NBER WORKING PAPER SERIES

TEACHER LABOR MARKET POLICY AND THE THEORY OF THE SECOND BEST

Michael D. Bates
Michael Dinerstein
Andrew C. Johnston
Isaac Sorkin

Working Paper 29728
http://www.nber.org/papers/w29728

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2022, Revised August 2024

Previously circulated as "Teacher Labor Market Equilibrium and Student Achievement." We thank Nikhil Agarwal, Claudia Allende, Peter Arcidiacono, Natalie Bau, Christina Brown, Jediphi Cabal, Dennis Epple, Felipe Goncalves, Caroline Hoxby, Peter Hull, Adam Kapor, Neale Mahoney, Richard Mansfield, Maxim Massenkoff, Derek Neal, Parag Pathak, Luigi Pistaferri, Jesse Shapiro, Camille Terrier, and numerous seminar and conference participants for helpful comments and conversations. Thanks to Ian Calaway for research assistance. We also thank representatives at an unnamed district, Kara Bonneau, and the North Carolina Education Research Data Center for data merging and access. Bates thanks the Blum Initiative and Sorkin thanks the Alfred P. Sloan Foundation for support. This paper subsumes drafts previously circulated under the titles “Teacher Labor Market Equilibrium and Student Achievement” and “Teacher Labor Market Equilibrium and the Distribution of Student Achievement.” This research was approved by the Stanford IRB (#43865), the UC Merced IRB (UCM2019-88), and the UC Riverside IRB (HS-16-094). It was deemed non-human subjects research by the University of Chicago IRB (IRB17-1651). Mistakes are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Teacher Labor Market Policy and the Theory of the Second Best
Michael D. Bates, Michael Dinerstein, Andrew C. Johnston, and Isaac Sorkin
NBER Working Paper No. 29728
February 2022, Revised August 2024
JEL No. I28, J08, J45

ABSTRACT

We estimate a matching model of teachers and elementary schools with rich data on teacher applications and principal ratings from a large, urban district in North Carolina. Both teachers' and principals' preferences deviate from those that would maximize the achievement of economically disadvantaged students: teachers prefer schools with fewer disadvantaged students, and principal ratings are weakly related to teacher effectiveness. In equilibrium, these two deviations combine to produce a surprisingly equitable current allocation, where teacher quality is balanced across advantaged and disadvantaged students. To close achievement gaps, policies that address deviations on one side alone are ineffective or harmful, while policies that address both could substantially increase the achievement of disadvantaged students.

Michael D. Bates
3135 Sproul Hall
Department of Economics
University of California, Riverside
Riverside, CA 92521
United States
michael.bates@ucr.edu

Michael Dinerstein
Department of Economics
Duke University
213 Social Sciences Bldg.
Box 90097
Durham, NC 27708
and NBER
michael.dinerstein@duke.edu

Andrew C. Johnston
Department of Economics
University of California, Merced
5200 N Lake Road
Merced, CA 95340
and NBER
acjohnston@ucmerced.edu

Isaac Sorkin
Department of Economics
Stanford University
579 Jane Stanford Way
Stanford, CA 94305
and NBER
sorkin@stanford.edu

A data appendix is available at http://www.nber.org/data-appendix/w29728

Achievement gaps between economically advantaged and disadvantaged children are large and have expanded over time (Reardon, 2013). These achievement differences in childhood predict significant differences in adult wages (Neal and Johnson, 1996). Recent work argues that the relationship between achievement and earnings is causal—teachers that causally improve achievement also improve adult income (Chetty, Friedman, and Rockoff, 2014b). Because teachers vary widely in their effect on student achievement (Hanushek, Kain, and Rivkin, 2004), adjusting the *allocation* of teachers to students may be one of the most potent policy instruments for helping the disadvantaged.

In this paper, we study the allocation of teachers to schools and assess the desirability of various policies. We combine rich data with the economics of two-sided markets to understand the current allocation and explore implications for policies.[1] We combine three empirical findings, each the subject of a large literature. First, within a district, teacher quality is balanced across disadvantaged and advantaged students. Second, teachers prefer to teach at schools with more advantaged students. Third, in hiring, principals mostly do not select high value-added teachers.

Conceptually, there is a tension between finding a balanced allocation and systematic teacher preferences for schools with advantaged students. We typically expect the entity that faces excess supply to have either lower prices or, if prices are restricted as in teacher labor markets, higher quality. In our setting, that intuition would imply that the schools with advantaged students would have better teachers. The third empirical finding offers a simple reconciliation: because principals do not hire (very much) on the basis of the value-added, they do not take advantage of the excess supply to hire better teachers.

This reconciliation is not just helpful in explaining the current allocation, it also provides useful guidance for policy. These results reflect "theory of the second best" (Lipsey and Lancaster (1956)) logic: when there are multiple deviations from policies that implement the first best, fixing any one can be ineffective or harmful. Suppose the planner wants to maximize the achievement of disadvantaged students. The district would approximately achieve the first best if teachers preferred teaching disadvantaged students and principals hired the best teachers. But when teachers' preferences and principal hiring both deviate from these benchmarks, policies that only target one side at a time may be ineffective or even harmful.

We empirically show that providing bonuses to teachers for teaching disadvantaged students as suggested by, e.g., Clotfelter, Ladd, and Vigdor (2011) and Goldhaber, Quince, and Theobald (2018), does not improve the quality of the teachers of disadvantaged students. The reason is simple: while bonuses increase the supply of teachers to schools with disadvantaged students, the principals do not take advantage of excess supply to hire better teachers. Similarly, providing bonuses or incentives to principals to hire based on value-added as suggested by, e.g., Ballou (1996) and Jacob et al. (2018a), backfires in that now the schools with advantaged students use the excess supply to hire better teachers, generating the unequal outcomes we expect based on the teacher preferences. Implementing these two types of policies in tandem, however,

[1] Historically in the United States, and in many international contexts, teacher labor markets were one-sided where principals had little discretion over which teachers to hire. But mutual consent, where both teacher and principal must agree to a match, has risen sharply in the US over the last few decades (Engel, Cannata, and Curran, 2017) such that 92% of large districts have now adopted it (National Council on Teacher Quality, 2022) and operate two-sided markets.

achieves an allocation close to the first best.

We begin with a model of the teacher labor market that enables us to specify a first-best benchmark and how the market clears. Teachers apply to vacancies and principals hire among applicants. We assume that the equilibrium allocations are pairwise stable (Roth and Sotomayor, 1992; Hitsch, Hortaçsu, and Ariely, 2010; Banerjee et al., 2013; Boyd et al., 2013).[2] Given an objective of maximizing the achievement of disadvantaged students, the model provides a policy benchmark: the first-best allocation can be approximately achieved if teachers prefer positions with the most disadvantaged students and principals prefer to hire the most effective teachers.[3]

To estimate the model, we use detailed data from a large urban school district in North Carolina and we focus on the market for elementary school teachers. To evaluate the current allocation, we observe teachers linked to their students' yearly test scores. To understand teachers' preferences, we observe the full set and timing of job applications that teachers submit as well as the timing of job postings. To understand principals' behavior, we see the full set of applications the principal receives, notes the principal records about applications, interviews, and offers.

We first specify our empirical model of how teachers affect student math achievement. For our baseline, we allow teachers to have different value-added with disadvantaged and advantaged students (Condie, Lefgren, and Sims, 2014; Delgado, 2023). But all of the results of the paper hold with a variety of more conventional homogeneous value-added models.

To identify teacher and principal preferences, we rely on relatively weak assumptions that allow for transparent identification. We focus on actions that are early in the process: the teacher decision to apply and the principal rating of applicants, rather than, say, the teacher decision to accept an offer or the principal decision to make an offer. This focus allows for strategy later on in the process. Based on institutional features and extensive analysis of applicant behavior, we argue that teachers apply non-strategically to vacancies when they are active, and principal rating behavior is non-strategic.

Our estimates of teacher and principal preferences broadly reproduce patterns in the literature that imply that teacher and principal preferences do not implement the first-best allocation. Relative to the literature, we have several advantages: we have actual choices, we observe choice sets, and we allow for preference heterogeneity. Consistent with the literature (e.g., (Greenberg and McCall, 1974; Antos and Rosen, 1975; Barbieri, Rossetti, and Sestito, 2011; Engel, Jacob, and Curran, 2014; Bonhomme, Jolivet, and Leuven, 2016; Fox, 2016; Johnston, 2024)), our key teacher finding is that teachers prefer schools with fewer disadvantaged students (though we find significant heterogeneity) so that teacher preferences differ from that would implement the first-best allocation in a market with uniform pay. Also consistent with the literature (e.g, (Ballou, 1996; Boyd et al., 2011; Jacob et al., 2018b; Jatusripitak, 2018; Hinrichs, 2021)), our key principal finding is that the principal's preferred candidate is rarely the one that is most effective at raising student test scores. We can also reject the vertical preference model typically assumed in settings where

[2] Our model fits in a recent literature considering allocation problems with non-choice outcomes (Agarwal, Hodgson, and Somaini, 2020; Ba et al., 2021; Cowgill et al., 2024; Dahlstrand, 2024).

[3] The statement is approximate because we allow for timing restrictions and comparative advantage.

choice sets are not observed (see Diamond and Agarwal, 2017). The lack of weight on test scores could reflect a lack of information or a lack of incentives, but regardless, principal behavior differs from that which implements the first-best allocation.

For our first main result, we show that advantaged and disadvantaged students have teachers of approximately equal strength. This pattern is present in raw test score gains, for a wide variety of value-added models, and for behavioral value-added. One notable feature of the current allocation is that disadvantaged students are more likely to have novice teachers for whom we cannot estimate value-added in our main models. We find a similar pattern when we rely either on a residual value-added estimator that just uses contemporaneous data and so also includes novice teachers, or when imputing value-added using observable characteristics.[4]

To understand why advantaged and disadvantaged students have teachers of approximately equal strength, we combine our estimates with the two-sided matching model. We show, consistent with the theory of the second best, that this result reflects multiple deviations combining to produce favorable allocations for disadvantaged students. If teachers only care about the number of disadvantaged students, then the allocation is little changed. The intuition is that because principals place little weight on value-added, they do not select more effective teachers from the larger applicant pool. Similarly, if principals only placed weight on value-added, then the outcome would be worse for disadvantaged students.

While these results explain the current allocation, they also inform the design of policies that the teacher preferences and principal hiring literatures recommend. Specifically, the basic message follows the theory of the second best: policies that only target one side of the market can be ineffective or even harmful.

While the current allocation is surprisingly balanced, it does not achieve the first-best allocation that maximizes the outcomes of disadvantaged students. We find that the first best would provide substantive gains: in a single year, reallocation could close one-fourteenth of the baseline achievement gap, while increasing average achievement. When extrapolating linearly, the achievement gap could largely be closed over twelve years of public education. Implementing extreme versions of teacher and principal policies jointly nearly achieves the first-best allocation.

We consider a number of extensions. We first show that our main results are robust to alternative assumptions about how the market clears—specifically, our estimates of the status quo are nearly identical

[4]Papers that study differences in teacher value-added by level of student disadvantage tend to find little to no difference. Glazerman and Max (2011) find no gap at the elementary level. Chetty, Friedman, and Rockoff (2014b) estimate a tiny correlation between teacher value-added and student income. Mansfield (2015) argues the allocation is fairly equitable, with a gap of 0.025 student standard deviations (σ) for high-school students. Isenberg et al. (2022) estimate a gap of 0.004σ. Other papers find mixed results that are sometimes sensitive to specification. Sass et al. (2012) estimate a gap of 0.04σ, though the gap varies with specification and disadvantaged students actually have larger year-to-year raw gains. Isenberg et al. (2013) estimate a 0.024σ gap that disappears when the value-added model controls for peer characteristics. Several papers use Washington state data and find larger gaps. For example, Goldhaber, Lavery, and Theobald (2015) estimate minimal gaps in fourth-graders' exposure to low-quality teachers though larger ones higher in the teacher quality distribution, and Goldhaber, Theobald, and Fumia (2022) find gaps of 0.02–0.03σ and 0.013–0.017σ for fourth and fifth grade, respectively. In the same Washington state context, Goldhaber, Lavery, and Theobald (2016) note that teacher transfer patterns differ depending on whether teacher experience offers transfer priority, consistent with our results that one-sided and two-sided markets deliver different levels of equity in their allocations. Angrist et al. (2024) find similar school value-added for advantaged and disadvantaged students.

when clearing the market at once or across multiple sub-periods. Most importantly, one argument for making principals more likely to hire high value-added teachers is that this affects the extensive margin of who works in the district. We therefore study an extension where we include teachers for whom we have to impute value-added. Making principals better at hiring does bring better teachers to the district, but given the structure of teacher preferences, advantaged (not disadvantaged) students largely benefit from the better teachers. Thus, the basic theory of the second best message persists even with an active extensive margin.

We also study more realistic teacher bonuses. We first consider one-sided bonuses that provide incentives for teachers to teach at schools with disadvantaged students. Such bonuses are only weakly effective because they do not affect how principals hire from the pool. If principals hire to maximize value-added (via some combination of information and incentives), then the teacher bonuses are effective.

To summarize, the unifying theme of this paper is the theory of the second best. Subsidizing one side of the market at a time can be ineffective or counterproductive, even when subsidizing both sides is beneficial, and the current allocation is balanced even though teachers' preferences suggest it would not be. Our results challenge the conclusions from the prior one-sided literatures by explaining an otherwise puzzling feature of the current allocation and reaching opposite conclusions about policy effectiveness. Reaching these conclusions requires rich data on the actions of both sides of the market.

This paper fits in a growing literature on equilibrium models of the teacher labor market. These papers tend to fall into two camps. In the first, the hiring side of the market faces constraints imposed by the government (e.g., they must hire the most experienced applicant) such that the market is essentially one-sided (Bobba et al., 2024; Combe et al., 2022; Elacqua et al., 2021; Tincani, 2021; Combe, Tercieux, and Terrier, 2022).[5] We instead focus on two-sided labor markets, which characterize nearly all teacher labor markets in the US and the hiring of permanent teachers in many non-US settings. In the second camp, several papers study two-sided markets but infer preferences from data on equilibrium allocations (Boyd et al., 2013; Bates, 2020; Biasi, Fu, and Stromme, 2021). We instead observe the actions of each side of the market, which allows us to relax the strong assumptions necessary for identification in the absence of such data. We show that these assumptions deliver misleading conclusions about the relationship between teacher quality and student disadvantage in equilibrium as well as the desirability of commonly-suggested policies. Like us, Davis (2022) and Ederer (2023) study two-sided markets with data on each side's actions.[6] Unlike these papers, we estimate teacher quality based on student test scores instead of relying on observable teacher characteristics. We find that restricting teacher quality to vary only with observable characteristics changes the assessment of equilibrium and policy conclusions; for example, we find that the allocation is not balanced across advantaged and disadvantaged students in terms of teacher observables (like experience) despite parity on multiple direct measures of effectiveness. This finding comes from our detailed data linking teachers to students, and their test scores, at a finer level than many papers in the literature.

Our study carries important lessons for the analysis of labor markets. Much of the labor literature,

[5] Bau (2022) studies an equilibrium model of school competition with school-student match effects.

[6] Laverde et al. (2023) also study a two-sided market with data on each side's actions.

on topics such as wage inequality (e.g., Card et al., 2018) and amenities (e.g., Sorkin, 2018), relies on matched employer-employee data where researchers only observe equilibrium allocations. These markets are two-sided, which forces researchers to rely on the same identifying assumptions that led to misleading conclusions in the teachers literature. Our findings thus reinforce Oyer and Schaefer (2011)'s call for labor economists to study how firms hire workers and Card et al. (2018)'s suggestion that the labor literature on imperfect competition would benefit from "IO-style" case studies of particular markets.

Finally, this paper relates to the industrial organization literature on information in matching markets. Work examining the centralized assignment of students to schools has found that incomplete information can limit student gains from being strategic (Kapor, Neilson, and Zimmerman, 2020) and can lead to costly search (Chen and He, 2021; Arteaga et al., 2022). We show that an information intervention that endows one side of the market with complete information can make that side worse off, especially when preferences are close to vertical.

1 An equilibrium model of the teacher labor market

Here, we write down an equilibrium model of the within-district teacher labor market. The model clarifies the set of factors shaping the equilibrium, allows us to define the first-best allocation, and explains when the decentralized equilibrium attains the first-best allocation.

1.1 Set-up

Teacher j derives utility u_{jk} from teaching at school k. School k's principal derives utility, v_{jk}, from hiring teacher j. Utility is non-transferable, as wages are set by the district and do not vary across assignments for a given teacher.[7]

A teacher-school assignment produces value-added VA_{jk}. Because we are interested in the achievement of disadvantaged and advantaged students, we allow the value-added to depend on the student type. Specifically, let μ_{jm} be teacher j's value-added with students of type m, where $m \in \{0,1\}$ indicates whether a student is disadvantaged. Let n_{km} be the number of students in school k of type m. Then:

$$VA_{jk} = n_{k0}\mu_{j0} + n_{k1}\mu_{j1}. \tag{1}$$

Finally, let $\mathcal{J}$ be the set of teachers, $\mathcal{K}$ be the set of schools, and assume for simplicity that the number of teachers and schools is the same. An assignment of teachers to classrooms is a one-to-one and onto function (bijection): $\phi : \mathcal{J} \to \mathcal{K}$ so that $\phi(j) = k$, the school k to which teacher j is assigned.[8] Denote by Φ the set of all possible assignments.

[7] This assumption also excludes transferable non-pecuniary benefits, such as favorable class assignments.

[8] For simplicity, we specify each school as having a single position. When we estimate the empirical model, schools may have multiple positions.

1.2 First-best allocation

We are interested in policies that increase the achievement of disadvantaged students. We take as given the set of teachers and positions the district has and ask how to assign them. In Section 8, we consider the set of teachers who apply in the transfer system and for whom we can estimate value-added: this set includes teachers who have previously taught anywhere in the state.[9]

The district values the achievement of disadvantaged students:[10]

$$\max_{\phi \in \Phi} \left\{ \sum_{j \in \mathcal{J}} n_{k1} \mu_{j1} \right\}. \tag{2}$$

The structure of the first-best allocation is simple: rank teachers in descending order by value-added with disadvantaged students and rank classrooms in descending order by the number of disadvantaged students. Then assign the strongest teacher to the classroom with the largest number of disadvantaged students and so on.[11]

Because the paper's goal is to study the allocation of teachers, and not how best to use existing dollars, we do not include a budget constraint in the district's problem. As cost is still a relevant consideration in evaluating allocations, in Section 9 we compare the effectiveness of policies that cost equal amounts.

1.3 Decentralized equilibrium

Our equilibrium concept is (timing-constrained) pair-wise stability. Schools meet with all teachers who are in the market at the same time. Under a stable allocation, no teacher and school pair would prefer to jointly deviate and match (Roth and Sotomayor (1992), Definition 2.3). Stability is a natural assumption in decentralized markets as it says that pairwise gains from trade have been exhausted (i.e., the set of stable allocations are the core).

To model the empirical status quo, we assume (1) teachers and principals have the preferences we estimate for them and (2) the timing of the market follows that which we observed in the administrative records, where not all matches are feasible. There is not necessarily a unique stable equilibrium. We model the status quo using the teacher-proposing deferred-acceptance algorithm (DA), which we use to find a stable equilibrium, not because DA is actually used.

When does the decentralized equilibrium correspond to the planner problem? Suppose that teachers rank schools according to the number of disadvantaged students ($u_{jk} \propto n_{k1} \forall j,k$) and principals rank teachers according to total output ($v_{jk} \propto VA_{jk} \forall j,k$). Then in the absence of comparative advantage or timing restrictions, the decentralized equilibrium—which is unique in this case—corresponds to the planner's solution. Notably, this combination of rankings is what the *joint* implementation of hard-to-staff school bonuses

[9] If we considered all possible teachers in the single district's problem (including potential teachers and those who do not apply to the district), then we would be ignoring how our focal district's behavior affects the allocation of teachers to and within other districts. The allocation problem then would no longer map into a social planner's problem.

[10] In Appendix B, we include advantaged students' achievement.

[11] Table 5 (Part 1) shows that our results are very similar if we hold class sizes constant.

and guided principal hiring would achieve. Of course, the theory of the second best says that aligning only the teacher or the principal with the planner may not improve outcomes.

