HCEO WORKING PAPER SERIES

Gains from Alternative Assignment? Evidence from a Two-Sided Teacher Market

Mariana Laverde
Elton Mykerezi
Aaron Sojourner
Aradhya Sood

Working Paper

2025-001

02-2025

HUMAN CAPITAL AND ECONOMIC OPPORTUNITY GLOBAL WORKING GROUP

The University of Chicago
1126 E. 59th Street Box 107
Chicago IL 60637

www.hceconomics.org

Gains from Alternative Assignment? 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

February 7, 2025

Abstract

The literature on assignment mechanisms largely focuses on efficiency based on agents' preferences, though policymakers may prioritize different goals. In assigning teachers to classrooms, a school district might prioritize student learning but must also consider teacher welfare. This paper studies the potential gains in student test scores from alternative within-district assignments of teachers to classrooms, using novel administrative data on teacher and school principal decisions from the district's internal transfer system (ITS) and student test scores under the observed assignments. To credibly predict student test scores under unrealized assignments, we jointly model student achievement and teacher and principal decisions, accounting for potential selection of teachers on test score gains. We estimate the variation in teachers' comparative advantage in producing learning to be one-ninth the magnitude of the variation in their general effectiveness. Further, teachers dislike comparative advantage–based assignments. Assignment of teachers to classrooms to maximize learning under the constraint of not reducing any assigned teacher's welfare would raise the average test score by 7% of a standard deviation (SD) relative to that under the observed assignment, with this effect driven mostly by assignment of teachers with higher general effectiveness to larger classrooms rather than by harnessing teachers' comparative advantage.

*This paper benefited from discussions with Nikhil Agarwal, Barbara Biasi, Jesse Bruhn, Chris Campos, William Delgado, Michael Dinerstein, Chao Fu, Michael Grubb, Peter Hinrichs, Caroline Hoxby, Adam Kapor, Charlie Murry, Minseon Park, Jonah Rockoff, Stephen L. Ross, Douglas Staiger, Sukjoon Son, Yang Song, Chris Walters, Roman A. Zarate, Seth Zimmerman and seminar participants at the International Industrial Organization Conference, NBER Education Meetings, Northeast Economics of Education Workshop, Stanford Institute for Theoretical Economics (SITE) Empirical Market Design Conference, Southern Economic Association Conference, Society of Labor Economists Conference, Washington University at St. Louis, and Western Economic Association Conference. This project relied on financial support from the McKnight Foundation. This study falls under the University of Minnesota IRB's project 1510S79046. Emails: laverdem@bc.com, myker001@umn.edu, sojourner@upjohn.org, and aradhya.sood@rotman.utoronto.ca.

1. Introduction

Teachers impact student learning more than do other school-based factors (Chetty et al., 2014b; Chetty and Hendren, 2018; Hanushek, 2020). Teachers are unequally effective in teaching all types of students: Their performance may vary with student race, income, or achievement (Delgado, 2023; Bates et al., 2024; Biasi et al., 2022), suggesting that assignment of teachers to students affects average student learning. What gains in student learning could districts achieve by harnessing teachers' comparative advantage? What distributional consequences would arise from these alternative assignments in terms of race- and income-based achievement gaps? How does the addition of constraints that would make reforms of the teacher assignment policy more politically feasible, such as not harming any teacher's welfare, affect the potential gains in student learning?

Teachers may have private information about their own comparative advantage, and school principals might be able to accurately identify the types of teachers who best fit their student populations. If the goal of principals and teachers is to maximize student learning, then the status quo mechanism, namely, decentralized assignment of teachers to classrooms, should lead to optimal student outcomes. However, if principals and teachers are not well informed about the quality of a teacher–classroom match, or if their objectives deviate (because, beyond student outcomes, teachers value other job amenities and principals value other aspects of the match), then the decentralized assignment may not be optimal for student learning. Understanding the potential gains from alternative assignment of teachers to classrooms and how these gains relate to teacher and principal decisions can inform policies that facilitate assignments that maximize student outcomes.

This paper quantifies the potential gains in student test scores from alternative assignments of teachers to classrooms within a school district. It compares the performance and distributional outcomes of counterfactual teacher–classroom matches to their observed counterparts and evaluates the impact of teacher and school principal decisions on student achievement. We find that assignment of teachers to classrooms under the constraint of not reducing any assigned teacher's welfare leads to an average test score gain of up to 7% of a standard deviation (SD) relative to the status quo, driven mostly by assignment of teachers with higher general effectiveness to larger classrooms rather than by gains from comparative advantage. We focus on within-district transfers because these are relatively easy to facilitate: They do not require the district to change its compensation policies or any changes to teachers' health insurance, tenure status, seniority, or pension rights.

A key challenge we face in estimating the test score gains under counterfactual teacher assignments is measuring teacher effectiveness in matches that have not actually occurred. Teachers' effectiveness in observed matches might not represent their counterfactual effectiveness elsewhere if teachers tend to apply to (or avoid) positions in which they have a comparative advantage or if principals tend to select teachers on the basis of such comparative advantage.[1] For an analysis of this kind, imposing the assumption of exogenous mobility of teachers across schools, as Chetty et al. (2014a) do, would be convenient, but some evidence weighs against it. For example, Feng and Sass (2017) find that a teacher's overall effectiveness plays a role in her decision to leave a position, and (Jackson, 2013) finds that teachers tend to systematically move to schools where they have better match quality. Rather than imposing this assumption, we leverage over 10 years of data from a large urban school district's internal transfer system (ITS) and longitudinal, student-level achievement data to estimate a joint model of student outcomes, teacher labor supply to schools, and school principal demand for teachers, allowing correlation between student potential outcomes and teacher and principal decisions, in the spirit of Roy (1951).

The model of student potential outcomes captures teacher effectiveness as the sum of a general-effectiveness component applying across all students and match effects between teachers and students. Similarly to how Agarwal et al. (Forthcoming) approach the kidney transplant market, we include a rich set of student, teacher, and match-specific observables and allow for match effects on unobservables that capture differences in effectiveness not traceable to observable student types. We model teachers' and principals' decisions using a latent variables framework and capture 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, which are

[1]Hiring authority rests with the principal. Though principals may delegate selection work to a team, they retain final decision-making power.

also part of the teacher and school principal decision models. Second, teacher and school principal decisions depend on teacher×school idiosyncratic tastes, which may correlate with the unobserved match effects in the student outcomes model. Both of these margins capture selection on test score gains. Last, principals' decisions may be correlated with teachers' general effectiveness.

The identification of our model relies on (1) the assumption of conditionally independent assignment of teachers to students within, but not across, schools and (2) two shifters that separate teacher and principal decisions from student outcomes.[2] We use the driving time from a teacher's home to each school as the shifter of teacher decisions because driving time is correlated with teacher supply decisions but is credibly independent of student outcomes and principal hiring decisions. For the shifter on principal decisions, we use a measure of unexpected need to hire. We assume that an unexpected increase in the need to hire reduces a principal's pickiness uniformly across teacher applicants but that student potential outcomes do not depend on a principal's need to hire.

We estimate this three-equation model using Bayesian inference and a Gibbs sampler as in Geweke et al. (2003). This allows us to leverage the data's nested structure to model teachers' general effectiveness and the unobserved teacher–school match effects using hierarchical priors so that information on student test scores can be shared among teachers. The 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 ordinary least squares (OLS) estimation.

We find that teachers' general effectiveness is approximately 9 times more important than teachers' comparative advantage in explaining student potential outcomes. The gains from comparative advantage stem mostly from match effects arising from teacher–student complementarities on observables, while we find that the role of match effects on unobservables is limited. Teachers tend to prefer schools with a higher share of high-income students, as found in Boyd et al. (2011, 2013), and are averse to assignments in the schools where they would be most effective. On the other hand, principals' decisions are

[2]This approach follows ideas and applications developed in Geweke et al. (2003), Heckman and Vytlacil (2005), Lewbel (2007), Hull (2020), and Agarwal et al. (Forthcoming).

uncorrelated with teachers' education, experience, or comparative advantage, but they do value teachers' general effectiveness.

We examine different counterfactual assignments of teachers across positions with attention to constraints that affect the political feasibility of the implied reforms. The counterfactuals studied are all constrained to alternative assignment only of the district's teachers, not teachers from outside the district, and to maintenance of any teacher's compensation constant regardless of assignment in the district or performance. In the counterfactual in which average student achievement is maximized under the further constraint that no assigned teacher's welfare is reduced (the "no-quits" counterfactual), average test scores could increase by up to 7% of an SD over the test scores under the observed assignments. This figure is only slightly below the potential gains without the no-quits constraint of 8% of an SD. Crucially, these gains come primarily from matching teachers with higher general effectiveness to larger classrooms rather than from leveraging teachers' comparative advantage with specific students.

The limited role of teacher comparative advantage in raising test scores is explained both by the larger impact of general effectiveness than of comparative advantage on student outcomes and by the constraint that teachers be reassigned given the observed student classroom compositions. While this assignment would raise average achievement in the school district, more-advantaged students would benefit more, which implies an efficiency–equity trade-off. All groups would experience increases in average test scores under this counterfactual assignment, but the gains would be larger for higher-achieving students and White students. In other words, maximizing achievement and raising all student groups' achievement would widen race- and income-based achievement gaps.

Literature Review: Our paper relates to a recent literature on teacher comparative advantage that uses various notions of optimality to evaluate observed teacher assignments to schools and classrooms (Condie et al., 2014; Ahn et al., 2024; Aucejo et al., 2022; Delgado, 2023; Biasi et al., 2022; Bates et al., 2024; Ahn et al., 2024; Umosen, 2024). While these papers typically estimate comparative advantage over a single, dichotomous observable student characteristic or do not allow for match effects from unobservables,[3] our model estimates a comprehensive measure of comparative advantage by allowing variation in effectiveness over a rich set of student observables and unobservables. Other papers addressing solutions to unequal access to teachers but that do not emphasize teacher comparative advantage include Combe et al. (2022) and Bobba et al. (2024).

Relative to the existing literature, our paper emphasizes the potential for bias in estimated outcomes of unrealized matches due to selection of teachers into schools. We leverage detailed data from a school district's ITS, which governs the assignment of teachers to positions in the district. Since we observe the set of open positions to which each teacher could apply, teachers' choices of whether to apply, principals' choices to offer interviews and jobs to applicants, and applicants' decisions of whether to accept each offer, i.e., we know each teacher's and principal's complete choice set and choices in each stage of the process (application, interview, offer, and acceptance), we have a rare opportunity to disentangle teacher decisions from principal decisions and their correlation with student outcomes. Stronger assumptions are required for identification if the econometrician observes only realized matches (He et al., 2024). Moreover, with credible estimates of teacher effectiveness and utility in hand, we can consider counterfactual assignments that would create value for both teachers and students, which would make such assignments implementable and politically feasible.

Much of the teacher value-added literature has focused on potential gains from terminating teachers revealed to be at the bottom of the general-effectiveness distribution (Staiger and Rockoff, 2010; Chetty et al., 2014a; Rothstein, 2015). By accounting for teacher comparative advantage with different kinds of students, beyond general effectiveness, and considering alternative assignments in addition to terminations, the counterfactuals modeled in this paper correspond to a more complex set of policy options. In addition, while a large literature focuses on assigning students to schools within a district, our paper looks at this assignment problem from the other side, i.e., moving teachers around the district while keeping the students fixed. This approach can be beneficial as moving students is

[3] Ahn et al. (2024) and Umosen (2024) estimate multidimensional value-added models that capture match effects over several observable student characteristics, but do not include unobservable match effects.

costly to families and school districts (Laverde, 2024; Angrist et al., 2024), and its impacts on student outcomes is mixed (Campos and Kearns, 2024; Angrist et al., 2024; Deming, 2011; Deming et al., 2014).

Similarly to Van Dijk (2019), Agarwal et al. (Forthcoming), Kapor et al. (2024), and Abdulkadiroğlu et al. (2017), our paper evaluates the outcomes-based performance of assignment mechanisms rather than focusing on welfare criteria alone. While traditional market design emphasizes welfare measures in evaluating assignment mechanisms, policymakers may have different considerations. In our case, school districts may want to affect student learning but, crucially, must consider teacher welfare when assigning teachers to classrooms. More generally, our paper relates to research on optimal assignment of talent within organizations and the public sector. This includes allocations of police officers to neighborhoods (Ba et al., 2021), of surgeons to hospitals (Mourot, 2025), and of managers, workers, and bureaucrats to teams and tasks (Osterman, 1984; Prendergast and Topel, 1996; Minni, 2024; Fenizia, 2022; Cowgill et al., 2024; Davis et al., 2023).

2. Background, Data, and Empirical Evidence

2.1 Institutional Details

Our study focuses on a large, diverse urban school district in the U.S. Midwest. According to our analysis from the National Center on Teacher Quality (2022), seven in 10 large U.S. school districts, including the district we study, delegate authority to school leadership to choose who will fill open teaching positions at a given school. Open teaching positions can arise because of the creation of new positions, retirements, or quits or firings of incumbent teachers. Incumbent teachers within the school district typically fill most open positions through the internal transfer mechanism. External candidates can also fill vacancies, but only after internal candidates are considered, as specified in the collective bargaining agreement of the school district.[4]

This paper focuses on the scope for student learning gains through within-district transfers of incumbent staff preserving the policy that teacher salary not vary across schools in

[4]At least 40% of large U.S. school districts explicitly prioritize internal transfer candidates over external hires when filling open positions (National Center on Teacher Quality, 2022).

the same district, one of the less disruptive and more politically feasible changes among the menu of possible reforms to teacher assignment policy. Many institutional forces make it unattractive to teachers to change districts. In addition to the usual disruption from changes to their health insurance and perhaps pension plans, teachers' seniority and tenure status may not transfer to another school district.[5]

Guided by a collective bargaining process with the teachers' union, the management of the school district that we study created a centralized ITS to govern its matching process and serve as a clearinghouse of its internal labor market. The process comprises two successive rounds of applications, interviews, and offers in the spring of each academic year. Each round involves the following steps. First, based on district projections of school enrollments and budgets and incumbent teachers' commitments to retire or take leave, each school posts its vacancies on the ITS for the coming year. Second, any incumbent teacher can apply to any vacancies within the school district for which her licenses qualify her.

