teacher-transfer system, and (3) find limited gains from reassignment driven by within-teacher heterogeneity in effectiveness. However, our paper differs from Bates et al.'s in key ways. First, while they allow productivity matches through one binary student observable characteristic, our model allows for match effects on observable and unobservable characteristics. While our model does not nest theirs, our approach allows us to capture heterogeneity in effectiveness that varies within observable student types. Having a model that allows for unobserved match effects but that, like Bates et al.'s model, does not find much evidence of gains explained by matching on comparative advantage offers key evidence in support of moderate heterogeneity in teacher effectiveness that is uncorrelated with general effectiveness. Second, our model estimates rich sources of correlation between the decisions of teachers and principals and the outcomes of students. We leverage these sources to generate unbiased estimates of teacher effectiveness in unobserved matches. This aspect of the model also offers novel evidence of the dimensions of teacher selection into schools and the drivers of the assignments observed in the data.

Rather than focusing on reallocation across schools within a district, Biasi et al. (2021) studies the potential for across-district teacher reallocations to improve student achievement using a model of hiring competition across districts. As in Bates et al. (2022), they allow for comparative advantage in teaching through one binary student observable characteristic. The model leverages information about observed assignments and pay, not applications, interviews, rankings, or offers. It also finds an efficiency-equity tradeoff and limited potential gains from matching on comparative advantage.

A few prior papers use internal-transfer-system data to model two-sided matching of workers to positions. Boyd et al. (2011) pioneered the use of data on teacher-application data, using New York City Public Schools teacher-application decisions and subsequently realized matches. However, it does not observe the intermediate step in which schools rank teacher applicants or make offers. It models teachers' decisions to apply and teachers' likelihood of being hired in two separate steps and does not offer counterfactuals. Bobba et al. (2021) studies application-, placement-, and school-year-level student-outcome data from Peru's national centralized teacher-placement system. Its counterfactuals focus on how to use higher pay to attract better teachers to the industry and to schools in less preferred locations.[4] Finally, Abdulkadiroğlu et al. (2017) and Abdulkadıroğlu et al. (2022) use assignment mechanisms of students to schools as experiments to study schools' student-achievement impacts.

The rest of the paper is organized as follows. Section 2 describes the institutional background and the data. The model, model identification, and estimation are described in Sections 3 and 4. Section 5 presents the estimation results. The counterfactuals are described in Section 6.

2. Data and Empirical Evidence

2.1 Institutional Background

Our study focuses on a medium-sized, diverse urban district in the US Midwest. Seven in 10 large school districts, including the district we study, delegate authority to school leadership to choose who will fill open teaching positions at each school, according to our analysis of data from National Center on Teacher Quality (2022). Open teaching positions arise because of creation of new positions, retirement, or firing of incumbent teachers. Other incumbent teachers typically fill many of the open positions, as they may prefer some of these vacancies over their current positions. External candidates can fill vacancies only after internal candidates are considered, as specified in the collective bargaining agreement.[5]

Guided by a collective bargaining process with the teachers' union, district manage-

[4]Aside from education, Haegele (2023) analyzes application data from a German private sector company's internal-transfer system to illuminate sources of gender disparity in promotions to management. Haegele seeks to explain application decisions rather than modeling a two-sided process and its effects. Cowgill et al. (2021) build a theoretical framework to study the match of workers to tasks within a firm, and they test their model using data from an organization's internal labor market.

[5]At least 40% of large school districts explicitly prioritize internal-transfer candidates over external hires when filling openings (National Center on Teacher Quality, 2022).

ment created a centralized ITS to govern the matching process and act as a clearinghouse. The process comprises two successive rounds of applications, interviews, and offers each year. Each round involves the following steps, with a date fixed between each. First, each school posts known vacancies on the ITS for the coming year. Second, any incumbent teacher can apply to any vacancies for which their licenses qualify them. After the application window closes, the central district checks each applicant's eligibility for their applied positions and the system automatically grants interviews to the four most senior applicants per the collective bargaining agreement. Then, for each vacancy, each school views its automatic interviewees and its remaining applicant pool and can choose up to four other applicants for interview, yielding a maximum of eight interviewees per position. Schools can abstain from interviewing any of the four most senior applicants, but they would lose the option to make an offer to any applicant outside that group. If they decide to interview only a subset of the most senior, they have to invite them by order of seniority. After interviews, each school can submit a ranking of up to four interviewees. Next, the system automatically and simultaneously emails offers to the first-ranked interviewee for each position. Any applicant may get zero, one, or multiple offers in this step. Offerees have 48 hours to accept any one offer. After 48 hours, for any remaining open position, the system automatically withdraws unaccepted offers and emails an offer to the second-ranked interviewee for each open position. This process repeats until each position's ranked list is exhausted. Within a round, no teacher can renege on a previously accepted offer.

After the first round is completed, any vacancies that remain or new vacancies that arise from transfers during the round can be posted in a second round. The whole application, interview, ranking, and offer process repeats a second time. After the second round, any vacancies become open to external and internal candidates. When schools evaluate an applicant, they can observe the applicant's CV. The district recommends a default format that includes information on the applicant's education, employment history, and other qualifications. In 2013 the district began generating four measures of teacher quality that were also available to schools during the selection process.

2.2 Data

We use data from the ITS from 2010 to 2019 and merge in additional data on student outcomes for these years. In particular, we observe the vacancy postings, applications, interviews, rankings, offers, and acceptance data. In addition, for each teacher, we observe seniority rank, experience, education, ethnicity, race, gender, and current position assignment every year. A position assignment is a school and position type, such as a third-grade math teacher. Furthermore, we observe teacher licenses, allowing us to measure each teacher's choice set in each year. Last, we observe each teacher's home address each year, allowing us to measure the commuting distance to each open position in a teacher's choice set.[6]

We restrict attention to grades 4–8 to measure teacher effectiveness most reliably.[7] This restriction is standard in the literature because (1) job assignment in these grades provides a strong match to tested students and (2) students are mandated to take standardized tests in that grade and the prior grade. We drop teachers working less than half time to ensure that a teacher spends a significant number of hours with matched students.

The teacher sample contains 823 teachers, and they are observed 4 years on average for a total of 3,268 teacher-year observations (Table 1, Panel A). Just over half the teachers ever participated in the centralized ITS by applying to at least one open position. Two-thirds are female. Most—about 80%—identify as white, followed by 9% identifying as Black. Teachers average 13 years of experience and 5 years of higher education. Teachers who apply for transfers are less experienced on average. Teachers who seek a transfer apply to an average of 7 positions per year-round out of 37 for which they are eligible (Panel B). On average, applicants get interviews for 3 positions, get ranked in the top 4 by selection committees for 1.5 positions, and get 0.8 offers.

On the other side of the market, the ITS hosted 972 vacancy postings for grades 4–8

[6] We geocode teacher and school addresses, and we measure driving times with Google Maps API.

[7] In later years, we also observe four types of district effectiveness ratings: measures of value added in math and in reading based on classroom rosters; measures without school fixed effects and without controls for student race/ethnicity; student survey results; and scores based on classroom observation by certified raters against a standardized rubric of effective instruction.

Table 1: Descriptive Statistics for Teachers & Internal Transfer System

	Mean or Percentage	
Panel A: Teacher Demographics	All Teachers	Teachers in ITS
% male	25.2	22.8
% Black	8.5	10.0
% Hispanic	3.3	2.9
% white	79.2	81.0
% Asian	4.4	4.3
Years of experience	13.4	11.2
Years of education	5.0	5.0
Years in sample	4.0	4.4
Teacher count	823	421
Teacher-year count	3,268	1,861
Panel B: ITS - Teachers		Teachers in ITS
Size of choice menu		36.5
Applications submitted		6.7
Number of interviews		3.4
Positions that ranked candidate		1.5
Number of offers		0.8
Panel C: ITS - Open Positions		Positions in ITS
Number of potential applicants		213.7
Number of applicants		4.8
Number of interviews		2.5
Number of offers		0.6
Position count		972

Note: Panels A and B show the mean or share of a teacher characteristic for all the teachers in the sample in the left column and for all the teachers who ever applied to a position in the ITS. For teachers who are ever found in the ITS, we average their characteristics for every year they are in the sample, including those years they are not found in the ITS. Panel C shows the mean of a position characteristic.

from 2010 to 2019.[8] Though 214 teachers are eligible for the average posting, positions

[8] Vacancy postings are also restricted to positions that have a requirement of more than half time schedule.

average 5 applicants (Panel C). Qualified teachers' propensity to not apply motivates the inclusion of an inertia cost in our model. Positions average 2.5 applicants interviewed and 0.6 offers.