Empirically, we are then interested in the extent to which teacher and principal preferences align with those that decentralize the planner's solution. We are also interested in whether the other factors we have abstracted from—timing and comparative advantage—affect the gap between the decentralized equilibrium and the planner's solution.

2 Data and institutional context

We use rich data on the labor market for elementary school teachers. Elementary schools are grades K to 5 (or sometimes 6). For the purpose of estimating the variance of classroom effects, we also use data from middle schools (grades 6 to 8) where teachers are more likely to teach multiple classrooms. The first type of data comes from the platform used to hire teachers in our focal district. We use this data to estimate teacher and principal preferences. The second type of data comes from staffing and achievement records from state accountability records. This data provides us with student-level test score data that we link to teachers and use to estimate value-added models. In addition, these records provide information about a variety of demographic characteristics of teachers and students as well as teachers' education and experience in the district. In this section, we briefly describe the data. See Appendix A for further details and Appendix Table A1 for summary statistics across samples.

2.1 Job application and vacancy data

We obtained application records from our focal district's system, which spans 2010 through 2019 and records 346,663 job applications. In the system, schools post job vacancies, and applicants apply for jobs. The system also records various actions that principals take.

For every posted position, the vacancy files indicate the school, position title, and whether the position is full-time or part-time. We use the detail on the position title to isolate non-specialized elementary school teacher jobs (i.e., we omit elementary school jobs such as "literary facilitator elementary").

We use two features of the teacher file. First, the file records which vacancies the candidate applied to, and when she submitted the application. The timing information allows us to construct choice sets, which we detail in Section 3. Second, the file records the city, zip code, and address where the teacher lives. This feature allows us to construct the commute time for each teacher-position combination.

We also have data in which principals record their assessments of teachers. Principals record their interest in different applicants, the equivalent of a "good" and a "bad" pile. Principals also record which candidates they invited to interview, which candidates were offered the position, and which candidates were hired.

2.2 Administrative data

We link the platform data to state administrative records on teachers and students. For teachers, we have their experience, salary, licensing, certification status, test scores, class assignments, and the school where they work. For students, we have scores on standardized exams, grades, race, sex, and whether they qualify as disadvantaged based on Federal programs. Records on class assignments allow us to link teachers to students.

The North Carolina Education Research Data Center (NCERDC) matched the data from the job-market platform to the state's administrative data, using names, birth dates, and the last four digits of teachers' social security numbers. For teachers who had a sufficiently good match (that is, a unique name-birth-year combination), we have a de-identified ID that allows us to connect their platform data to their staffing records and students' achievement. Appendix Table A2 shows the share of newly hired teachers in the district that we find in our job market platform data. The lowest rate is 94% and in our focal year it is within rounding error of 100%.

The data show that student types vary considerably across teachers, which is driven by the sorting of students across schools. Appendix Figure A1 plots the fraction of a teacher's students that are economically disadvantaged. Almost a third of teachers have classrooms with almost entirely economically disadvantaged students. In Appendix Table A3, we show that this pattern reflects sorting of students across schools rather than across classrooms within schools. Specifically, the adjusted-R^2 of a regression predicting disadvantaged students is 0.4 using either school or classroom dummies. Similarly, the peer share of disadvantaged students that are disadvantaged is around 70% when using schools or classrooms. Given such student sorting across schools, different allocations of teachers to schools have the potential to yield very different learning gains for disadvantaged students.

2.3 Market overview

Our district organizes a decentralized hiring and transfer process in which teachers choose where to apply and principals choose whom to hire. External and internal (transfer) applicants are pooled into one market. Here we describe the basic market structure.

Market organization: The school district runs a centralized online hiring platform, where each school posts openings. Teachers choose whether to apply to each posting.

Timing: We examine the "on-cycle" part of the market, which dictates hiring and transfers between school years. It begins in the winter, when the district notifies each school of known and expected attrition among the school's work force and of how many positions that school may hire. It ideally ends with filled positions by late August before the new school year. Similar to what Papay and Kraft (2016) find, some schools are unable to fill all positions by the start of the new school year.

Postings: The number of postings at a school reflects a combination of enrollment, budget, and the number of teachers who leave. All three pieces of information are not necessarily known before the main hiring season starts. This information delay generates variation within and across schools in the timing of postings. For example, late information about enrollment or budget fluctuations often necessitates late posting. Or if there is mid-year attrition, then the school would know long before hiring season started that there would be a vacancy, which allows for early posting.

Applications: An application consists of a variety of documents, including teacher certification and a brief diversity statement. The same set of documents applies to all positions. Thus, a prospective teacher faces a fixed cost of preparing materials but little marginal cost to apply to an additional posting.

Evaluation and hiring: When a teacher applies to a position, the hiring school receives her application materials through the platform. For teachers who previously taught in the district, principals may request or teachers may disclose a district-calculated measure of value-added.[12] The school's principal may then rate the applications and choose to interview applicants on a rolling basis. For known positions at the beginning of the hiring period, there is a short window during which only transfers from within the district are able to apply. Schools can either hire from this pool or wait and consider more applicants.

If the principal wants to hire the candidate, she extends a job offer. The candidate has 24 hours to accept the offer, during which she might contact other schools that have shown interest. If the teacher accepts the offer, she commits to not accepting an alternate offer in the same cycle.

With a few small exceptions, teacher pay is determined by a mechanical formula that depends on degrees, certifications, and experience. These costs are borne by the district, so hiring a more experienced or credentialed teacher does not cost the principal more.

Eligibility: Teachers are eligible for positions if they have the necessary certification. We focus on the market for elementary-school classroom teachers because the common certification allows us to reliably classify which teachers are eligible. We can also infer elementary school teachers' quality from systematic gains in their students' test scores because teachers in these positions are typically responsible for instruction in the tested subjects.

3 The vacancy posting and application process

In this section, we describe our model of teacher and principal actions in the labor market. We specify our model assumptions, consider how violations of the assumptions might manifest in the data, and show empirical evidence consistent with the assumptions. Our empirical analysis includes robustness checks

[12] While raw growth scores have been available in North Carolina since 1997, the district also began using Education Value-Added Assessment System (EVAAS) measures of value added since 2013. All teachers with such measures have an opportunity to reveal these evaluations with their initial application though some may not choose to do so.

around possible alternate assumptions. We defer a discussion of the pair-wise stability assumption until Section 8.

Our models of teacher preferences and principal behavior assume that there is an action that reflects preference orderings, which leads to transparent arguments for how we identify preferences. There are several actions that teachers and principals take in order to form a match. A teacher decides to apply, the principal views the application and assigns a rating, the principal decides to select the teacher for an interview, the teacher accepts the interview, the principal decides to offer the teacher the job, and finally the teacher accepts the job. We use the earliest action we observe on both sides of the market to infer preference orderings: the teacher decision to apply and the principal rating decision. Conceptually, early actions are lower-stakes and so are less likely to be done strategically. Moreover, by using the earliest action we allow later actions to be strategic.

3.1 The teacher perspective

3.1.1 How we model applications

The district's labor market consists of potential teachers, indexed by j, and a set of positions, indexed by p. Each position is associated with a specific school, $k = k(p)$, and may be assigned to at most one teacher. The exception is the outside option ($p = 0$), which includes leaving the district or teaching and has unlimited capacity.

At the beginning of year t, each teacher has an assignment, denoted by c. For teachers new to the district, this assignment is the outside option ($c = 0$), while for incumbent teachers, the assignment is j's position in the prior year, $c = p(j, t-1)$. Teachers may always keep their initial assignment. On an exogenous date $r = r(j, t)$, teacher j enters the transfer system.[13] If she enters, then she is active in the transfer system until an exogenous end date, $r' = r'(j, t)$.

If the teacher enters the transfer system, then she may apply to any position p that is active at some point between r and r'. These positions comprise her choice set, $\mathcal{P}_{jt}$. There is no marginal cost to applying and there is no limit on the number of applications she can submit within the choice set. Let a_{jpt} be an indicator for whether teacher j applied to position p in year t. A teacher's application a_{jpt} is known only to position p and teacher j.

These assumptions lead teachers to treat the application process non-strategically by applying to any position with utility higher than her current position and the outside option. A teacher submits an application to position p if:

$$a_{jpt} = \mathbf{1}\{u_{jpt} > \max\{u_{jct}, u_{j0t}\}\}, \tag{3}$$

where u_{jpt} is teacher j's utility from working at position p in time t.

[13] We assume entry into the system is exogenous. We discuss selection into the system in Appendix C.

3.1.2 Model assumptions

There are three key assumptions that underlie this model of teacher application behavior. First, applications are non-strategic: if a position is more appealing than the outside option and current position, then the teacher applies. Second, the teacher considers all vacancies that overlap with her timing. Third, the set of positions the teacher sees is conditionally exogenous.

Non-strategic applications: Assuming non-strategic applications is reasonable because of three institutional features. First, we focus on applications rather than interviews or the decision to accept the job. The application stage is less susceptible to strategic considerations than later stages because the teacher does not have to commit to an interview or accepting the job.[14] Second, the marginal cost of applying to a vacancy is effectively zero (it just requires clicking submit given already uploaded materials) so it is reasonable that a teacher just compares a given position to the outside option. Third, principals do not see the teacher's other applications, which limits complicated signaling stories.

Several empirical patterns are also consistent with non-strategic behavior. If teachers were instead strategic in submitting applications, then most models would imply a dynamic portfolio strategy where teachers might delay when they apply to a vacancy. We empirically investigate the frequency of delayed applications by constructing a measure of a teacher's wait time to apply to a vacancy. We calculate the time elapsed between the first day a teacher could have applied to a vacancy and the day the teacher actually applied to the vacancy, where we assume that the teacher only learns that a vacancy is available on days she logs into the system and applies.[15] The top panel of Figure 1 shows that the median wait time to apply to vacancies that were already posted on the first day the teacher logged into the system (the "stock" of vacancies) is 0 days. The bottom panel shows that the median wait time to apply to vacancies that were posted after the first day the teacher applies (the "flow" of vacancies) is also 0 days. We thus find minimal waiting to apply to positions, such that teachers are unlikely to be engaging in dynamic portfolio strategies.

We similarly find no evidence of strategic delays resulting in non-applications.[16] Strategic non-applications imply asymmetric behavior according to market conditions. When a school posts two vacancies in a cycle, delaying an application is more useful for an early posting than a late posting. If applicants are trying to delay, then we might see higher application rates for the latter of the two vacancies. Appendix Table A6 shows that both the conditional (applied to the other position) and unconditional application rates are very similar for the earlier and later vacancies. This symmetry thus provides further evidence against the presence of

[14] Even if teachers wanted to avoid the chance of having to take a future costly action (interview or offer), these actions are extremely rare. The mean number of interviews for a teacher is 0.2 (and 0.3 for an internal teacher), and a given position interviews on average only 2 teachers (Appendix Table A4). Thus, because a "successful" application is quite rare, it is hard for strategic considerations to enter.

[15] Let $\mathcal{A}_{jt}$ denote the set of days where teacher j applied to at least one vacancy in year t, with $a_{jt} \in \mathcal{A}_{jt}$ in days. Let b_{kt} be the day that position k's vacancy is posted, and let c_{jkt} be the day that teacher j applies to position k. For every application j sent in year t, we define wait time w_{jkt} as: $w_{jkt} \equiv c_{jkt} - \min_{a_{jt} \in \mathcal{A}_{jt} : a_{jt} \geq b_{kt}} a_{jt}$.

[16] A strategic non-application requires that the vacancy closes while the applicant is waiting. But Appendix Table A5 shows that vacancies clear very slowly, especially early in the cycle.

strategic non-applications.[17]

Teachers consider all available vacancies: It is reasonable to assume teachers consider all available vacancies because teachers appear not to delay applications. If teachers were unaware of some open vacancies, then we would expect teachers to apply frequently after the first opportunity to do so. We see little evidence of such delayed applications. This pattern could reflect teachers missing a vacancy when it is posted and never searching for older vacancies. But we see the opposite – on the first day of applying, teachers apply to old and new vacancies, with a mean vacancy length of 23 days (Appendix Table A7, panel B).[18]

(Conditionally) Exogenous choice sets: Choice sets reflect a teacher's time in the market, which is an equilibrium outcome related to our behaviors of interest. For example, as we will find below, principals are slightly more likely to hire high value-added teachers, which would remove them from the market faster than teachers who do not receive offers. We therefore do not expect that choice sets will be identical on average across teacher types. Rather, we assume that conditional on observable characteristics, variation in teachers' choice sets, which our model links to variation in market entry and exit dates, is unrelated to teachers' idiosyncratic preferences for certain positions or position types.

We assess this assumption by examining teacher entry and exit patterns and vacancy posting patterns. On the teacher side, we find little evidence of strategic timing in entering the market. Comparing teachers with above and below median value-added, we find that they apply for positions at similar times in the cycle (see Appendix Table A9b). While a discernible relationship between value-added and entry timing could still be consistent with choice set exogeneity, the lack of a relationship suggests that entry may be close to random. Further, as we previously described, when entering the market teachers tend to apply to many jobs immediately, which suggests that teachers were not timing their entry for specific jobs. But in case teachers were targeting their entry for when an idiosyncratically desirable set of positions are posted, we conduct a robustness check (see Table 5 (Part 2)) where we estimate teacher preferences leaving out all vacancies that were posted within one week of when the teacher first started applying. Thus, these preference estimates reflect application behavior to positions posted well before or well after the day the applicant first applies.

The case for conditionally exogenous market exit is more complicated because one reason for exit—receiving and accepting a job offer—is possibly related to idiosyncratic preferences. But teachers exit the market for multiple reasons, and indeed we see that many teachers—including those who do not successfully

[17]Even if non-applications were common, they would reduce our information about preferences but would not necessarily affect our results. Consider three types of positions: those the applicant likes enough to apply immediately, those the applicant likes somewhat and may not apply to for strategic reasons, and those the applicant does not like. Only the first group would receive applications, but all positions receiving applications would still be preferred to those not. In Table 5 (Part 2), we report a robustness check where we use the baseline estimates to simulate teacher utilities for each position. Among the positions each teacher actually applied to, we then convert the least preferred 20% of these to non-applications, provided there is at least one application remaining. We re-estimate the teacher preference model with the altered applications and find nearly identical results to the baseline.

[18]Appendix Table A8 shows statistics about the distribution of the time between when a vacancy is first posted and when a teacher applies for all applicants (median of 7 days), hired candidates (median of 5 days), interviewed candidates (median of 5 days) and positively assessed candidates (median of 7 days).

transfer—stop applying long before the end of the hiring season (9% in April or before, 15% in May, 21% in June; see Appendix Table A7, panel C). This pattern suggests that much of exit is driven by shocks unrelated to accepting a job, or to the nature of the jobs being posted. Even for the teachers who leave the market by accepting a job, the job offer often comes well after the teacher applied.[19] This delay leaves a long period when the teacher may keep applying to more positions even while her preferred position is sitting on her application. To avoid further any potential relationship between when applicants leave the market and their idiosyncratic preferences for the positions available at that time, we conduct a robustness check (see Table 5 (Part 2)) where we estimate teacher preferences based only on vacancies that were available the day the teacher first applied for jobs that cycle.

On the school side, vacancy posting is spread throughout the hiring season. We split schools into Title I and non-Title I (Title I schools are high-poverty schools). Given results elsewhere in the paper, Title I schools on average are less sought after schools. Appendix Figure A2 looks at the distribution of first and last posting dates by type. The main feature of the graph is that postings are spread throughout the hiring season. Even within school, there is vast variation in the timing of postings across years: pooling across the years in our data, 85% of schools that post jobs in July also post jobs in April, and a similar pattern holds for schools with April postings (see Appendix Table A9c). The secondary feature of Appendix Figure A2 is that if anything more sought after vacancies (non-Title I) are active later in the hiring season.[20] Appendix Table A11 confirms this broad pattern for a variety of other student demographic characteristics. If we zoom in on multiple vacancies posted within school, then Appendix Table A12 shows that the earlier vacancy is more likely to hire a teacher with non-missing value-added but conditional on hiring a teacher with non-missing value-added, the later vacancy hires slightly better teachers. Combined, these features suggest that there is likely little correlation between teacher characteristics and the set of vacancies that they see.[21]

As a result of the preceding discussion, we construct a teacher's start (r) and end (r') (search) date as the dates of her first and last application, respectively. We thus estimate fairly large choice sets out of which teachers make a large number of choices, which helps us estimate preference heterogeneity. Specifically, the mean choice set size is 159 (median: 139), and the mean number of applications is 23 (median: 8).

3.2 The principal perspective

3.2.1 How we model principal behavior

Each position p is associated with a principal with the same index. Principal p derives value v_{jpt} from teacher j holding the position in year t. We model a principal as giving teacher j a positive rating ($b_{jpt} = 1$) if the value is positive: $v_{jpt} > 0$. A positive rating is at least one positive outcome: recording a positive note about the application, offering an interview, or extending a job offer. While we will often refer to values

[19] In Appendix Table A10, panel C, we show that 10% teachers are still applying to positions 23 days after the hired teacher did.

[20] Appendix Figure A3 shows that vacancy fill rates do not differ very much over the cycle or between Title I and non-Title I schools.

[21] Table 5 (Part 2) shows robustness to a seven-day buffer on both ends or to dropping teachers who only apply to one school. If choice sets are restricted, then fixing the deviations is further from first-best.

as reflecting utilities, principals may rank a teacher higher because of poor information rather than a utility comparison. Either interpretation is consistent with the paper's results and only affects the labeling of the hypothetical policy that would alter principals' choices.

3.2.2 Model assumptions

There are two assumptions underlying our model of principal behavior. First, principals value applicants who receive a positive outcome more than those who do not. Second, principals consider all applicants.

Principals value applicants with a positive outcome more than those without: The note-taking system is supportive of the first assumption. Principals may be strategic in deciding on interviews or offers if such actions are costly and a preferred teacher may have a low probability of accepting. Because the note-taking system allows principals to rate applicants with no direct consequences, principals can reveal their preferences while remaining strategic in consequential actions.[22]

Principals consider all applicants: The second assumption is reasonable because we see no relationship between when an applicant applied and the applicant's outcome. The applications that receive ratings are similar in timing to those that the principals do not rate (see Appendix Table A10).[23]

4 Production of student achievement

In this section, we first lay out the production model, which specifies teacher output at each school. Second, we describe our three-step estimation procedure and discuss parameter estimates. Third, we present a range of validation checks.

4.1 Model

Given our interest in outcomes for disadvantaged students, we allow teacher value-added to differ between advantaged and disadvantaged students.[24] This choice follows the quickly expanding literature documenting match effects or allowing for comparative advantage (Dee, 2004, 2005; Condie, Lefgren, and Sims, 2014; Jackson, 2013; Aucejo et al., 2022; Delgado, 2023; Graham et al., 2023; Biasi, Fu, and Stromme, 2021; Bau, 2022). We show below (in Table 5 (Part 6)), however, that all of our conclusions are unchanged if we estimate the homogeneous model that is standard in the literature.

[22] While our assumptions allow for strategic interviews and offers, we do not find evidence that strategic behavior is common enough to affect our conclusions. Table 5 (Part 3) shows that results are robust to instead modeling principal behavior with a rank-order logit (where hires imply larger utilities than interviews, etc.), including where we restrict to only active choices (i.e., drop applications with no records in the note-taking system).

[23] Table 5 (Part 4) shows that results are robust to varying which applicants we assume principals consider.

[24] In robustness checks in Table 5 (Part 5), we consider two alternative splits of students: race and lagged student achievement. We find that our substantive conclusions are nearly identical.

We use notation that follows Chetty, Friedman, and Rockoff (2014a) and Delgado (2023). Let i index students and t index years, where t refers to the spring of the academic year, e.g., 2016 refers to 2015-2016. Each student i has an exogenous type $m(i,t) \in \{0,1\}$ in year t (whether the student is economically disadvantaged). Student i attends school $k = k(i,t)$ in year t and is assigned to classroom $c = c(i,t)$. Each classroom has a single teacher $j = j(c(i,t))$, though teachers may have multiple classrooms.