After the application window closes, the school district checks each applicant's eligibility for the positions to which she has applied, and the system automatically grants interviews to the four most senior applicants for each vacancy, per the collective bargaining agreement. Then, for each vacancy, each school principal reviews these automatic interviewees and the remaining applicant pool and can choose up to four additional applicants for interview, yielding a maximum of eight interviewees per position. While schools can abstain from interviewing any of the four most senior applicants, they would lose the option to make an offer to any applicant outside this group. If they decide to interview only a subset of the most senior applicants, they have to invite them in order of seniority. When principals 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.

After interviews, each principal can submit a ranking of up to four interviewees to the

[5]In contrast, Biasi et al. (2022) consider cross-district teacher mobility and differentiation of a teacher's pay across schools in the same district. Bates et al. (2024) also study the effect of introducing bonus pay in a similar two-sided matching model.

district via the ITS. Next, the system automatically and simultaneously emails offers to the first-ranked interviewee for each position across all principals in the district. Thus, an applicant can receive zero, one, or multiple offers at this stage. Offerees have 48 hours to accept up to one offer. After 48 hours, the ITS 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 or all positions are filled. 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 the second round. The whole application, interview, ranking, and offer process repeats again. After the second round, any vacancies become open to both external and internal candidates.

2.2 Administrative 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, interview decisions, rankings, offers, and acceptances. In addition, for each teacher, we observe seniority rank, experience, education, ethnicity, race, gender, and current position assignment for every year. We also observe each teacher's home address each year, allowing us to measure the driving time to each open position in a teacher's choice set.[6]

A position assignment is a school–grade–position type combination, such as "third-grade math teacher." Because we observe teacher licenses, we can accurately measure each teacher's choice set each year. Accurately defining the choice set is important, as emphasized by Almagro and Sood (2025), who show that inaccurately specifying choice sets leads to biased utility estimates. To decide which positions to include, we use the district's internal formal value-added system as a guide, including positions that the district evaluates for math and reading effectiveness and excluding others.[7]

We restrict attention to grades 4–8 to measure teacher effectiveness most reliably. This

[6]We geocode teacher and school addresses and measure driving times using the Google Maps API.

[7]This leads us to include the following position types: 7th Grade Math, Elementary Math, Algebra/ Integrated Math, General Elementary Education, Elementary Reading, Secondary Reading, and Comprehensive Language Arts.

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 position choice menu		36.5
Applications submitted		6.7
Number of interviews		3.4
Number of times ranked		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: Panel A shows the mean or percentage 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 in the right column. For teachers who are ever found in the ITS, we average their characteristics for every year they are in the sample between 2010 and 2019, including those years they are not found in the ITS. Panel B shows the mean characteristic for teachers who applied to an open position in the sample, while Panel C shows the mean characteristic of an open position.

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 these grades and the prior grades. We drop teachers working less than half time to ensure that a teacher spends a significant number of hours with the matched students.

The sample contains 823 teachers. They are observed in 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 between 2010 and 2019. Two-thirds of the teachers are women. Most—about 80%—are White, followed by 9% who are 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 positions for which they are eligible (Panel B). On average, applicants receive interviews for 3 positions, are ranked by principals for 1.5 positions, and receive 0.8 offers.

On the other side of the market, the ITS included 972 vacancy postings for grades 4–8 from 2010 to 2019.[8] Though 214 teachers are eligible for the average posting, each position on average attracted approximately 5 applicants (Panel C). The propensity of qualified teachers not to apply motivates the inclusion of an inertia term in our model. An average of 2.5 applicants were interviewed and 0.6 offers made for each position (Panel C).

For each student in each academic year, we observe state-mandated standardized test scores in math and reading and their grade, school, demographic characteristics such as race, ethnicity, and gender, indicators of English language learner status, special education status, and eligibility for free or reduced-price lunch (FRPL), which proxies for low household income. The sample includes 35,608 students in grades 4 through 8 in 54 schools between 2009 and 2019. Approximately a third of the students are Black, and a third are White. Hispanic students constitute approximately 20% of the sample. In contrast, more than 80% of teachers are White. Last, for each school–year, we observe teacher, school, principal, and student characteristics that potential hires may value or that may influence the principal's taste for hire types.[9]

2.3 Empirical Evidence

To motivate our model and counterfactual exercises, we begin by examining the observed assignment of teachers to schools and the transfer patterns of teachers over their careers.

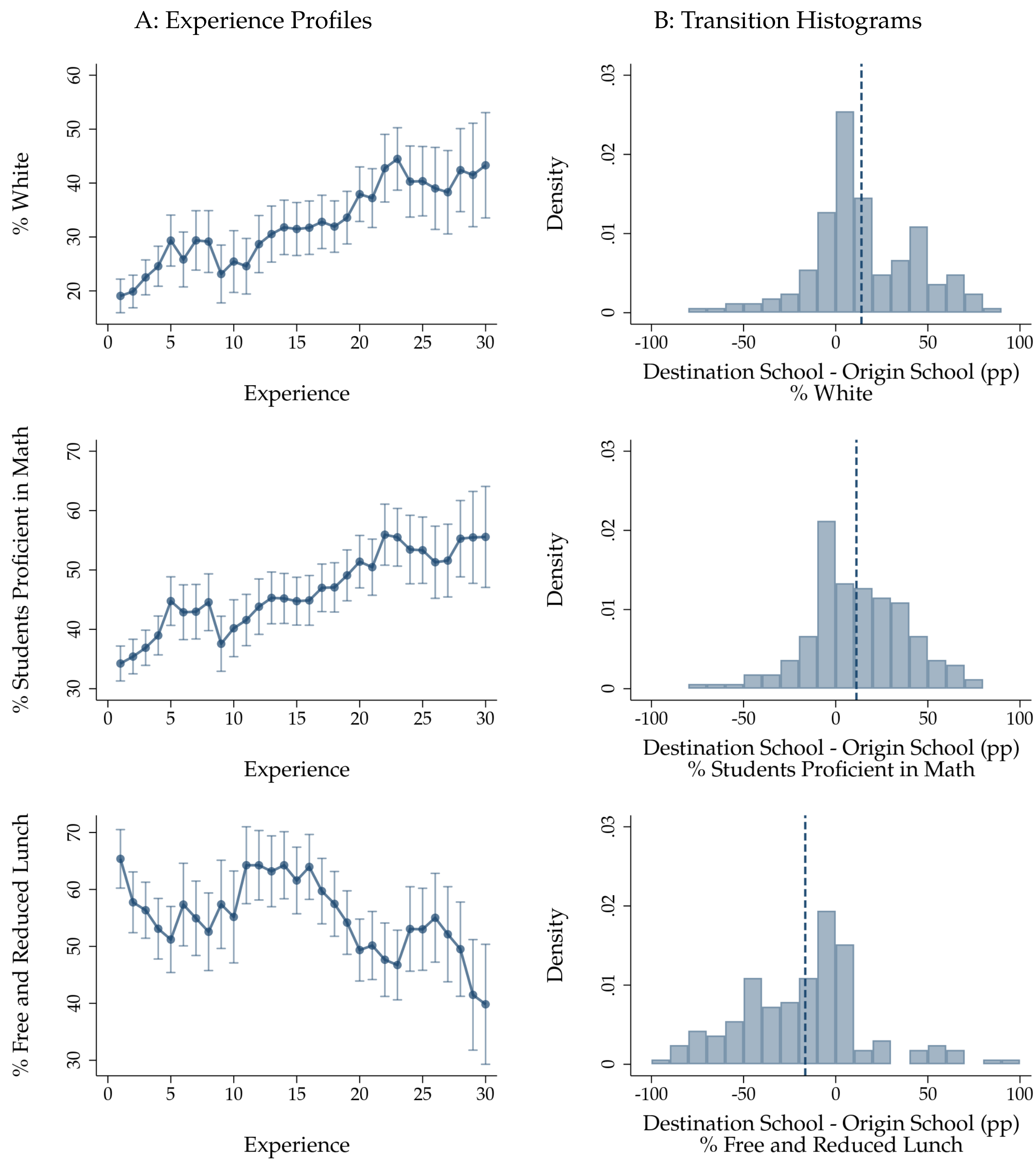
[8] Vacancy postings are restricted to positions requiring more than a half-time schedule.
[9] See Appendix Table D.1 for additional school-level descriptive statistics.

In Figure 1, Panel A, we find that 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 mathematics, and a lower share of low-income students. The decisions of teachers and schools during the transfer process partially explain the patterns observed in the figure, although other processes, such as initial matches and differential retirements, may also contribute to these patterns.

Focusing on teachers who transferred schools between 2010 and 2019, we observe that teachers tend to move to schools with a higher share of White students, a higher share of students proficient in mathematics, and a lower share of low-income students than those of their prior school (Figure 1, Panel B), although there is considerable heterogeneity across these measures. This finding is consistent with the evidence in the literature that teachers of more disadvantaged students are likelier to transfer to other schools (Boyd et al., 2005; Scafidi et al., 2007; Goldhaber et al., 2011; Isenberg et al., 2022; Goldhaber et al., 2019).

The observed teacher–school assigned 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. Without a structural model, the role of teachers' decisions cannot be disentangled from that of principals' decisions in producing the observed assignments reflected in Figure 1. In section 3, we build a two-sided matching model to disentangle the joint decisions of principals and teachers in the transfer process and their correlation with student outcomes. In addition, even if teachers have incentives to sort into positions in which they have the most impact and principals want to hire the teachers with the largest impact on their students, these are unlikely to be the only considerations valued by either of the agents. Principals' lack of discretion over the setting of wages within the school district or laying-off of less effective tenured teachers may also push the realized assignments away from the student achievement–maximizing assignments. Our model below quantifies the degree to which sorting patterns are correlated with student gains in this market, what an assignment that maximizes student achievement would look like, and how test score gains can be produced under alternative assignments. Our two-sided

Figure 1: Teacher Transition Characteristics



Note: Panel A shows the mean and 95% confidence intervals of a school characteristic at the school to which teachers are assigned with each year of experience. Panel B restricts the sample to teachers who changed schools between 2010 and 2019, and it plots the histogram of the difference in school characteristics between the destination and origin schools. The vertical line is the mean difference. "pp" stands for percentage points.

matching model jointly models student achievement and teacher and principal decisions to determine the alternative teacher assignments that maximize student outcomes.

3. Model

We jointly model students' outcomes and teachers' and school principals' decisions to incorporate the potential for selection of teachers into schools. Equation 1 describes the learning outcome of student $k \in \mathcal{K}$ when she is assigned to teacher $i \in \mathcal{I}$ at school $s \in \mathcal{S}$.

$$Y_{ki} = f_Y(w_k, v_i, x_s, \theta_i, \eta^y_{is}) \tag{1}$$

where w_k, v_i, and x_s are the observable student, teacher, and school characteristics, respectively. The model captures the effects of teacher–student matching on observables by including interactions between the characteristics in w_k and v_i. θ_i is teacher i's unobserved general (context-independent) effectiveness net of the effect of the characteristics in v_i. η^y_{is} is the match effect on unobservables between teacher i and the students in 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 school s. We assume that the teacher–school match effects are constant over time.

We model the decisions of teachers and principals in the transfer process using a latent variables framework, representing both teacher utility and principal expected utility. Equation 2 describes teacher i's utility from an assignment in school s. U_{is} is a function of the observable teacher and school characteristics and the interactions between these. z^u_{is} is a vector of the observable teacher–school-level characteristics that are excluded from the outcomes model and the principal decisions model. η^u_{is} represents an idiosyncratic taste shock of teacher i for school s.

$$U_{is} = f_U(v_i, x_s, z^u_{is}, \eta^u_{is}) \tag{2}$$

$$D_{is} = f_D(v_i, x_s, z^v_{is}, \varphi_i, \eta^d_{is}) \tag{3}$$

Equation 3 describes school s's willingness to make a job offer to teacher i. D_{is} depends on observable teacher and school characteristics and interactions between them. z_{is}^{v} is a vector of observable teacher–school characteristics that are excluded from the outcomes model. This model also includes a teacher-level unobservable term, φ_i, capturing the unobserved attractiveness of teacher i that is common across schools and time, and a shock to school s's taste for teacher i, η_{is}^{d}.

Sources of Selection. The interactions between teacher and school observable characteristics in both of the choice models (Equations 2 and 3) capture the heterogeneity in the same dimensions as the model of student outcomes (Equation 1). The parameters associated with these interactions allow us to capture selection on student gains via several observable components. For example, the models allow for a differential impact of teachers on students depending on whether both share the same minority status. Similarly, the model allows school principals to have a taste for teachers of the same minority status as the majority of the student body and for teachers to prefer to teach at schools with larger shares of students who share a teacher's minority status.

Similarly to Agarwal et al. (Forthcoming) but in contrast to the previous education literature, on the unobserved side, we allow for correlations between the teacher–school match effects in student outcomes and idiosyncratic shocks in teacher and principal decision models ($\eta_{is}^{y}, \eta_{is}^{u}, \eta_{is}^{v}$). It is important to allow for match effects from unobservables in the model because while it includes a rich set of characteristics to capture match effects from observables, these may still miss some unobserved dimensions that give rise to comparative advantage. A positive correlation between η_{is}^{y} and η_{is}^{u} would imply that teachers tend to value schools at which they have a comparative advantage. A positive correlation between η_{is}^{y} and η_{is}^{v} would imply that principals tend to make job offers to teachers who are especially effective at teaching their particular students. Moreover, we allow for a correlation between a teacher's general effectiveness, θ_i, and principals' common unobserved taste for the teacher, φ_i. A positive correlation between θ_i and φ_i would indicate that principals tend to value teachers with higher levels of general effectiveness.

3.1 Model Parametrization

The parametrized version of Equation 1 is

$$y_{kit} = w_{kt}\alpha + v_{it}\mu^y + x_{st}\zeta^y + c_{kit}\beta^y + \theta_i + \eta^y_{is} + \varepsilon^y_{kt} \tag{4}$$

where y_{kit} is the standardized test score of student k in year t when she is assigned to teacher i. Similarly to the traditional value-added models, student outcomes are a function of a rich set of observable student and school characteristics that may change over time (w_{kt}, x_{st}). w_{kt} includes a flexible polynomial on the student's past test scores and demographic characteristics, and x_{st} includes averages of classroom characteristics.[10]

The model also includes observable teacher characteristics (v_{it}), namely, teacher experience, education, and race or ethnicity. The vector c_{kit} includes interactions between selected characteristics in w_{kt} and v_{it} and captures the match effects in teaching arising from observable teacher and student characteristics.