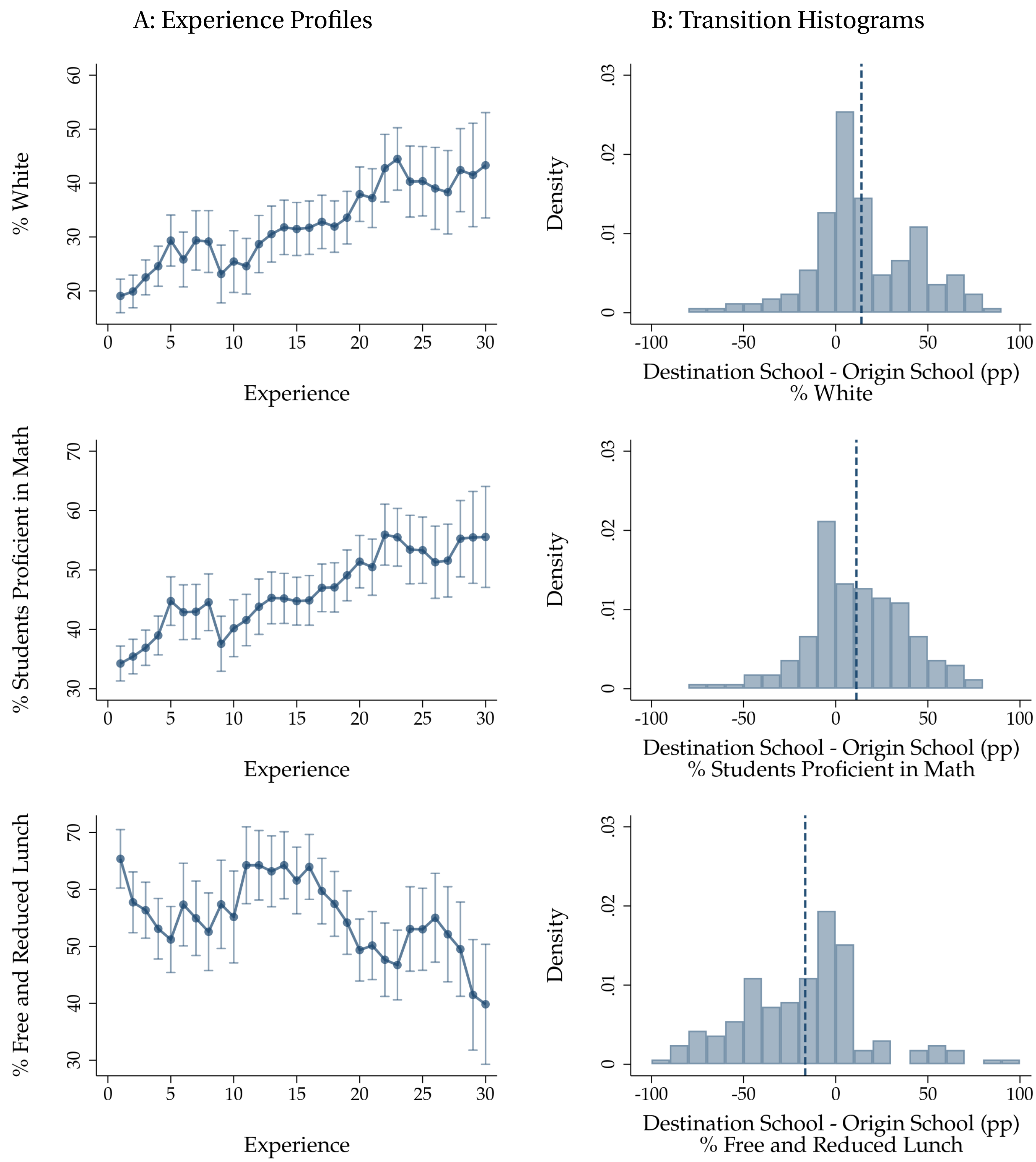
For each student in each year, we observe state-mandated standardized-test scores in math and reading as well as their grade, their school, their demographic characteristics such as race and gender, indicators of English-language-learner status, their eligibility for free or reduced-price (FR) lunch, which proxies for low household income, and their special education status. Our sample includes 35,608 students in grades 4 through 8 in 54 schools between 2009 and 2019. About a third of students identify as Black and a third as white. Hispanic students constitute about 20% of the sample. In contrast, more than 80% of teachers identify as white. Last, for each school year, we observe teacher, school, principal, and student characteristics that potential hires may value or that may influence the school's preferences for hire types.

2.3 Empirical Evidence

To motivate our model and counterfactual exercises, we examine the observed assignment and transfer patterns of teachers over their careers. Teachers with more experience tend to serve in positions at schools with a higher share of white students, a greater share of students proficient in math, and a lower share of low-income students (Figure 1, Panel A). The decisions of teachers and schools in the transfer process at least partially explain these patterns. Focusing on teachers transferring schools, we observe that teachers tend to move to schools with a higher share of white students, a higher share of students proficient in math, and a lower share of low-income students than their prior school (Panel B), although there is considerable heterogeneity. This is consistent with outside evidence that teachers of more disadvantaged students tend to be more likely to transfer across schools (Boyd et al., 2005; Scafidi et al., 2007; Goldhaber et al., 2011; Isenberg et al., 2022; Goldhaber et al., 2019).

The observed equilibrium arises from the joint decisions of schools and teachers in the transfer process, and it may or may not be aligned with teachers' comparative advantage. While teachers may have incentives to sort into positions in which they

Figure 1: Teacher-Transition Characteristics



Note: Panel A shows the mean and 95% confidence intervals of a school characteristic at the school assigned to teachers with each level of experience. Panel B restricts the sample to teachers who changed schools, and it plots the histogram of the difference in school characteristics between the destination and origin school. The vertical line is the mean difference.

have the most impact and school principals may want to hire the teachers that have the largest impact on their students, these are unlikely to be the only things valued. The lack of discretion schools have to set wages and to lay off less effective tenured teachers may push realized assignments away from the achievement-maximizing assignment. Our analysis quantifies the degree to which sorting patterns are correlated with student gains in this market, how an assignment that maximizes student achievement would look, and how much test score gains can be compelled.

3. Model

We jointly model student outcomes and teachers' and school leaders' decisions to incorporate the potential for selection into schools. Equation 1 describes student outcomes, where y_{kt} is the standardized-test score of student k in year t. We denote by s the school of each student and by i each teacher.[9]

Similarly to traditional value-added models, student outcomes are a function of a rich set of observable student and school characteristics that may change over time (ω_{kt}, x_{st}). The model also includes observable teacher characteristics (τ_{it}). To capture observable match effects in teaching, we incorporate $C_0(\omega_{kt}, x_{st}, \tau_{it})$, a flexible function of student, school, and teacher characteristics and of interactions between teacher and student characteristics.[10]

$$y_{kt} = C_0(\omega_{kt}, x_{st}, \tau_{it})\alpha^y + \theta_i + \eta^y_{is} + \varepsilon^y_{kt} \tag{1}$$

In addition to interactions on observables, the model includes teacher-level and teacher-school-level unobservables (θ_i, η^y_{is}). θ_i is teacher i's general (context independent) effectiveness net of the effect of characteristics in τ_{it} and interactions. η^y_{is} is the unobservable match effect between teacher i and school s. This match effect reflects that some teachers may thrive in environments with certain leadership styles or with students of a specific unobservable type that are more prominent at s. θ_i and η^y_{is} are

[9]For each (k,t), $s(k,t)$ is the school of k in year t and $i(k,t)$ is a teacher assigned to k in t. We abuse notation and simply use s and i to refer to the named school and teacher.

[10]Here we also include school and year fixed effects.

normally distributed with variances to be estimated. We further assume teacher-school match effects are constant over time. ε^y_{kt} is a shock that is specific to the teacher, student, and year.

$$u_{ist} = C_1(x_{st}, \tau_{it}, z^u_{ist})\alpha^u + \gamma I_{ist} + \eta^u_{is} + \varepsilon^u_{ist} \tag{2}$$

$$v_{ist} = C_2(x_{st}, \tau_{it}, z^v_{ist})\alpha^v + \varphi_i + \eta^v_{is} + \varepsilon^v_{ist} \tag{3}$$

Preferences and decisions of teachers and schools in the transfer process are modeled using a random-utility framework. Equation 2 describes teacher i's utility for school s. u_{ist} is a function of observable teacher and school characteristics and interactions between these, captured by the function $C_1(x_{st}, \tau_{it}, z^u_{ist})$, where z^u_{ist} is a vector of observable teacher-school-level characteristics that are excluded from the outcomes model. We also include a teacher-school-level unobservable match in teacher tastes, η^u_{is}. η^u_{is} is normally distributed with a variance to be estimated. The model also includes school fixed effects, capturing each school's unexplained attractiveness that is common across teachers and time. Because teachers face the decision to apply for a transfer each year, we include an inertia term in teacher utilities that captures the value to teachers of not changing jobs. This includes the value teachers place on not having to change their routine, form new networks and friendships, and generate new teaching material in a new environment. We assume the value of inertia is the same for all teachers. The model expresses inertia by the parameter γ. Here I_{ist} is an indicator variable that turns on for a teacher's current position. Finally, ε^u_{ist} is a shock specific to a teacher, school, and year.

Equation 3 describes school s's willingness to hire teacher i. v_{ist} depends on observable teacher and school characteristics and interactions between these, captured by the function $C_2(x_{st}, \tau_{it}, z^v_{ist})$, in which z^v_{ist} is a vector of observable characteristics at the teacher and school level that are excluded from the outcomes model. The model also includes a teacher-school-level unobservable for match in school tastes, η^v_{is}, and a teacher-level unobservable, φ_i, capturing unobserved attractiveness of each teacher common across schools and time. η^v_{is} and φ_i are normally distributed with variances to be estimated.

The interactions between teacher and school observables in both choice models capture heterogeneity in the same dimensions as the model of student outcomes. For example, we allow for a same-race effect in student outcomes and also allow teachers to prefer schools with a greater share of same-race students. Likewise, schools may prefer teachers who share the race/ethnicity of the majority of the school's student body. The parameters associated with these interactions allow us to capture selection by student gains via several observable components.

On the unobserved side, we allow for, and estimate, the correlation between a teacher's general effectiveness, θ_i, and schools' common unobserved taste for the teacher, φ_i. A positive correlation between θ_i and φ_i indicates that schools tend to value teachers with greater general effectiveness. We also allow for correlations between the teacher-school match effects in student-achievement production, in teacher's tastes, and in school's tastes $(\eta^y_{is}, \eta^u_{is}, \eta^v_{is})$. A positive correlation between η^y_{is} and η^u_{is} implies teachers tend to value schools at which they have a comparative advantage. A positive correlation between η^y_{is} and η^v_{is} implies schools tend to value teachers who are especially effective at teaching their students.