Student achievement depends on observed student characteristics, teacher value-added, school effects, time effects, classroom-student-type effects, and an error term. Formally, we model student achievement A^*_{it} as:

$$A^*_{it} = \beta_s X_{it} + \nu_{it}, \tag{4}$$

where X_{it} is a set of observed determinants of student achievement and

$$\nu_{it} = f(Z_{jt};\alpha) + \mu_{jmt} + \mu_k + \mu_t + \theta_{cmt} + \tilde{\varepsilon}_{it}. \tag{5}$$

Here, Z_{jt} is teacher experience (and f maps experience into output) and μ_{jmt} is teacher j's value-added in year t for student type m, excluding the return to experience. As in Chetty, Friedman, and Rockoff (2014a), we allow a teacher's effectiveness to "drift" over time. μ_k captures school factors, such as an enthusiastic principal, while μ_t are time shocks. θ_{cmt} are classroom shocks specific to a student type, and $\tilde{\varepsilon}_{it}$ is idiosyncratic student-level variation. We make three standard assumptions to identify the model (see Appendix D).

Our object of interest is a forecast of teacher j's value-added from a hypothetical assignment to a new classroom (or set of classrooms) in school k. Define p_{kmt} as the proportion of type-m students in school k in year t. Given our model of match effects, a teacher's predicted mean value-added at school k in year t is:

$$VA^p_{jkt} = p_{k0t}\mu_{j0t} + p_{k1t}\mu_{j1t} + f(Z_{jt};\alpha), \tag{6}$$

such that a teacher's total value-added for n_{jkt} students is $VA_{jkt} = n_{jkt} VA^p_{jkt}$. We use data through $t-1$ from the whole state to forecast VA^p_{jkt} for assignments we see in the data and for counterfactual assignments.[25]

4.2 Estimation

We estimate our model in three steps using math scores and data from the whole state.[26] In the first step, we estimate the coefficients on student characteristics by regressing test scores (standardized at the state-level to have mean 0 and standard deviation 1 in each grade-year) on a set of student characteristics and classroom-

[25]Our match effects model is sparse, to reflect the amount of variation we have in the data, and thus unlikely captures all forms of match effects. But because we specify match effects to vary at the same level as the social planner's objective – i.e., based on whether students are economically disadvantaged – any remaining orthogonal match effects do not affect the results.

[26]Focusing on a single subject allows us to rank all possible levels of output. We follow Biasi, Fu, and Stromme (2021) in choosing math because it is typically more responsive to treatment (e.g., Rivkin, Hanushek, and Kain (2005), Kane and Staiger (2008), and Chetty, Friedman, and Rockoff (2014a) for evidence). In Section 8 we show robustness to including a teacher's value-added on behavioral outcomes.

student-type fixed effects. In the second step, we project the residuals (A_{it}) onto teacher fixed effects, school fixed effects, year fixed effects, and the teacher experience return function. In the final step, we form our estimate of teacher j's value-added in year t for type m (μ_{jmt}) as the best linear predictor based on the prior data in our sample (this prediction includes the experience function). Since in this final step we shrink the estimates, we understate the dispersion in match effects relative to the true dispersion. Using shrunken estimates and prior data implies that we use the information available to policy-makers. See Appendix D.2 for estimation details and a discussion of what variation pins down parameters.

Alternative value-added models: We consider four alternative value-added models. The first is a homogeneous effects model, where we assume that teachers' effects on students are type-invariant, rather than allowing for comparative advantage. This model is restrictive relative to our baseline model, but increases our forecast precision. This model tests whether our results rely on comparative advantage or reduced forecast precision. The second model estimates the school effects differently: rather than including school fixed effects (as in, e.g., Jackson (2018) or Mansfield (2015)), we include school-level means of all of the covariates (as in, e.g., Chetty, Friedman, and Rockoff (2014b)). This model tests whether our results depend on how we decompose effects into school and teacher components. Third, we use the Chetty, Friedman, and Rockoff (2014b) estimator. Unlike our "homogeneous" value-added model, this model (a) forecasts using future test scores in addition to past test scores, (b) includes classroom controls like peer mean characteristics rather than school fixed effects, and (c) residualizes test scores using a teacher fixed effect rather than a teacher-year fixed effect. This model tests whether our results are robust to a more "standard" estimator. Finally, we also consider a simple residual estimator where just residualize contemporaneous test score gains for student characteristics. This estimator has the benefit that we can compute value-added for all teachers in the district in a given year and so allows us to directly address concerns about differential missingness of value-added between students of different types. See Appendix D.3 for details.[27]

4.3 Validation of the match effects model

To validate our value-added model, we use a version of Chetty, Friedman, and Rockoff (2014a)'s test for mean forecast unbiasedness. We predict a teacher j's value-added in school k in year t (μ_{jkt}) using data from all years prior to t. We then regress the realized mean student residuals in year t ($\bar{A}_{jt}$) on the prediction and test whether the coefficient on our prediction equals 1. Column (1) of Appendix Table A13 shows that the math value-added estimate is an unbiased predictor of residualized output, with a tight confidence interval around 1.06. Appendix Figure A4 shows that forecast unbiasedness holds throughout the distribution of teacher value-added. As our exercise will involve assigning teachers to new schools, forecast unbiasedness across "nearby" assignments may be weak validation for predicting output in "far away" assignments; for example, a teacher's ability with disadvantaged students estimated in a school with a small number of disadvantaged students might be a poor guide to their ability with disadvantaged students in a school with a

[27] Table 5 (Part 6) shows that our central conclusions do not depend on which value-added model we use.

large number of disadvantaged students. Therefore, we conduct additional tests that validate our estimates over moves similar to those in our counterfactuals. Column (4) of Appendix Table A13 shows mean forecast unbiasedness nearly holds for transferring teachers (with a coefficient of 0.98, not statistically different from 1) while the last two columns show mean forecast unbiasedness even for cases where teachers switch between classrooms with very different compositions or sizes.

We conduct a similar test for the comparative advantage component of value-added. In column (2) we compare our forecast of the difference in a teacher's value-added across (economically) disadvantaged and advantaged students with the realized test score difference. Again, we find that our estimates are nearly forecast unbiased. Appendix Figure A5 shows that forecast unbiasedness holds throughout the distribution. Appendix D.4 further assesses the validity of the comparative advantage component of value-added, providing inference around relevant structural parameters (the estimated correlation between teacher value-added with advantaged and disadvantaged students is 0.86), likelihood tests, and additional validation around transferring teachers.

5 Teacher preferences

5.1 Parameterization

We adopt a characteristics-based representation of teacher utilities over positions, which helps us to estimate preference heterogeneity. Teacher utilities over positions are:

$$u_{jpt} = -\gamma d_{jpt} + \pi_j \hat{VA}_{jpt} + \beta_j X_{pt} + \eta_{jt} + \varepsilon_{jpt}. \tag{7}$$

Teacher utility for the outside option is $u_{j0t} = \varepsilon_{j0t}$. d_{jpt} is the one-way commute time (in minutes) between the teacher and the position and will serve as a numeraire for exposition. VA_{jpt} is teacher j's total value added at position p in year t.

Predicted value-added, $\hat{VA}_{jpt}$, combines absolute and comparative advantage. We define a teacher's absolute advantage to be her predicted value-added at a representative school: $AA_{jt} = n_{0t}\hat{\mu}_{j0t} + n_{1t}\hat{\mu}_{j1t}$, where n_{mt} is the average number of type m students in a classroom in the district. Comparative advantage, CA_{jpt}, at a specific position is then the difference between predicted value-added at school $k(p)$ and absolute advantage: $CA_{jpt} = VA_{jpt} - AA_{jt}$. Because we control for absolute advantage in the person-time effects, the coefficient on $\hat{VA}_{jpt}$, π_j, captures the strength of teachers' preferences for schools where their comparative advantage is high. We allow for preference heterogeneity by including a random coefficient in π_j:

$$\pi_j = \bar{\pi} + \sigma^{VA} \nu_j^{VA}, \tag{8}$$

where $\nu_j^{VA} \sim^{iid} N(0,1)$.

X_{pt} is a vector of observed characteristics of positions: the fraction of a school's students that are (1) above the median in prior year math test scores (s), (2) Black (b), and (3) Hispanic (h), and (4) the average number of students in a class at the school that are economically disadvantaged (e). We allow for heterogeneous preferences:

$$\begin{aligned}\beta_j^e &= \beta_0^e + \beta_1^e AA_{jt} + \sigma^e \nu_{jt}^e \\ \beta_j^b &= \beta_0^b + \beta_1^b AA_{jt} + \beta_{j2}^b Black_j + \sigma^b \nu_{jt}^b\end{aligned} \tag{9}$$

where $Black_j$ is an indicator for teacher race category and ν is a vector of independent, standard normal random coefficients, which captures the standard deviation of idiosyncratic preferences. The equations for lagged achievement and Hispanic are parallel.[28]

We follow Mundlak (1978) and Chamberlain (1982) and model η_{jt} using correlated random effects. We model teacher-year unobserved heterogeneity in preferences for teaching in the district as the sum of several components:

$$\eta_{jt} = \lambda Z_{jt} + \rho CM_{jt} + \sigma^\eta \nu_{jt}^\eta. \tag{10}$$

Z_{jt} are teacher-year characteristics – whether the teacher is in the district, whether the teacher is Black, whether the teacher is Hispanic, whether the teacher is female, the teacher's predicted value-added for economically disadvantaged students, the teacher's predicted value-added for non-economically disadvantaged students, and dummy variables for whether the teacher has 2-3 years of prior experience, 4-6 years of prior experience, or more than 6 years of prior experience. CM_{jt} is a set of teacher-year averages of the variables that vary across the job postings within teacher-year (value-added, commute time, interactions of teacher and school characteristics). Through CM_{jt}, we allow unobserved heterogeneity to be correlated with CA_{jpt} and X_{pt}. Finally, ν_{jt}^η is an independent standard normal random effect.[29]

ε_{jpt} is an iid Type I extreme value error. Let $V_{jpt} = u_{jpt} - \varepsilon_{jpt}$ be j's representative value for position p in year t. Then the distributional assumption on ε_{jpt} implies that:

$$Pr(a_{jpt} = 1) = \frac{exp(V_{jpt})}{1 + exp(V_{jct}) + exp(V_{jpt})} \text{ and } Pr(a_{jpt} = 1) = \frac{exp(V_{jpt})}{1 + exp(V_{jpt})}, \tag{11}$$

for teachers already in the district and teachers new to the district, respectively.

5.2 Estimation and Identification

The data we use to estimate teacher preferences are applications to positions, and the method we use is maximum simulated likelihood, where we simulate from the normal distributions of the random coefficients. Let n index each simulation iteration and let $A_{jptn}(\theta)$ be the model-predicted probability that j applies to

[28] Table 5 (Part 7) shows that our results are robust to allowing for correlation in the random coefficients.

[29] In Table 5 (Part 8), we consider binary logits, and show that our results are robust to either omitting random effects, or to including various combinations of teacher and school random and fixed effects.

position p in year t in simulation iteration n at parameter vector θ. For each teacher j in year t, we construct the simulated likelihood as:

$$L_{jt} = \frac{1}{500}\sum_{n=1}^{500}\prod_{p\in\mathcal{P}_{jt}}\left(a_{jpt}A_{jptn}(\theta)+(1-a_{jpt})(1-A_{jptn}(\theta))\right), \tag{12}$$

where a_{jpt} is an indicator for whether j applied to p in the data. Our full simulated log likelihood function is:

$$l = \frac{1}{J}\sum_{j} logL_{jt}. \tag{13}$$

In Section 3 we argued that the institutions and data are consistent with teachers applying non-strategically. Under this assumption, the choices that teachers make identify preferences and preference heterogeneity. Heuristically, if within her choice set a teacher is more likely to apply to positions with a particular characteristic than a position without this characteristic, then we infer that the teacher has a preference for schools with this characteristic. Similar reasoning applies for mean coefficients, and observed and unobserved preference heterogeneity.

We seek to predict teachers' valuations over positions rather than causal effects of changes in characteristics on choices. In counterfactuals, we give utility bonuses as a function of school characteristics and so do not assume that teachers value money or these characteristics. As a convenient way to interpret magnitudes, we sometimes convert utility to minutes of commute time, which requires the stronger assumption that commute time is exogenous. We do not rely on having consistently estimated the causal effect of commute time, however, because we only make relative comparisons of the costs of various policies.

5.3 Teacher Preference Estimates

Table 1 presents the teacher preference estimates. First, teachers prefer positions with more advantaged students. Second, teachers dislike positions with longer commutes. Finally, teachers have only slight preference toward positions where they have higher value-added.[30]

Responsiveness to school and match characteristics varies with observable and unobservable heterogeneity. For example, teachers with higher absolute advantage have relatively lower preferences for schools with more disdvantaged students. We also find a large positive same-race premium for Black teachers and schools with large fractions of Black students. In terms of unobserved heterogeneity, we typically find substantial dispersion in the random coefficients. For example, a standard deviation of the random coefficients on the number of disadvantaged is about the same as the mean valuation.

To help interpret the strength of—and heterogeneity in—some of these relationships, Panels (a) through (c) of Figure 2 show how the average rank of positions in teachers' preferences change as single character-

[30]We use the value-added forecast, $\hat{VA}_{jt}$, in our preference model. In Table 5 (Part 9), we show robustness to excluding value-added derived variables in our preference model.

istics change, as well as the 10th and 90th percentile of these positions in teachers' rankings. We do not hold other characteristics fixed so that, for example, when we study commute time, other characteristics of schools are potentially changing. The figure emphasizes that commute time is a powerful predictor of rankings: changing commute time from 5 minutes to 25 minutes decreases the average rank of a position (for the average teacher) from about the 70th percentile to the 50th percentile. Similarly, the fraction of students that are disadvantaged is a powerful predictor of ranking: across the support, the mean ranking moves by over 20 percentiles. In contrast, while teachers do pursue comparative advantage (see the coefficient in Table 1), this relationship is quite weak: across the support of the data, varying teachers' comparative advantage only increases the rank of a position by a couple of percentiles. The figures also emphasize that there is substantial heterogeneity in teachers' rankings of positions: across the support of these characteristics, the range from the 10th percentile in the teacher distribution to the 90th is very large.

Hence, not only do teacher preferences deviate from those that would decentralize the planner's solution, they are negatively correlated. With minimal assumptions and data on real choices, we confirm the findings of the teacher preference literature regarding mean preferences but estimate considerable heterogeneity.

6 Principal behavior

6.1 Parameterization and identification

We adopt a characteristics-based model and parameterize v_{jpt} to be a linear function of position and teacher characteristics, a random effect, and an idiosyncratic teacher-position error:

$$v_{jpt} = \alpha_p W_{jpt} + \sigma_\kappa \kappa_{pt} + \upsilon_{jpt}. \tag{14}$$

To allow principal behavior to possibly align with output, W_{jpt} includes j's total predicted value-added at school $k(p)$.[31] We further include teacher characteristics: teacher prior experience (in bins of 2-3 years, 4-6 years, and 7+ years), whether the teacher has a Masters degree, whether the teacher is licensed, whether the teacher is certified, the teacher's Praxis score, whether the teacher is Black, whether the teacher is Hispanic, and whether the teacher is female.[32] Finally, we include a constant and interact whether the teacher is Black with the fraction of the school's students that are Black and whether the teacher is Hispanic with the fraction of the school's students that are Hispanic. We exclude salary because principals in our empirical context do not have to pay teacher salaries out of a school budget. We allow principals to have heterogeneous valuations over teachers based on W_{jpt} by letting α_p vary with whether the school has Title I status.

To capture heterogeneous outside options and variation in propensity to assign ratings, κ_{pt} is a normally distributed random effect. Finally, υ_{jpt} is i.i.d. Type I extreme value.

[31] We include predicted value-added, rather than realized value-added, in W_{jpt} so that principals only incorporate the information available at the time the application was received.

[32] We also include indicators for whether each covariate is missing. The Praxis test is a standardized teacher certification test administered by the Educational Testing Service.

As with teachers, identification is straightforward given our characterization of the process in Section 3. We observe the set of applications that a principal receives and we observe whether a principal gives an application a positive outcome. We interpret the decision to give an application a positive rating as a non-strategic and costless action. This interpretation allows us to infer principal valuations from their choices in a straightforward way: those that are rated positively are preferred to those that are not. Because we observe the ratings, even if *interviewing is costly* and so principals are strategic at this stage, then our identification assumption still holds. One might also worry that assigning a rating is costly, and so it is done strategically. To alleviate this concern, we show below that if we restrict attention to applications where a principal assigned a rating (either positive or negative), then our results are quantitatively identical (see Table 5 (Part 3)).

6.2 Estimates

Before presenting estimates from our baseline model, we consider what types of characteristics determine principal ratings. Appendix Table A14 presents the changes in pseudo-R^2s from including different sets of observable teacher characteristics. The main set of characteristics that explain ratings decisions are various observable characteristics of teachers: experience, licensing, certification, and Praxis scores. While one might think that these characteristics would predict value-added, in Appendix Table A15 we show that they have very limited predictive power. Indeed, value-added by itself or in addition to other characteristics generates very small changes in model fit.[33]

Despite the small explanatory power of value-added in principal decisions, Table 2 shows that principals do favor teachers with higher value-added in our baseline model.[34] We also observe significant heterogeneity, as Title I school principals rate Black and Hispanic teachers more positively than non-Title I principals do. To help interpret the strength of the value-added relationship, Panel (d) of Figure 2 shows that the mean percentile of teachers in principals' ratings goes from the 35th percentile to the 60th percentile across the support of projected value-added. Consistent with the idea that observed characteristics poorly predict value-added, Appendix Figure A6 shows that if we omit value-added from the principal model then the relationship dramatically flattens. Nevertheless, the strength of this relationship is difficult to directly interpret. To assess the extent to which value-added explains principal decisions, in the model section below, we compare the allocations achieved using the estimated principal behavior to those with random principal behavior.

The relationship between value-added and principal rankings could reflect preferences or information. Distinguishing between these does not affect analysis of the current allocation (or the counterfactuals) because we will compare how principals currently act with a proposed alternative ranking.[35] But if incomplete

[33]EVAAS, the state of North Carolina's value-added measure, has even less explanatory power. As principals have access to this information, it is unlikely that the estimated weights principals place on value-added are due to measurement error in our estimates of value-added. Our results are quantitatively robust to significant amounts of attenuation. See Table 5 (Part 10).

[34]See Appendix E for the likelihood, which closely parallels the one for teachers.

[35]Appendix Table A14 shows that the model's explanatory power actually decreases when using the readily-available EVAAS

information explains principal rankings, our empirical strategy might use the data differently. For example, we use principals' notes for identification because we can then allow for strategic interviewing or offering. But if interview or offer decisions deviate from notes because information resolves (rather than strategy), then we would want to use interviews and offers and not the notes. We show, however, in Table 5 (Part 3) that principal models estimated using only offers delivers nearly identical results.

Hence, consistent with the previous literature, principal valuations deviate from those that would implement the planner's solution, as principals rank teachers based on predictable and unpredictable factors not related to value-added. Whether the positive relationship between rankings and value-added is strong enough to generate allocations close to the planner's solution depends on how both sides combine in equilibrium.

7 The current allocation

With our model estimates, specifically of teacher value-added, we now discuss the current allocation of teachers across schools.

Student and teacher characteristics: Table 3 presents properties of the current allocation where we report student-weighted results when we split students by our measure of economic disadvantage. We report results in our focal district, as well as in all other districts in North Carolina. Disadvantaged students are more likely to be minorities. Disadvantaged students also have teachers with worse observed characteristics. Specifically, they are less experienced, less likely to have a graduate degree, a regular license, be certified, and have lower Praxis test scores (a standardized test).

Test scores and teacher value-added: Between advantaged and disadvantaged students, there are large achievement gaps in levels. But in gains, we see no gaps. This "raw" data fact hints that there are not large differences in learning across schools, which suggests that the average quality of teachers is likely similar.

Looking across a variety of measures of teacher value-added, the broad pattern is that disadvantaged students have teachers of similar strength to advantaged students. This pattern is true both in our focal district, as well as in the rest of North Carolina. As we mentioned in the introduction, this finding is not new to us and has been found in many districts across the United States (see footnote 4). Specifically, with our baseline value-added model, we find equivalent value-added with advantaged and disadvantaged students among teachers of advantaged and disadvantaged students. Our alternative value-added models find similar patterns: with homogeneous value-added, advantaged students have a slight advantage and this grows slightly with the Chetty, Friedman, and Rockoff (2014b) estimator and an estimator that uses school mean characteristics rather than school fixed effects. The estimators that find a slight advantage for advantaged students in our focal district tend to find smaller differences in the rest of North Carolina. The table also

measure.

reports measures of behavioral value-added (see Appendix D.5 for details on how we construct behavioral value-added) and shows that they are approximately balanced across advantaged and disadvantaged students.