In addition to the interactions on observable characteristics, the model includes the unobserved teacher general effectiveness, in θ_i, and the match effects from unobservables at the teacher–school level, η^y_{is}. θ_i and η^y_{is} are assumed to be normally distributed with a mean of zero and variances that we estimate. ε^y_{kt} is a shock specific to the student and year. $\chi^y = (\alpha, \mu^y, \zeta^y, \beta^y)$ denotes the vector of coefficients in the outcomes model.

The parametrized version of Equations 2 and 3 is given by

$$u_{ist} = x_{st}\zeta^u + q_{ist}\beta^u + z^u_{ist}\phi^u + \gamma I_{ist} + \eta^u_{is} + \varepsilon^u_{ist} \tag{5}$$
$$d_{ist} = v_{it}\mu^d + q_{ist}\beta^d + z^d_{ist}\phi^d + \varphi_i + \eta^d_{is} + \varepsilon^d_{ist} \tag{6}$$

where u_{ist} is a function of the school's observable characteristics. The interactions between school and teacher characteristics in the vector q_{ist} map to interactions in the model of outcomes. z^u_{ist} is the driving time between teacher i's home and a school s. This driving time serves as a supply shifter and is included in the model of teacher decisions but is excluded from the model of student outcomes and the model of principal decisions. Section 4 discusses the assumptions and identification arguments behind this shifter. The model also includes school fixed effects, capturing each school's unexplained attractiveness to

[10] A list of all the variables included in the model can be found in Appendix Table D.4.

teachers that is common across teachers over time.

Because teachers face the decision of whether to apply for a transfer within the school district each year, we include an inertia term in teacher utility that captures the cost to teachers of changing jobs. This includes the value teachers place on not having to update their CV, change their routines, 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 inertia cost is given by the parameter γ. 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. Let $\chi^u = (\zeta^u, \beta^u, \phi^u, \gamma)$ denote the vector of coefficients in the teacher's decisions model.

d_{ist} depends on the observable teacher characteristics and interactions between school and teacher characteristics in the vector q_{ist}. The unobservables η^d_{is} and φ_i are assumed to be normally distributed with a mean of 0 and variances to be estimated. Let $\chi^d = (\mu^d, \beta^d, \phi^d)$ denote the vector with the coefficients in the teacher decision model.

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

The parameters to be estimated are the coefficients of the observable covariates in χ^y, χ^u, and χ^d, the variance–covariance matrices $\Sigma_{\theta\varphi}$ and Σ_η, and the variance $\sigma^2_{\varepsilon y}$.

4. Identification and Estimation

4.1 Mapping the Model to the Data

We estimate the model parameters using the decisions of teachers and principals in the transfer system and the test scores of students under the observed teacher–classroom matches.[11] While we have described our model as one in which teachers and principals

[11]The data actually reveal matches only at the school–grade–year level. If this level contains observations from multiple classrooms, all the assigned teachers, that is, the school–grade–year team, contribute to the

are on each side of the market, in practice, each school can have more than one position open in the ITS concurrently. Consequently, we estimate a model at the position level such that teachers have preferences over positions within a school and school principals' preferences over teachers are position specific. Since all the parameters in our 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, indicating she finds one position more appealing than the other within that school.

Teachers' Decisions. At the end of each academic year t, every teacher who had an assigned teaching position within the school district 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 of hiring each year, we build at most 20 menus per teacher and can observe the decisions in each case. On average, teachers' school menus have 37 schools, and conditional on their applying at all, teachers apply to an average of 7 positions in each round (Table 1, Panel B).

Note that 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 certain level, requiring a reduction in budget and staffing. If this occurs, the principal chooses the teacher who loses her assignment, which is called being "excessed." In principle, seniority protects teachers from losing their assignment, and more junior teachers are at risk of excessing. Principals also have the option to push probationary teachers out of their school each year. A teacher who loses her school assignment and wants to find a new one within the district goes to the ITS to search for a new position for the next academic year. Our data identify which ITS applicants had lost their prior assignment and which did not. In our sample, approximately 41% of the 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 process ends is the expected value of the match they expect to find in the scramble

students' learning gains. The changes in teachers' teams identify their individual effectiveness.

round, in which every unmatched teacher must match with a remaining position, or the next round (only in the case of round 1). In contrast, a teacher who did not lose her assignment has her current position as her fallback option. Consistent with this part of the ITS process, we assume that, in each round, a teacher applies to every position on her choice menu that she prefers to her fallback option, net of inertia. The inertia cost applies only to teachers who did not lose their position.

At the application stage, we assume that teachers consider all the positions on their choice menu. This means they are aware of all these positions and can compare them. The teaching positions in our sample tend to follow standard descriptions and are differentiated mainly by the school and position type. Both are easily observable characteristics that a teacher probably has an understanding of before coming to the ITS in a given year. We also assume that there are no application costs for teachers. This assumption is motivated by the infrastructure of the platform that collects applications. An applicant uploads a resume in standardized format and the other required information to the centralized system once and clicks to apply to each position of interest. The system then distributes the applicant's packet to the hiring team for each position to which she has applied.

In the data corresponding to the stage at which teachers receive offers, we observe a subset of teachers who receive multiple concurrent offers and choose one. We assume teachers choose their preferred offer. We use this choice as an additional source of variation to estimate teacher preferences. Given these data, the strategic environment, and the counterfactuals we aim to model, we will interpret our model of teacher decisions as a model of teacher preferences.

School Principals' Decisions. For each open position in school s, we use the data on the position-specific rankings of interviewed teachers by school principals to estimate Equation 6 and assume d_{ist} is decreasing in the rankings, with candidates interviewed but not ranked having the lowest values of d_{ist}[12]:

$$d_{R_1 s} \geq d_{R_2 s} \geq \cdots \geq d_{R_k s} \geq 0, \quad \text{and} \quad d_{R_j s} \geq d_{R_0 s} \quad \text{for } 1 \leq j \leq 4$$

[12]We omit the subscript t in the following inequalities for simplicity, in addition to omitting the position subscript.

where $k \leq 4$ and $R_k \in \mathcal{I}_s$ is the teacher ranked in k^{th} position by s for the open position in question and $\mathcal{I}_s$ is the set of teachers interviewed by s for that position. $R_0 \in \mathcal{I}_s$ represents any candidate interviewed by s for the position but not ranked. We further assume that if s did not rank all interviewed candidates and the principal ranked fewer than 4 candidates for a position, then $d_{R_0 s} \leq 0$. This represents the case of a principal who decided not to rank an interviewed candidate even when the rank-ordered list had not been exhausted.

We do not interpret the model of principal decisions as a model of principal utility because principals may have strategic considerations when deciding how to rank teachers. Importantly, modeling principal decisions is sufficient for our purposes. The model and counterfactual policies do not require separate identification of principal utility and beliefs. Given the mechanism design, a principal may optimally choose to skip attractive candidates if she believes the candidates are unlikely to accept their offer because this will reduce her chance of securing lower-ranked candidates in that round. To capture selection of teachers into schools, the ideal variation would pin down the correlation between the teacher and principal decisions that determine matches and student outcomes. If principals' utility deviates from their decisions because principals act strategically, then a model of principal utility is ill suited for dealing with the selection issues inherent in the model. Rankings, on the other hand, map into matches given how the offer process is structured. Thus, we interpret the latent variable in the model of principal decisions as principal expected utility combining both preferences and subjective probabilities of attracting a candidate.

4.2 Identification

To identify the joint distribution of decisions and outcomes, we need variation that lets us separate teacher general effectiveness, defined as the average impact of each teacher on student test scores, from any teacher–position match effects. The central challenge for identification relates to whether teachers are systematically matched to schools with which they have high (or low) match effects. To address this challenge, our identification strategy uses shifters for teacher and principal decisions that are unrelated to students' potential outcomes. By introducing a degree of randomness (variation conditionally independent of students' potential outcomes) into these decisions, the shifters help identify $\mathrm{var}(\eta^y_{is})$ and pin down the distribution of match effects from the population of quasi–randomly assigned teachers, using cross-teacher variation in effectiveness that is not constant across schools. Selection on unobservables is pinned down by $\mathrm{cov}(\eta^y_{is}, \eta^u_{is})$ and $\mathrm{cov}(\eta^y_{is}, \eta^v_{is})$. We identify these quantities by comparing the outcomes of quasi–randomly assigned and selected teachers.

A second source of selection in our model arises because principal decisions are observed only for (a subset of) the teachers who applied for a position at the school. Because teacher decisions to apply are not random, inferring principal preferences and their correlation with student outcomes from decisions over a selected sample of teachers may not extrapolate to other teachers. To account for any correlation between teacher and principal preferences that do not map onto the observables in our model, we introduce variation in teacher utility that is excluded from principal decisions. Teachers exposed to this variation are likely to apply to schools by virtue of the shifter and, from the perspective of schools, represent a quasi-random sample of teachers because they are not especially attractive or unattractive candidates, which allows us to extrapolate to the aggregate teacher sample.

Formally, we require both shifters, z^u_{ist} and z^d_{ist}, to be conditionally independent of $(\eta_{is}, \theta_i, \varphi_i)$ and the error terms $(\varepsilon^u_{ist}, \varepsilon^d_{ist}, \varepsilon^y_{kt})$. Moreover, z^u_{ist} is assumed independent of z^d_{ist}. This implies that the shifters do not affect the distribution of potential outcomes but only affect observed outcomes by affecting the assignment of teachers to classrooms. Similarly, z^u_{ist} does not affect the distribution of d_{ist} but only affects the realized d_{ist} by affecting the set of teachers whom a principal is able to interview.

This strategy builds on a large literature on selection models that jointly model outcomes and choices to estimate treatment effects corrected for selection bias (Geweke et al., 2003; Heckman and Navarro, 2007; Lewbel, 2007; Hull, 2020). More closely, our strategy builds on the work by Walters (2018) and Agarwal et al. (Forthcoming), who use shifters to evaluate outcomes from assignment mechanisms. Because we allow the teacher effects to be school specific, these match-specific benefits result in a large number of treatments. Our strategy helps identify the distribution of match-specific treatment effects, as in Agarwal et al. (Forthcoming). Our identification strategy extends the ideas in these papers to a situation where an assignment is produced by the joint and possibly selected decisions of two agents on both sides of the market. The exogenous supply shifter separates teachers' decisions from student outcomes and from school principals' decisions. The exogenous shifter of principal decisions separates these decisions from student outcomes. Each shifter helps identify the correlation between decisions and outcomes and does so by introducing a degree of randomness in decisions.

Teachers' Utility Shifter: We use the driving time in minutes from a teacher's residence to each school in her choice set as a shifter of teacher utility (z_{ist}^{u}).[13] We expect that teachers place a higher value on job opportunities closer to their homes, as these represent lower commuting costs. Plotting the probability of application against driving time reveals that, conditional on teachers' applying to at least one job, their probability of applying to a school significantly decreases as driving time increases (see Figure 2).

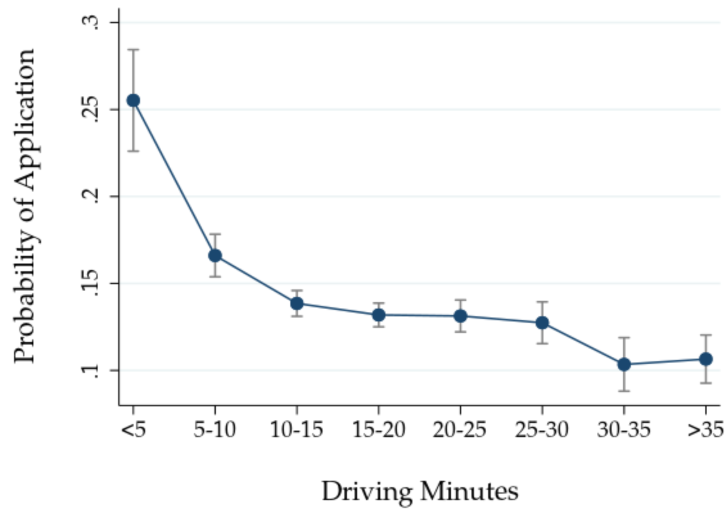
For exclusion, we require the shifter to be independent of student potential outcomes and principal decisions. We assume z_{ist}^{u} is conditionally independent of the unobservables in the models of outcomes and principal decisions. This implies that teachers do not systematically choose to live near the schools where they have an unobservable comparative advantage. Similarly, teachers do not select their residences based on their attractiveness to potential principals, and principals do not consider teachers' residential locations when making offers.[14] In terms of exclusion from principal decisions, the audit study of Hinrichs (2021) provides evidence supporting the assumption that school principals do not have preferences over teachers' commuting time.[15]

[13]The mean driving time between teacher residences and all open positions is 19 minutes in the sample, far below the 2017 U.S. average work-to-residence commuting time of 27 minutes. The U.S. median was 18 minutes, and the 95$^{\text{th}}$ percentile driving time is 36 minutes. Source: 2018 American Community Survey 1-year estimates.

[14]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 to teach at the school are assumed to be driven more by proximity than by selection on effectiveness, with the effectiveness of teachers living farther from their school, whose choices are assumed to be more driven by selection on effectiveness. A finding that the former are less effective than the latter would be interpreted as positive selection.

[15]Hinrichs (2021) randomly assigns fictitious new teacher applicant characteristics to resumes submitted in application for teacher job openings and measures the effects on actual principals' positive responses, such as callbacks. Although the published version of the paper does not contain this result, Hinrichs states in email correspondence that he controlled for a cubic of distance along with an indicator for a teacher's

Figure 2: Probability of application as a function of driving time from home to school



Note: This figure plots the probability of application by teachers as a function of the driving time from their home to the school with open positions, conditional on having applied to at least one open position.

School Principals' Decisions Shifter: We use a measure of the principal's unexpected need to hire (z^v_{ist}) as a plausibly exogenous shifter of the principal's willingness to hire a teacher. The ideal shifter of principal choices z^v_{ist} would move the principal's selectivity threshold but not the students' potential outcomes. We measure the urgency of the principal's need to hire as the difference between the number of positions a school posts in a given round and the number posted in the first round each year. For the first round, the shifter is always zero, and for the second round, it takes the value of the difference between second-round and first-round openings.

A large difference between second-round and first-round postings can arise when a principal is surprised with a larger than expected number of open positions that she needs to fill and cannot be as selective as in other years or as other schools in that year. It represents an unexpected surge or sag in the need to hire driven in large part by first-round selection processes at *other* schools. Other principals' decisions dictate the other offers available to the school's first-round offerees and to the school's incumbents

being in-state. In the public school subsample, none of the three distance coefficients was significant, and an F-test for the joint significance of the three distance coefficients gave a p-value of 0.635.