We do not impose restrictions on the correlation structure of the unobservables, and we further assume $(\theta_i, \varphi_i) \overset{iid\ i}{\sim} N(0, \Sigma_{\theta\varphi})$ and $\eta_{is} = (\eta^y_{is}, \eta^u_{is}, \eta^v_{is}) \overset{iid\ is}{\sim} N(0, \Sigma_\eta)$. Also, $\varepsilon^y_{kt} \overset{iid\ kt}{\sim} N(0, \sigma^2_{\varepsilon y})$, $\varepsilon^u_{ist} \overset{iid}{\sim} N(0,1)$, $\varepsilon^v_{ist} \overset{iid}{\sim} N(0,1)$. The latter assumptions imply that any correlations in time over teacher or school preferences are captured by observables in each model or by the variables η^u_{is} and η^v_{is}.

The parameters to be estimated are the coefficients of the observable covariates $(\alpha^y, \alpha^u, \alpha^v, \gamma)$, the variance-covariate matrices $\Sigma_{\theta\varphi}$ and Σ_η, and the variance of the error in the model of outcomes, $\sigma^2_{\varepsilon y}$.

3.1 Mapping the Model to the Data

We estimate the parameters of the model using the decisions of teachers and schools in the transfer process and the test scores of students from observed matches. While we have described our model as one in which teachers and schools are on each side of the market, in practice each school can have more than one open position in the ITS concurrently. Consequently, we estimate a model in which teachers have preferences over positions and in which preferences over teachers are position specific. Since all the parameters in the random-utility models are school specific, any within-school position-level disagreements will be captured by variation in the error terms. These situations may capture instances in which, for example, a teacher applies to only one of two open positions at a school, which would indicate she finds one position description more appealing than the other.

Teacher decisions. At the end of each academic year t, every teacher who had an assigned teaching position can decide to apply for a transfer. We observe the licenses of every teacher each year and the licenses required for every open position in the ITS. With this information, we create a menu of positions for each teacher. Since we observe 10 years of data and two rounds each year, we build at most 20 menus per teacher and observe the decisions in each case.

At the end of each year, a number of teachers who had an assigned teaching position learn that they lost their assignment. This happens if enrollment at a school falls below a level that requires an adjustment of the budget. In this case, the principal chooses the teacher that must lose her assignment. In principle, seniority protects teachers from demotions, and only the more junior teachers are at risk of losing their assignments in these situations. A teacher who loses her assignment needs to go to the ITS to search for a new position and has no fallback option. Our data allows us to identify these teachers. In our restricted sample, about 41% of teachers searching for a match in the ITS each year had lost their assignment. For these teachers, the value of remaining unassigned after the ITS clears is the expected value of the match they expect to find in the scramble round, in which every unmatched teacher must match with a remaining position, or the next round. In contrast, a teacher who did not lose her assignment has her current position as her fallback option.

Consistently with this, we assume that in each round a teacher applies to every position on her choice menu she prefers to her fallback option, net of the inertia term.

Inertia represents the average value of not changing jobs and only applies to teachers who did not lose their position.

We assume that at the application stage, teachers consider all positions on their choice menu. This means they are aware of all these positions and can compare them. Given that in each round, each teacher has an average of 37 positions to consider, we think this is a reasonable assumption. Moreover, the teaching positions in our sample tend to follow standard descriptions and are mainly differentiated by the school and grade. Both are easily observable characteristics that a teacher probably has a taste for before coming to the ITS. We also assume there are no application costs for teachers. This assumption is motivated by the infrastructure of the platform that collects applications. Similarly to the job market for economics PhDs, an applicant must attach application materials for each posting and these materials do not need to be customized for each position. We use a second source of variation to estimate teacher preferences. After teachers receive offers, we observe a subset of teachers who receive multiple concurrent offers and choose one. We assume teachers choose their preferred offer.

School decisions. We observe schools making two types of decisions. They first choose which candidates to interview out of those who applied for a position, and later they select which interviewees to rank and in which order. We assume that after the interviews are carried out and schools learn information about the candidates, schools' ranking of candidates expresses their willingness to hire them. We do not use interview decisions to estimate school hiring teams' preferences but instead assume that at this stage schools may eliminate candidates that are perceived as poor matches for the job or candidates that schools believe they have no chance of getting if an offer were made.

Our model assumes that ranked candidates are acceptable to schools. This means the utility a school would derive from getting any of those teachers is positive. We assume ranked candidates are listed in preference order, although we do not make any assumptions about candidates that are not ranked after the interview phase. Some of these candidates may be unacceptable; others may be perceived as unreachable. In consequence, we only use data on the set of ranked candidates from each position to infer school preferences.

A few aspects of our setting and observations from our data motivate our assumptions. First, schools are invited to rank teachers who they know have expressed interest in their position in at least two stages: initially, by applying to the position, and later by attending an interview. The latter is presumably more costly, and both signal how serious a teacher is about a position. Second, if schools skip candidates they like but perceive as unreachable by not ranking them at all, that is not a worry; nevertheless, if schools skip them by ranking them lower than less preferred but safer choices, this may pose a problem. We find that 82% of positions with more than one candidate to rank could have included more candidates but did not. This means the overwhelming majority of hiring teams skip candidates after the ranking stage. Moreover, we find evidence that schools are choosing to not rank candidates who are popular: we find that candidates who are ranked first by multiple positions are more likely to not be ranked by other positions they interviewed for (Figure XX). Finally, we take data from the comments section on each interview, which has no impact on the assignment and is designed for internal communication between the hiring team and the school leadership, and we do not find any reference to concerns about the feasibility of attracting candidates. Instead, hiring teams typically comment on the match quality of a candidate or their overall quality.[11]

4. Identification and Estimation

4.1 Identification

To identify the joint distribution of utilities and outcomes, we need variation that lets us separate teacher effectiveness, defined as the average impact of a teacher on student test scores, from teacher-effectiveness heterogeneity that is potentially correlated with the decisions of teachers and schools in the transfer process. To do this we use two shifters. The first separates teachers' decisions from student outcomes; the second separates school leaders' decisions from outcomes. We use the distance between a

[11] Although adding comments is not mandatory, more than 80% of the positions in our sample have at least one comment

teacher's home and every school as a shifter of teacher preferences (z^u_{ist}), and we use the share of same-race teacher peers in a school as a shifter of a school's willingness to hire a teacher (z^v_{ist}).

Formally this requires two sets of assumptions. First, both z^u_{ist} and z^v_{ist} are assumed to be conditionally independent of (η_{is}, θ_i, φ_i) and the error terms (ε^u_{ist}, ε^v_{ist}, ε^y_{kt}) conditional on the controls. Moreover, z^u_{ist} is assumed to be conditionally independent of z^v_{ist}. This implies that the shifters do not affect the distribution of potential outcomes but only affect observed student outcomes via the assignment of teachers to classrooms. Similarly, z^u_{ist} does not affect the distribution of school utilities but only affects the utilities of observed decisions by affecting the set of teachers that a school gets to interview and rank.

A second set of assumptions relates to the impact of the instruments on teacher and school choices as their values approach their limits. We assume that as the distance from home to a school, z^u_{ist}, approaches zero, the probability that a teacher will apply to a position in that school, and choose it among concurrent offers, approaches one. Also, as the share of same-race teacher peers, z^v_{ist}, approaches one, the probability that a school will rank the teacher as their first choice is assumed to approach one. Moreover, given the assumption that z^u_{ist} and z^v_{ist} are independent, it follows that as $(z^u_{ist}, 1 - z^v_{ist}) \rightarrow (0, 0)$, the probability that teacher i will be assigned to s approaches one. In consequence, each instrument independently affects the decisions of teachers and schools, and together they shift the probability of assignment.

Intuitively, the instruments help separate teacher effectiveness from selection on student gains by allowing a comparison of (1) teachers that are close to being randomly assigned to s by means of the random shifter and (2) teachers that are in s as a result of selection. For instance, suppose teachers who live closer to school s tend to have students who underperform relative to the students of teachers who live farther. The first set of teachers is closer to being randomly assigned to s by virtue of the shifter. Then we would extrapolate that there is positive sorting toward s in the sense that teachers who are more likely to be good matches for s are more likely to apply to s.

A second source of selection in our model comes from the fact that schools' revealed preferences are only observed for (a subset of) the teachers that applied for a position at the school. Since application decisions are not random, inferring school preferences from decisions over a selected sample might not extrapolate to other sets of teachers. To account for any correlation between teacher and school preferences that does not map onto the observables in our model, we exploit the fact that z_{ist}^{u} shifts teacher decisions but is omitted from school decisions. That is, we assume that schools do not value teachers' residential location when choosing who to interview or offer a job to. The same logic and intuition apply to this shifter in relation to school decisions.