Other student classifications: In Appendix Tables we present similar sets of summary statistics for a wide variety of alternative "splits" of students: splitting students by race (Appendix Table A16), by lagged achievement (Appendix Table A17), by a measure of persistent disadvantage (Appendix Table A18, and see Appendix Table A19 for the relationship between disadvantage and persistent disadvantage in our sample), and splitting by school characteristics (high-poverty vs. not) rather than by student characteristics (Appendix Table A20). The basic patterns persist across all these variants. We emphasized in Section 2.2 that there is minimal within-school sorting of students across classrooms. Validating the lack of within-school sorting, Appendix Table A21 shows similar patterns when we measure the advantaged-disadvantaged gap using school averages of teacher characteristics.[36]

Missing value-added: One critique of this finding is that it refers to teachers for whom we can estimate value-added, and disadvantaged students are especially likely to have inexperienced teachers for whom we cannot estimate value-added. Table 3 shows that disadvantaged students are more than twice as likely to have a teacher for whom we cannot estimate value-added.

To address the concern about differential missingness, we report results of the residual value-added estimator (the teacher's mean of A_{it}, in the notation of Appendix D), which only uses data from the current year and so can be estimated for all teachers. This value-added estimator finds similar patterns. As an alternate measure, we impute value-added for the teachers for whom we cannot estimate value-added. At a high level, we use the set of observed characteristics in the top portion of the table (Appendix F details the exact imputation process). Naturally, since one of the themes of this paper is that observed characteristics poorly predict value-added, there is a limit to how good the imputation model can be, though this may simply reflect that principals also have limited information. The main finding is that including inexperienced teachers does not alter the central message of the table that disadvantaged and advantaged students have teachers of similar strength.

8 Understanding the current allocation

To understand how the current allocation is generated, we simulate the market equilibrium by combining the estimated market timing from Section 3, the estimated match-specific output from Section 4, the estimated teacher preferences from Section 5, and the estimated principal valuations from Section 6.

[36]Appendix Figure A7 shows the result visually. If we classify schools by their mean teacher value-added, we find that the share of disadvantaged students is weakly increasing in the school's mean teacher value-added. This pattern holds for three different value-added measures.

8.1 Simulation details

We consider allocating the set of *teachers who apply for positions* in the district in the 2015-2016 cycle, including teachers who are not currently in the district. We restrict attention to the teachers for whom we can compute value-added, which includes teachers who have previously taught anywhere in the state. This restriction drops a large number of teachers: we end up with 178 elementary school teachers and 296 positions. To avoid the possibility of artificial imbalance playing a role in our estimates (see Ashlagi, Kanoria, and Leshno (2017)), in each of 400 simulation runs we randomly drop positions so that there are the same number of teachers and positions. In Section 8.7.1, we study an extension where teachers outnumber positions.

While we estimate a distribution of random coefficients, in simulations we use the single draw of the random coefficients per teacher and principal that maximizes the likelihood for the teacher or principal. We draw i.i.d. type I errors for ε_{jpt} and υ_{jpt}.

In using DA to find stable allocations, we have teachers and principals submit rankings according to their true preferences. If there are multiple equilibria, then for one side of the market it is not a dominant strategy to report truthfully. Below we show, however, that the equilibrium is essentially always unique and so truthful reporting is a dominant strategy.

For teachers and vacancies that are not in each other's choice sets, we assign a large negative number to the valuations. We do not include an outside option when we run DA. Given that we impose balanced markets, all teachers are hired and all positions are filled (in Section 8.7.1 some teachers are not hired).

8.2 Model fit

We now turn to the fit of the model under the status quo. Because we estimate several model components fairly directly from data, fit largely highlights how well our market equilibrium assumption (pairwise stability) performs. Figure 3 shows that the model matches the basic qualitative patterns in the data: schools with a larger share of disadvantaged students have teachers (a) with stronger absolute advantage, (b) with comparative advantage in teaching economically disadvantaged students, (c) less likely to be experienced, and (d) more likely to be Black. Quantitatively, the model almost exactly matches the slope for teacher experience and whether teachers are Black. The model underpredicts the slope in absolute advantage.[37]

Figure 4 (and Table 4) shows that in the estimated status quo, disadvantaged students are assigned slightly *better* teachers than advantaged students. This feature matches the data.

8.3 The importance of second-best reasoning

In the last section, we documented that advantaged students have no more effective teachers than disadvantaged students. Relative to the structure of teacher preferences, this balance is surprising in that teachers'

[37]We also find our model fits better than models with alternate equilibrium assumptions: a teacher serial dictatorship ordered by absolute advantage or experience or a principal serial dictatorship ordered by fraction of students that are economically disadvantaged. Results are available upon request.

revealed preference is strongly averse to teaching at schools with disadvantaged students. In this section, we explain this result through the economics of two-sided markets and the theory of the second best.

A couple of subtle explanations play no role in explaining the current allocation. First, there is no room for equilibrium selection. Changing the equilibrium from the teacher-proposing equilibrium to the school-proposing equilibrium has no effect on the allocation. Second, timing has little role. Changing timing so that all vacancies and teachers are active at the same time increases output slightly for advantaged students and barely decreases it for disadvantaged students. We show these and other allocations in Figure 4 and Table 4.

Aligning teacher and principal preferences with the planner's solution shows that there are important interactions between both sides of the market, such that thinking about one side at a time leads to ineffective or harmful policy ideas. First, if teachers had preferences that would decentralize the planner's solution—they only care about the number of disadvantaged students in a school—then the allocation is little changed. Thus, a natural teacher-side policy is ineffective. Second, if principals had preferences that would decentralize the planner's solution—they only care about the output in the match—then the allocation is worse for equity and resembles what we might expect based on the structure of teacher preferences.[38] Thus, a natural policy based on one-sided reasoning is harmful.

One-sided reasoning is misleading here because of the theory of the second best: preferences on both sides of the market deviate from the preferences that decentralize the planner's solution, but these deviations interact to generate surprisingly favorable allocations. Were we to eliminate the deviation on the principal side of the market and have principals order teachers by value-added, then the strongest teachers would reach their most preferred schools. Given the structure of teacher preferences, this change would lead advantaged students to have much more effective teachers. Hence, by placing weight on factors other than value-added, principals "push back" on teacher preferences and overcome differences in applicant pools across positions.

Reaching these conclusions required an equilibrium model and data to identify preferences from actions rather than equilibrium assignments. With data only on equilibrium assignments, typically one assumes that one side of the market has vertical preferences, which fills in the choice sets for the other side of the market (see Diamond and Agarwal, 2017). If we had (incorrectly) assumed principals have vertical preferences in value-added, then we would have concluded that the status quo was very unfavorable to disadvantaged students, and teacher bonus policies by themselves were effective.

Figure 4 (and Table 4) shows that there are substantial gains in the first-best. Disadvantaged students gain about 0.06σ, or about one-fourteenth of the unconditional achievement gap that we document in Table 3. While these numbers refer only to teachers in the transfer system, in Appendix Table A22 we show that these gains are similar if we look at all teachers in the district. Naturally, these gains are not costless—they come somewhat at the expense of advantaged students, whose teacher quality suffers, but total output still increases (by about 0.021σ in the transfer sample, which is similar to the 0.016σ in the whole sample).

Finally, Figure 4 (and Table 4) shows that the combination of the two policies mentioned above—

[38]The allocation is also worse for efficiency: per student output declines by about 0.009σ.

teachers rank schools based on the number of disadvantaged students and principals rank teachers based on projected output—comes close to decentralizing the first best (it achieves 94% of the first-best, the remaining gap is due to comparative advantage and timing). Thus, in Section 9 we study policies that move us closer to this point.

8.4 Parameterizing teacher preferences and principal behavior using the model

In the previous subsection we emphasized stylized features of teacher preferences and principal behavior to explain our results. First, teachers prefer not to teach at schools with disadvantaged students. Second, principals do not place very much weight on value-added.

To more directly parameterize the inequity for disadvantaged students implied by the structure of teacher preferences, in Appendix Table A23 we display the results of two exercises. First, we ignore market clearing and assign each teacher her preferred position so some positions have multiple teachers and some have none. Conditional on receiving a teacher, disadvantaged students do as well as, or better than, advantaged students. The assignment rate, however, is dramatically different across advantaged and disadvantaged students. Thus, if we follow teachers' preferences and ignore market clearing, then few disadvantaged students would receive teachers. Second, we impose a market clearing mechanism that lets teachers' preferences matter the most. Specifically, we clear the market using a serial dictatorship ordered by teacher's value-added with disadvantaged students. Here, disadvantaged students do dramatically worse.

To more directly parameterize principal behavior, we simulate equilibrium allocations where we give principals random preferences over teachers. Table 4 shows that this allocation is very similar for disadvantaged (and advantaged students) to the status quo. Thus, the loose heuristic that principals hire essentially randomly is a decent approximation to the data. Relatedly, Appendix Table A24 shows that principals make mistakes in the sense that they have much better teachers in their choice set than the ones they either rate positively, interview, or hire.

8.5 Different objectives

Our social planner maximizes disadvantaged students' output. Here, we consider how our results might change with alternate objectives.

First, the social planner may place weight on other forms of output, not just math test scores. We estimate teachers' value-added on an index of behavioral outcomes and find that behavioral value-added is still balanced across advantaged and disadvantaged students (Table 3, see Appendix D.5 for details on how we construct behavioral value-added).

Second, the social planner may place weight on other agents, not just disadvantaged students. First, the social planner may place equal weight on all students. We formalize this objective in Appendix B. Row 11 of Table 5 shows that aligning principals' preferences with the social planner's objective function still lowers total academic achievement. Aligning teachers' preferences with the social planner's, though, can lead to some total output gains.[39] Second, the social planner may place weight on teacher utility. In Section 9, we constrain the policies we consider to make each teacher weakly better off than in the status quo.

8.6 Robustness

In Table 5 (panels 1 through 11), we report the robustness checks we have mentioned throughout the text. The following three basic findings are robust across all of these alternatives: first, there are large gains from moving to the first-best; second, fixing one of the deviations from what decentralizes first best (making teachers value the number of disadvantaged students or principals maximize value-added) is either ineffective or harmful; and third, that fixing both comes close to implementing the first-best (the exception is when we restrict teacher choice sets to the first day because the timing constraints bite more).

8.7 Extensions: the extensive margin and timing

8.7.1 Extensive margin

In our baseline, we only include the teachers for whom we can estimate value-added, which means that there are fewer teachers than vacancies. In reality, there are more applicants than positions. To see how having additional teachers affects allocations, we impute value-added for the teachers missing value-added. Our imputation model was discussed above. We refer to the teachers with the imputed value-added as imputed teachers and the remaining teachers as the non-imputed teachers.

To include the imputed teachers, we clear the market in two stages. First, we find the stable allocation using all positions and the non-imputed teachers. Second, among the remaining open positions and the imputed teachers, we find the stable allocation. This two-stage process aproximates the institutional reality that the vast majority of the non-imputed teachers are already teaching in the district. Thus, because their outside option includes retaining their current position, they will still be teaching in the district even if they are not hired in the transfer system.

We consider two variants of this market-clearing protocol. In the first variant, we randomly drop imputed teachers so that the overall market is balanced (though in the first stage of market clearing the market is not balanced). In the second variant, we keep all imputed teachers so that positions are short in the market and there is an active extensive margin.

Panel 12 of Table 5 shows that our main results are quantitatively robust. For both the balanced market protocol and the unbalanced market protocol (where teachers are long), we find that one-sided interventions are either ineffective or harmful. Similarly, relaxing timing restrictions has very small effects on the disadvantaged students. The exception is that combining teachers maximizing the number of disadvantaged students and principals maximizing value-added does not get as close to the first best as in our baseline results.

[39] Appendix B discusses the gains to match-specific prices with comparative advantage and an efficiency objective.

A natural intuition is that improving information or incentives for principals to maximize value-added will affect the quality of teachers hired into the district. From this perspective, the very weak effect of having principals maximize value-added on the quality of teachers for disadvantaged students should be surprising. Appendix Table A25 shows that there is an operative extensive margin effect, but it serves to only benefit advantaged students. Thus, second best reasoning is operative even with an extensive margin.

8.7.2 Timing

In our baseline results, we show that relaxing timing restrictions does not generate large changes in the allocation. Within our baseline results, however, there are interesting patterns across the hiring season. We divide vacancies by when they were posted into three subperiods: April, May/June, and July/August. The first three rows of Panel 13 of Table 5 show that vacancies that post early hire better teachers for disadvantaged students than vacancies that post late. Indeed, the gap in value-added for disadvantaged students is as large as the distance from the status quo to the first best. Appendix Table A26 shows that there are similar magnitudes for advantaged students and that these declines are quantitatively similar for Title I and non-Title I schools.

Conceptually, this decline might raise questions about the exogenous entry assumption in that one possible explanation is that the positions that are posted late are different than those posted early. The fourth row of Appendix Table A26 Panel A, however, shows that with no timing restrictions there is little temporal pattern in the quality of hires across the hiring season. Relatedly, Appendix Table A27 shows that when we relax timing restrictions there is little temporal pattern in whether teachers end up in more preferred positions. Thus, exogenous entry is approximately satisfied in that there is little temporal variation in the desirability of positions or teachers that enter.

To show that our way of modeling timing restrictions captures the data well, we consider an alternate market-clearing protocol. We split the market into the three sub-periods: April, May/June, and July/August. We date vacancy and teacher entry by when the vacancy is posted or when the teacher first applied. We then clear the market subperiod by subperiod, allowing for imbalance within each period but imposing overall balance. The fourth row of Panel 13 of Table 5 shows that clearing the market in one or three periods results in very similar allocations, with the value added for disadvantaged students within rounding error in the two ways of clearing the market.

9 Teacher bonus counterfactuals

In this section, we consider policies that may move the allocation closer to the first-best. We compare teacher bonus policies that cost the district equivalent amounts while holding all teachers harmless. We then interact these bonuses with principal-side policies.

9.1 Implementation details

The district offers a two-part bonus on the basis of a teacher-position characteristic, z_{jpt}, where each teacher receives a lump-sum amount, b_0, and a bonus b_1 per unit of characteristic, z_{jpt}. Teacher j's utility for teaching at position p in year t is

$$\tilde{u}_{jpt} = u_{jpt} + \gamma(b_0 + b_1 z_{jpt}), \tag{15}$$

where we multiply by the commute time coefficient (γ) to express bonus spending in minutes of commute time. For each b_1, we solve for the teacher-optimal stable equilibrium assignments, where $p^*(j)$ is j's assigned position, given the bonus size and the object that generates the bonus. Thus, because we give teachers utility directly for the characteristic, we do not use our estimated coefficients on the characteristics.[40]

To focus on policies that are likely to receive teachers' support, we make each teacher weakly better off than in the status quo equilibrium.[41] We set the transfer such that the teacher with the worst change is indifferent. This lump-sum transfer can be either positive or negative. Thus, the district's total bonus to a teacher depends both on the choice of how much to compensate for the characteristic and how it changes the allocation.[42]

We examine bonus schemes over two objects. First, we study bonuses based on the number of disadvantaged students the teacher has ($n_{k(p)1t}$). These bonuses mimic the hard-to-staff school bonuses that some districts have piloted. Second, we interact school and teacher characteristics by considering bonuses based on a teacher's absolute advantage times the number of disadvantaged students: $((p_{0t}\hat{\mu}_{j0t} + (1-p_{0t})\hat{\mu}_{j1t})n_{k(p)1t})$.[43]

9.2 Results

Panel (a) of Figure 5 shows the effect of these two bonus schemes on disadvantaged students' test scores when principals hire according to their estimated preferences. The top line shows achievement in the the first-best allocation. To allow for comparisons across bonus schemes, the horizontal axis is the total realized spending per teacher (normalized to be in minutes of commute time per teacher).

We have three results, all of which reflect the theory of the second best. First, untargeted bonuses for teaching disadvantaged students are relatively ineffective in raising disadvantaged students' test scores. Second, targeted bonuses that pay the best teachers more for teaching disadvantaged students are more effective than untargeted bonuses because they jointly address deviations on both sides of the market. Specifically, the applicant pool only expands among the best teachers, so then the principals' difficulties in identifying good

[40] We compare the effectiveness of bonuses with equivalent utility costs. Because we use the same conversion factor for all schemes, the conversion factor does not affect the comparisons.

[41] This assumption also allows us to consider a fixed set of teachers rather than model attrition.

[42] Formally, let $\Delta\tilde{u}^{b_1}_{jpt} = (u_{jp^*(j)t} - u_{jpt}) + \gamma b_1 z_{jp^*(j)t}$ be the change in teacher j's utility (excluding the transfer) between the zero-bonus and the b_1 bonus equilibria. The transfer is: $b_0 = -\min_j \Delta\tilde{u}^{b_1}_{jpt}$. The total bonus to teacher j is $b_0 + b_1 z_{jp^*(j)t}$.

[43] We implement the bonus schemes starting from the status quo with relaxed timing constraints. When we implemented bonuses with timing constraints, with large bonuses we found that our algorithm sometimes produced matches that should have been ruled out by timing constraints, which made the results hard to interpret.

teachers matters less (a random teacher in the new applicant pool is better). Third, the bonuses eventually become less effective as they grow larger. Here, larger bonuses expand the applicant pool for disadvantaged schools, but the larger pool causes the deviation in principal preferences to matter more.

Effective policy needs to address the deviations jointly. In Panel (b) of Figure 5 we consider the effect of the teacher bonus schemes when principals hire according to value-added. Such hiring rules may be induced by a combination of an information intervention and principal bonuses for hiring effective teachers.

We again have three results. First, as in the prior section, if teacher bonuses are small such that estimated teacher preferences largely guide applications, then principals hiring according to value-added leads to large decreases in disadvantaged students' test scores. Fixing the deviation on the principal side, but hardly closing the teacher deviation, has a large negative effect relative to the status quo. Second, as teacher bonuses get larger, principals hiring according to value-added make the teacher bonuses particularly effective. At the equivalent of about 50 minutes of commute time per teacher, the bonuses have nearly reached the first-best. That teacher bonus effectiveness is increasing in principal bonuses (or information interventions) reflects the interaction of the two sides of the market. Third, for some levels of spending untargeted bonuses now outperform targeted bonuses. Because the principal deviation has been closed, the targeting of bonuses is no longer needed. In fact, such targeting is now counter-productive.

10 Discussion

We have studied the equity consequences of the within-district allocation of teachers to schools. We consider both the current allocation and alternative policies. To approximately decentralize the first-best that maximizes disadvantaged students' achievement, teachers would need to prefer schools with more disadvantaged students and principals would need to prefer higher value-added teachers. Using rich data from the teacher transfer system that allows us to observe actions, we show that both sides' preferences deviate from these. Nonetheless, and consistent with the theory of the second best, these two deviations interact to generate a surprisingly equitable allocation, where disadvantaged students do not have worse teachers than advantaged students. In terms of policy, and again consistent with the theory of the second best, fixing one deviation at a time is either ineffective or harmful. Fixing both deviations could close about a fourteenth of the achievement gap per year.

More broadly, this paper has demonstrated the value of using rich data to study the functioning of particular labor markets. Our data allows us to estimate the behavior of the main agents in the market, rather than relying on strong assumptions to infer these from the observed equilibrium. In so doing, we have arrived at surprising conclusions about the determinants of the equilibrium and the design of policies. Presumably, other labor markets would also benefit from such analysis.

References

Agarwal, Nikhil, Charles Hodgson, and Paulo Somaini. 2020. "Choices and Outcomes in Assignment Mechanisms: The Allocation of Deceased Donor Kidneys." Working Paper 28064, National Bureau of Economic Research.

Angrist, Joshua, Peter Hull, Parag A Pathak, and Christopher R Walters. 2024. "Race and the mismeasure of school quality." *American Economic Review: Insights* 6 (1):20–37.

Antos, Joseph R. and Sherwin Rosen. 1975. "Discrimination in the market for public school teachers." *Journal of Econometrics* 3:123–150.

Arteaga, Felipe, Adam J Kapor, Christopher A Neilson, and Seth D Zimmerman. 2022. "Smart matching platforms and heterogeneous beliefs in centralized school choice." *The Quarterly Journal of Economics* 137 (3):1791–1848.