Table 2: Expected Position in the Principal Ranking and Principal Decisions Shifter

	Interview Ranking			
	(1)	(2)	(3)	(4)
Delta ITS rounds	-0.074	-0.078	-0.083	-0.082
	(0.026)	(0.027)	(0.031)	(0.031)
Observations	1075	1075	1075	1075
Teacher FE	Y	Y	Y	Y
Year FE	N	Y	N	Y
School FE	N	N	Y	Y

Note: This table shows coefficients from a linear regression of the instrument Delta ITS rounds on the ranking of interviewed candidates. The data includes every teacher-position combination where a teacher was interviewed and ranked. We include teacher fixed effects in all specifications. Columns (2) and (4) also include year fixed effects, while columns (3) and (4) include school fixed effects.

who sought internal transfer in the first round. While, for each school, the number of first-round postings is the result of known information,[16] an unusually high number of second-round postings relative to the number of first-round postings signals a shock to the principal's hiring urgency. This occurs if a principal unexpectedly lost incumbent teachers to other schools in the first round or if she was not successful in filling positions in the first round. With the number of applicants held fixed, a principal with more vacancies in the second round cannot be as choosy.

We assume that, as the principal's unexpected need to hire in the second round (z^v_{ist}) increases, so does the probability that she will rank a given applicant higher or will rank a candidate at all. The average difference in postings between rounds two and one is 0.3 positions with an SD of 1.2 and a range from -4 to 7. A regression of each candidate's position in a school's ranking as the unexpected need to hire increases, that includes

[16]Before the first round, each principal goes through a process to determine how many hires she needs to make. Well before the round opens, incumbent teachers must declare whether they will retire or take a leave of absence next year. Incumbents can, but are not required to, disclose whether they intend to seek a transfer. Central administration projects student enrollment and budget and informs each principal how many teacher positions it will have funded in the next year. The principal decides on any discretionary dismissals of probationary teachers and any teacher layoffs due to closed positions (excesses). The difference between budgeted positions and continuing incumbents is the number of openings. The principal must post the openings in the first round so incumbents at other schools have the opportunity to apply for transfer before external hires are considered.

teacher, school, and year fixed effects yields an estimated coefficient on the shifter of -0.08 (Table 2, column 4). Thus, all else equal, a candidate is expected to be ranked higher as the schools' need to hire increases. These results are robust to alternative specifications with fewer fixed effects.

For exclusion, we assume that teachers do not value principals' unexpected need to hire when making their application or acceptance decisions. We assume that, if principal pickiness changes, all applicants' probability of hire changes and that the change is independent of the applicants' general effectiveness θ_i and comparative advantage η^y_{is}.

4.3 Estimation Strategy

We estimate the joint distribution of χ^y, χ^u, χ^d and the variance–covariance matrices $\Sigma_{\theta\varphi}, \Sigma_\eta$, and $\sigma^2_{\varepsilon y}$ using a Gibbs sampler and assuming conjugate uninformative priors (Gelman et al., 2013). Using this method, we generate draws of the joint distribution of the parameters and latent variables in the model. We draw 110,000 iterations of the sampler, burn 50,000 initial iterations, and keep only 1 of every 10 draws to reduce autocorrelation in the chains. We inspect the chains for convergence before reporting the results.

5. Results

5.1 Estimates

We now present the mean and SD of the posterior empirical distribution of the parameters in the models of student outcomes, teacher preferences, and principal decisions, as shown in Tables 3, 4, and 5.[17] Although we estimate the models using a comprehensive set of characteristics for teachers, students, and schools, the tables above display only the parameter estimates for a selected subset of these characteristics. For parameter estimates on all the characteristics, please refer to Appendix Tables D.9, D.10, and D.11.

Student Outcomes. Our first key finding, displayed in Table 3, is that a teacher's general effectiveness has a significantly larger effect on student outcomes than does her observed

[17]Our estimation derives the empirical joint posterior distribution of the model parameters. We provide the mean and SD of the *normal* marginal distribution for each parameter. In contrast to the frequentist approach, which emphasizes statistical significance, presenting the mean and SD allows us to visualize the complete distribution of the posterior.

and unobserved comparative advantage. Panel A shows that students with lower past test scores, male students, low-income students, and non-White students have lower test scores in a given year. Teacher gender and experience also affect student test scores, but their predicted effect is approximately an order of magnitude smaller than that of students' own characteristics (Panel B). A lower low-income student share in the school also increases the student test scores (Panel C). The effects from teacher–student match on observables also affect student test scores (Panel D). We estimate that a match between teacher and student on gender improves student scores slightly. The literature has found mixed results on this question (Dee, 2007; Antecol et al., 2015). A teacher–student match on minority status (here defined as their being non-White) has a positive and smaller estimated effect. We also find that more educated and experienced teachers have a comparative advantage in teaching minority and low-achieving students, in contrast to the equilibrium assignment patterns we observe in the data (Figure 1).

Importantly, the student outcomes model estimates imply that a 1 SD improvement in teacher general effectiveness (θ_i) increases the normalized student test scores by approximately 0.08 SD. Our teacher general effectiveness estimates are slightly smaller than the estimates in Chetty et al. (2014a), but they do not account for match effects in their model, which might drive this variation in their estimates with respect to ours. The impact of match effects from unobservables (η_{is}^{y}) on student test scores is orders of magnitude smaller than the match effects from observables (Panel D) and teacher general effectiveness.[18]

Teacher Utility. Our second key finding, displayed in Table 4, is that teachers value the school's student characteristics and is consistent with the suggestive evidence (Figure 1) that teachers, especially more experienced ones, prefer schools with fewer low-income and minority students. Teachers prefer working in a school with a high fraction of experienced colleagues. In addition, teachers strongly 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, more-educated teachers

[18]Note that we report the SD, not the variance, of θ_i as well as η_{is}^{y} for easier comparison with estimates from the literature.

Table 3: Estimated Parameters of the Student Outcomes Model

	Mean	*Std. Dev.*
Panel A: Student Characteristics		
Previous score	0.770	0.001
Race – Black	-0.126	0.003
Male	-0.012	0.002
Low income	-0.142	0.005
Panel B: Teacher Characteristics		
Male	-0.020	0.005
Education	-0.003	0.002
Experience 2 to 3	0.020	0.004
Experience 4 to 6	0.022	0.005
Experience 7+	0.027	0.005
Panel C: School Characteristics		
% Low income	-0.023	0.018
Panel D: Student–Teacher Interactions		
Match minority student	0.003	0.002
Same-gender student	0.005	0.002
Teacher education * Student minority	0.005	0.002
Teacher education * Student previous score	-0.003	0.001
Teacher experience * Student minority	0.002	0.002
Teacher experience * Student previous score	-0.002	0.001
Std. Dev. of teacher general effectiveness (θ_i)	0.0815	
Std. Dev. of teacher–school unobservables match effectiveness (η^y_{is})	3.5e-04	

Note: The table shows the mean and the standard deviation of the estimated chains of the parameters of the student outcomes model (Equation 4). Panel A presents the parameters associated with student's own characteristics. For the race dummy, the omitted category is White. Panels B and C present the parameters associated with the teacher and school characteristics. Panel D presents the parameters associated with the interaction between student and teacher characteristics. Minority is defined as non-White. Note that we report the standard deviation, not the variance, of θ_i as well as η^y_{is}. See Appendix Table D.9 for the parameter estimates for the full set of characteristics.

prefer teaching higher-achieving students.

As expected, teachers dislike teaching farther from their homes, and the driving time parameter—the teacher utility shifter—is negative. Additionally, teachers face significant inertia costs when faced with the option of a transfer. The mean inertia parameter estimate

Table 4: Estimated Parameters of Teacher Utility

	Mean	*Std. Dev.*
Panel A: School Characteristics – Students		
% Low income	-0.100	0.049
Average test scores	0.010	0.120
% Black	-0.219	0.348
Panel B: School Characteristics – Teachers		
% Black	0.290	0.369
Average teacher experience	0.047	0.020
Inertia	3.582	0.034
Panel C: Teacher–School Interactions		
% of students match minority	0.499	0.049
% of teachers match minority	0.454	0.049
Teacher education * Average test scores	0.119	0.028
Teacher education * % minority	0.098	0.022
Teacher experience * Average test scores	-0.062	0.037
Teacher experience * % minority	-0.318	0.026
Teacher experience * Average teacher experience	0.008	0.013
Panel D: Shifter		
Driving time	-0.010	0.001
Std. Dev. of idiosyncratic taste shock for schools (η^u_{is})	0.493	

Note: The table shows the means and standard deviations of the estimated chains of the parameters of the teacher utility model (Equation 5). Panels A and B present the parameters associated with the school's student and teacher characteristics, respectively. Panel C presents the parameters associated with the interaction between the teacher and school characteristics. Minority is defined as non-White. In Panel D we present the shifter. Driving time is measured in minutes. We report the standard deviation, not the variance, for η^u_{is}. See Appendix Table D.10 for the parameter estimates for the full set of characteristics.

(3.6) is several orders of magnitude larger than the values on any of the other observable school characteristics or match effects from observables. Last, while the teacher–school match effects from unobservables (η^y_{is}) matter significantly less for student outcomes, they (η^u_{is}) play a role in teachers' preferences over schools.

School Principal Decisions. Our third key finding, displayed in Table 5, is that while observable teacher characteristics such as education and experience are not crucial in explaining the choices of school principals, principals prefer to hire teachers whose minority

Table 5: Estimated Parameters of School Principal Decisions

	Mean	*Std. Dev.*
Panel A: Teacher Characteristics		
Male	0.205	0.079
Education	-0.250	0.305
Experience	0.236	0.336
% Black	-0.783	0.139
Panel B: Teacher–School Interactions		
% of students match minority	1.230	0.119
Teacher education * Average test scores	0.130	0.236
Teacher education * % minority	0.424	0.462
Teacher experience * Average test scores	0.051	0.260
Teacher experience * % minority	-0.463	0.503
Teacher experience * Average teacher experience	0.033	0.041
Panel C: Shifter		
Delta ITS rounds	0.037	0.026
Std. Dev. of teacher effects (φ_i)	0.119	
Std. Dev. of idiosyncratic taste shock for teachers (η_{is}^{v})	0.020	

Note: The table shows the means and standard deviations of the estimated chains of the parameters of the school principal decision model. Panel A presents the parameters associated with the teacher characteristics. Panel B presents the parameters associated with the interaction between the teacher and school characteristics. Minority is defined as non-White. Panel C presents the shifter. We report the standard deviation, not the variance, for φ_i and η_{is}^{v}. See Appendix Table D.11 for the parameter estimates for the full set of characteristics.

status matches a greater share of the school's student minority status and to hire male teachers. Principals value an SD of teacher effectiveness (φ_i) less than teacher gender or a match on minority status. In addition, the positive posterior mean value of the *Delta ITS rounds* parameter—the principal decision shifter–implies that when principals are surprised by the need to fill an unexpectedly large number of positions, all else equal, they are likelier to rank a candidate whom they would otherwise leave unranked or to rank the candidate highly.[19]

[19] The posterior SD of the coefficient on the delta rounds parameter implies that 92% of the estimated posterior takes positive values.

5.2 Impact of Teacher Effectiveness on Decisions and Student Outcomes

To quantify the importance of the three different aspects of teacher effectiveness—general effectiveness (θ_i), match effects from observables, and match effects from unobservables (η^y_{is})—for student outcomes and teacher and principal decisions, Table 6 shows the change in student test scores, teacher decisions, and principal decisions with a rise in teacher general effectiveness or match effects from very low to very high values. To do so, we simulate the change in these quantities when the value of θ_i or η^y_{is} goes from the 1$^{\text{st}}$ to the 99$^{\text{th}}$ percentile. We also draw from the underlying distribution of teacher and student observable characteristics to generate the distribution of match effects from observables and evaluate the changes as the match quality goes from the 1$^{\text{st}}$ to the 99$^{\text{th}}$ percentile of the distribution.

Table 6: Impact of Teacher Effectiveness on Decisions and Student Outcomes

	Outcomes	Principal Decisions	Teacher Utility
	Δy_{kt}	Δv_{ist}	Δu_{ist}
From percentile 1 *to* 99			
Teacher general effectiveness, θ_i	0.387	0.469	
	(0.017)	(0.214)	
Match effects from unobservables, η^y_{is}	0.002	-0.067	-1.041
	(1.6e-05)	(0.046)	(0.399)
Match effects from observables	0.036	-0.042	-0.063
	(0.005)	(0.833)	(0.031)

Note: The table shows the results from simulated changes in student outcomes, principal decisions, and teacher utility with a rise in teacher general effectiveness (θ_i), teacher–school match effectiveness from unobservables (η^y_{is}), and match effectiveness from observables from the 1$^{\text{st}}$ to the 99$^{\text{th}}$ percentile. Changes are expressed in standard deviations. Standard deviations are shown in parentheses.

Teacher general effectiveness has greater explanatory power than match effects for student outcomes.[20] While the substitution of a teacher at the bottom of the general effectiveness distribution with a teacher at the top of the distribution would improve student test scores by 0.39 SD, a move from the lowest- to the highest-quality match on observables would improve test scores by a ninth of that (0.04 SD). The impact of the quality of the match on observables aggregates the effects of a student's being matched

[20] The first paper to use internal transfer data to separate teacher and school preferences found evidence schools valued proxies for general teacher effectiveness Boyd et al. (2011).

with a teacher of the same (or different) gender, and of the same (or different) minority status, and of low-achieving students' and minority students' being matched with more (or less) experienced and educated teachers. The match effects from unobservables drive a gain of only approximately 0.002 SD when a teacher at the bottom of the distribution of match effects from unobservables is substituted by a teacher at the top. All these effects are statistically different from zero across our simulations.

Taken together, the match effects from observables and unobservables account for approximately 9% of the impact of teacher general effectiveness on student outcomes, which is in the middle of prior estimates. Relative to the calculations in Jackson (2013) and Ahn et al. (2024), our estimated share of 9% is smaller than the estimated share of two-thirds in Jackson and the estimated share of 12% and 25% in Ahn et al. for math and reading, respectively. Our findings align more closely with those of Delgado (2023), who finds that comparative advantage effects account for approximately 6% of the overall effectiveness in math and 13% in reading.[21] By studying a large, urban school district operating within a unionized teaching context, our study resembles Delgado (2023) but differs from much of this literature that studies North Carolina (Jackson, 2013; Bates et al., 2024; Ahn et al., 2024), where collective bargaining agreements are not legal.[22]

Substituting a teacher from the bottom of the teacher general effectiveness distribution with one from the top of the distribution raises principal expected utility by 0.47 SD. However, we observe no change in principal expected utility when we consider the replacement of a teacher at the bottom with one at the top of the match quality distributions for that school. On the other side of the market, we find that teachers are averse to positions in which they would have high match effectiveness. This is especially true for the portion of match effectiveness not captured by observable characteristics. For example, our model predicts that teachers with more years of education have a stronger preference for teaching high-achieving students, but their comparative advantage lies in teaching lower-achieving

[21]These estimates are based on our calculations from estimates in Delgado (2023) (Appendix C).