The identification strategy in this paper builds on strategies used in papers that evaluate outcomes resulting from assignment mechanisms. Our case extends the analysis to a situation in which an assignment is produced by the joint, and possibly selected, decisions of two agents on either side of a market. Agarwal et al. (2020) models a situation in which a patient's decision to accept a kidney transplant may be correlated with outcomes. As the authors show, identification is obtained from two sources: an instrument that shifts patients' decisions and is uncorrelated with the distribution of potential outcomes, and the assumption of quasi-randomness in the types of kidneys being offered to a patient. Similarly, Walters (2018) estimates the gains from charter school attendance in a model in which decisions of families who seek admission may be correlated with student outcomes. Identification is achieved with the use of a shifter of families' preferences and the fact that charters' decisions are random and hence uncorrelated with student outcomes. Our identification strategy extends these ideas and requires a shifter for the preferences of agents on each side of the market. Each shifter helps tease out the correlation between decisions and outcomes and does so by introducing a degree of randomness in supply and demand. Other papers that jointly model and estimate correlated decisions and outcomes using instruments for decisions include Hull (2018), Van Dijk (2019), Kapor et al. (2022), and Geweke et al. (2003). More generally, this set of papers builds Lewbel (2007), Heckman (1990), and Heckman and Vytlacil (2005).

4.2 Estimation Strategy

While we have described our model as one in which teachers and schools are each on one side of the market, in practice each school can have more than one open position in the ITS concurrently. Consequently, we estimate a model in which teachers have preferences over positions, and preferences over teachers are position specific. Since all the parameters in the random-utility models are school specific, any within-school position-level disagreements will be captured by variation in the error terms. These situations may capture instances in which, for example, a teacher applies to only one of two open positions at a school, which would indicate she finds one position description more appealing than the other.

The observable teacher characteristics include years of education, race/ethnicity, gender, and indicators capturing returns to experience over teachers' careers. The observable student characteristics include demographics and past test scores. The observable school characteristics include structural characteristics such as average student demographic characteristics at the school-year level, and average experience levels of school teachers.[12]

We estimate the joint distribution of the parameters using a Gibb's sampler and assuming conjugate uninformative priors (Gelman et al. 2013). This methodology has been used by Geweke et al. (2003), Agarwal et al. (2020), and Kapor et al. (2022). Using this method we generate draws of the joint distribution of the parameters and latent variables in the model. We draw 112,000 iterations of the sampler, burn a number of initial iterations, and keep only 1 of every 10 draws to reduce autocorrelation in the chains. We inspect chains for convergence.

5. Results

Student Outcomes. Table 2 displays the posterior mean and standard deviation of the parameters in the student-outcomes model. As in traditional teacher value-added models, our specification contains a rich set of student, teacher, and school observable char-

[12] Appendix Table A3 provides a complete list of observable teacher, student, and school characteristics.

acteristics, including past student test scores and student and school demographic information. However, we depart from traditional value-added models and include a rich set of teacher characteristics and interactions between teacher and student characteristics that capture observable match effects in outcomes as well as unobservable school-teacher match effectiveness.

Students with lower past test scores who are male, low income, and non-white have lower test scores (Panel A). For example, FR students have test scores, on average, 0.14 standard deviations lower. Teacher gender and experience also affect student test scores, but their predicted effect is about an order of magnitude smaller than that of students' characteristics (Panel B). For example, female teachers are associated with test scores 0.02 standard deviations higher than male teachers, equivalent to the return of two years of teaching experience at the beginning of a teacher's career.

Teacher and student observable match effects (Panel D) affect final test scores. Having a teacher of the same sex improves student outcomes, but having a teacher of the same minority status does not always do the same. We define minority status here as being non-white. While the estimated same-sex parameter is remarkably close to others found in the literature, the same-race effect is not (Delgado 2022). Our results indicate that more educated and experienced teachers have a comparative advantage in teaching minority and low-achieving students, contrary to the observed equilibrium assignment patterns we observe in the data (Figure 1).

Teacher Supply. Table 3 shows key moments of the posterior distribution of parameters in the model of teacher preferences. Teachers value characteristics of students at a school. Consistently with Figure 1, teachers prefer schools with fewer low-income and minority students. This is especially true for more experienced teachers. In addition, teachers prefer teaching students of their own race. While teachers have no preference over the overall share of minority teachers in a school, they value working with colleagues of the same race. Furthermore, while the results indicate that teachers with little education and experience do not prefer schools with a larger share of higher-achieving students, more educated teachers prefer teaching higher-achieving students.

Table 2: Parameters of the Outcomes Model

	Mean	*Std. Dev.*
Panel A: Student Characteristics		
Prev. score	0.770	0.001
Male	−0.012	0.002
Low income	−0.142	0.005
English language learner	−0.052	0.002
Race - Black	−0.127	0.003
Race - Hispanic	−0.067	0.003
Panel B: Teacher Characteristics		
Male	−0.022	0.005
Education	−0.003	0.002
Experience 2 to 3	0.020	0.005
Experience 4 to 6	0.022	0.005
Experience 7+	0.028	0.005
Panel C: School Characteristics		
% Low income	−0.022	0.018
% English language learner	−0.041	0.013
% Black	0.036	0.021
% Hispanic	0.125	0.022
Low income*% Low income	0.071	0.007
Panel D: Student-Teacher Interactions		
Match minority	0.003	0.003
Match gender	0.005	0.002
Minority*Education	0.005	0.002
Prev score*Education	−0.003	0.001
Minority*Experience	0.002	0.002
Prev score*Experience	−0.002	0.001
Class size	0.036	0.030
Std. Dev. of Teacher General Effectiveness (θ_i)	0.0817	
Std. Dev. of Teacher-School Match Effectiveness (η^y_{is})	0.0004	

Note: The table shows the mean and the standard deviation of the last 200 realizations from the chains of each estimated parameter.

Last, teachers prefer working in a school with a high fraction of experienced colleagues.

Table 3: Parameters of Teacher Supply

	Mean	*Std. Dev.*
Panel A: School Characteristics - Students		
% Free or reduced-price lunch	−0.111	0.059
% Black	−0.162	0.358
% Hispanic	−1.309	0.334
Avg. test scores	0.026	0.112
Panel B: School Characteristics - Teachers		
% Black	0.247	0.376
% Hispanic	−0.538	0.534
Average teacher experience	0.044	0.019
Panel C: Teacher-School Interactions		
% of students match minority	0.498	0.047
% of teachers match minority	0.440	0.050
Education * Avg prev score	0.125	0.025
Education * % minority	0.104	0.019
Experience * Avg prev score	−0.068	0.040
Experience * % minority	−0.317	0.030
Experience * Avg teacher experience	0.008	0.013
Panel D: Teacher Characteristics		
Driving time	−0.010	0.002
Inertia	3.582	0.034
Std. Dev. of Teacher-School Match Effects (η^u_{is})	0.495	

Note: The table shows the mean and the standard deviation of the last 200 realizations from the chains of each estimated parameter.

As expected, teachers dislike teaching farther from their homes, so longer driving times come at a cost to teachers. Teachers experience inertia when faced with the choice of a transfer. The mean inertia parameter estimate is several orders of magnitude larger than any of the observable school characteristics or the observable match effects.

School Demand. Table 4 shows the posterior mean and standard deviation of the parameters in the model of school demand. While observable teacher characteristics such as sex, education, and experience are not crucial in explaining the school choices of teachers, schools prefer to hire teachers whose race matches a greater share of the school's teachers.

Table 4: Parameters of School Demand

	Mean	*Std. Dev.*
Panel A: Teacher Characteristics		
Male	0.072	0.148
Education	−0.340	0.504
Experience	0.374	0.777
Panel B: Teacher-School Interactions		
% Match race - students	−0.359	0.351
% Match race - teachers	4.061	0.258
Education * Avg prev score	0.027	0.420
Education * % Minority	0.577	0.772
Experience * Avg prev score	−0.463	0.633
Experience * % minority	−0.717	1.174
Experience * Avg experience	0.032	0.093
Std. Dev. of Teacher Effects (φ_i)	0.153	
Std. Dev. of Teacher-School Match Effects (η^v_{is})	0.013	

Note: The table shows the mean and the standard deviation of the last 200 realizations from the chains of each estimated parameter.

Correlation between Decisions and Outcomes. To quantify the importance of the three different aspects of teacher effectiveness—general (θ_i), match on observables, and unobservable match (η^y_{is})—for student outcomes and teacher and school utilities, Table 5 shows the change in student test scores, teacher utilities, and school utilities if teacher effectiveness or the match effects changed from very low to very high values. We simulate the change in these quantities when the value of θ_i or η^y_{is} goes from the 1st to the 99th percentile. We also draw from the underlying distribution of teacher and student observables to generate the distribution of observable match effects and evaluate changes as match quality goes from the bottom to the top of the distribution.

Table 5: Correlations between decisions and outcomes

	Outcomes	School Utility	Teacher Utility
	Δy_{kt}	Δv_{ist}	Δu_{ist}
From percentile 1 *to* 99			
Teacher general effectiveness, θ_i	0.386	0.618	
	(0.023)	(0.298)	
Unobservable match effects, η^y_{is}	0.002	0.002	-0.541
	(0.00002)	(0.049)	(0.031)
Observable match effects	0.037	0.602	-0.017
	(0.005)	(1.676)	(0.017)

Note: The table shows results from simulated changes in student outcomes, school utility, and teacher utility if teacher general effectiveness (θ_i), unobserved teacher-school match effectiveness (η^y_{is}), and overall observed match effectiveness went from percentile 1 to 99. Changes are expressed in standard deviations. For example, matching with a teacher with a value of θ_i in the 99th percentile raises student test scores by 0.4 standard deviations relative to being matched with a teacher in the 1st percentile. Schools, on the other side, value effective teachers. Hiring a teacher in the 99th percentile of efficiency, relative to percentile 1, raises schools' utilities by 0.6 standard deviations. Standard deviations are shown in parentheses.