Ashlagi, Itai, Yash Kanoria, and Jacob D Leshno. 2017. "Unbalanced random matching markets: The stark effect of competition." *Journal of Political Economy* 125 (1):69–98.

Aucejo, Esteban M, Patrick Coate, Jane Cooley Fruehwirth, Sean Kelly, and Zachary Mozenter. 2022. "Teacher Effectiveness and Classroom Composition: Understanding Match Effects in the Classroom." *Economic Journal* 132 (648):3047–3064.

Ba, Bocar, Patrick Bayer, Nayoung Rim, Roman Rivera, and Modibo Sidibé. 2021. "Police Officer Assignment and Neighborhood Crime." Working Paper 29243, National Bureau of Economic Research.

Ballou, Dale. 1996. "Do public schools hire the best applicants?" *The Quarterly Journal of Economics* 111 (1):97–133.

Banerjee, Abhijit, Esther Duflo, Maitreesh Ghatak, and Jeanne Lafortune. 2013. "Marry for what? Caste and mate selection in modern India." *American Economic Journal: Microeconomics* 5 (2):33–72.

Barbieri, Gianna, Claudio Rossetti, and Paolo Sestito. 2011. "The determinants of teacher mobility: Evidence using Italian teachers' transfer applications." *Economics of Education Review* 30 (6):1430–1444.

Bates, Michael. 2020. "Public and private employer learning: Evidence from the adoption of teacher value added." *Journal of Labor Economics* 38 (2):375–420.

Bau, Natalie. 2022. "Estimating an equilibrium model of horizontal competition in education." *Journal of Political Economy* 130 (7):1717–1764.

Biasi, Barbara, Chao Fu, and John Stromme. 2021. "Equilibrium in Teachers' Labor Markets with Flexible Pay." Working Paper 28530, National Bureau of Economic Research.

Bobba, Matteo, Tim Ederer, Gianmarco Leon-Ciliotta, Christopher Neilson, and Marco Nieddu. 2024. "Teacher compensation and structural inequality: Evidence from centralized teacher school choice in Perú." Working Paper 29068, National Bureau of Economic Research.

Bonhomme, Stephane, Gregory Jolivet, and Edwin Leuven. 2016. "School characteristics and teacher turnover: assessing the role of preferences and opportunities." *Economic Journal* 126:1342–1371.

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

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and Jim Wyckoff. 2013. "Analyzing the determinants of the matching of public school teachers to jobs: disentangling the preferences of teachers and employers." *Journal of Labor Economics* 31 (1):83–117.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and labor market inequality: Evidence and some theory." *Journal of Labor Economics* 36 (S1):S13–S70.

Chamberlain, Gary. 1982. "Multivariate regression models for panel data." *Journal of Econometrics* 18 (1):5–46.

Chen, Yan and Yinghua He. 2021. "Information acquisition and provision in school choice: an experimental study." *Journal of Economic Theory* 197:105345.

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

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014b. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American Economic Review* 104 (9):2633–79.

Clotfelter, Charles T, Helen F Ladd, and Jacob L Vigdor. 2011. "Teacher mobility, school segregation, and pay-based policies to level the playing field." *Education Finance and Policy* 6 (3):399–438.

Combe, Julien, Umut Dur, Olivier Tercieux, Camille Terrier, and M Utku Unver. 2022. "Market Design for Distributional Objectives in (Re) assignment: An Application to Improve the Distribution of Teachers in Schools." Working paper.

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

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

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

Dahlstrand, Amanda. 2024. “Defying Distance? The Provision of Services in the Digital Age.” Working paper.

Davis, Jonathan. 2022. “The Short and Long Run Impacts of the Deferred Acceptance Algorithm: Evidence from Matching Teach For America Teachers to Schools.” Working paper.

Dee, Thomas S. 2004. “Teachers, race, and student achievement in a randomized experiment.” *Review of Economics and Statistics* 86 (1):195–210.

———. 2005. “A teacher like me: Does race, ethnicity, or gender matter?” *American Economic Review* 95 (2):158–165.

Delgado, William. 2023. “Disparate Teacher Effects, Comparative Advantage, and Match Quality.” Working paper.

Diamond, William and Nikhil Agarwal. 2017. “Latent indices in assortative matching models.” *Quantitative Economics* 8 (3):685–728.

Ederer, Tim. 2023. “Labor market dynamics and teacher spatial sorting.” Tech. rep.

Elacqua, Gregory, Anne Sofie Westh Olsen, Santiago Velez-Ferro et al. 2021. “The Market Design Approach to Teacher Assignment: Evidence from Ecuador.” Tech. rep., Inter-American Development Bank.

Engel, Mimi, Marisa Cannata, and F Chris Curran. 2017. “Principal influence in teacher hiring: Documenting decentralization over time.” *Journal of Educational Administration* 56 (3):277–296.

Engel, Mimi, Brian A Jacob, and F Chris Curran. 2014. “New evidence on teacher labor supply.” *American Educational Research Journal* 51 (1):36–72.

Fox, Lindsay. 2016. “Inferring school desirability based on teacher transfer requests in New York City.” Working paper.

Glazerman, Steven and Jeffrey Max. 2011. “Do Low-Income Students Have Equal Access to the Highest-Performing Teachers?” *National Center for Education Evaluation and Regional Assistance* .

Goldhaber, Dan, Lesley Lavery, and Roddy Theobald. 2015. “Uneven playing field? Assessing the teacher quality gap between advantaged and disadvantaged students.” *Educational Researcher* 44 (5):293–307.

———. 2016. “Inconvenient truth? Do collective bargaining agreements help explain the mobility of teachers within school districts?” *Journal of Policy Analysis and Management* 35 (4):848–880.

Goldhaber, Dan, Vanessa Quince, and Roddy Theobald. 2018. “Has it always been this way? Tracing the evolution of teacher quality gaps in US public schools.” *American Educational Research Journal* 55 (1):171–201.

Goldhaber, Dan, Roddy Theobald, and Danielle Fumia. 2022. "The role of teachers and schools in explaining STEM outcome gaps." *Social Science Research* :102709.

Graham, Bryan S., Geert Ridder, Petra M. Thiemann, and Gema Zamarro. 2023. "Teacher-to-classroom assignment and student achievement." *Journal of Business and Economic Statistics* 41 (4):1328–1340.

Greenberg, David and John McCall. 1974. "Teacher mobility and allocation." *Journal of Human Resources* 9 (4):480–502.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2004. "Why public schools lose teachers." *Journal of Human Resources* 39 (2):326–354.

Harris, Douglas N and Tim R Sass. 2011. "Teacher Training, Teacher Quality and Student Achievement." *Journal of Public Economics* 95 (7-8):798–812.

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

Hitsch, Gunter J, Ali Hortaçsu, and Dan Ariely. 2010. "Matching and sorting in online dating." *American Economic Review* 100 (1):130–63.

Isenberg, Eric, Jeffrey Max, Philip Gleason, and Jonah Deutsch. 2022. "Do Low-Income Students Have Equal Access to Effective Teachers?" *Educational Evaluation and Policy Analysis* 44 (2):234–256.

Isenberg, Eric, Jeffrey Max, Philip Gleason, Liz Potamites, Robert Santillano, Heinrich Hock, and Michael Hansen. 2013. "Access to Effective Teaching for Disadvantaged Students. NCEE 2014-4001." *National Center for Education Evaluation and Regional Assistance* .

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

Jackson, C Kirabo. 2018. "What do test scores miss? The importance of teacher effects on non–test score outcomes." *Journal of Political Economy* 126 (5):2072–2107.

Jacob, Brian A, Jonah E Rockoff, Eric S Taylor, Benjamin Lindy, and Rachel Rosen. 2018a. "Teacher applicant hiring and teacher performance: Evidence from DC public schools." *Journal of Public Economics* 166:81–97.

———. 2018b. "Teacher applicant hiring and teacher performance: Evidence from DC public schools." *Journal of Public Economics* 166:81–97.

Jatusripitak, Napat. 2018. "Three Essays in Applied Microeconomics." .

Johnston, Andrew C. 2024. "Preferences, Selection, and the Structure of Teacher Pay." *American Economic Journal: Applied Economics* .

Kane, Thomas J and Douglas O Staiger. 2008. "Estimating teacher impacts on student achievement: An experimental evaluation." Working Paper 14607, National Bureau of Economic Research.

Kapor, Adam J, Christopher A Neilson, and Seth D Zimmerman. 2020. "Heterogeneous beliefs and school choice mechanisms." *American Economic Review* 110 (5):1274–1315.

Laverde, Mariana, Elton Mykerezi, Aaron Soujourner, and Aradhya Sood. 2023. "Gains from Reassignment: Evidence from a Two-Sided Teacher Market." Working paper.

Lipsey, Richard G. and Kelvin Lancaster. 1956. "The general theory of second best." *The Review of Economic Studies* 24 (1):11–32.

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

Michelmore, Katherine and Susan Dynarski. 2017. "The gap within the gap: Using longitudinal data to understand income differences in educational outcomes." *AERA Open* 3 (1).

Mundlak, Yair. 1978. "On the pooling of time series and cross section data." *Econometrica* 46 (1):69–85.

National Council on Teacher Quality. 2022. "Teacher contract database."

Neal, Derek A and William R Johnson. 1996. "The role of premarket factors in black-white wage differences." *Journal of Political Economy* 104 (5):869–895.

Oyer, Paul and Scott J Schaefer. 2011. "Personnel economics: Hiring and incentives." *Handbook of Labor Economics* 4 (PART B):1769–1823.

Papay, John P and Matthew A Kraft. 2016. "The productivity costs of inefficient hiring practices: Evidence from late teacher hiring." *Journal of Policy Analysis and Management* 35 (4):791–817.

Reardon, Sean F. 2013. "The widening income achievement gap." *Educational leadership* 70 (8):10–16.

Rivkin, Steven G, Eric A Hanushek, and John F Kain. 2005. "Teachers, schools, and academic achievement." *Econometrica* 73 (2):417–458.

Roth, Alvin E. and Marilda A. Oliveira Sotomayor. 1992. *Two Sided-Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press.

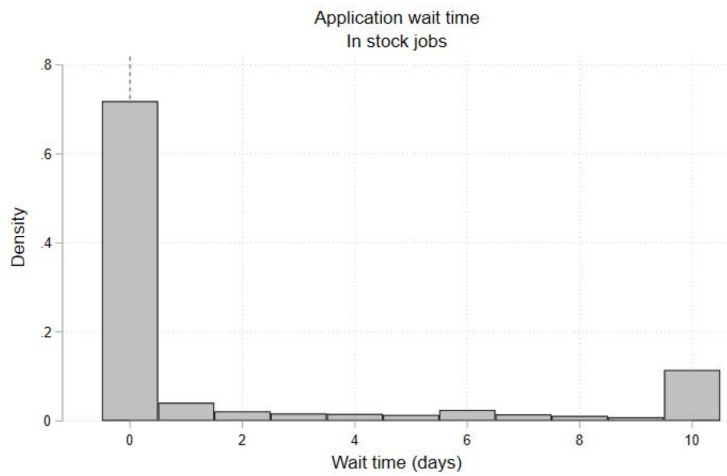
Sass, Tim R., Jane Hannaway, Zeyu Xu, David N. Figlio, and Li Feng. 2012. "Value added of teachers in high-poverty schools and lower poverty schools." *Journal of Urban Economics* 72:104–122.

Sorkin, Isaac. 2018. "Ranking firms using revealed preference." *Quarterly Journal of Economics* 133 (3):1331–1393.

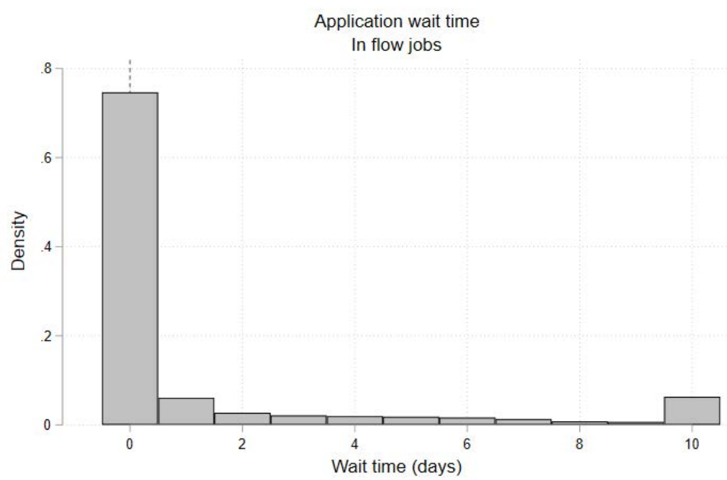
Tincani, Michela M. 2021. "Teacher Labor Markets, School Vouchers, and Student Cognitive Achievement: Evidence from Chile." *Quantitative Economics* 12 (1):173–216.

Figure 1: Wait time to apply to vacancies

(a) Stock of vacancies



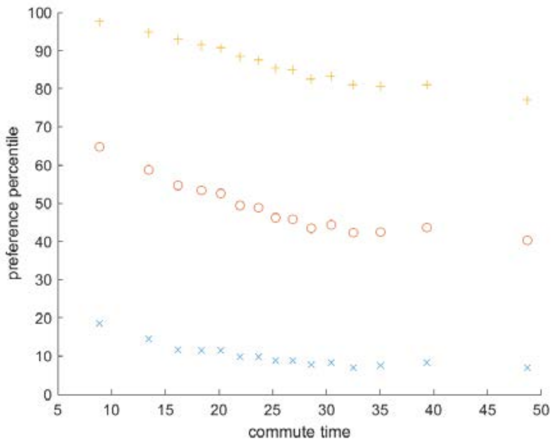
(b) Flow of vacancies



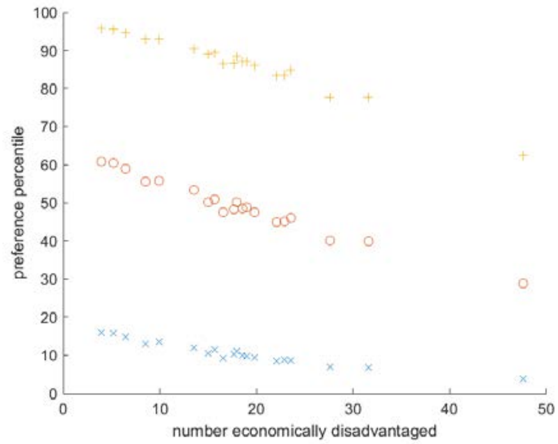
The figures show the wait time for applicants to apply to vacancies. In Panel A, we look at vacancies that were "in stock" (already posted) on the day the teacher first applied on the platform. We plot the "leave one out" wait time, where we omit one job the teacher applied to on the first day. In Panel B we look at the wait time to apply to vacancies that were posted after the teacher first applied on the platform. We measure wait time as the time from when the teacher first applied to another job (once the focal position is posted) until they apply to the posted job. The final category corresponds to waiting at least 10 days. The median wait time is zero in both figures.

Figure 2: Bivariate preference relationships

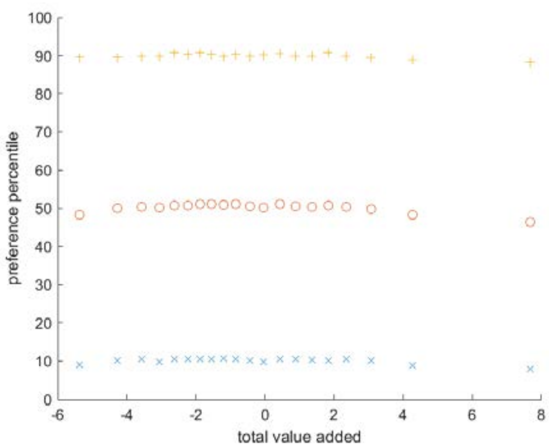
(a) Teachers: commute time



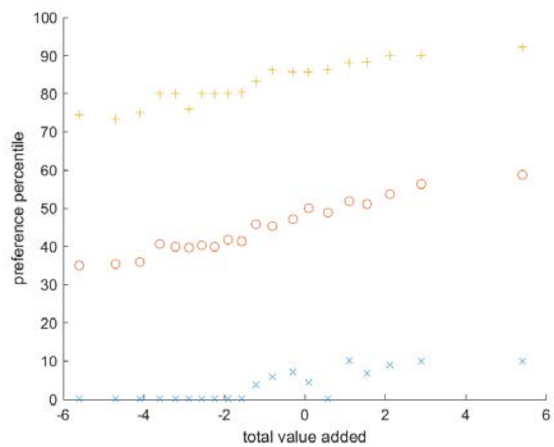
(b) Teachers: number disadvantaged



(c) Teachers: output



(d) Principals: output



This figure shows binscatters of bivariate relationships between characteristics and preferences. The middle set of points (red circle) is the mean percentile, while the top (orange cross) and bottom (blue x) sets of points are the pointwise 10th and 90th percentiles, respectively. In Panels (a)-(c), we show the bivariate relationship between characteristics in the teacher preference model and how teachers rank positions by estimating each teacher's ranking over positions and ordering positions from a teacher's most preferred (100) to least preferred (0). In Panel (d), we estimate show the bivariate relationship between characteristics in the principal model and principal rankings. We estimate each principal's ranking over teachers and order teachers from a principal's most preferred (100) to least preferred (0).

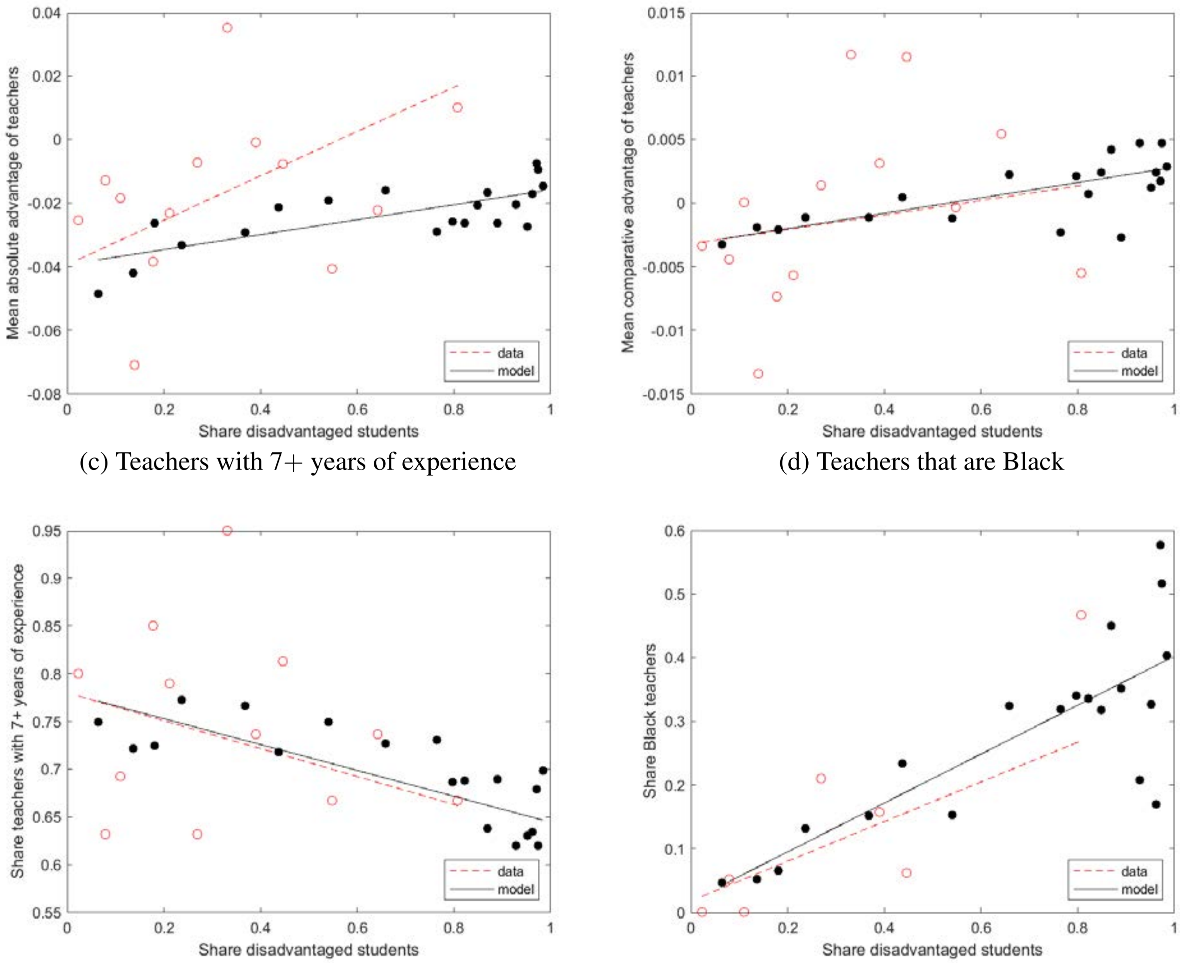
Figure 3: Model fit

(a) Teacher absolute advantage

(b) Teacher comparative advantage

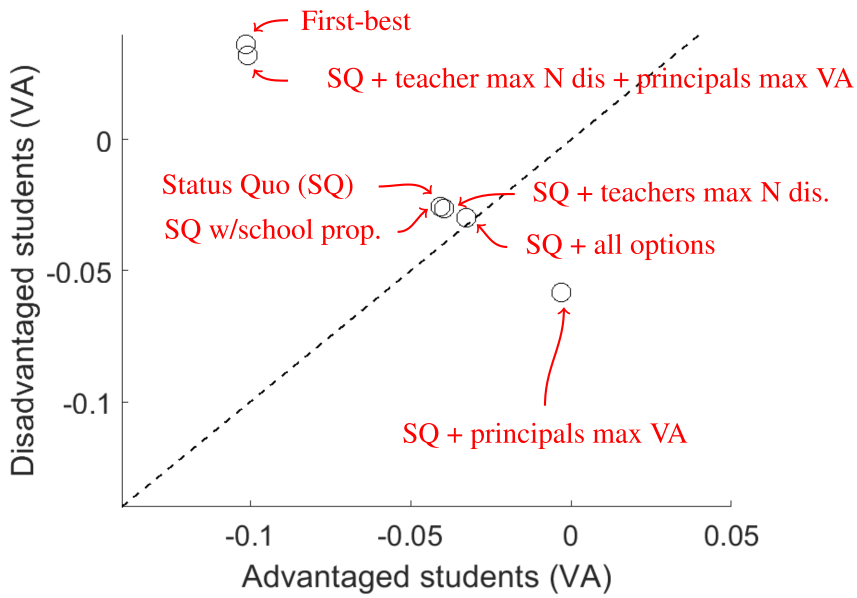
(c) Teachers with 7+ years of experience

(d) Teachers that are Black



This figure compares the allocations implied by the model to the allocations we observe in the data. The solid line presents the fitted values and the dots represent the binscatter. The data refers to all teachers in the district. The model refers to the teachers who apply in the transfer system for whom we have value-added scores. Positions are sorted on the x-axis by the share of disadvantaged students in the school. The intercepts are normalized to be equal. In panel (a), absolute advantage is the average of the teacher's ability with advantaged and disadvantaged students, weighted by the share of these students in the district as a whole. In panel (b), comparative advantage is the difference between value-added with disadvantaged and advantaged students. In panel (c), the outcome is the share of teachers with 7 or more years of experience in the state of the North Carolina. In panel (d), the outcome is the share of teachers at the school that are Black.