[22]Labor relations law and policy can affect the educational production function, teacher and district incentives and constraints, and selection into a district by teachers and students. North Carolina stands out nationally by making collective bargaining by public school teachers illegal, with teacher strikes punishable by jail time (Schlemmer, 2024). Teacher collective bargaining is legal in 47 states in the U.S.

students.

5.3 Robustness

Alternative Teacher Effectiveness Measures: 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 (see Appendix Table D.2).[23]

Alternative Principal Decision Shifter: We use the proportion of teachers at a school who share the minority status of a candidate teacher as an alternative shifter of principal decisions. The choice of this alternative shifter is based on two key observations. First, hiring committees may be subject to homophily, whereby they form stronger connections with candidates of similar race or ethnicity. This would lead to such candidates being ranked higher regardless of their qualifications or fit for the school. Second, principals aim to hire teachers who demonstrate high retention potential. Irrespective of their effectiveness, teachers who share cultural traits with existing staff may be perceived as likelier to remain at the school.[24] Consistent with this, we show that teacher candidates who share the minority status of a majority of the teaching staff are more likely to be ranked in top positions than other candidates interviewed (see Appendix Table D.5). Notably, we find no evidence that racial congruity influences the decision to interview a candidate. This suggests that the interview process—where candidates interact with the hiring committee—acts as a mediator in this effect, likely because of enhanced communication between the candidates and committee members. For exclusion, we rule out any peer effects in teaching that are mediated by the race of teachers. This means that a teacher is

[23]These district evaluation measures began to be collected after 2012. The correlations are estimated with the observations from 2013 to 2019.

[24]In a similar vein, in a Missouri and Tennessee longitudinal analysis, Black school principals increased the share of teachers of color in their school more than other principals via greater probabilities of hiring and retention (Bartanen and Grissom, 2023). This also echoes the finding from Grissom and Keiser's (2011) cross-sectional analysis that teachers whose race is congruent with their principal's are likelier to retain employment than are colleagues of noncongruent race in the same school.

not more effective as a result of her having a large share of colleagues of her own race. Any correlation between teacher effectiveness and the share of same-race teacher peers that is mediated by the racial composition of students would not violate the exclusion restriction. The model estimates derived with this shifter are largely similar to those presented in the baseline (see Appendix Tables D.12, D.13, and D.14).

Additional Controls for Teacher Utility Shifter: If teachers are systematically more effective at teaching students with whom they share a cultural background and are more likely to reside in neighborhoods with these families (an effect not fully accounted for by race interactions), then the exclusion restriction for our teacher utility shifter may be violated. To address this concern, we incorporate school attendance boundaries into our outcomes model with an indicator that identifies whether each school serves the teacher's neighborhood of residence. Since the schools attended by the children of teachers are likely to serve students who share unobservable cultural traits with those teachers, the coefficient on this indicator variable will reflect the impact of cultural familiarity on student test scores. We find that the results are robust to our including this control (see Appendix Tables D.16, and D.17).

External Validity: The student and school characteristics of the district we study strongly resemble those of the population of U.S. 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 value between 0 and 1. We compare the schools in our sample to 15,389 U.S. schools in the population inference sample in which 5th, 6th, 7th, or 8th grade is taught, that are in an urban locale, and that are not charter schools. Based on parameters such as school size, percentage of low-income students, percentage girls, percentage White, percentage Black, percentage Hispanic, percentage U.S. citizens, and median family income, our analysis implies a generalizability index value of 0.81, which the tool characterizes as a high level of generalizability.

6. Counterfactual Teacher Assignments

To quantify the potential gains associated with a reallocation of teachers to alternative classrooms, we study the assignment of teachers to the set of all positions we can associate with observed test scores during our 10-year study period. We call this set $\mathcal{P}$. For each position in $\mathcal{P}$, we observe the teacher assigned to it and the set of students associated with it in a given year.[25] The set of teachers who are alternative candidates for an assignment to the positions in $\mathcal{P}$ in each year includes the teachers assigned to a position in $\mathcal{P}$ in that year and the set of teachers who applied to any open position in the ITS that year but were not assigned a position in $\mathcal{P}$. We further restrict the menu of positions for an alternative assignment for each teacher by using data on the licenses each teacher holds and the licenses required for each position, such that only licensed math teachers can be assigned to math positions and so on. Ignoring these constraints would lead the model to consider many infeasible assignments and bias the estimated scope for gains (Almagro and Sood, 2025).

We simulate counterfactual assignments of teachers to alternative classrooms with a view toward two policy objectives. The first is maximizing average student test scores in the school district. The second is maximizing the percentage of proficient students. Both objectives are widely used by school districts, parents, and policymakers to evaluate school district performance around the country. We study each policy objective under four counterfactual assignment scenarios: 1) *unconstrained*, 2) retained teachers unharmed or *no quits*, 3) scenario 2) + *no additional layoffs*, and 4) scenarios 2) and 3) + *same school*. These are explained in detail below.

First, we consider the counterfactual scenario in which teachers are assigned to classrooms to maximize each of the two policy objectives given the positions in $\mathcal{P}$ and the set of teachers described above without any additional constraints. We refer to this scenario as the *unconstrained* counterfactual. We generate this assignment by solving a linear program

[25] A position in our sample is associated with a grade, school, and year. We do not observe classroom identifiers, so we keep the grade–school–year as the minimum level of aggregation. By construction, each position in our sample is associated with a single year. Even if a teacher–classroom pair remains matched for many years, we treat each annual observation as distinct.

in which we restrict each position to be filled by exactly one teacher and each teacher is assigned to at most one position.[26] Because some assignments generated under the unconstrained counterfactual may be unacceptable to a teacher without an adjustment in pay, the second counterfactual scenario restricts each teacher's menu of possible assignments to only those that the teacher would weakly prefer over her observed assignment based on our simulation of teacher utility using the model parameters.[27] We refer to the second scenario as the *retained- teachers-unharmed* or *no-quits* counterfactual because it guarantees that any assigned teacher does not have lower utility than that under her observed assignment, though teachers who are left unassigned may have lower utility. Under the first two counterfactual scenarios, the pool of teachers who are candidates for assignment is larger than the set of positions because assignments can go to either a) the teachers observed to be assigned to a position in our sample or b) the teachers who unsuccessfully sought a new position in our sample through the ITS.[28]

The third counterfactual scenario restricts alternative assignments to only those teachers observed as assigned to a position in $\mathcal{P}$ in each year, such as group a) as described in the previous paragraph. We refer to this third scenario as + *no additional layoffs* counterfactual. It has the same restriction as the second counterfactual but further excludes the b) group of teachers not observed in any assigned position in the data. In this counterfactual, the number of available teachers and available positions is the same. It does not allow a change in who is teaching. It simply allows reshuffling of the observed teachers across positions in each year.

The fourth counterfactual scenario further restricts the potential position assignments for each teacher to only within-school assignments. We refer to this fourth scenario as + *same school* counterfactual. Thus, in this counterfactual, teachers can no longer be assigned to positions across schools, only to positions in the same school in which they already

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

[27]Because the observed assignment is feasible under this restriction, the set of solutions is not empty.

[28]Teachers not observed to be assigned to one of the positions in our sample may be assigned to other positions outside our sample within the school district. These are the positions either not associated with test scores or with less than 0.5 full-time equivalence required.

teach that year. Finally, we benchmark the outcomes under the four scenarios against those under the observed assignment and under a random assignment of the teachers whom we observe assigned to positions in $\mathcal{P}$.

We view the counterfactual assignments as assignments made by the school district under perfect foresight, taking the estimated teacher-specific general effectiveness values, the estimated parameters on match effects from observables, and the estimated distribution of match effects from unobservables as given. We do not include teacher inertia costs in any of the counterfactuals as we consider assignments of teachers as alternatives to the observed ones rather than as transitions from a past position.[29] The outcomes under the counterfactual assignments do not reflect the dynamic gains generated by an alternative assignment in year t for test scores in the subsequent years; instead, they are the outcomes under one-shot alternative assignments in each year independently. That is, the counterfactual results quantify the average one-year gains from an alternative assignment of teachers to classrooms. Last, for each counterfactual and for the observed and random assignment, we generate 100 draws of the parameters in our model and compute the mean and SD of the gains across these simulations in each case.

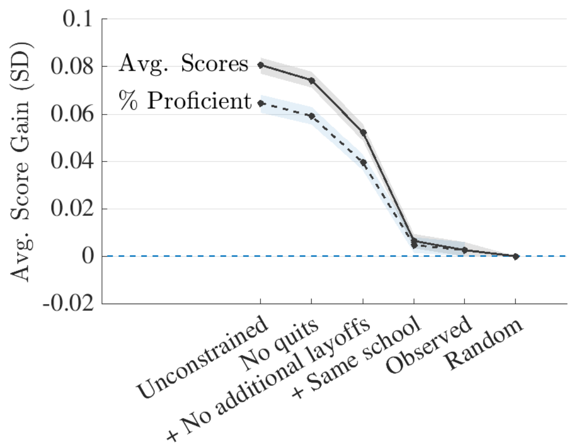
6.1 Policy Objective: Maximizing Average Student Achievement

As seen in Figure 3, under the first counterfactual scenario of unconstrained assignment of teachers to positions, a policymaker seeking to maximize the average student test scores can push the average scores up by 8% of an SD relative to those under the observed assignment (see Appendix Table D.6 for details). Under the second counterfactual scenario, where no retained teacher is harmed (*no quits*), we find that most of the gains in the average student test scores can still be realized. In this case, test scores would increase by 7% of an SD relative to those under the observed assignment. Compared with the unconstrained assignment, this assignment is more politically feasible as the retained teachers' welfare does not fall.

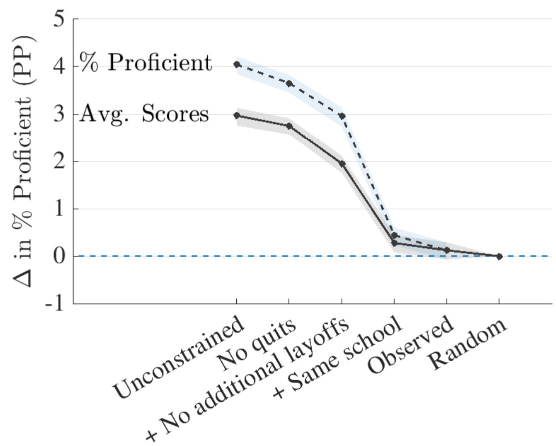
In the third counterfactual scenario, where we restrict the teacher pool to include only

[29] It remains important to account for the inertia costs in estimation to avoid biasing our estimates of teacher preferences over school characteristics.

Figure 3: Counterfactual Gains in Test Scores and Percentage Proficient by Policy Objective



(a) Average Test Score Gains by Objective



(b) Gains in Percentage Proficient by Objective

Note: Panel (a) shows the gains in average test scores in standard deviation (SD) terms under four counterfactual scenarios that aim to maximize average test scores (solid line) and percentage proficient (dashed line) relative to the outcomes under the observed assignment and a random assignment. Panel (b) shows the gains in percentage points (PP) under the same counterfactuals.

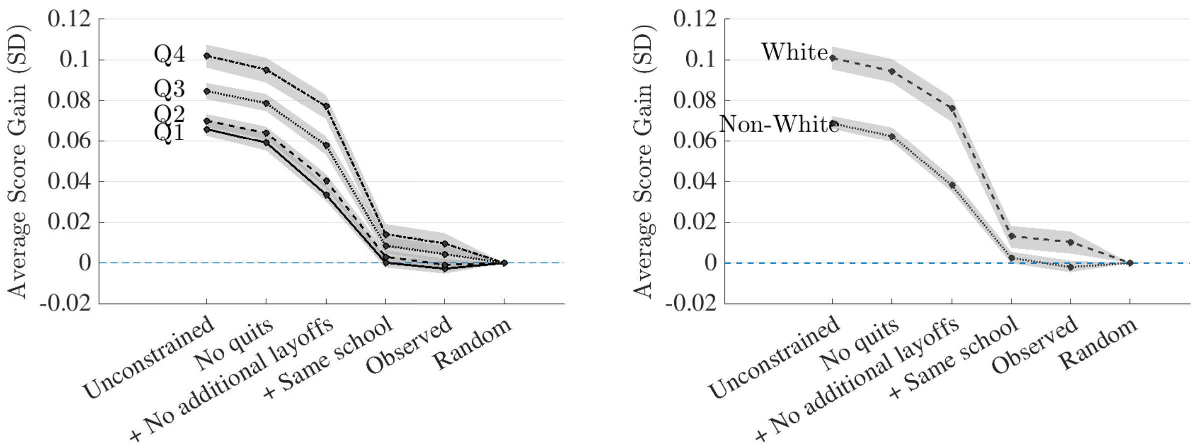
those observed assigned to full-time math and reading positions and keep the no-quits constraint, we still find significant gains in test scores of 5% of an SD relative to those under the observed assignment. Relative to this third scenario, scenarios 1 and 2 select teachers with higher general effectiveness to teach in the set of positions $\mathcal{P}$, and thus, these scenarios result in higher gains in average student achievement.

Finally, in the fourth scenario, where we consider only assignments within a teacher's current school, the potential gains in student test scores fall dramatically to 0.4% of an SD relative to those under the observed assignment. This shows that most of the potential gains in student test scores come from an assignment of teachers across schools within a district, which can be realized by policy interventions using the district's ITS.[30]

Figure 4 shows that all four counterfactual assignments that aim to maximize the av-

[30]Note that the observed equilibrium in our setting is different from that found in Bates et al. (2024), in which teacher quality is equally distributed across these groups (Table D.3). The difference in the two settings may be explained by differences in the policy objectives of the two school districts or the fact that, for seven years in our sample period, information about teacher effectiveness measured via classroom observations was observable to principals when they were hiring, likely creating a positive correlation between teacher effectiveness and principal preferences.