Teacher general effectiveness has greater explanatory power than changes in match effects on student outcomes. While substituting a teacher who is at the bottom of the general-effectiveness distribution for a teacher at the top of the distribution improves student test scores by 0.4 standard deviations, moving from the lowest to the highest quality match on observables pushes test scores up by 1/10th of that (0.04 standard deviations). Unobservable match effects can drive only a gain of about 0.002 standard deviations. All these are statistically different from zero across simulations. The impact of match quality on observables aggregates the effects of being matched with a teacher of the same sex and of matching minority students and low-achieving students with more experienced and educated teachers.

Substituting a teacher from the bottom of the teacher general-effectiveness distribution for one at the top of the distribution raises school utilities by 0.6 standard deviations. However, we observe no change in school utility when we move a teacher from the bottom to the top of the match-quality distributions. On the other side, we find that teachers are averse to positions in which they would have high observable or unobservable match effectiveness. This is especially true for the portion of match effectiveness not captured by observable characteristics.[13] Teachers appear averse to acting on their comparative advantage; schools do not.

5.1 Robustness

As a validation exercise for the student-outcome model, we compare our teacher-effectiveness measure with multiple teacher-effectiveness measures used by the school district, including value-added estimates for math and reading, a student survey-based measure, and a score based on a standardized rubric of effective instruction through classroom observation by certified peer raters. We find that our paper's estimated teacher general-effectiveness measure strongly correlates with the four other measures of teacher effectiveness (Appendix Table A1).[14]

In addition, there might be a concern about the generalizability of the evidence from this district. The district we study has student and school characteristics that strongly resemble the population of US urban schools. We use the Generalizer tool specifically designed to quantify the degree of generalizability between a sample of studied K-12 schools and a target inference population of schools (Tipton and Miller, 2022). The Generalizer tool uses propensity scores to measure the similarity between the sample and inference population, yielding a generalizability index between 0 and 1. We compare the schools in our sample to 15,389 US schools in the population inference sample that teach 5th, 6th, 7th, or 8th grade, are in an urban locale, and are not charter schools. Based on parameters such as school size, percentage of FR students, female percentage, white percentage, Black percentage, Hispanic percentage, US-citizen percentage, and median family income, our analysis implies a generalizability index of 0.81, which the tool characterizes as high generalizability.

[13]The change in utility associated with changes in the observable component of match effects is negative in most simulations but exhibits a positive right tail. This may be evidence of heterogeneity in the valuation for observable match effects.

[14]The four district evaluation measures started after 2012. The correlations are estimated from 2013 to 2019.

6. Counterfactual Teacher Assignments

6.1 Description of Counterfactuals

To quantify the potential gains associated with a reallocation of teachers to classrooms, we consider a reassignment of teachers to the set of active positions associated with test scores during our 10-year study period. We call this set $\mathcal{P}$. By construction, each position in our sample is associated with a single year, meaning that even when some teacher-school pairs remain matched for many years, we treat each as a different position. For each position in $\mathcal{P}$, we observe the teacher assigned to it and the set of students associated with the position. The set of teachers that are candidates for assignment to positions in $\mathcal{P}$ in each year includes the set of teachers that were assigned to a position in $\mathcal{P}$ in that year and the set of teachers that applied to a position in the ITS that year but were not assigned a position in $\mathcal{P}$. Using data on the licenses each teacher holds and the licenses required for each position, we further restrict the menu of positions for each teacher each year.

We consider two policy objectives. The first is to maximize average student test scores; the second is to maximize the share of proficient students. Both measures are widely used by school districts, parents, and policy makers to evaluate district performance. We first consider the reassignment of teachers to classrooms to maximize each policy objective, taking the positions in $\mathcal{P}$ and the set of teachers described above, without further constraints. We refer to this counterfactual as the unconstrained counterfactual. We generate this assignment by solving a linear program in which we restrict each position to be filled by exactly one teacher and each teacher to be assigned to at most one position.[15] When we generate this assignment, we do not consider the dynamic gains generated by a reassignment in year t for test scores in subsequent years; instead we consider the one-shot reassignment in each year independently. In consequence, the counterfactual results quantify the average one-year gains from a reassignment of teachers to classrooms.

[15] We solve the following linear program, in which $l \in \mathcal{P}$, and i is a teacher: $\max_a \sum_{i,k \in l} a_{il} \cdot y_{ki}$ s.t. $a_{il}(1 - c_{li}) = 0, \sum_i a_{il} \leq 1, \sum_l a_{il} = 1$. Further, $a_{il} = 1$ if i is assigned to l and $c_{il} = 1$ if i is feasible for l; both are zero otherwise.

Because some assignments generated under the unconstrained counterfactual may be unacceptable to teachers under current pay schemes, a second counterfactual further restricts each teacher's menu to positions that are weakly preferred to their observed assignment.[16] Here we simulate teacher utilities using model parameters. We refer to this as the counterfactual that guarantees no teacher is harmed. We do not include inertia costs here because we consider reassignments as alternatives to the observed ones rather than as transitions from a past position.

Next, we add two additional restrictions on top of not allowing teacher harm. A third counterfactual restricts the set of teachers to be those assigned to positions in $\mathcal{P}$ in each year. This means, for this counterfactual, we have the same number of teachers and positions and we are simply allowing for a reshuffling of the teachers we observe assigned in each year.[17] A fourth counterfactual further restrict assignments to only within-school reassignments. Here, we only allow teachers to be reassigned to positions in the same school in which they already teach that year.

Finally, we use the model parameters to predict outcomes under a random reassignment of the teachers observed assigned to positions in $\mathcal{P}$ and under the observed assignment. Both these will be used as benchmarks. For each counterfactual, and for the observed and random assignments, we simulate data from 100 realizations of the parameters in our model and compute the mean and standard deviation of the gains across these simulations in each case.

6.2 Counterfactual Results

The observed assignment generates remarkably small improvements in average student test scores and the share of proficient students relative to a random assignment (Figure 2). This is perhaps not surprising given the low correlation between teacher and school preferences and their match effects discussed in Section 5. When we look into students by baseline achievement and race/ethnicity, we find that the observed assignment relative to a random assignment benefits higher-achieving students and white

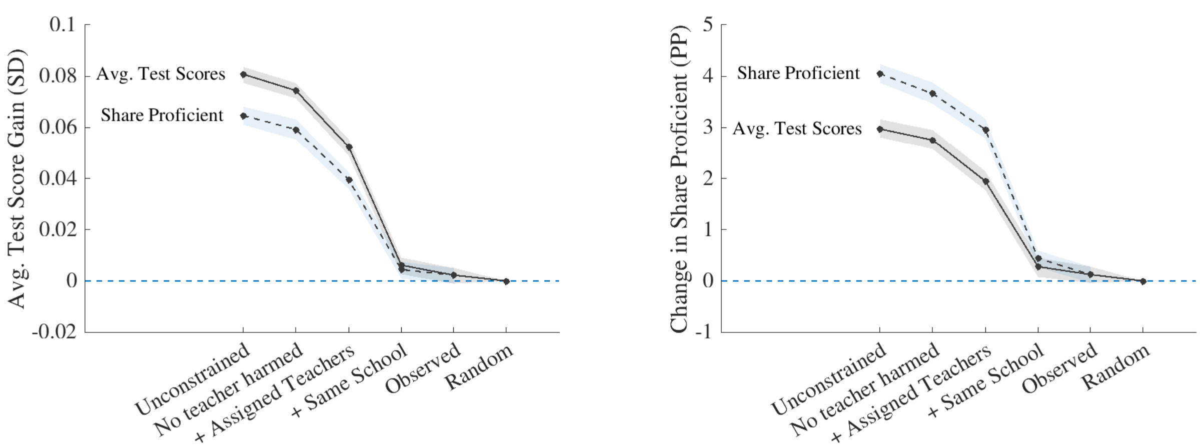
[16] Because the observed assignment is feasible, the set of solutions is not empty.

[17] Teachers that apply to positions but wind up unassigned are excluded from consideration.

students while generating no improvements for the rest (Figure 3). This is because we observe white students and higher-achieving students assigned to more effective teachers (Table A2) and is consistent with model parameters that show that teachers have a preference for teaching higher-income, high-achieving, and white students and that schools value generally effective teachers. This observed equilibrium is different from that found in Bates et al. (2022), in which teacher quality is equally distributed across these groups. The difference in the two settings may be explained by the fact that for seven years in our sample, information about teacher effectiveness was observable to schools favoring a correlation between effectiveness and school preferences.

Figure 2: Counterfactual Gains in Test Scores and Share Proficient by Policy Objective

(a) Average Test-Score Gains (b) Gains in Share Proficient

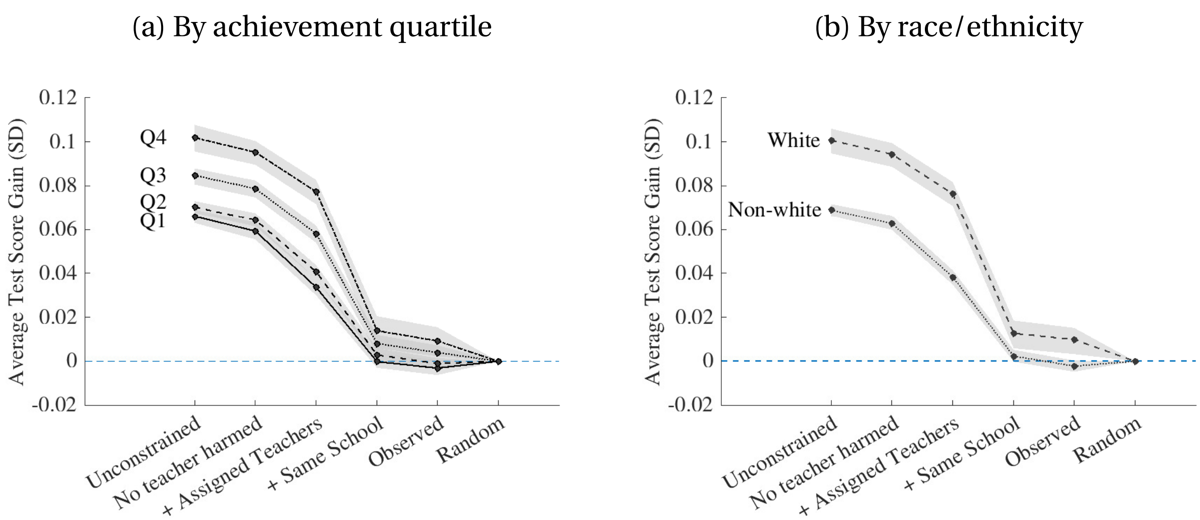


Note: Panel (a) shows gains in average test scores relative to random assignment in counterfactuals that aim to maximize average test scores (solid line) and share proficient (dashed line) under a set of constraints. Panel (b) shows gains in share proficient under the same counterfactuals.