Figure 4: Current allocation, alternative policies, and first-best

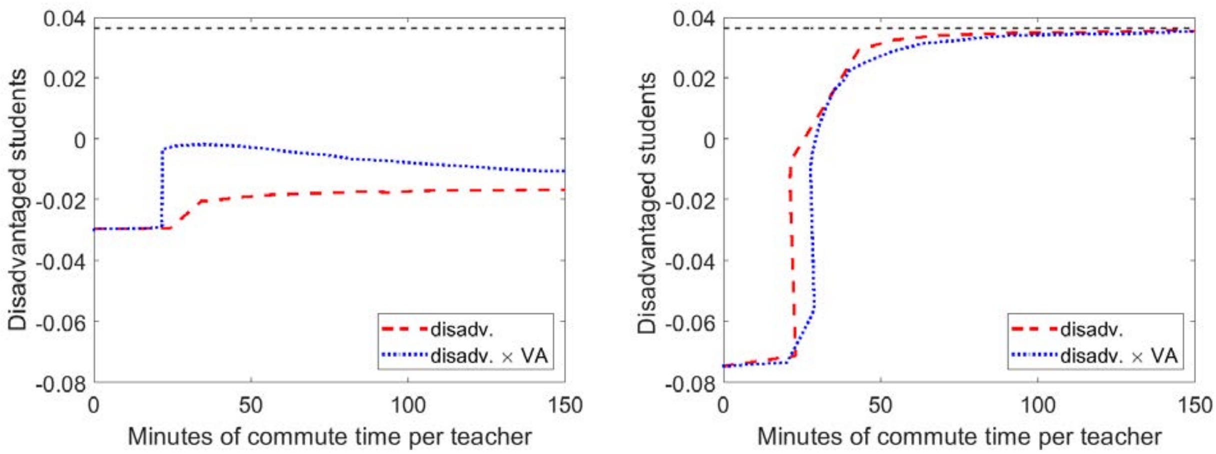


This figure simulates the trade-off between student achievement for economically advantaged and disadvantaged students. The axes refer to per student achievement. The status quo uses teacher and principal estimated preferences and restricted choice sets, and solves for the teacher proposing stable allocation. The status quo with school proposing replaces the teacher proposing deferred acceptance algorithm with school proposing; this point is the only one in the figure that uses school proposing deferred acceptance. The status quo plus all options relaxes the timing restrictions and allows teachers to match with any position; this point is the only one in the figure that relaxes timing constraints. The status quo plus teachers max N disadvantaged replaces the estimated teacher preferences with the assumption that teachers seek to maximize the number of disadvantaged students they teach. The status quo plus principals maximize value-added replaces the estimated principal behavior with the assumption that principals seek to hire teachers to maximize the achievement of their students. The status quo plus teachers maximize the number of disadvantaged and principals maximize value-added replaces estimated with teacher and principal preferences with these assumptions. The first-best is the allocation where the planner maximizes the achievement of disadvantaged students. The Figure plots averages over 400 simulations.

Figure 5: Teacher bonus schemes

(a) Principals' estimated preferences

(b) Principals value output



This figure shows the effect of teacher bonus schemes on the achievement of disadvantaged students. The y-axis is per student achievement. The x-axis shows the cost of the policy per teacher, which we express in minutes of commute time per teacher. The y-axis shows the benefits in terms of achievement per disadvantaged student. We consider two policies: subsidizing the position based on the number of disadvantaged students in the position, and subsidizing the position based on number disadvantaged interacted with the teacher's absolute advantage. In the left panel, we take as the baseline allocation the status quo without timing constraints. In the right panel, we take as the baseline allocation one where principals maximize output without timing constraints. The horizontal dashed lines show the output in the first-best.

Table 1: Teacher preference estimates

	Estimate	Standard Error
Constant	4.370	3.803
Commute Time	-0.058	0.001
Commute Time Missing	-1.415	0.033
Value Added	0.075	0.021
St Dev Value Added RC	0.017	0.009
School Characteristics and Interactions		
N Disadv.	-0.034	0.003
Frac. Black	-0.656	0.123
Frac. Hispanic	0.141	0.122
Frac. Above Med. Achiev.	0.350	0.137
Abs Adv x N Disadv.	-0.060	0.033
Abs Adv x Frac. Black	-0.206	1.062
Abs Adv x Frac. Hispanic	1.251	1.128
Abs Adv x Frac. Above Med. Achiev.	-1.181	1.259
Black x Frac. Black	1.431	0.191
Hispanic x Frac. Hispanic	0.697	0.829
St Dev N. Disadv. RC	0.032	0.002
St Dev Frac. Black RC	1.478	0.067
St Dev Frac. Hispanic RC	1.513	0.090
St Dev Frac. Above Med. Achiev. RC	1.749	0.053
Teacher Characteristics		
VA Non-Disadv. Students	-0.583	0.415
VA Disadv. Students	0.373	0.487
In District	-0.039	0.078

	Estimate	Standard Error
Black	0.924	1.423
Hispanic	5.430	4.672
Female	-0.091	0.093
Experience 2-3	0.077	0.137
Experience 4-6	-1.166	0.125
Experience 7+	-1.240	0.112
St Dev Random Effect	1.558	0.038
Chamberlain-Mundlak Device		
N. Disadv. Mean	0.397	0.153
Commute Time Mean	0.011	0.005
Commute Time Missing Mean	0.741	0.190
Value Added Mean	0.107	0.768
Frac. Black Mean	-4.218	3.695
Frac. Hispanic Mean	-13.589	3.879
Frac. Above Med. Ach. Mean	10.079	5.127
Abs Adv x N Disadv. Mean	-0.434	0.805
Abs Adv x Frac. Black Mean	-0.162	17.202
Abs Adv x Frac. Hispanic Mean	3.557	18.942
Abs Adv x Frac. Above Med. Achiev. Mean	1.428	20.823
Black x Frac. Black Mean	-4.718	3.274
Hispanic x Frac. Hispanic Mean	-20.910	19.121
Number of Students Mean	-0.440	0.123
Sample Size: N Applicants	866	
Sample Size: N Obs	128,264	

The two columns of the table report coefficients from the same model. The table shows teacher preference coefficients, estimated using maximum simulated likelihood. We model the probability that a teacher applies to a position where the alternate options are not teaching in the district or keeping the current position. Random coefficients ("RC") are independent and simulated from the standard normal distribution. We model unobserved teacher-year heterogeneity using a Mundlak (1978) and Chamberlain (1982) device, taking the mean of each covariate across an applicant's choices. Commute time is measured in minutes, value added is total predicted output. Experience below 2 years is the omitted category.

Table 2: Principal valuation estimates

	Estimate	Standard Error		Estimate	Standard Error
Characteristics			Female	-0.009	0.107
Constant	-5.553	0.529	Female x Title I	0.071	0.134
St Dev Random Effect	1.398	0.021	Gender Missing	0.854	0.480
Title I	0.611	0.682	Gender Missing x Title I	-0.576	0.647
Value Added	0.095	0.027	Race Missing	-0.454	0.227
Value Added x Title I	0.036	0.036	Race Missing x Title I	0.330	0.270
Experience 2-3	0.360	0.131	VA Missing	0.488	0.090
Experience 2-3 x Title I	-0.043	0.167	VA Missing x Title I	-0.201	0.126
Experience 4-6	0.182	0.119	Praxis	0.169	0.054
Experience 4-6 x Title I	0.080	0.162	Praxis x Title I	0.007	0.068
Experience 7+	0.037	0.095	Praxis Missing	-0.139	0.066
Experience 7+ x Title I	-0.315	0.127	Praxis Missing x Title I	0.121	0.083
Experience Missing	-0.356	0.068	Grad Deg	0.157	0.069
Experience Missing x Title I	0.437	0.092	Grad Deg x Title I	-0.234	0.088
Masters	0.055	0.112	Grad Deg Missing	-0.138	0.731
Masters x Title I	0.258	0.142	Grad Deg Missing x Title I	-0.415	0.834
Black	-0.972	0.235	Certified	0.998	0.678
Black x Title I	1.773	0.475	Certified x Title I	-1.015	0.811
Black x Frac. Black	0.646	0.280	Certified Missing	0.244	0.671
Black x Frac. Black x Title I	-0.512	0.532	Certified Missing x Title I	-0.792	0.801
Hispanic	-0.651	0.456	Licensed	0.955	0.429
Hispanic x Title I	0.502	0.566	Licensed x Title I	0.462	0.457
Hispanic x Frac. Hispanic	2.277	2.230	**Sample Size: N Positions**	1,824	
Hispanic x Frac. Hispanic x Title I	-1.773	2.364	**Sample Size: N Obs**	343,161	

The two columns of the table report coefficients from the same model. The table shows principal valuation coefficients, estimated using maximum simulated likelihood. We model the probability that a principal submits a positive outcome (hire, interview, positive rating) for an application. Random effects are simulated from the normal distribution. Experience below 2 years is the omitted category. Value-added is total predicted output.

Table 3: Summary statistics for 2015-16, by economic disadvantage

	Focal, Adv	Focal, Disadv	Other, Adv	Other, Disadv
Students				
White (%)	64.61	9.11	75.58	35.09
Black (%)	17.04	51.78	9.54	32.63
Hispanic (%)	6.77	32.58	6.00	23.90
Student performance (level scores)				
Math	0.70	-0.16	0.43	-0.30
Student performance (gain scores)				
Math	0.07	0.07	-0.01	0.00
Teachers				
Experience (% of teachers)				
0 years	4.32	10.99	3.35	4.85
1-2 years	10.45	17.20	6.90	9.80
3-5 years	17.32	19.30	11.21	12.84
6-12 years	29.48	23.01	26.72	26.19
13 or more years	38.43	29.49	51.82	46.32
Graduate degree (%)	45.20	43.34	39.65	37.43
Regular license (%)	94.69	85.60	95.69	93.44
NBPTS certified (%)	16.08	6.82	14.27	9.95
Praxis score	0.37	0.03	0.29	0.13
Age	39	37	41	40
Mean math value-added				
Baseline, econ disadv	0.02	0.02	-0.01	-0.01
Baseline, econ adv	0.01	0.01	-0.01	-0.01
Homogeneous	0.02	0.01	-0.00	-0.01
CFR	0.09	0.07	0.01	0.01
Using school means	0.15	0.13	0.07	0.08
Imputed, econ disadv	-0.03	-0.01	-0.01	-0.01
Imputed, econ adv	0.01	0.01	-0.01	-0.01
Fraction imputed	0.16	0.41	0.14	0.22
Residual	0.02	0.02	-0.03	0.00
Mean behavioral value-added				
Baseline	-0.01	0.01		
Sample size				
Number of students	12,329	22,628	122,903	197,028

The table shows mean student and teacher characteristics in our sample for the 2015-16 school year. We split the sample into whether the student is in our focal district ("Focal") or in the rest of North Carolina ("Other") and whether he or she is economically advantaged ("Adv") or disadvantaged ("Disadv"). Math scores are standardized to have mean 0 and standard deviation 1 at the state-grade-year level. The alternate VA estimators are (a) a homogeneous value-added model with constant effects across student types, (b) using the estimator from Chetty, Friedman, and Rockoff (2014a), and (c) a model that uses school mean characteristics rather than school fixed effects. Imputed value-added predicts value-added based on how observable teacher characteristics predict value-added. "Residual" is the unshrunken value-added for 2015-16, which has no missing values.

Table 4: Current allocation, alternative policies, and first-best

Description	VA disadv.	VA adv.	mean VA
Status quo	-0.026	-0.040	-0.031
Noisy hiring	-0.025	-0.037	-0.029
School propose	-0.026	-0.040	-0.031
All options (timing)	-0.030	-0.033	-0.031
Principals rank by VA	-0.058	-0.003	-0.040
Teachers rank by N disadv.	-0.026	-0.041	-0.031
Previous two changes	0.032	-0.101	-0.013
First best	0.036	-0.101	-0.010

This table displays numbers corresponding to the allocations plotted in Figure 4, as well as the overall achievement per student. The status quo uses teacher and principal estimated preferences and restricted choice sets, and solves for the teacher proposing stable allocation. Noisy hiring maintains the status quo but replaces the estimated principal behavior with a random valuation of applicants. School propose takes the status quo and instead uses the school proposing stable allocation. All options relaxes the timing constraint in the status quo. Teachers rank by N disadvantage changes the teacher preferences in the status quo. Principals rank by VA changes the principal preferences in the status quo. Previous two changes takes the status quo and replaces the teacher preferences with teachers ranking on the number of disadvantaged students and principals ranking by value added. The first-best is the allocation where the planner maximizes the achievement of disadvantaged students. We report averages over 400 simulations.

Table 5: Robustness: disadvantaged achievement

	Status quo	All options	Principal Max VA	Teach Max N Dis	Previous Two	First best
Baseline	-0.026	-0.030	-0.058	-0.026	0.032	0.036
1. Hold class sizes constant: baseline uses class size						
Constant class size	-0.027	-0.030	-0.049	-0.027	0.017	0.020
Constant class size (CFR)	-0.026	-0.029	-0.051	-0.034	0.019	0.022
2. Vary choice set construction for teachers						
7 day buffer	-0.028	-0.031	-0.062	-0.024	0.032	0.036
First day choice sets only	-0.030	-0.031	-0.046	-0.026	0.015	0.036
Drop single app. teachers	-0.025	-0.029	-0.055	-0.025	0.030	0.035
Donut	-0.034	-0.036	-0.062	-0.032	0.024	0.031
Drop 20 percent of apps.	-0.027	-0.031	-0.060	-0.026	0.032	0.036
3. Estimate principal preferences using rank order logit: baseline is binary logit						
All data	-0.027	-0.030	-0.058	-0.031	0.032	0.036
Active choices	-0.024	-0.026	-0.059	-0.036	0.032	0.036
Hire outcome only	-0.025	-0.028	-0.058	-0.027	0.032	0.036
4. Vary window in which we estimate principal preferences: baseline is all applications						
W/in 2 weeks of hire	-0.026	-0.030	-0.059	-0.026	0.032	0.036
First half	-0.027	-0.029	-0.058	-0.025	0.032	0.036
Second half	-0.025	-0.027	-0.058	-0.031	0.032	0.036
5. Vary student type split: baseline is economic disadvantage						
Achievement	-0.026	-0.031	-0.055	-0.025	0.028	0.034
Race	-0.026	-0.029	-0.054	-0.027	0.024	0.029
6. Alternative value-added models						
Homogenous	-0.026	-0.032	-0.052	-0.024	0.032	0.044
Using school means	-0.026	-0.031	-0.055	-0.033	0.037	0.050
CFR	-0.026	-0.028	-0.061	-0.034	0.035	0.044
7. Allow for correlated random coefficients in teacher preferences						
Corr. R.C.	-0.026	-0.029	-0.061	-0.026	0.032	0.036
8. Vary teacher preference specification to use binary logit						
No REs or FEs	-0.026	-0.028	-0.049	-0.026	0.032	0.036
Teacher REs, School FEs	-0.026	-0.028	-0.043	-0.026	0.032	0.036
Teacher FEs, School FEs	-0.027	-0.028	-0.046	-0.026	0.032	0.036
9. Omit value-added from teacher preferences						
No VA	-0.027	-0.030	-0.061	-0.026	0.032	0.036
10. Multiply value-added coefficients by 10 in principal model						
Multiply by 10	-0.029	-0.034	-0.059	-0.021	0.032	0.036
11. Efficiency objective: outcome is mean achievement						
	-0.026	-0.027	-0.035	-0.006	-0.000	0.011
12. Impute value-added for teachers without value added						
Balanced	-0.026	-0.029	-0.039	-0.041	-0.015	0.019
Teachers long	-0.026	-0.022	-0.019	-0.048	-0.019	0.031
13. The role of timing						
One period, period 1	-0.008	N/A	N/A	N/A	N/A	N/A
One period, period 2	-0.018	N/A	N/A	N/A	N/A	N/A
One period, period 3	-0.066	N/A	N/A	N/A	N/A	N/A
Three periods, overall	-0.027	N/A	N/A	N/A	N/A	N/A

The table shows robustness checks for our main results. Each of the robustness checks have been explained and motivated throughout the text and prior footnotes. The columns show the value-added of teachers assigned to disadvantaged students. The status quo uses teacher and principal estimated preferences and restricted choice sets, and solves for the teacher proposing stable allocation. All options relaxes the timing constraint in the status quo. Teachers rank by N disadvantage changes the teacher preferences in the status quo. Principals rank by VA changes the principal preferences in the status quo. Previous two changes takes the status quo and replaces the teacher preferences with teachers ranking on the number of disadvantaged students and principals ranking by value added. The first-best is the allocation where the planner maximizes the achievement of disadvantaged students. For comparability, we normalize the status quo outcome to be the same as in the baseline for rows where the value-added model, population or outcome is different: Panel 1 (row 2), Panels 5, 6, 11, and 12. We report averages over 400 simulations.

Online Appendix

Teacher labor market policy and the theory of the second best

Michael Bates, Michael Dinerstein, Andrew Johnston, and Isaac Sorkin

A Data Appendix

A.1 Student-level data

We use student records from the NCERDC over the years of 2006-2007 through 2017-2018 to measure multi-dimensional teacher productivity in raising math test scores. This provides 8,177,312 student-year observations. We focus on math teachers in grades 4 through 8 to capture the majority of teachers with prior performance data who enter the applicant pool. We use third to seventh-grade math and reading scores as lagged achievement. Test score data as well as student demographics such as ethnicity, gender, gifted designation, disability designation, whether the student is a migrant, whether the student is learning English, whether the student is economically disadvantaged, test accommodations, age, and grade come from the NCERDC master-build files. We use only data from standard end-of-grade exams. This leaves us with 5,322,896 student-year observations.