Figure 4: Differential Gains in Test Scores by Achievement and Race Under the Policy Objective of Maximizing Average Achievement



(a) By Achievement Quartile

(b) By Race and Ethnicity

Note: Panels (a) and (b) show the average test score gains in standard deviation (SD) terms relative to those under the observed and a random assignment by baseline student achievement quartile and student race and ethnicity, under the four counterfactual scenarios that maximize average test scores.

erage student test scores disproportionately benefit higher-achieving and White students, even though students of all achievement levels and non-White students experience gains relative to their outcomes under the observed assignment. This implies that all four counterfactual assignments would widen the racial achievement gap and inequality in the overall achievement level in the school district, pointing to an equity–efficiency trade-off under all of these counterfactual assignments.

Counterfactual Decomposition: To understand what drives the gains in each of the four scenarios, we decompose the gains in the average test scores additively into the portions attributable to the match effects and to teacher general effectiveness. Gains from teacher general effectiveness can arise from two sources: keeping the least effective teachers out of classrooms and matching the most effective teachers to larger classrooms.

Although match effects do matter for student test scores (as seen in Table 6), realizing those gains through assignment of teachers to classrooms is less effective because each student's best teacher match is different and not every student can be assigned her best-matched teacher simultaneously without a change to the grouping of students across classrooms and schools.[31] Consequently, the decomposition in Table 7 shows that the gains from alternative assignment are almost entirely explained by retention of the most effective teachers in full-time teaching positions and their assignment to larger classrooms, rather than by matching of teachers to students based on comparative advantage.

Table 7: Decomposition: Gains Under Policy Objective of Maximizing Average Student Test Scores

	Decomposition of Gains Relative to the Observed Assignment (SD)		
	Total Effect	*Effectiveness*	*Match Effects*
Unconstrained	0.078	0.073	0.005
	(0.001)	(0.001)	(0.005)
No quits	0.072	0.067	0.005
	(0.001)	(0.001)	(0.004)
+ *No additional layoffs*	0.050	0.046	0.004
	(0.001)	(0.001)	(0.003)
+ *Same school*	0.004	0.004	0.0002
	(0.0004)	(0.0004)	(0.0002)

Note: This table presents the decomposition of the gains in average test scores in standard deviation (SD) terms under the four counterfactual scenarios relative to the outcomes under the observed assignment. Standard errors in parentheses.

Counterfactual Extension: While the model of student outcomes includes a class-size effect, it does not allow for heterogeneity in teacher effectiveness by class size. To ensure that we can identify the gains explained by comparative advantage if they exist, we consider an alternative assignment that maximizes average *classroom* test scores instead of average *student* test scores. By doing this, we eliminate the scope for gains from the matching of more effective teachers to larger classrooms. In this counterfactual extension, we find gains relative to the scores under the observed assignment of approximately half the magnitude of the baseline gains in the unconstrained counterfactual: namely, 3.7% of an SD (see Appendix Table D.8). In the second scenario, where no retained teacher is harmed (*no quits*), we find a 3.4% SD improvement in scores over those in the observed assignment. These gains, again, come almost exclusively from retention of the most effective teachers in full-time teaching positions. When we consider assignment of only

[31] Similar results are found in Umosen (2024), who studies the role of classroom segregation in explaining gains from alternative teacher assignments.

the teachers whom we observe assigned to full-time positions in our sample (an extension of the third counterfactual), we find no gains, providing further evidence that gains from matching on comparative advantage alone are limited.[32]

6.2 Policy Objective: Maximizing Percentage of Proficient Students

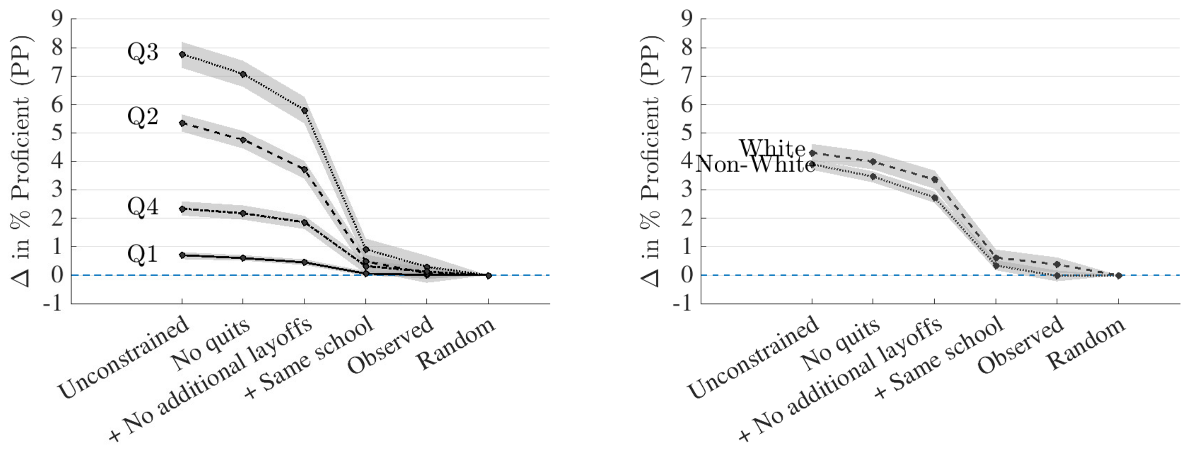
When we consider a policymaker objective of maximizing the percentage of proficient students in the school district instead of average student test scores, we find the percentage of proficient students would increase by 4 percentage points (PP) under the unconstrained assignment relative to that under the observed assignment (Figure 3).[33] As before, almost all the gains in the percentage of proficient students can still be realized if we impose the no-quits constraint (3.5 PP). Further restricting the sample to only teachers observed assigned reduces the gains to 3 PP relative to the share under the observed assignment, while allowing only within-school assignments shrinks the gains significantly to 0.3 PP relative to the share under the observed assignment. As with the gains in average test scores, most of the gains in the percentage of proficient students come from between-school assignments within a district.

While the four counterfactual scenarios that maximize the average test scores in the school district benefit higher-achieving and White students more than other groups, the gains are less disproportionate when the four scenarios maximize the percentage of proficient students in the district (Figure 5). These scenarios benefit students in the middle of the achievement distribution the most (quartiles 2 and 3), while students in the two tails of the distribution experience lower gains. This is intuitive and consistent with findings by Neal and Schanzenbach (2010): Classrooms with a larger share of students in the middle of the achievement distribution are close enough to the proficiency threshold that there are marginally more gains from a reallocation of teachers with higher general effectiveness toward them. In contrast, classrooms with many students in the top quartile are for the

[32]The third scenario in the counterfactual extension is similar to the scenario in the study by Biasi et al. (2022), though we focus on within-district teacher assignments and they focus on cross-district assignments. Neither allows gains from assignment of teachers to larger classrooms, and both restrict the counterfactuals only to allocations of teachers with an observed assignment. Similarly to us, the authors find minimal potential gains.

[33]See Appendix Table D.7 for details.

Figure 5: Differential Gains in Percentage Proficient by Achievement and Race Under the Policy Objective of Maximizing Percentage with Proficient Achievement



(a) By Achievement Quartile

(b) By Race and Ethnicity

Note: Panels (a) and (b) show the gains under the four counterfactual scenarios that maximize the percentage of proficient students in the percentage of proficient students, measured in percentage points (PP), relative to the shares under the observed and a random assignment by baseline student achievement quartile and student race and ethnicity.

most part already proficient, and classrooms with many students in the bottom quartile would require a more radical alternative assignment or other interventions to increase the percentage of proficient students. In addition, the gains for White and non-White students are similar across the four scenarios that maximize the percentage of proficient students in a district.

6.3 Comparison of Outcomes Under the Observed and a Random Assignment

A random assignment would generate average student test scores and a percentage of proficient students in the district that are statistically similar to their counterparts under the observed assignment (Figure 3). This is not surprising given that the match effects are less salient to student outcomes than is teachers' general effectiveness, as discussed in Section 5. When we disaggregate by baseline student achievement quartile and student race and ethnicity, we find that, relative to a random assignment, the observed assignment benefits higher-achieving students in the fourth quartile and White students while generating effects statistically similar to those of the random assignment for non-higher-achieving and non-White students (Figure 4). This is because we observe White students and higher-achieving students assigned to more effective teachers in the observed assignment (see Appendix Table D.3) and is consistent with the model parameter results that show that teachers prefer to teach high-achieving and White students and that principals value effective teachers.

7. Conclusion

We used novel administrative data from a market for teacher transfers that allowed us to track the decisions of both teachers and principals in the within–school district ITS and the test scores of students under the observed teacher assignments. We jointly modeled student outcomes and the decisions of teachers and principals in this labor market. This model allowed us to account for potential correlation between student outcomes and these teacher and principal decisions, account for selection biases in the observed matches, and, consequently, predict teacher effectiveness in unobserved matches. We found that teachers' general effectiveness (that is, effectiveness across all students) impacts student learning approximately nine times more than teachers' comparative advantage (match effects from interactions between student and teacher observables and from quality of the match on unobservables). The match effects are nonnegligible but matter much less than teacher general effectiveness. We found evidence that more experienced teachers have a comparative advantage in teaching minority and low-achieving students.

Our estimates show that principals 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 school districts value effectiveness but not teacher education or experience. Teachers, on the other side, are estimated to dislike placements based on their match effectiveness: They do not value working in schools in which they have a comparative advantage. Many teachers tend to prefer to work in schools with fewer minority students and low-income students.

Finally, we found that an alternative assignment of teachers to classrooms could 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 significant gains. Under a counterfactual in which the percentage of proficient students is maximized, the gains for students in the middle of the test score distribution are larger than those of students in the top and bottom quartiles. Across racial and ethnic groups, the gains are similar. Under these counterfactual assignments, most of the gains come from the assignment of more effective teachers to larger classrooms, and little is explained by matching based on comparative advantage. While match effects in teaching exist, our results suggest that realizing gains from such effects can be 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.

Agarwal, Nikhil, Charles Hodgson, and Paulo Somaini, "Choices and outcomes in assignment mechanisms: The allocation of deceased donor kidneys," *Econometrica*, Forthcoming.

Ahn, Tom, Esteban M. Aucejo, and Jonathan James, "The Importance of Student-Teacher Matching: A Multidimensional Value-Added Approach," *manuscript*, 2024.

Almagro, Milena and Aradhya Sood, "Sorting by Choice? Separating Preferences from Discrimination in Housing Markets," *manuscript*, 2025.

Angrist, Joshua, Guthrie Gray-Lobe, Clemence M Idoux, and Parag A Pathak, "Still worth the trip? school busing effects in boston and new york," Technical Report, National Bureau of Economic Research 2024.

Antecol, Heather, Ozkan Eren, and Serkan Ozbeklik, "The effect of teacher gender on student achievement in primary school," *Journal of Labor Economics*, 2015, *33* (1), 63–89.

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*, 2022.

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.

Bartanen, Brendan and Jason A Grissom, "School principal race, teacher racial diversity, and student achievement," *Journal of Human Resources*, 2023, *58* (2), 666–712.

Bates, Michael, Michael Dinerstein, Andrew C Johnston, and Isaac Sorkin, "Teacher Labor Market Policy and the Theory of the Second Best*," *The Quarterly Journal of Economics*, 2024, p. qjae042.

Biasi, Barbara 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 2022.

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 2024.

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.

Campos, Christopher and Caitlin Kearns, "The Impact of Public School Choice: Evidence from Los AngelesâĂŹs Zones of Choice," *The Quarterly Journal of Economics*, 2024, *139* (2), 1051–1093.

Chetty, Raj and Nathaniel Hendren, "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects*," in "in," Vol. 133 02 2018, pp. 1107–1162.

__ , 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.

__ , __ , and __ , "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood," *American economic review*, 2014, *104* (9), 2633–2679.

Combe, Julien, Olivier Tercieux, and Camille Terrier, "The design of teacher assignment: Theory and evidence," *The Review of Economic Studies*, 2022, *89* (6), 3154–3222.

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 MV Davis, B Pablo Montagnes, and Patryk Perkowski, "Stable Matching on the Job? Theory and Evidence on Internal Talent Markets," *Management Science*, 2024.

Davis, Jonathan MV, Kyle Greenberg, and Damon Jones, "An Experimental Evaluation of Deferred Acceptance: Evidence from Over 100 Army Officer Labor Markets," Technical Report, National Bureau of Economic Research 2023.

Dee, Thomas S, "Teachers and the gender gaps in student achievement," *Journal of Human resources*, 2007, *42* (3), 528–554.

Delgado, William, "Heterogeneous Teacher Effects, Comparative Advantage, and Match Quality," Technical Report, Working Paper 2023.

Deming, David J, "Better schools, less crime?," *The Quarterly Journal of Economics*, 2011, *126* (4), 2063–2115.

_ , Justine S Hastings, Thomas J Kane, and Douglas O Staiger, "School choice, school quality, and postsecondary attainment," *American Economic Review*, 2014, *104* (3), 991–1013.

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.*

Feng, Li and Tim R Sass, "Teacher quality and teacher mobility," *Education Finance and Policy*, 2017, *12* (3), 396–418.

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.

Grissom, Jason A and Lael R Keiser, "A supervisor like me: Race, representation, and the satisfaction and turnover decisions of public sector employees," *Journal of Policy Analysis and Management*, 2011, *30* (3), 557–580.

Hanushek, Eric A, "Education production functions," in "The economics of education," Elsevier, 2020, pp. 161–170.

He, YingHua, Shruti Sinha, and Xiaoting Sun, "Identification and Estimation in Many-to-one Two-sided Matching without Transfers," *Econometrica*, 2024, *92.*

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

_ and Salvador Navarro, "Dynamic discrete choice and dynamic treatment effects," *Journal of Econometrics*, 2007, *136* (2), 341–396.

Hinrichs, Peter, "What kind of teachers are schools looking for? Evidence from a randomized field experiment," *Journal of Economic Behavior & Organization*, 2021, *186*, 395–411.

Hull, Peter, "Estimating hospital quality with quasi-experimental data," *unpublished manuscript*, 2020.

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," *Journal of Political Economy*, 2024, *132* (7), 2346–2395.

Laverde, Mariana, "Distance to schools and equal access in school choice systems," *Unpublished Manuscript*, 2024.

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

Minni, Virginia, "Making the invisible hand visible: Managers and the allocation of workers to jobs," *unpublished manuscript*, 2024.