Under the unconstrained reassignment of teachers, a policy maker seeking to maximize average student test scores could push scores up by 8% of a standard deviation relative to the observed assignment. Though we predict teachers have strong preferences regarding student demographics and commute time, we find most of the gains can still be realized under constraints that ensure no teacher is harmed. In this case, test scores would increase by 7% of a standard deviation (Table A4).

Under both these counterfactuals, the pool of teachers that are candidates for reassignment is larger than the set of positions because we consider both teachers who are observed to be assigned to a position in our sample the next year and teachers who unsuccessfully sought a new position in our sample through the ITS.[18] If we restrict the teacher pool to include only those that we observe assigned and keep the no-teacher-harm constraint, we find the gain in test scores is 5% of a standard deviation. Finally, if we only allow for reassignments within schools, the gains fall to 0.4% of a standard deviation. This shows that most of the gains come from a reallocation of teachers across schools and hence can be realized by intervening in the transfer system (Table A4).

Figure 3: Differential Gains in Test Scores by Achievement and Race When Maximizing Average Achievement



Note: Panels (a) and (b) show average test-score gains relative to a random assignment, by student baseline-achievement quartile and race/ethnicity, under a counterfactual that maximizes average test scores.

Relative to the scenario in which only assigned teachers are considered, the unconstrained counterfactual selects teachers with higher general effectiveness. Their effectiveness is validated by schools in the ITS, as these teachers are more likely to receive an offer and are ranked higher than teachers who are displaced in the unconstrained counterfactual. Also, looking at state employment records, we find these teachers are

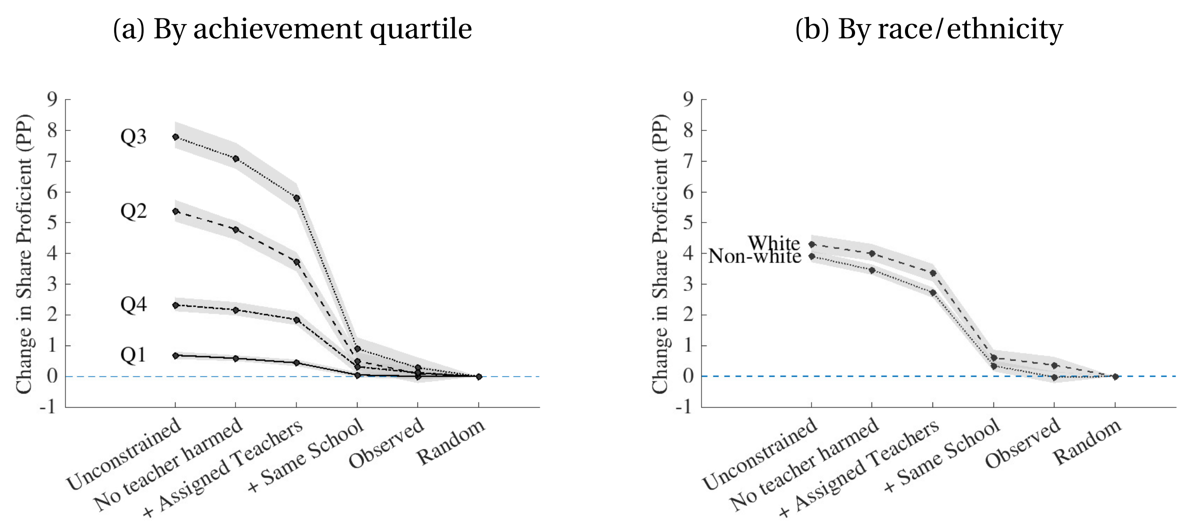
[18]Teachers who we do not observe assigned to one of the positions in our sample are observed assigned to other positions outside our sample. These are positions that are not associated with math or reading test scores or positions with less than 0.5 full-time equivalence required.

more mobile across districts in the state but are less likely to ever leave the teaching profession than teachers displaced in the counterfactual.[19]

Counterfactual assignments that maximize average student test scores benefit higher-achieving students and white students more, even when all achievement levels and non-white students experience gains. This means these assignments widen the racial achievement gap and achievement inequality overall (Figure 3).

If we assume the policy maker's objective is to maximize the share of proficient students, we find this quantity would increase by 4 percentage points (pp) under the unconstrained reassignment relative to the observed assignment. Similarly to before, almost all gains can be realized if we impose a no-teacher-harm constraint (3.5 pp). Further restricting the sample to assigned teachers reduces the gains to 3 pp, while only allowing for within-school reassignments shrinks them by an order of magnitude to 0.3 pp. Again, most of the gains come from between-school reassignments (Table A5).

Figure 4: Differential Gains in Share Proficient by Achievement and Race When Maximizing Share Proficient Achievement



Note: Panels (a) and (b) show gains in the share of proficient students relative to a random assignment, by student baseline-achievement quartile and race/ethnicity, under a counterfactual that maximizes the share of proficient students.

[19]We define leaving the teaching profession as not being assigned to any school in the state for two or more consecutive years and not being assigned to any position in the state thereafter. We use assignment data from 2009 to 2019 and consider all teaching positions in the state, not only those in our sample.

While counterfactuals that maximize average test scores benefit higher-achieving students and white students more than other groups, the gains are less disparate under counterfactuals that maximize the share of proficient students. These counterfactuals benefit students in the middle of the achievement distribution the most, while students in the two tails of the achievement distribution experience lower gains. This is intuitive and consistent with Neal and Schanzenbach (2010), as classrooms with larger numbers of students in the middle of the achievement distribution are close enough to the achievement threshold and so there is a more marginal gain from reallocating resources toward them. Classrooms with many students in the top quartile are for the most part already proficient, and classrooms with many students in the bottom quartile are harder to turn around. In this case, the achievement gap will shrink for the top three-fourths of the achievement distribution but increase relative to students in the bottom fourth. On the other side, gains for white and non-white students are close across the counterfactuals (Figure 4).

Counterfactual Decomposition. To illuminate what drives the gains in each counterfactual, we decompose the gains in average test scores additively into a portion that is explained by match effects or by teacher general effectiveness. Gains from teacher general effectiveness come via two sources: keeping the least effective teachers out of the classroom and matching the most effective teachers to larger classrooms.

Although match effects matter for student test scores (Table 5), realizing those gains by reassigning teachers to classrooms is unfeasible. Because each student's best teacher match is different and not every student can be assigned their best teacher match simultaneously. We keep students grouped together the same way. Consequently, the decomposition shows that the gains are entirely explained by keeping the most effective teachers and assigning them to larger classrooms, not by matching teachers via their comparative advantage.

While the model of outcomes includes a class-size effect, it does not allow for heterogeneity in teacher effectiveness by class size. To give match effects the best chance, we consider the reassignment that maximizes average classroom outcomes, and hence we

Table 6: Decomposition: Maximize Average Student Test Scores

	Decomposition of Gains Relative to the Observed Assignment (SD)		
	Total Effect	*Effectiveness*	*Match Effects*
Unconstrained	0.079	0.073	0.005
	(0.001)	(0.001)	(0.005)
No teacher harmed	0.072	0.068	0.005
	(0.001)	(0.001)	(0.005)
+ *Assigned teachers only*	0.050	0.047	0.004
	(0.001)	(0.001)	(0.004)
+ *Same school only*	0.004	0.004	0.0002
	(0.0004)	(0.0004)	(0.0002)

Note: Decomposition of gains in average test scores relative to the observed assignment. Standard errors in parentheses.

eliminate any gains from class size. We find half the gain, 4% of a standard deviation, in the unconstrained counterfactual. These gains, again, come exclusively from selecting the most effective teachers. When we consider reassigning only assigned teachers, we find no possible gain at all (Table A6).[20]

7. Conclusions

We used novel data from a market for teacher transfers that allowed us to track both the decisions of teachers and schools in the within-school-district transfer process and the test scores of students from observed assignments. We jointly modeled student outcomes and the decisions of teachers and schools in the transfer system. This model allowed us to account for potential correlation between student outcomes and the decisions of teachers and schools, to account for selection, and to predict teacher effectiveness in unobserved matches. We found some degree of match effects that contribute to about a 10th of the contribution of teacher effectiveness to student test scores. Match

[20]This counterfactual is the one most similar to that in the study by Biasi et al. (2021), though we focus on within-district reassignments and they focus on cross-district ones. Neither allows gains from matching more effective teachers with more students, and both restrict attention to reallocations of teachers in observed assignments. While we use different machinery, we aim at similar objectives and they would produce similar allocations. We directly choose the allocation to maximize average student achievement. They have districts in a flexible-wage competitive equilibrium, maximizing district utility that almost exclusively weights student achievement. A similar objective function that would result in a similar allocation. Like us, they find a very small potential gain here.

effects are non-negligible but matter less than teachers' overall quality. We found evidence that more experienced teachers have a comparative advantage in teaching minority and low-achieving students. Also, we found that having a teacher of the same sex can contribute to learning.