Beginning in the 2006-2007 school year, the state began recording course membership files linking students directly to courses and instructors. Prior to this change, teachers were linked to students through data on the proctors of the end-of-course exams. The new course membership files provide stronger teacher–subject-student links than the previous system, in which teachers were more frequently linked to the wrong subject (Harris and Sass, 2011).

With the course membership files, we still must determine which teacher is most responsible for teaching math. We use a tiered system. We use course codes (starting with "20") and course names (including the text "math," "alg," "geom," and "calc") to do so. We also want to prioritize standard classes as opposed to temporary or supplemental instruction (course names including text such as "study," "special," "resource," "pullout," "remed," "enrich," "indiv," and "except"). We assign students to the teacher most likely to be the math teacher according to the following rules: (1) Students are assigned first to a high-certainty math teacher (the course code and title indicate a standard math class without mention of supplemental instruction). (2) Students with self-contained teachers are assigned to that teacher if there is no high-certainty math teacher present. (3) Students with course codes and course titles indicating math teachers but no self-contained teachers or high-certainty math teachers are assigned to those middle-certainty math teachers. (4) Students with a teacher of a course that either has a math code or a math course title but no other math course or self-contained teacher are assigned to those low-certainty math teachers. (5) Students with a science course code but no math course or self-contained courses are assigned to their science teachers to accommodate recent trends in math and science block scheduling. We exclude classes in which more than half the class requires special accommodations. Ultimately, our sample for constructing teacher value-added measures is composed of 5,159,337 student-year observations providing measures for 38,566 teachers.

A.2 Application and vacancy data

Our application and vacancy data cover the 2010-2019 cycles. We restrict our sample to applications and vacancies for on-cycle, standard elementary school positions. We show how these restrictions change the sample in Appendix Table A1.

We define on-cycle as positions that receive their first applications of a cycle between April 1 and August 15.

We select standard elementary school positions by filtering on the vacancy type (”instructional”) and the vacancy title. Seventy percent of posted vacancies are for instructional positions. We require that the position indicate elementary school grades by having at least one of the following text strings in the title: “k-”, “3rd”, “4th”, “5th”, “-5”, “-6”, “4-6”, or ”elem”. 39% of vacancies include at least one of these strings in the title.

We then exclude positions with specific subjects mentioned in the title or indications that the position is non-standard (“specialized”, “end of year”, “interim”, “assistant”, “virtual”, “resource”, “itinerant”, “exchange”, “extensions”, “immersion”, “academic support”, “temporary”, “continuous”, “early end”, “interventionist”, or “substitute”). With all of the restrictions above, our final sample consists of 20% of the full set of applications, 25% of the full set of applicants, and 7% of the full set of vacancies.

We code the application’s outcome into whether the candidate is hired (“Accepted-Pending Licensure”, “Hired”, “Hiring Request in Process”, “Offer Accepted”), declines an offer (“Offer Declined”), offered an interview (“Completed BEI Interview”, “Contact for Interview”, “Interview Scheduled”, “Invited to Complete Virtual Interview”, “Invited to Interview”, “Recommended for Interview (By Request)”), or given a positive rating (“1st Choice”, “2nd Choice”, “Highly Recommend for Interview”, “Recommend”, “Recommend for Interview”, “Recommendation Accepted”, “Strong Candidate”). These categories are encodings of a single variable, so they are mutually exclusive (i.e., if a candidate is hired, the prior outcome may be overwritten). For robustness analysis, we also split up the remaining applications into middle ratings (“Attended Info Session/Class”, “Hold for Later Consideration”, “Invited to Info Session/Class”, “Possible recommend for interview”, “Recommend with Hesitation”), negative ratings (“Failed Job Questionnaire”, “Incomplete Application”, “Ineligible Selection”, “Not Good Fit”, “Not Qualified”, “Pool - Ineligible”, “SS - INELIGIBLE”, “Screened - Not Selected”), withdrawals (“Candidate Withdrew Interest”), or no evaluation (“Eligible Selection”, “New”, “Pool - Eligible”, “Pool Candidate”).

A.3 Matching across datasets

For this project the North Carolina Education Research Data Center (NCERDC) combined records held there on teacher work histories, school characteristics, and student achievement with data provided by a large urban school district containing further personnel files, open positions within the school district, and applications for those positions. They performed an interactive fuzzy match using the last four digits of social security numbers, names, and birth dates. For teachers who had a sufficiently good match, we have a de-identified ID that allows us to connect their platform data to their staffing records and students' achievement.

The NCERDC reports that of the 74,395 applicants to positions, 29,008 are matched to NCERDC records. Many of these applicants never teach in the state and thus would not be expected to match. Of the 26,983 employees listed within the district, 20,966 are matched to NCERDC records. However, the match rate is much better among personnel who teach tested subjects. Of the 13,982 teachers with EVAAS scores in the district, 13,865 are matched to the NCERDC data.

A.4 Sample characteristics

Returning to Appendix Table A1, we see how the sample's characteristics vary with sample restrictions. The "Elementary Sample" restricts to on-cycle elementary school instructional positions without specialization, the "Value-Added Sample" further restricts to teachers with value-added forecasts based on prior years, and the "2015 Sample" further restricts to the 2015 application cycle. We use the "Elementary Sample" for estimating principal preferences, the "Value-Added Sample" for estimating teacher preferences, and the "2015 Sample" for estimating counterfactual allocations.

We see a few expected patterns based on the sample restrictions. For the last two columns, we require teachers to have value-added forecasts based on data from prior years. This restriction leads us to a more experienced sample of teachers. These teachers are more likely both to already be in the district and to transfer to a new school (from a prior school or from out of district). We also see these teachers have lower application rates, perhaps because many already have in-district placements. We see little change in the teacher sample's mean value-added (by student type or at a representative school) or choice set size. The mean characteristics in the positions sample also change minimally with the sample restrictions.

B Efficiency: district objective is maximizing average achievement

In the body of the paper, we consider a district that has the objective of maximizing the achievement of disadvantaged students. In this Appendix, we show our main results when instead the district cares about maximizing the achievement of all students. Many of the same messages apply, but there are a few differences. Notably, when we consider overall achievement as an objective, there is an important limitation of relying on common prices for output. Instead, there are benefits to more flexible pricing.

B.1 The objective

In the body of the paper, we consider a district with the objective function:

$$\max_{\phi \in \Phi} \left\{ \sum_{j \in \mathcal{J}} n_{k1} \mu_{j1} \right\}. \tag{A1}$$

In words, the district picks the assignment to maximize the achievement of disadvantaged students. Here we consider the objective function:

$$\max_{\phi\in\Phi}\left\{\sum_{j\in\mathcal{J}}(n_{k0}\mu_{j0}+n_{k1}\mu_{j1})\right\}. \tag{A2}$$

The difference is that now the planner also values the achievement of the non-disadvantaged students.

B.2 Allocations and results

We consider a set of allocations that parallel those we discussed in the body of the paper. The only difference is that rather than considering a teacher who seeks to maximize the number of disadvantaged students, we consider a teacher who seeks to maximize the output in the match, when considering both advantaged and disadvantaged students in the class.

Panel 11 of Table 5 shows the results. The basic pattern is similar to our results about equity. Specifically, timing plays a minor role in limiting output and having principals maximize output actually reduces output. One difference is that when teachers seek to maximize output, this change moves the allocation closer to the first-best than when we study equity.

Finally, while when we study equity if we "fix" both principals and teachers then we are very close to the first best levels of achievement for disadvantaged students, for efficiency this property does not hold. In particular, the output is only 70% of the way from the status quo to the first best, whereas when we study equity it is 94% of the way there.[44] The reason is the lack of personalized pricing. Specifically, making preferences be based on output is equivalent to assigning a price per unit of output. This pricing scheme is equivalent to pricing based on absolute advantage. To maximize output, however, the planner wants to price based on comparative advantage. The only way to price based on comparative advantage is to "personalize" prices by allowing them to depend on the specification combination of value-added with advantaged and disadvantaged students. Thus, for a planner who cares about efficiency, the lack of personalized pricing in the teacher labor market is an important barrier to achieving desirable allocations.

C Selection into the transfer market

To examine the selection of teachers into the transfer market, we first look at four cohorts, 2010-2013, such that we can follow them for five years. We further restrict attention to those for whom we can measure productivity, leaving us with 553 teachers who entered the state's data during those years. Of those, 207 applied to transfer at some point during the first five years. Only 124 remain in their original school and have not applied to transfer within five years of entering the district. The remaining 287 leave the district. Ap-

[44] The calculations are: $0.70 = 1-\frac{0.011}{0.011+0.026}$ and $0.94 = 1-\frac{0.036-0.032}{0.036+0.026}$.

pendix Table A28 shows that there is very little difference in comparative advantage and absolute advantage between teachers who applied to transfer and the teachers who did not.

D Omitted details on value-added model

D.1 Formal statement of assumptions for value-added model

Here we formally state the assumptions that were informally discussed in Section 4.

Assumption 1 (Exogeneity and stationarity of classroom and student-level shocks)**.** *Classroom-student-type shocks (θ_{cmt}) are independent across classrooms and independent from teachers and schools. Classroom-student-type shocks follow a stationary process:*

$$\mathbb{E}\left[\theta_{c0t}|t\right] = \mathbb{E}\left[\theta_{c1t}|t\right] = 0 \tag{A3}$$

$$Var\left(\theta_{c0t}\right) = \sigma^2_{\theta_0},\ Var\left(\theta_{c1t}\right) = \sigma^2_{\theta_1},\ Cov\left(\theta_{c0t}, \theta_{c1t}\right) = \sigma_{\theta_0\theta_1} \tag{A4}$$

for all t.

Student-level idiosyncratic variation is independent across students and independent from teachers and schools. Student-level shocks follow a stationary process depending on the student's type:

$$\mathbb{E}\left[\tilde{\varepsilon}_{it}|t\right] = 0 \tag{A5}$$

$$Var\left(\tilde{\varepsilon}_{it}\right) = \sigma^2_{\varepsilon m} \text{ for } m = 0, 1 \tag{A6}$$

for all t.

Assumption 2 (Joint stationarity of teacher effects)**.** *The non-experience part of teacher value-added for each student type follows a stationary process that does not depend on the teacher's school. The covariances between the teacher's value-added across student types depend only on the number of years elapsed:*

$$\mathbb{E}\left[\mu_{j0t}|t\right] = \mathbb{E}\left[\mu_{j1s}|t\right] = 0 \tag{A7}$$

$$Var\left(\mu_{j0t}\right) = \sigma^2_{\mu_0},\ Var\left(\mu_{j1t}\right) = \sigma^2_{\mu_1},\ Cov\left(\mu_{j0t}, \mu_{j1t}\right) = \sigma_{\mu_0\mu_1} \tag{A8}$$

$$Cov\left(\mu_{j0t}, \mu_{j0,t+s}\right) = \sigma_{\mu_0 s},\ Cov\left(\mu_{j1t}, \mu_{j1,t+s}\right) = \sigma_{\mu_1 s} \tag{A9}$$

$$Cov\left(\mu_{j0t}, \mu_{j1,t+s}\right) = \sigma_{\mu_0\mu_1 s} \tag{A10}$$

for all t.

Assumption 3 (Independence of drift and school effects)**.** *Let $\bar{\mu}_{jm}$ be teacher j's mean value-added for student type m. Let k be j's assigned school in year t. Then:*

$$(\mu_{jmt} - \bar{\mu}_{jm}) \perp \mu_k \ for\ m = 0, 1. \tag{A11}$$

D.2 Additional details on estimation

In the first step, we estimate β_l by regressing test scores (standardized to have mean 0 and standard deviation 1 in each grade-year) on a set of student characteristics (X_{it}) and classroom-student-type fixed effects:

$$A^*_{it} = \beta_s X_{it} + \lambda_{cmt} + \upsilon_{it}. \tag{A12}$$

For characteristics, we include ethnicity, gender, gifted designation, disability designation, whether the student is a migrant, whether the student is learning English, whether the student is economically disadvantaged, test accommodations, age, and grade-specific cubic polynomials in lagged math and lagged reading scores. We subtract the estimated effects of the student characteristics to form the first set of residuals, $\hat{\upsilon}_{it}$:[45]

$$\hat{\upsilon}_{it} = A^*_{it} - \hat{\beta}_s X_{it}. \tag{A13}$$

These student-level residuals include teacher, school, and classroom components, as well as idiosyncratic student-level variation.

In the second step, we project the residuals onto teacher fixed effects, school fixed effects, and the teacher experience return function. Following the literature, we specify the experience return function as separate returns for every level of experience up to 6 years, and then a single category of experience of at least 7 years:

$$\hat{\upsilon}_{it} = \sum_{e=1}^{6} \alpha^e \mathbb{1}\{Z_{jt} = e\} + \alpha^7 \mathbb{1}\{Z_{jt} \geq 7\} + \mu_{jm} + \mu_k + \mu_t + \varepsilon_{it}, \tag{A14}$$

where $\varepsilon_{it} = (\mu_{jmt} - \mu_{jm}) + \theta_{cmt} + \tilde{\varepsilon}_{it}$. We then form a second set of student-level residuals by subtracting off the estimated school and experience effects:

$$A_{it} = \hat{\upsilon}_{it} - \left(\sum_{e=1}^{6} \hat{\alpha}^e \mathbb{1}\{Z_{jt} = e\} + \hat{\alpha}^7 \mathbb{1}\{Z_{jt} \geq 7\} + \hat{\mu}_k \right). \tag{A15}$$

We aggregate these student-level residuals into teacher-year mean residuals for each student type: $\bar{A}_{jmt}$. Let $\mathbf{A}_j^{-t}$ be a vector of mean residuals for each student type-year that j teaches in the data, prior to year t.

In the final step, we follow Delgado (2023) and form our estimate of teacher j's value-added (net of experience effects) in year t for type m as the best linear predictor based on the prior data in our sample:

$$\hat{\mu}_{jt} \equiv \mathbb{E}^* \left[\mu_{jt} | \mathbf{A}_j^{-t} \right] = \psi' \mathbf{A}_j^{-t}, \tag{A16}$$

[45] Here we deviate from the standard notation, by introducing $\hat{\upsilon}_{it}$. Our procedure has two residualization steps because we include classroom-student type fixed effects in the first step, which would subsume the teacher and school fixed effects. We thus decompose student residuals into teacher and school components in a second step.

where μ_{jt} is a $(2x1)$ vector for the teacher's output across the two student types and ψ is a $2(t-1)x(t-1)$ matrix of reliability weights where $t-1$ is the number of years of prior data. These weights minimize the mean squared error between the estimate of the teacher's value-added and our forecast based on prior data:

$$\hat{\psi}^{'} = argmin \sum_{j} (\bar{A}_{jt} - \psi^{'} \mathbf{A}_{j}^{-t})'(\bar{A}_{jt} - \psi^{'} \mathbf{A}_{j}^{-t}). \tag{A17}$$

We estimate ψ following Delgado (2023). Here we describe how we estimate the structural parameters: $\sigma_{\varepsilon 0}, \sigma_{\varepsilon 1}, \sigma_{\theta 0}, \sigma_{\theta 1}$, $cov(\theta_{c0t}, \theta_{c1t}), \sigma_{\mu 0}, \sigma_{\mu 1}$, $cov(\mu_{j0t}, \mu_{j1t}), cov(\mu_{j0t}, \mu_{j0s})$, $cov(\mu_{j1t}, \mu_{j1s}), cov(\mu_{j0t}, \mu_{j1s})$.

- $\hat{\sigma}_{\varepsilon m} = \frac{1}{N_c} \sum_{c=1}^{N_c} \frac{1}{n_{cm}-1} \sum_{n=1}^{n_{cm}} (\hat{v}_{it} - \frac{1}{n_{cm}} \sum_{n=1}^{n_{cm}} \hat{v}_{it})$
- $\hat{\sigma}_{\theta m} = Var(\bar{A}_{jmtc}) - \hat{\sigma}_{\mu m} - \frac{1}{N_{cm}} \sum_{i=1}^{N_{cm}} \frac{\hat{\sigma}_{\varepsilon m}}{n_{cm}}$
- $c\hat{o}v(\theta_{c0t}, \theta_{c1t}) = cov(\bar{A}_{j0tc}, \bar{A}_{j1tc}) - c\hat{o}v(\mu_{j0t}, \mu_{j1t})$
- $\hat{\sigma}_{\mu m} = \sqrt{cov(\bar{A}_{jmtc}, \bar{A}_{jmtc'})}$, where $c \neq c^{'}$
- $c\hat{o}v(\mu_{j0t}, \mu_{j1t}) = cov(\bar{A}_{j0tc}, \bar{A}_{j1tc'})$, where $c \neq c^{'}$
- $c\hat{o}v(\mu_{j0t}, \mu_{j0s}) = cov(\bar{A}_{j0t}, \bar{A}_{j0s})$
- $c\hat{o}v(\mu_{j1t}, \mu_{j1s}) = cov(\bar{A}_{j1t}, \bar{A}_{j1s})$
- $c\hat{o}v(\mu_{j0t}, \mu_{j1s}) = cov(\bar{A}_{j0t}, \bar{A}_{j1s})$

where N_c is the number of classes, N_{cm} is the number of classes times student types, and n_{cm} is the number of students in class c of type m,

Our estimate of teacher j's composite value-added at school k in year t is:

$$\widehat{VA}_{jkt} = p_{k0t}\hat{\mu}_{j0t} + p_{k1t}\hat{\mu}_{j1t} + f(Z_{jt}; \hat{\alpha}). \tag{A18}$$

Variation in the data: We now discuss the variation in the data that pins down key parameters. The coefficient on student characteristics uses how test scores vary with within-classroom-student type variation in student characteristics.[46] The school effects use the change in (student) output when teachers switch schools, beyond what would be predicted by drift and by the change in student-type composition. Heuristically, if teachers' output regularly increases when teachers transfer to a certain school, then we would estimate a high school effect. The teacher mean effects for each student type are pinned down by relative increases in students' (residualized) test scores across different teachers. We are able to rank teachers both within and across schools, provided teachers and schools are in a set connected by transfers so that we can identify the school effects.

[46] Because we include classroom-student-type fixed effects, our model allows for an arbitrary correlation between students' characteristics and the quality of their assigned teachers. Allowing such correlation is important in a context where teachers have some control over where they work.

Finally, we identify the parameters of the teacher value-added distribution and the drift process based on the stationarity assumptions and the observations of teachers across years, classrooms, and student types. As an example, the variance of the teacher effects for student type m is identified by the covariance between a teacher's mean student residuals for student type m in two different classrooms in the same year. In our setting many elementary school teachers have students from multiple classes. The prevalence of multiple classrooms is increasing over time (Appendix Table A29). With our assumptions that classroom and student shocks are uncorrelated across classrooms, the only reason a teacher's students would have similar (residualized) outcomes is the teacher's value-added.

Appendix Table A30 presents parameter estimates. The first key parameter estimate is the significant dispersion in value-added for both student types of about 0.24σ. The second key parameter estimate is the strong correlation of 0.86 between the teacher's value-added with the two types of students. We find large returns to experience in the first year, and then a profile that flattens out after about four years of experience.

D.3 Alternative estimators

In our analysis, we explore the robustness of our results to elements of our value-added model. We focus on four variations from our baseline model.

Homogeneous: We estimate a model where teachers have a homogeneous effect on students' test scores and classroom shocks are not specific to student type:

$$\begin{aligned} \mu_{j0t} &= \mu_{j1t} = \mu_{jt} \\ \theta_{c0t} &= \theta_{c1t} = \theta ct. \end{aligned} \tag{A19}$$

This imposes an economic restriction common to the literature and increases our forecast precision. This alternative estimator lets us test whether our results are sensitive to modeling comparative advantage or reduced forecast precision.

Chetty, Friedman, & Rockoff 2014 (CFR): We estimate value-added using the Chetty, Friedman, and Rockoff (2014a) model. This model (a) has homogeneous value-added (teachers have a homogeneous effect on students' test scores and classroom shocks are not specific to student type), (b) has no school fixed effects (we instead include school-level means for all variables in X_{it}, (c) residualizes test scores using a restricted teacher fixed effect (rather than a teacher-year-student type fixed effect), and (d) incorporates future test scores when forecast prior value-added. This alternative estimator lets us test whether our results are robust to using the standard estimator from the literature. We estimate using the vam.ado file in Stata.