Mourot, Pauline, "Should Top Surgeons Practice at Top Hospitals? Sorting and Complementarities in Healthcare," *University of Chicago, mimeo*, 2025.

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.

Osterman, Paul, *Internal labor markets*, Vol. 1, The MIT Press, 1984.

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

Rothstein, Jesse, "Teacher quality policy when supply matters," *American Economic Review*, 2015, *105* (1), 100–130.

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.

Schlemmer, Liz, "In 1959, NC banned public workers from bargaining contracts–what it means for teachers today," *WUNC*, 2024.

Staiger, Douglas O and Jonah E Rockoff, "Searching for effective teachers with imperfect information," *Journal of Economic perspectives*, 2010, *24* (3), 97–118.

Tipton, E. and K. Miller, "Generalizer [Web-tool]," https://thegeneralizer.org 2022. Accessed: 2022-11-23.

Umosen, Ini, "Teacher Comparative Advantage, Achievement Gains, and Classroom Segregation," *manuscript*, 2024.

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

ONLINE APPENDIX

Gains from Alternative Assignment: Evidence from a Two-Sided Teacher Market

Mariana Laverde, Elton Mykerezi, Aaron Sojourner, & Aradhya Sood

A. Data Details

Teacher Data: The data on teachers comes from publicly available personnel data from the state's Department of Education. The dataset is a teacher level panel containing each teacher-year's assignments, demographics, qualifications and credentials, and pay. Each assignment includes a code linking the teacher to specific grades at a specific school within the district and the role performed by the teacher (e.g., 5th grade general education teacher). A teacher assignment can be less than full time and each teacher can hold multiple assignments in a given year. We use this data to identify the teachers who are primarily responsible for students' math and reading instruction in a given year and tested grades (starting at grade 3). Teacher demographics include gender, race, ethnicity, and age. Teacher credentials include formal education and years of experience (which together largely determine their pay within a district) and state license information. The license data are used to determine which open positions each teacher would be eligible to apply for.

District's Internal Teacher Transfer System: Data on position postings and the teacher transfer process are proprietary administrative data provided by the district to us under the IRB protocol 1510S79046. The dataset consists of two files: an open positions file that lists each open position available to incumbent teachers (before they can be posted for external applicants) and a transaction file that documents each decision taken by a teacher candidate, a district, or a school official. The transaction file includes applications, invitations for interviews, post interview rankings, position offers, and offer acceptance/rejection decisions. The district also provided us a bridge file that links internal personnel data, including the teachers' home addresses, to the publicly available teacher data from the state's Department of Education.

Student Test Scores and Characteristics: Student test scores and characteristics are provided by the State's Department of Education as an anonymized panel at the individual student level that includes the entire history of their standardized test sores as well as demographic and socio-economic characteristics, including race, ethnicity, gender, a proxy for low family income (free- and reduced-price lunch), English Language Learner (ELL), and special education status.

School Characteristics: The school's address is collected from publicly available data from the state's Department of Education. School×year average student and teacher characteristics are computed from their respective datasets. We compute shares of students in each ethnic and racial category, the share of free- and reduced-price lunch (FRL), ELL, and special education students, and the average class size.

Teacher Utility Shifter: The teacher utility shifter is the travel time between a teacher's home address and a school's address. To construct the shifter, we geocode teacher and school addresses and measure the driving times using the Google Maps API. We measure one way drive times on September 12, 2023 at 8 am.

School Principal Decisions Shifter: The principal decisions shifter is the difference in the number of position openings that the school posts between the ITS's second and first rounds in a given year. We calculate this using the district's administrative ITS data.

School Attendance Zones Data: The school attendance zoning information comes from the School Attendance Boundary Survey conducted by the National Center for Education Statistics with assistance from the U.S. Census Bureau to collect school attendance boundaries for the 2015-2016 school year. We use the school-level attendance zone shapefiles to identify all the unique elementary, middle, and high school attendance boundaries in our school district. We then identify the assigned schools for all teachers whose home address is located within our school district by matching the geo-coordinates of the addresses to the school zones spatially.

B. Model and Estimation Details

Students are denoted by k, year by t, teacher by i, school by s. Let K, T, N, and P be the sizes of the sets, respectively. The model consists of the following equations. The first one describes students outcomes y_{kit}, the second describes teacher utility u_{ist} and the third describes school principal decisions d_{ist}:

$$y_{kit} = w_{kt}\alpha + v_{it}\mu^y + x_{st}\zeta^y + c_{kit}\beta^y + \theta_i + \eta_{is}^y + \varepsilon_{kt}^y \tag{7}$$

$$u_{ist} = x_{st}\zeta^u + q_{ist}\beta^u + z_{ist}^u\phi^u + \gamma I_{ist} + \eta_{is}^u + \varepsilon_{ist}^u \tag{8}$$

$$d_{ist} = v_{it}\mu^d + q_{ist}\beta^d + z_{ist}^d\phi^d + \varphi_i + \eta_{is}^d + \varepsilon_{ist}^d \tag{9}$$

We rewrite η_{is} and (θ_i, φ_i) as

$$\begin{aligned}
\eta_{is}^y &= f_{is,1} \\
\eta_{is}^u &= \beta_1^u f_{is,1} + f_{is,2} \\
\eta_{is}^v &= \beta_1^v f_{is,1} + \beta_2^v f_{is,2} + f_{is,3} \\
\varphi_i &= f_{i,4} \\
\theta_i &= \beta_4 f_{i,4} + f_{i,5}
\end{aligned}$$

where $f_{is,1} \sim \mathcal{N}(0, \sigma_1^2)$, $f_{is,2} \sim \mathcal{N}(0, \sigma_2^2)$, $f_{is,3} \sim \mathcal{N}\left(0, \sigma_3^2\right)$, $f_{i,4} \sim \mathcal{N}(0, \sigma_4^2)$, $f_{i,5} \sim \mathcal{N}\left(0, \sigma_5^2\right)$. Then the model parameters are:

$$\begin{aligned}
\kappa &= \left(\alpha^y, \alpha^u, \alpha^v, \alpha^\pi, \gamma, \beta_1^u, \beta_1^v, \beta_2^v, \beta_4\right) \\
\sigma &= \left(\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_{\varepsilon y}\right)
\end{aligned}$$

Gibbs Sampler

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

Step 1: Data Augmentation

In this step we update the values of the latent variables u^1 and d^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 principals and teachers.

Step 2: Update κ conditional on u^1, d^1, f^0, σ^0

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

Step 4: Update the f's

Step 5: Update the $\{\sigma_i\}_{i=1}^5$

C. Comparing Results to the Existing Literature

In Section 5, we compare the estimated results from our model with the estimates from the literature, focusing on Jackson (2013), Ahn et al. (2024), and Delgado (2023). While it is easy to compare our estimates with Jackson (2013) and Ahn et al. (2024), a comparison with Delgado (2023) is not straightforward. In this section, we describe how to translate results from Delgado (2023) to compare with results from our model. In particular, we are interested in comparing the ratio of the standard deviation in match effectiveness to the standard deviation of general effectiveness in each model. This is a measure of the size of the estimated match effect in each case.

Most papers that study teacher comparative advantage characterize it over one observable binary student category (Biasi et al., 2022; Bates et al., 2024; Delgado, 2023). Following their model specification, let $j \in \{0, 1\}$ index student type, which can be low and high-income, low and high-achieving, or Black and non-Black categories. Using these student categories, the model of teacher effectiveness comes from a student test score model where k indexes a student, t a year, $i = i(kt)$ is a teacher and $j(k) = j$ is student k's type, assumed to be constant over time:

$$y_{kt} = M_{kt} + \mu_{ij} + \epsilon_{kt},$$

where M_{kt} includes the impact of student, classroom, and school characteristics on student test scores[1], and μ_{ij} is teacher i's effectiveness with students of type j. In this kind of model, *each teacher has two effectiveness parameters*, μ_{i0} and μ_{i1}, one for each student type.

Instead, we capture comparative advantage using the following formulation:

$$y_{kt} = M_{kt} + \theta_i + \delta_{ij} + \epsilon_{kt},$$

where $j(k) = j$ is student k's type. In our model, *the effectiveness of teacher i with students of type j is $\theta_i + \delta_{ij}$*. Using this notation, δ_{ij} is assumed to captures both observable

[1] M_{kt} can also include teacher experience

and unobservable match effects. Whereas other papers create two student-type-specific effectiveness measures for each teacher, we estimate one general effectiveness measure for each teacher that is the average across all students and define type-specific effectiveness added to that. In our application, the student type set is not binary.

Table C.1: Comparison of Match Effects Size

	Elementary Math	Elementary Reading
Panel A: Parameters reported in Delgado (2022)		
Variance of Black VA (μ_{i0})	0.053	0.028
Variance of non-Black VA (μ_{i1})	0.054	0.017
Covariance of Black and non-Black VA	0.048	0.022
Share of Black student (w_0)	0.370	0.370
Variance of CA ($\mu_{i0} - \mu_{i1}$)	0.003	0.005
Panel B: Our calculations		
Implied std(θ_i)	0.226	0.145
Implied std(δ_{ij})	0.013	0.019
Ratio: std($\delta_{ij})/std(\theta_i)$)	0.06	0.13

We describe how to translate between models to compare estimates—in particular how to compare the ratio of the standard deviation in match effectiveness to the standard deviation of general effectiveness. If the set of student types under both models is the same and described by $j \in \{0, 1\}$, then the effectiveness of teacher i with students of type j is,

$$\underbrace{\theta_i + \delta_{ij}}_{\text{our model}} = \underbrace{\mu_{ij}}_{\text{their models}}$$

$$= \begin{cases} \underbrace{\left(\omega_0\mu_{i0} + (1-\omega_0)\mu_{i1}\right)}_{\theta_i} + \underbrace{\omega_o(\mu_{i1} - \mu_{i0})}_{\delta_{i1}} & \text{if } j = 1 \\ \underbrace{\left(\omega_0\mu_{i0} + (1-\omega_0)\mu_{i1}\right)}_{\theta_i} + \underbrace{(\omega_o - 1)(\mu_{i1} - \mu_{i0})}_{\delta_{i0}} & \text{if } j = 0, \end{cases}$$

where ω_0 is the share of students of type 0. Here, $\theta_i = \omega_0\mu_{i0} + (1-\omega_0)\mu_{i1}$ is the mean impact of teacher i across student types, and δ_{ij} are deviations from the average effect for each student type.

Following this,

$$\begin{aligned} \text{var}(\theta_i) &= \omega_0^2\text{var}(\mu_{i0}) + (1-\omega_0)^2\text{var}(\mu_{i1}) + 2\omega_0(1-\omega_0)\text{cov}(\mu_{i0}, \mu_{i1}), \quad \text{and} \\ \text{var}(\delta_{ij}) &= (2\omega_0 - 1)^2\text{var}(CA_i) \end{aligned}$$

where $CA_i = \mu_{i0} - \mu_{i1}$ is the comparative advantage of i with students of type $j = 0$.

Taking estimates reported in Delgado (2023) we compute the values of $\text{var}(\theta_i)$ and $\text{var}(\delta_{ij})$ that are consistent with our model, and Delgado's data and student type space. Table C.1 reports Delgado's relevant estimates and our calculations. In Delgado's context match effects are about 6% of general effectiveness for Elementary Math and 13% for Elementary Reading. Our model predicts match effects are 9% of general effectiveness.

D. Additional Results

Table D.1: Descriptive Statistics of Schools

	Mean	*St. dev*	*p5*	*p95*
Share low-income	55.6	29.7	0.0	96.8
Share English language learner	23.8	19.3	1.7	61.2
Share Black	37.3	22.9	6.5	82.3
Share Hispanic	17.3	17.9	2.5	58.2
Share White	32.2	25.1	2.6	75.2
Share female	48.9	2.9	44.3	53.3
Share proficient - math	46.7	22.1	15.0	85.1
Share proficient - reading	46.9	22.8	13.6	84.5
Share female teachers	27.0	11.8	12.7	45.9
Average teacher experience	13.7	4.0	8.0	21.0
Average teacher age	43.1	4.1	36.0	50.0
School-year observations	504	504	504	504

Note: The table shows the mean, standard deviation, 5th percentile (*p5*), and the 95th percentile (*p95*) for various school level characteristics.

Table D.2: Correlation of Teacher Effectiveness and District Evaluation Measures

	District Evaluations			
	Math VA	*Reading VA*	*Student Surveys*	*Classroom Observations*
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 four measures of teacher effectiveness built by the district. Standard errors are in parenthesis. These include the district's own value-added (VA) measures constructed using classroom identifiers, as well as a measure based on student surveys and one based on scores on an established rubric from four classroom observations by trained peer teachers each year.

Table D.3: 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: In this table, 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 D.4: List of All Variables Used in Model Estimation

Student Outcomes	**Teacher Utility**	**School Principal Decisions**
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	Delta ITS rounds
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: This table presents a list of variables used in the estimation. For race & ethnicity, the ommited category is White. In the middle of our sample, the standardized test was redesigned. We capture this change with the dummy variable MCAS III. Minority is defined as not identifying White.

Table D.5: Correlation between share of teachers of same minority status and school decisions

	Position in Ranking	Pr(ranked first)	Pr(ranked)	Pr(interviewed)
	(1)	(2)	(3)	(4)
Share match-minority teachers	-0.383	0.078	0.129	-0.027
	(0.123)	(0.030)	(0.037)	(0.025)
Constant	3.818	0.146	0.341	0.500
	(0.096)	(0.023)	(0.029)	(0.020)
Observations	2,224	2,224	2,224	4,633

Note: The table shows regression of the share of peer teachers that share the minority status of a candidate teacher with the position of a teacher in the school's ranking. Panel (1) - (3) include every teacher-position pair where a teacher was interviewed, and panel (4) includes every teacher-position pair where the teacher applied for the position.

Table D.6: Counterfactual Summary: Maximize Average Test Scores

	Average Test Score Gains Relative to the Observed Assignment (SD)			
	Unconstrained	No Quits	+No additional Layoffs	+Same School
All	0.078	0.072	0.050	0.004
	(0.001)	(0.001)	(0.001)	(0.000)
By Achievement				
First quartile	0.069	0.062	0.036	0.003
	(0.001)	(0.001)	(0.001)	(0.001)
Second quartile	0.071	0.065	0.041	0.004
	(0.001)	(0.001)	(0.001)	(0.001)
Third quartile	0.080	0.074	0.054	0.004
	(0.001)	(0.001)	(0.001)	(0.001)
Fourth quartile	0.092	0.086	0.068	0.005
	(0.002)	(0.002)	(0.002)	(0.001)
By Race				
Non-white	0.071	0.064	0.040	0.005
	(0.001)	(0.001)	(0.001)	(0.000)
White	0.091	0.084	0.066	0.003
	(0.002)	(0.002)	(0.002)	(0.000)

Note: This table presents the average test score gains relative to the observed assignment in standard deviation terms (SD) for all students as well as for students by baseline achievement and race and ethnicity in a counterfactual where average student test scores are maximized. Standard errors in parenthesis.