Our estimates show that schools value teachers with general effectiveness but do not value teacher education, experience, nor match effectiveness with their school. This is consistent with Biasi and Sarsons's (2022) estimates that districts value effectiveness but not teacher education nor experience. Teachers, on the other side, are estimated to be averse to their match effectiveness, meaning they place value on working in schools in which they do not have a comparative advantage. Many teachers tend to prefer to work in schools with fewer minority students and low-income students. This, combined with the strong preference schools have for effective teachers, generates unequal assignments. Higher-achieving and white students tend to be assigned more effective and more experienced teachers.

Finally, we found that a reassignment of teachers to classrooms can achieve test-score and proficiency gains. Under a counterfactual in which average student test scores are maximized, high-achieving students and white students benefit more than their counterparts, although all groups experience gains. Alternatively, under a counterfactual in which the share of proficient students is maximized, gains for students in the middle of the test-score distribution are larger than those of students in the top and bottom quarters. Across racial and ethnic groups, gains are similar. Under these counterfactuals, most gains come from assigning more effective teachers to larger classrooms, and little is explained by matching based on comparative advantage. While match effects in teaching exist (see Table 5), our results suggest realizing gains from such effects is challenging.

References

Abdulkadiroğlu, Atila, Joshua D Angrist, Yusuke Narita, and Parag A Pathak, "Research design meets market design: Using centralized assignment for impact evaluation," *Econometrica*, 2017, *85* (5), 1373–1432.

Abdulkadıroğlu, Atıla, Joshua D Angrist, Yusuke Narita, and Parag Pathak, "Breaking ties: Regression discontinuity design meets market design," *Econometrica*, 2022, *90* (1), 117–151.

Agarwal, Nikhil, Charles Hodgson, and Paulo Somaini, "Choices and outcomes in assignment mechanisms: The allocation of deceased donor kidneys," Technical Report, National Bureau of Economic Research 2020.

Aucejo, Esteban, Patrick Coate, Jane Cooley Fruehwirth, Sean Kelly, and Zachary Mozenter, "Teacher effectiveness and classroom composition: Understanding match effects in the classroom," *The Economic Journal*, 2021.

— , — , — , — , and — , "Teacher effectiveness and classroom composition: Understanding match effects in the classroom," *The Economic Journal*, 2022, *132* (648), 3047–3064.

Ba, Bocar, Patrick Bayer, Nayoung Rim, Roman Rivera, and Modibo Sidibé, "Police officer assignment and neighborhood crime," Technical Report, National Bureau of Economic Research 2021.

Bates, Michael D, Michael Dinerstein, Andrew C Johnston, and Isaac Sorkin, "Teacher Labor Market Equilibrium and the Distribution of Student Achievement," Technical Report, National Bureau of Economic Research 2022.

Bergeron, Augustin, Pedro Bessone, John Kabeya Kabeya, Gabriel Z Tourek, and Jonathan L Weigel, "Optimal assignment of bureaucrats: Evidence from randomly assigned tax collectors in the DRC," Technical Report, National Bureau of Economic Research 2022.

Biasi, Barbara, "The labor market for teachers under different pay schemes," *American Economic Journal: Economic Policy*, 2021, *13* (3), 63–102.

— and Heather Sarsons, "Flexible wages, bargaining, and the gender gap," *The Quarterly Journal of Economics*, 2022, *137* (1), 215–266.

— , Chao Fu, and John Stromme, "Equilibrium in the market for public school teachers: District wage strategies and teacher comparative advantage," Technical Report, National Bureau of Economic Research 2021.

Bobba, Matteo, Tim Ederer, Gianmarco Leon-Ciliotta, Christopher Neilson, and Marco G Nieddu, "Teacher compensation and structural inequality: Evidence from centralized teacher school choice in Perú," Technical Report, National Bureau of Economic Research 2021.

Boyd, Don, Hamp Lankford, Susanna Loeb, Matthew Ronfeldt, and Jim Wyckoff, "The role of teacher quality in retention and hiring: Using applications to transfer to uncover preferences of teachers and schools," *Journal of Policy Analysis and Management*, 2011, *30* (1), 88–110.

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff, “Explaining the short careers of high-achieving teachers in schools with low-performing students,” *American economic review*, 2005, *95* (2), 166–171.

— , — , — , and — , “Analyzing the determinants of the matching of public school teachers to jobs: Disentangling the preferences of teachers and employers,” *Journal of Labor Economics*, 2013, *31* (1), 83–117.

Chetty, Raj, John N Friedman, and Jonah E Rockoff, “Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates,” *American Economic Review*, 2014, *104* (9), 2593–2632.

Condie, Scott, Lars Lefgren, and David Sims, “Teacher heterogeneity, value-added and education policy,” *Economics of Education Review*, 2014, *40*, 76–92.

Cowgill, Bo, Jonathan Davis, B Pablo Montagnes, and Patryk Perkowski, “Matchmaking Principals: Theory and Evidence from Internal Labor Markets,” 2021.

Delgado, William, “Heterogeneous Teacher Effects, Comparative Advantage, and Match Quality,” Technical Report, Working Paper 2022.

Dijk, Winnie Van, “The socio-economic consequences of housing assistance,” *University of Chicago Kenneth C. Griffin Department of Economics job market paper, 0–46 i–xi*, 2019, *36.*

Fenizia, Alessandra, “Managers and productivity in the public sector,” *Econometrica*, 2022, *90* (3), 1063–1084.

Gelman, Andrew, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin, *Bayesian data analysis*, CRC press, 2013.

Geweke, John, Gautam Gowrisankaran, and Robert J Town, “Bayesian inference for hospital quality in a selection model,” *Econometrica*, 2003, *71* (4), 1215–1238.

Goldhaber, Dan, Betheny Gross, and Daniel Player, “Teacher career paths, teacher quality, and persistence in the classroom: Are public schools keeping their best?,” *Journal of Policy Analysis and Management*, 2011, *30* (1), 57–87.

— , Vanessa Quince, and Roddy Theobald, “Teacher quality gaps in US public schools: Trends, sources, and implications,” *Phi Delta Kappan*, 2019, *100* (8), 14–19.

Haegele, Ingrid, “The Broken Rung: Gender and the Leadership Gap,” Technical Report, Working Paper 2023.

Hanushek, Eric A and Steven G Rivkin, “Constrained job matching: Does teacher job search harm disadvantaged urban schools?,” Technical Report, National Bureau of Economic Research 2010.

Heckman, James, "Varieties of selection bias," *The American Economic Review*, 1990, *80* (2), 313–318.

Heckman, James J and Edward Vytlacil, "Structural equations, treatment effects, and econometric policy evaluation 1," *Econometrica*, 2005, *73* (3), 669–738.

Hull, Peter, "Estimating hospital quality with quasi-experimental data," *Available at SSRN 3118358*, 2018.

Isenberg, Eric, Jeffrey Max, Philip Gleason, and Jonah Deutsch, "Do low-income students have equal access to effective teachers?," *Educational Evaluation and Policy Analysis*, 2022, *44* (2), 234–256.

Jackson, C Kirabo, "Match quality, worker productivity, and worker mobility: Direct evidence from teachers," *Review of Economics and Statistics*, 2013, *95* (4), 1096–1116.

Kapor, Adam, Mohit Karnani, and Christopher Neilson, "Aftermarket frictions and the cost of off-platform options in centralized assignment mechanisms," Technical Report, National Bureau of Economic Research 2022.

Lewbel, Arthur, "Endogenous selection or treatment model estimation," *Journal of Econometrics*, 2007, *141* (2), 777–806.

Mansfield, Richard K, "Teacher quality and student inequality," *Journal of Labor Economics*, 2015, *33* (3), 751–788.

National Center on Teacher Quality, "Teacher Contract Database," https://www.nctq.org/contract-database/home 2022. Accessed: 2022-11-23.

Neal, Derek and Diane Whitmore Schanzenbach, "Left behind by design: Proficiency counts and test-based accountability," *The Review of Economics and Statistics*, 2010, *92* (2), 263–283.

Prendergast, Canice and Robert H Topel, "Favoritism in organizations," *Journal of Political Economy*, 1996, *104* (5), 958–978.

Rockoff, Jonah E, "The impact of individual teachers on student achievement: Evidence from panel data," *American Economic Review*, 2004, *94* (2), 247–252.

Roy, Andrew Donald, "Some thoughts on the distribution of earnings," *Oxford economic papers*, 1951, *3* (2), 135–146.

Scafidi, Benjamin, David L Sjoquist, and Todd R Stinebrickner, "Race, poverty, and teacher mobility," *Economics of education review*, 2007, *26* (2), 145–159.

Tipton, E. and K. Miller, "Generalizer [Web-tool]," https://thegeneralizer.org 2022. Accessed:

2022-11-23.

Walters, Christopher R, “The demand for effective charter schools,” *Journal of Political Economy*, 2018, *126* (6), 2179–2223.

Gains from Reassignment: Evidence from a Two-Sided Teacher Market

by Mariana Laverde, Elton Mykerezi, Aaron Sojourner, and Aradhya Sood

ONLINE APPENDIX

A. Additional Figures and Tables

Table A1: Correlation of Teacher Effectiveness and District Evaluation Measures

	District Evaluations			
	Math VA	*Reading VA*	*Survey-based Evaluation*	*Standard of Instruction*
Teacher Effectiveness (θ_i)	8.40	5.96	1.47	3.52
	(0.53)	(0.67)	(0.58)	(0.70)
Observations	441	334	432	398

Note: The table shows regressions where the dependent variable is the estimated teacher effectiveness, and the independent variables are 4 measures of teacher quality built by the District. These include their own VA measures constructed using classroom identifiers, as well as measures based on student surveys, and evaluations by peer teachers.

Table A2: Effectiveness and Student Characteristics under the Observed Assignment

	Previous Test Score	*Non-white*	*FRL status*
Teacher Effectiveness	0.50	-0.30	-0.36
	(0.03)	(0.02)	(0.02)
Observations	167,174	167,174	167,174

Note: Each column shows a regression where the independent variable is the estimated teacher effectiveness and the dependent variable is a student characteristic. The data has every teacher-student pair in the observed assignment each year. Standard errors in parenthesis.

Table A3: Model Variables

Student Outcomes	Teacher Utilities	School Utilities
Panel A: Student Characteristics	***Panel A: School Characteristics***	***Panel A: Teacher Characteristics***
Prev. score	Race/Ethnicity students- Share Black	Sex - male
Prev. score sq	Race/Ethnicity students - Share Hispanic	St. Education
Prev. score cube	Race/Ethnicity students- Share other	Race/Ethnicity - Black
Race/Ethnicity - Black	Share FRL	Race/Ethnicity - Hispanic
Race/Ethnicity - Hispanic	Share ELL	Race/Ethnicity - other
Race/Ethnicity - other	Share special education	St. Experience
Sex - male	Average prev. test scores	***Panel B: Teacher-School Interactions***
FRL	Race/Ethnicity teachers- Share Black	Match minority students
ELL	Race/Ethnicity teachers - Share Hispanic	Match minority teachers
Special education	Race/Ethnicity teachers- Share other	Teach St. Educ * Share stud minority
Panel B: Teacher Characteristics	Average teacher experience	Teach St. Educ * Average prev. test score
Sex - male	***Panel B: Teacher-School Interactions***	Teach St. Exp * Share stud minority
St. Education	Match minority students	Teach St. Exp * Average prev. test score
Race/Ethnicity - Black	Match minority teachers	Teach St. Exp * Average st. experience
Race/Ethnicity - Hispanic	Teach St. Educ * Share stud minority	
Race/Ethnicity - other	Teach St. Educ * Average prev. test score	
Experience - 2 and 3 years	Teach St. Exp * Share stud minority	
Experience - 4 and 6 years	Teach St. Exp * Average prev. test score	
Experience - $\geq$7 years	Driving minutes	
Panel C: School Characteristics	***Panel C: Other***	
Race/Ethnicity - Share Black	Inertia	
Race/Ethnicity - Share Hispanic	School FE	
Race/Ethnicity - Share other		
Share FRL		
Share ELL		
Share special education		
FRL*Share FRL		
Classsize: Tot FTE per student		
Panel D: Student-Teacher Interactions		
Match minority		
Match gender		
Teach St. Educ * Stud minority		
Teach St. Educ * Stud prev. score		
Teach St. Exp * Stud minority		
Teach St. Exp * Stud prev. score		
Panel E: Other		
MCAS III		
Year FE		
School FE		

Note: For race/ethnicity the ommited category is white. In the middle of our sample the standardized test was redesigned, we capture the change with the dummy variable MCAS III. Minority is defined as not identifying white.

Table A4: Counterfactual Summary: Maximize Average Test Scores

	Scenario 1	*Scenario 2*	*Scenario 3*	*Scenario 4*
Constraints				
No Teacher is harmed	N	Y	Y	Y
Only assigned teachers	N	N	Y	Y
Same school assignments only	N	N	N	Y
All	0.079	0.072	0.05	0.004
	(0.001)	(0.001)	(0.001)	(0.0004)
By Achievement				
First quartile	0.069	0.062	0.037	0.003
	(0.001)	(0.001)	(0.001)	(0.001)
Second quartile	0.071	0.065	0.042	0.004
	(0.001)	(0.001)	(0.001)	(0.001)
Third quartile	0.081	0.075	0.054	0.004
	(0.001)	(0.001)	(0.001)	(0.001)
Fourth quartile	0.093	0.086	0.068	0.005
	(0.002)	(0.002)	(0.002)	(0.001)
By Race/Ethnicity				
Non-white	0.071	0.065	0.041	0.005
	(0.001)	(0.001)	(0.001)	(0.001)
White	0.091	0.084	0.066	0.003
	(0.002)	(0.002)	(0.002)	(0.0004)

Note: Average test score gains relative to the observed assignment for all students, and students by baseline achievement and race/ethnicity, in a counterfactual where average student test scores are maximized. Standard errors in parenthesis.

Table A5: Counterfactual Summary: Maximize Share Proficient

Constraints				
No Teacher is harmed	N	Y	Y	Y
Only assigned teachers are considered	N	N	Y	Y
Same school assignments only	N	N	N	Y
All	3.93	3.54	2.84	0.32
	(0.07)	(0.07)	(0.07)	(0.03)
By Achievement				
First quartile	0.69	0.59	0.44	0.05
	(0.06)	(0.06)	(0.05)	(0.02)
Second quartile	5.3	4.69	3.64	0.41
	(0.15)	(0.15)	(0.13)	(0.06)
Third quartile	7.51	6.82	5.55	0.62
	(0.18)	(0.19)	(0.2)	(0.08)
Fourth quartile	2.2	2.04	1.73	0.19
	(0.11)	(0.1)	(0.1)	(0.05)
By Race				
Non-white	3.92	3.48	2.75	0.36
	(0.09)	(0.09)	(0.08)	(0.03)
White	3.93	3.62	3	0.24
	(0.13)	(0.12)	(0.13)	(0.04)

Note: Gains in share proficient relative to the observed assignment for all students, and students by baseline achievement and race/ethnicity, in a counterfactual where the share of proficient students is maximized. Standard errors in parenthesis.

Table A6: Decomposition: Maximize Average Classroom Test Scores

	Decomposition of Gains Relative to the Observed Assignment (SD)		
	Total Effect	*Effectiveness*	*Match Effects*
Unconstrained	0.038	0.032	0.006
	(0.001)	(0.002)	(0.005)
No teacher harmed	0.035	0.03	0.005
	(0.001)	(0.001)	(0.005)
+ *Assigned teachers only*	0.002	−0.002	0.004
	(0.002)	(0.002)	(0.004)
+ *Same school only*	−0.0004	−0.0006	0.0003
	(0.0005)	(0.0005)	(0.0003)

Note: Decomposition of gains in average test scores relative to the observed assignment. Standard errors in parenthesis.

B. Model and Estimation

Students are denoted by k, year by t, teacher by i, school by s, and grade-school by c. And let K, T, N, and P be the size of the sets respectively.

The model consists of the following equations. The first one describes students outcomes y_{kt}, the second describes teacher supply u_{ist}, the third describes school demand v_{ist}, and the last one describes school beliefs,

$$y_{kt} = C_0(\omega_{kt}, x_{st}, \tau_{it})\alpha^y + \theta_i + \eta^y_{is} + \varepsilon^y_{kt} \tag{4}$$

$$u_{ist} = C_1(x_{st}, \tau_{it}, z^u_{ist})\alpha^u + \eta^u_{is} + \gamma I_{ist} + \varepsilon^u_{ist} \tag{5}$$

$$v_{ist} = C_2(x_{st}, \tau_{it}, z^v_{ist})\alpha^v + \varphi_i + \eta^v_{is} + \varepsilon^v_{ist} \tag{6}$$

We rewrite η_{is} and (θ_i, φ_i) as

$$\begin{aligned}
\eta^y_{is} &= f_{is,1} \\
\eta^u_{is} &= \beta^u_1 f_{is,1} + f_{is,2} \\
\eta^v_{is} &= \beta^v_1 f_{is,1} + \beta^v_2 f_{is,2} + f_{is,3} \\
\varphi_i &= f_{i,4} \\
\theta_i &= \beta_4 f_{i,4} + f_{i,5}
\end{aligned}$$

where $f_{is,1} \sim N(0, \sigma^2_1)$, $f_{is,2} \sim N(0, \sigma^2_2)$, $f_{is,3} \sim N(0, \sigma^2_3)$, $f_{i,4} \sim N(0, \sigma^2_4)$, $f_{i,5} \sim N(0, \sigma^2_5)$.

Then the model parameters are:

$$\begin{aligned}
\kappa &= (\alpha^y, \alpha^u, \alpha^v, \alpha^\pi, \gamma, \beta^u_1, \beta^v_1, \beta^v_2, \beta_4) \\
\sigma &= (\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_{\varepsilon y})
\end{aligned}$$

Gibbs Sampler

Start with values of κ^0 and σ^0 from diffuse conjugate priors, and values f^0, u^0, v^0, π^0 and ξ^0 for the latent variables. u^0, v^0, and π^0 must be consistent with ranking, interview, and application decisions.

Step 1: Data Augmentation

In this step we update the values of the latent variables u^1 and v^1 given the values of the parameters of the model, the rest of the realizations of the latent variables, and the application, interview and ranking decisions of schools and teachers.

Step 2: Update κ conditional on $u^1, v^1, f^0, \sigma^0, y^1$

Step 3: Update $\sigma^2_{\varepsilon y}$

Step 4: Update the f's and ν_c

Step 5: Update the $\{\sigma_i\}^5_{i=1}$ and σ^2_c