Table D.7: Counterfactual Summary: Maximize Percent Proficient

	Gains in Share Proficient Relative to the Observed Assignment (SD)			
	Unconstrained	No Quits	+No additional Layoffs	+Same School
All	3.92	3.53	2.83	0.32
	(0.08)	(0.08)	(0.08)	(0.03)
By Achievement				
First quartile	0.69	0.59	0.44	0.05
	(0.06)	(0.05)	(0.05)	(0.02)
Second quartile	5.30	4.69	3.64	0.42
	(0.16)	(0.16)	(0.15)	(0.06)
Third quartile	7.48	6.78	5.52	0.61
	(0.21)	(0.20)	(0.20)	(0.09)
Fourth quartile	2.21	2.04	1.73	0.19
	(0.09)	(0.10)	(0.10)	(0.05)
By Race				
Non-white	3.92	3.48	2.75	0.36
	(0.09)	(0.09)	(0.09)	(0.04)
White	3.92	3.60	2.98	0.24
	(0.13)	(0.13)	(0.13)	(0.05)

Note: This table presents the gains in the percent of proficient students relative to the observed assignment for all students as well as for students by baseline achievement and race & ethnicity, in a counterfactual where the percentage of proficient students is maximized. Standard errors in parenthesis.

Table D.8: Decomposition: Maximizing the Average Classroom Test Scores

	Gains Relative to the Observed Assignment (SD)		
	Total Effect	Effectiveness	Match Effects
Unconstrained	0.037	0.032	0.006
	(0.001)	(0.001)	(0.005)
No Quits	0.035	0.030	0.005
	(0.001)	(0.001)	(0.005)
+ No additional Layoffs	0.001	−0.003	0.004
	(0.002)	(0.002)	(0.004)
+ Same School	0.000	−0.001	0.000
	(0.000)	(0.000)	(0.000)

Note: This table presents the decomposition of gains in the average student test scores relative to the observed assignment. Standard errors in parenthesis.

Table D.9: Estimated Parameters of the Student Outcomes Model - All Variables

	Mean	*Std. Dev.*
Panel A: Student Characteristics		
Previous score	0.770	0.001
Previous score sq	0.002	0.001
Previous score cube	-0.003	0
Race – Black	-0.126	0.003
Race - Hispanic	-0.066	0.004
Race - other	-0.061	0.004
Male	-0.012	0.002
Low income	-0.142	0.005
English language learner	-0.052	0.003
Special needs	-0.172	0.003
Panel B: Teacher Characteristics		
Male	-0.020	0.005
Education	-0.003	0.002
Black	0.006	0.008
Hispanic	0.001	0.009
Other	-0.002	0.007
Experience 2 to 3	0.020	0.004
Experience 4 to 6	0.022	0.005
Experience 7+	0.027	0.005
Panel C: School Characteristics		
Race - % Hispanic	0.037	0.019
Race - % Black	0.128	0.021
Race - % other	0.167	0.023
% Low income	-0.023	0.018
% English language learner	-0.040	0.014
% Special needs	0.034	0.019
Low Income * % Low income	0.071	0.007
Panel D: Student–Teacher Interactions		
Match minority student	0.003	0.002
Same-gender student	0.005	0.002
Teacher education * Student minority	0.005	0.002
Teacher education * Student previous score	-0.003	0.001
Teacher experience * Student minority	0.002	0.002
Teacher experience * Student previous score	-0.002	0.001
Std. Dev. of teacher general effectiveness (θ_i)	0.0815	
Std. Dev. of teacher–school unobservables match effectiveness (η^y_{is})	3.5e-04	

Note: The table shows the mean and the standard deviation of the estimated chains of the parameters of the student outcomes model (Equation 4), on the full set of characteristics. Panel A presents parameters associated with student's own characteristics. For the race dummy, the omitted category is White. Panels B and C present the parameters associated with the teacher and school characteristics. Panel D presents parameters associated with the interaction between student and teacher characteristics. Minority is defined as non-White. Note that we report the standard deviation, and not the variance of θ_i as well as η^y_{is}.

Table D.10: Estimated Parameters of Teacher Utilities - All variables

	Mean	*Std. Dev.*
Panel A: School Characteristics – Students		
% Low income	-0.100	0.049
% English language learner	0.074	0.270
% Special Needs	0.975	0.430
Average test scores	0.010	0.120
% Hispanic	-1.298	0.375
% Black	-0.219	0.348
% other	2.237	0.513
Panel B: School Characteristics – Teachers		
% Hispanic	-0.543	0.449
% Black	0.290	0.369
% other	1.221	0.464
Average teacher experience	0.047	0.020
Inertia	3.582	0.034
Panel C: Teacher–School Interactions		
% of students match minority	0.499	0.049
% of teachers match minority	0.454	0.049
Teacher education * Average test scores	0.119	0.028
Teacher education * % minority	0.098	0.022
Teacher experience * Average test scores	-0.062	0.037
Teacher experience * % minority	-0.318	0.026
Teacher experience * Average teacher experience	0.008	0.013
Panel D: Shifter		
Driving time	-0.010	0.001
Std. Dev. of idiosyncratic taste shock for schools (η^u_{is})	0.493	

Note: The table shows the mean and the standard deviation of the estimated chains of the parameters of the teacher utility model (Equation 5), on the full set of characteristics. Panels A and B presents parameters associated with the school's students and teachers characteristics, respectively. Driving time is measured in minutes. Panel C presents parameters associated with the interaction between the teacher and school characteristics . Minority is defined as non-White. We report the standard deviation, and not the variance for η^u_{is}.

Table D.11: Estimated Parameters of School Principal Decisions - All Variables

	Mean	*Std. Dev.*
Panel A: Teacher Characteristics		
Male	0.205	0.079
Education	-0.250	0.305
Experience	0.236	0.336
% Hispanic	-0.534	0.298
% Black	-0.783	0.139
% other	-0.489	0.143
Panel B: Teacher–School Interactions		
% of students match minority	1.230	0.119
Teacher education * Average test scores	0.130	0.236
Teacher education * % minority	0.424	0.462
Teacher experience * Average test scores	0.051	0.260
Teacher experience * % minority	-0.463	0.503
Teacher experience * Average teacher experience	0.033	0.041
Panel C: Shifter		
Delta ITS rounds	0.037	0.026
Std. Dev. of teacher effects (φ_i)	0.119	
Std. Dev. of idiosyncratic taste shock for teachers (η_{is}^{v})	0.020	

Note: The table shows the mean and the standard deviation of the estimated chains of the parameters of the school principal decision model on the full set of characteristics. Panel A presents parameters associated with teacher characteristics. Panel B presents parameters associated with the interaction between the teacher and school characteristics. Minority is defined as non-White. We report the standard deviation, not the variance, for φ_i and η_{is}^{v}.

D.1 Alternative School Principal Decisions Shifter

Table D.12: Estimated Parameters of the Outcomes Model - Alternative Shifter

	Mean	*Std. Dev.*
Panel A: Student Characteristics		
Previous score	0.77	0.001
Race - Black	-0.127	0.003
Male	-0.012	0.002
Low income	-0.142	0.005
Panel B: Teacher Characteristics		
Male	-0.022	0.005
Education	-0.003	0.002
Experience 2 to 3	0.02	0.004
Experience 4 to 6	0.022	0.005
Experience 7+	0.028	0.005
Panel C: School Characteristics		
% low income	-0.024	0.018
Panel D: Student-teacher Interactions		
Match minority student	0.003	0.002
Same gender student	0.005	0.002
Education * Minority	0.005	0.002
Education * Previous Score	-0.003	0.001
Experience * Minority	0.002	0.002
Experience * Previous Score	-0.002	0.001
Std. Dev. of Teacher General Effectiveness (θ_i)	0.0815	
Std. Dev. of Teacher-School Unobservable Match Effectiveness (η^y_{is})	3.5e-04	

Note: The table shows the mean and the standard deviation of the last realizations from the chains of each estimated parameter when we use the alternative school principal shifter as described in Section 5.3. Panel A presents parameters associated with student's own characteristics. For the race dummy, the omitted category is White. Panels B and C present the parameters associated with the teacher and school characteristics. Panel D presents parameters associated with the interaction between student and teacher characteristics. Minority is defined as non-White. Note that we report the standard deviation, and not the variance of θ_i as well as η^y_{is}.

Table D.13: Estimated Parameters of Teacher Utility - Alternative Shifter

	Mean	*Std. Dev.*
Panel A: School Characteristics - Students		
% Low Income	-0.100	0.049
Average test scores	0.012	0.121
% Black	-0.224	0.347
Panel B: School Characteristics - Teachers		
% Black	0.308	0.375
Average teacher experience	0.046	0.020
Driving time	-0.010	0.001
Inertia	3.584	0.034
Panel C: Teacher-School Interactions		
% of students match minority	0.497	0.049
Education * Average test scores	0.120	0.028
Education * % minority	0.099	0.021
Experience * Average test scores	-0.063	0.036
Experience * % minority	-0.318	0.026
Experience * Average teacher experience	0.008	0.013
Std. Dev. of Teacher-School Unobservable Match Effects (η^u_{is})	0.492	

Note: The table shows the mean and the standard deviation of the last realizations from the chains of each estimated parameter when we use the alternative school principal shifter as described in Section 5.3.

Table D.14: Estimated Parameters of Principal Decision - Alternative Shifter

	Mean	*Std. Dev.*
Panel A: Teacher Characteristics		
Male	0.027	0.082
Education	-0.297	0.31
Experience	0.22	0.346
Race - Black	0.663	0.183
Panel B: Teacher-School Interactions		
% of students match minority	-0.341	0.177
Education * Average test scores	0.098	0.242
Education * % minority	0.467	0.470
Experience * Average test scores	-0.003	0.266
Experience * % minority	-0.149	0.521
Experience * Average teacher experience	0.010	0.041
Panel C: Shifter		
% of teachers match minority	1.136	0.097
Std. Dev. of Teacher Effects (φ_i)	0.185	
Std. Dev. of Teacher-School Unobservable Match Effects (η^v_{is})	0.013	

Note: The table shows the mean and the standard deviation of the last realizations from the chains of each estimated parameter when we use the alternative school principal shifter as described in Section 5.3.

Table D.15: Correlations between Decisions and Outcomes - Alternative Shifter

	Outcomes Δy_{kt}	School Decision Δv_{ist}	Teacher Utility Δu_{ist}
From percentile 1 *to* 99			
Teacher general effectiveness, θ_i	0.388	0.832	
	(0.021)	(0.20)	
Unobservable match effects, η^y_{is}	0.002	0.001	-1.269
	(0.0002)	(0.052)	(0.402)
Observable match effects	0.036	0.127	-0.081
	(0.005)	(0.745)	(0.041)

Note: This table presents the decomposition of the gains in average test scores in standard deviation (SD) terms under the four counterfactual scenarios relative to the outcomes under the observed assignment when we use the alternative school principal shifter as described in Section 5.3. Standard errors in parentheses.

D.2 Additional Controls for Teachers' Utility Shifter

Table D.16: Estimated Parameters of the Student Outcomes Model - Model with Attendance Boundary Dummy

	Mean	*Std. Dev.*
Panel A: Student Characteristics		
Previous score	0.770	0.001
Race – Black	-0.127	0.003
Male	-0.012	0.002
Low income	-0.142	0.005
Panel B: Teacher Characteristics		
Male	-0.022	0.005
Education	-0.003	0.002
Experience 2 to 3	0.019	0.005
Experience 4 to 6	0.021	0.005
Experience 7+	0.026	0.005
Attendance Boundary dummy	-0.032	0.009
Panel C: School Characteristics		
% low income	-0.024	0.018
Panel D: Student–Teacher Interactions		
Match minority student	0.003	0.003
Same-gender student	0.005	0.002
Teacher education * Student minority	0.005	0.002
Teacher education * Student previous Score	-0.003	0.001
Teacher experience * Student minority	0.002	0.002
Teacher experience * Student previous Score	-0.002	0.001
Std. Dev. of teacher general effectiveness (θ_i)	0.0811	
Std. Dev. of teacher–school unobservables match effectiveness (η^y_{is})	3.5e-04	

Note: The table shows the mean and the standard deviation of the estimated chains of the parameters of the student outcomes model (Equation 4), including attendance boundary dummies in the outcomes equation. Panel A presents parameters associated with student's own characteristics. For the race dummy, the omitted category is White. Panels B and C present the parameters associated with the teacher and school characteristics. Panel D presents parameters associated with the interaction between student and teacher characteristics. Minority is defined as non-White. Note that we report the standard deviation, and not the variance of θ_i as well as η^y_{is}.

Table D.17: Estimated Parameters of Teacher Utilities - Model with Attendance Boundary Dummy

	Mean	*Std. Dev.*
Panel A: School Characteristics – Students		
% Low Income	-0.100	0.047
Average test scores	0.011	0.121
% Black	-0.211	0.333
Panel B: School Characteristics – Teachers		
% Black	0.293	0.374
Average teacher experience	0.049	0.021
Inertia	3.585	0.034
Panel C: Teacher–School Interactions		
% of students match minority	0.500	0.051
% of teachers match minority	0.455	0.048
Teacher education * Average test scores	0.122	0.029
Teacher education * % minority	0.099	0.022
Teacher experience * Average test scores	-0.065	0.037
Teacher experience * % minority	-0.319	0.026
Teacher experience * Average teacher experience	0.007	0.014
Panel D: Shifter		
Driving time	-0.01	0.001
Std. Dev. of idiosyncratic taste shock for schools (η_{is}^{u})	0.494	

Note: The table shows the mean and the standard deviation of the estimated chains of the parameters of the teacher utility model (Equation 5), including attendance boundary dummies in the outcomes equation. Panels A and B presents parameters associated with the school's students and teachers characteristics, respectively. Driving time is measured in minutes. Panel C presents parameters associated with the interaction between the teacher and school characteristics . Minority is defined as non-White. We report the standard deviation, and not the variance for η_{is}^{u}.