Using school means: In our baseline model, we include school fixed effects: μ_k. For robustness, instead of including μ_k in Equation A14, we include school-level means for all of the variables in X_{it}. Note that this will not deliver identical estimates because we do not include school-level means of the teacher fixed effects. This alternative estimator lets us test whether our results depend on how we decompose effects into school and teacher components.

Residual: In our baseline model, we shrink teachers' residual output to create measures that are forecast unbiased at the teacher level. For robustness, we estimate a teacher's value-added as the mean residual across j's students in year t: $\bar{A}_{jt}$. This alternative estimator lets us test whether our results are affected by missing data because the mean residual is available for all teachers. While we apply the estimator to characterizing the current allocation, our counterfactual estimates rely on forecast unbiased measures at the teacher level.

D.4 Testing for comparative advantage

Our measures forecast teachers' future value-added without bias. Our high estimated correlation between a teacher's effectiveness with the two student types raises the question of whether our estimates of comparative advantage simply reflect statistical noise. Beyond the exercise presented in Appendix Figure A5, we present two additional ways of testing our multi-dimensional value-added model versus a single-dimensional model.

First, we estimate standard errors and confidence intervals for the structural parameters in our production model. The estimated correlation in teacher value-added across student types is 0.86. We can, however, decisively reject a correlation of 1 as the 95% confidence interval is $(0.73, 0.87)$ (Appendix Table A30).

Second, we perform a likelihood-ratio test comparing our model with a model with one-dimensional teacher value-added. We take the mean residuals at the level of the teacher-classroom-student type, $\bar{A}_{jcmt}$, and collect a teacher's mean residuals across classrooms and student types, which come from a normal distribution:

$$\begin{pmatrix} \bar{A}_{jc1t} \\ \bar{A}_{jc'2t} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{\mu_1} + \sigma^2_{\theta_1} + \frac{\sigma^2_{\varepsilon_1}}{N_{jc0t}} & \sigma_{\mu_1\mu_2} \\ \sigma_{\mu_1\mu_2} & \sigma^2_{\mu_2} + \sigma^2_{\theta_2} + \frac{\sigma^2_{\varepsilon_2}}{N_{jc2t}} \end{pmatrix} \right). \tag{A20}$$

We compare the likelihoods across our baseline model and an alternate model of homogeneous value-added where $\sigma^2_{\mu_1} = \sigma^2_{\mu_2}$, $\sigma^2_{\theta_1} = \sigma^2_{\theta_2}$, $\sigma^2_{\varepsilon_1} = \sigma^2_{\varepsilon_2}$, and $\sigma_{\mu_1\mu_2} = 0$. Our likelihood-ratio test has 4 degrees of freedom, and we reject the homogeneous value-added model in favor of the heterogeneous model, with a test statistic of 610, so the p-value is arbitrarily small ($p < 0.0001$).[47]

Third, we fix a teacher's type according to whether she is above or below the median in comparative advantage in teaching economically disadvantaged students in pre-transfer schools. We then test whether changes in the share of economically disadvantaged students differentially predict changes in student test score residuals ($\hat{\nu}_{it}$ from equation A15) in post-transfer schools by teacher-type. If our estimated comparative advantage is meaningful, as the share of disadvantaged students rises, teachers with a comparative advantage in teaching disadvantaged students should see gains in average productivity relative to teachers with a comparative advantage in teaching economically advantaged students. Appendix Table A31 shows that for teachers with a comparative advantage in teaching advantaged students, productivity falls as the share of disadvantaged students rises (p-value=0.043). In contrast, for teachers with a comparative advantage in teaching disadvantaged students, productivity rises as the share of disadvantaged students rises (p-

[47] We restrict the sample to one randomly-chosen vector of mean residuals per teacher so that the observations in our likelihood are independent. We also find a similar test statistic when we use mean residuals, $\bar{A}_{jcmt}$, from a model where the fixed effects in the residualizing steps are not separated by student type.

value=0.014). These findings indicate that comparative advantage is persistent across settings and predictive of match-specific productivity.

D.5 Behavioral value-added

In robustness checks, we incorporate a measure of a teacher's value-added on behavioral outcomes (Jackson, 2018). Because we focus on elementary school teachers, we have fewer outcomes available (e.g., no grades). We thus measure teachers' effects on a student's log absence rate, whether the student has any in-school suspension, and whether the student has any out-of-school suspension. We recover the first principal component and use this as our outcome.

We estimate two-dimensional behavioral value-added with identical methods to those we use for math value-added. When controlling for lagged student outcomes, we use the lagged value of the first principal component.

E Principal preferences estimation

We estimate principal preferences via maximum simulated likelihood, where we simulate from the normal distributions of the random effect at the level of the position-year. Let n index each simulation iteration and let $B_{jptn}(\theta)$ be the model-predicted probability that p rates j positively in year t in simulation iteration n at parameter vector θ. For each position p in year t, we construct the simulated likelihood as:

$$L_{pt} = \frac{1}{500}\sum_{n=1}^{500}\prod_{j\in\mathcal{J}_{pt}}(b_{jpt}B_{jptn}(\theta)+(1-b_{jpt})(1-B_{jptn}(\theta))), \tag{A21}$$

where $\mathcal{J}_{pt}$ is the set of teachers who applied to a position p in year t and b_{jpt} is an indicator for whether p rated j positively in the data. Our full simulated log likelihood function is:

$$l = \frac{1}{P}\sum_{p} logL_{pt}. \tag{A22}$$

F Imputed value-added

Most of our analysis considers test score output among teachers with value-added forecasts. But teachers may be missing value-added forecasts, and thus not be included in our main analysis, for several reasons. The primary reason is that we forecast value-added using a teacher's prior output. Thus, prospective teachers new to North Carolina public schools would not have a forecast. Furthermore, some teachers may have experience in non-tested grades or subjects and later seek to switch into the grades and subjects we consider. Finally, some teachers may already be teaching in our focal district, in the relevant grades and subjects, but have classrooms with only one student type. Our baseline value-added model includes heterogeneous effects across student types and thus drops teachers who have never taught a student of a certain type.

Several of our analyses are robust to these missing forecasts. In assessing the observed allocation (Table 3), the "Homogeneous" measure includes teachers who have only taught one student type. The "CFR" measure further includes all other teachers who teach at least two years in North Carolina public schools.[48] And the "Residual" measure includes all teachers.

We also show that our model baseline and counterfactual estimates are robust to using several of these measures.[49] But unlike the observed allocation, the model analysis could in principle incorporate prospective teachers who apply for positions but are not hired during our sample. To show robustness to including these teachers, we estimate a model that we use to impute their value-added.

We specify a prospective teacher j's (experience-adjusted) value-added with student type m in year t as:

$$A_{jmt} = \pi_m \tilde{W}_{jt} + \eta_{jmt}. \tag{A23}$$

In W_{jt} we include a constant, the prospective teacher's Praxis score, whether the prospective teacher has a graduate degree, whether the prospective teacher is NBPTS certified, and whether the prospective teacher has a regular license. We also include indicator variables for whether each of these characteristics are missing in the data. These comprise the same set of observable characteristics that are readily available for hiring principals and that have the most explanatory power for principals' ratings (Appendix Table A14). Thus, we will view our imputed value-added as reflecting similar levels of information as hiring principals might have.

For A_{jmt}, we use the mean student-level test score residual, where the mean is taken over all of teacher j's type m students in year t. This mean residual comes from estimating Equation A15 on our full sample of teachers and is available for all teachers who end up teaching in the district, even many teachers for whom we lack value-added forecasts. The residual has already removed any experience effects, which we return to below. The residual is a noisy estimate of a teacher's value-added.

To estimate these regressions, we seek a sample of teachers most similar to the ones we are imputing value-added for. We therefore use the sample of teachers in our focal district who do not have value-added forecasts.

Our imputations rely on two main assumptions. First, we are extrapolating from teachers in the data (i.e., who have been hired) to teachers who have not been hired. We thus assume that on average hired novice teachers are no different from non-hired applicants, conditional on the observable characteristics $\tilde{W}_{jt}$. Second, we are imputing all novice teachers as having identical value-added, conditional on $\tilde{W}_{jt}$. We thus assume if novice teachers' value-added actually differs, these differences do not affect the allocation. Roughly, this means that principals have no private information about these differences and the differences do not correlate with teacher preferences or timing. While we cannot test these assumptions in the relevant

[48]The "CFR" estimator leaves out data from the focal year, such that a second year is necessary for forecasting a teacher's value-added.

[49]We do not show robustness to the "Residual" measure because unlike mean value-added in the current allocation, we require forecast unbiased estimates of each teacher's value-added.

population, among the teachers with value-added forecasts we show in our principal preferences model that conditional on $\tilde{W}_{jt}$, principal hiring does not vary much with a teacher's forecasted value-added. Further, teacher preferences across schools are weakly related to value-added forecasts.

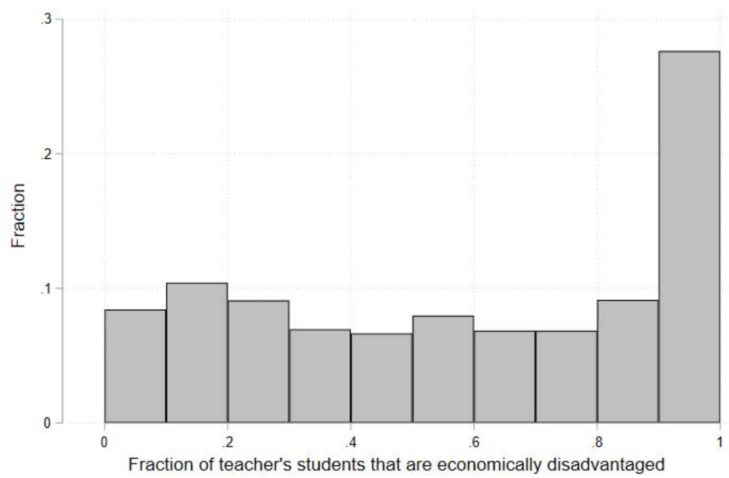
We present the estimates in Appendix Table A32. We also include a third column showing the model with a teacher's comparative advantage (difference in mean residuals) as the outcome. We see that higher Praxis scores and certification (for teaching disadvantaged students) are statistically significant predictors of teacher value-added. As expected from Appendix Table A15, these characteristics as a whole have a limited relationship with teacher value-added such that there is little pre-hiring information about novice teachers to determine which are likely to be more effective. The mean residual, listed at the bottom of the table, stands out, as this sample of teachers has some meaningful comparative advantage in teaching disadvantaged students. This result distinguishes the sample from the teachers for whom we have value-added forecasts, where comparative advantage is fairly limited.

We investigate why the value-added residuals differ among teachers with and without value-added forecasts with a Oaxaca-Blinder decomposition. We present the results for advantaged student value-added in Appendix Table A33, for disadvantaged student value-added in Appendix Table A34, and for comparative advantage in Appendix Table A35. For all three measures, differences in coefficients, rather than differences in W_{jt}, account for the differences in residuals. The estimated constant, especially, stands out such that some of these cross-sample differences are not explained by observable characteristics.

After estimating these regressions, we construct imputed value-added as: $VA^{imp}_{jmt} = \hat{\pi}_m \tilde{W}_{jt} + f(Z_{jt};\hat{\alpha})$. For most teachers, we are adding back the experience effect associated with novice teachers, though for teachers who have experience but do not have forecasts for other reasons, we add a more positive experience effect. We then use these imputed measures for teachers lacking value-added forecasts.

For our model counterfactuals, we have also considered robustness to two other sets of imputation assumptions. First, we impute with a model that includes year fixed effects and is estimated on the sample of teachers with value-added forecasts. Second, we impute with a model like Equation A23 but estimated using the sample of teacher with value-added forecasts and using these forecasts, rather than residuals, on the left-hand-side. These alternative estimators serve two purposes. First, they show whether our results are sensitive to extrapolating from different sets of teachers (those with and without value-added forecasts). Second, they show how incorporating information across a teacher's career affects the results. Our results are robust to these alternative imputation procedures and are available upon request.

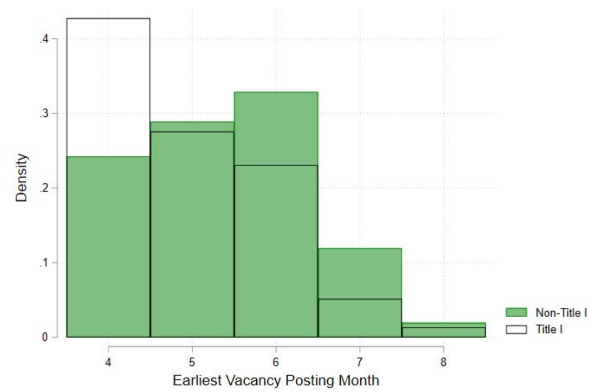
Figure A1: Teacher's fraction of students that are disadvantaged



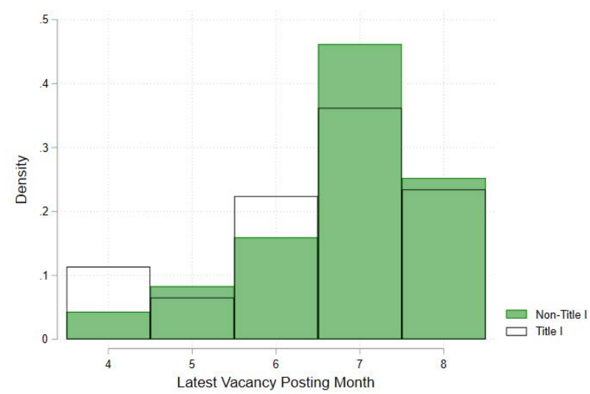
The figure is a histogram of the distribution of student composition varies across teachers. An observation is a teacher-year in our focal district. Student composition is measured as the fraction of the teacher's students in that year that are economically disadvantaged.

Figure A2: Distribution of first and last vacancy posted in cycle

(a) First vacancy posting month

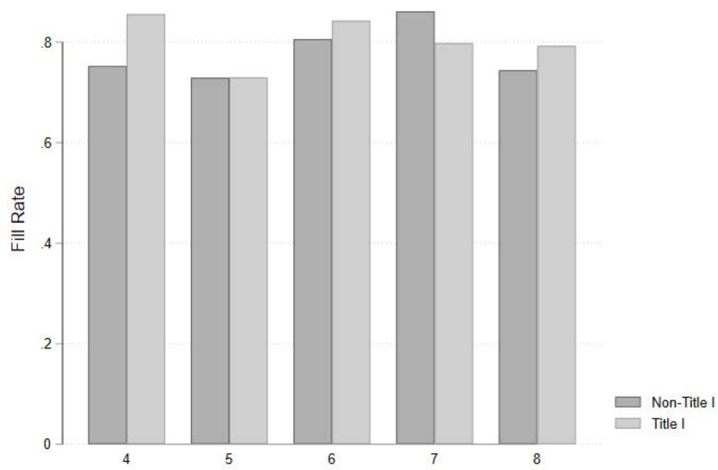


(b) Last vacancy posting month



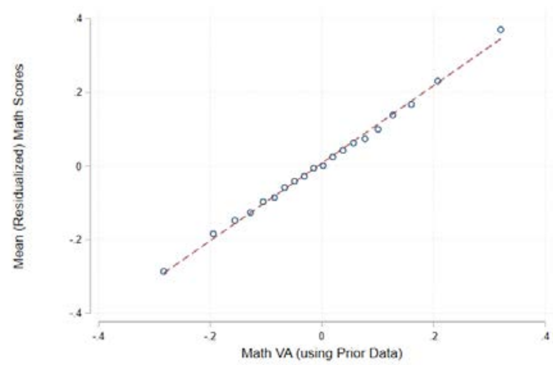
The figure is a histogram of which month (e.g., April = 4, May = 5, etc.) a school first posts a vacancy in each cycle and last posts a vacancy. A value of "6" means that the school's first vacancy in a cycle was posted in June. The shaded histogram is for non-Title I schools and the outlined histogram is for Title I schools. The sample consists of vacancies in our focal district in the 2011-2016 application cycles.

Figure A3: Fill rate by posting period



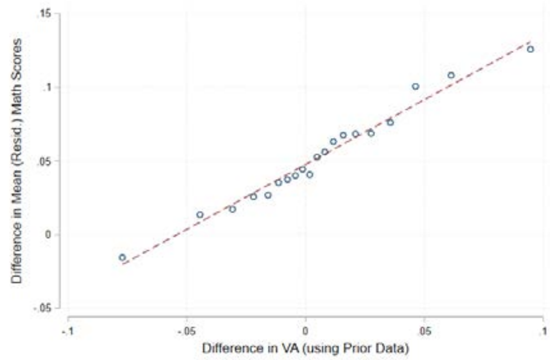
The figure shows the probability a vacancy leads to a hire ("Fill rate"), separated by month (e.g., April = 4, May = 5, etc.) in which the vacancy was posted. The darker columns are for non-Title I schools and the lighter columns are for Title I schools. The sample consists of vacancies in our focal district in the 2011-2016 application cycles.

Figure A4: Math Value-Added Forecast Unbiasedness



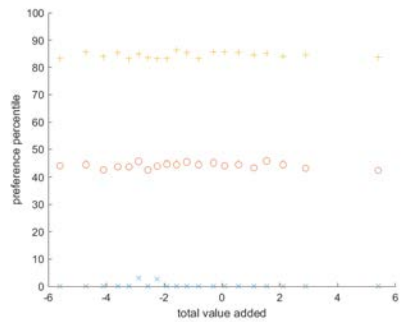
The figure is a binscatter, where an observation is a teacher-year and math value-added estimates are predictions using data from prior years. Units are student standard deviations. The y-axis is the mean student math test score, residualized by student demographics including lagged scores, school fixed effects, and teacher experience measures. The mean is taken over all students for a given teacher-year.

Figure A5: Math Comparative Advantage Forecast Unbiasedness



The figure is a binscatter, where an observation is a teacher-year and "Difference in VA" is the difference in a teacher's math value-added between economically disadvantaged and advantaged students. Value-added estimates are predictions using data from prior years. Units are student standard deviations. The y-axis is the difference in mean student math test score, residualized by student demographics including lagged scores, school fixed effects, and teacher experience measures. The mean is taken over all students (of a given type) for a given teacher-year and the difference is between a teacher's economically disadvantaged and advantaged students.

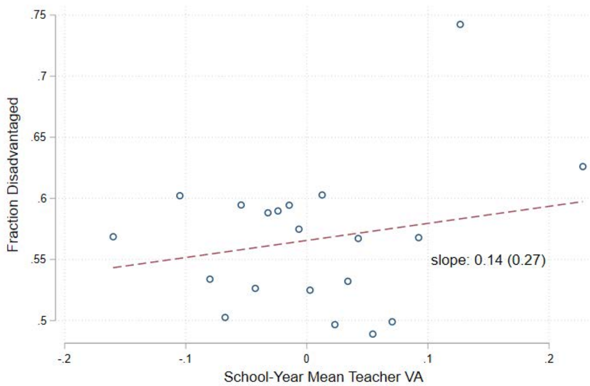
Figure A6: Bivariate preference relationship – principal model without output



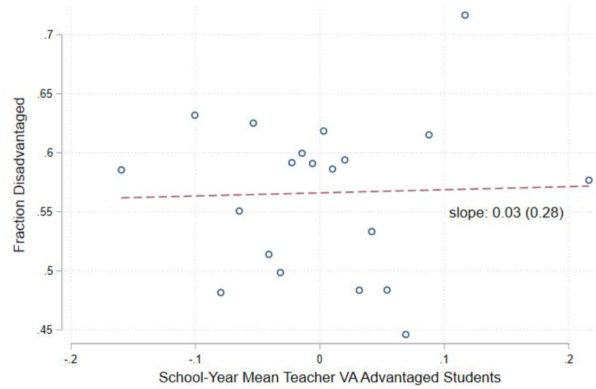
This figure shows a binscatter of the bivariate relationships between teacher output and principal preferences. We estimate each principal's ranking over teachers and order teachers from a principal's most preferred (100) to least preferred (0). The estimated model does not include value-added as a characteristic. The figure shows the bivariate relationship between the teacher's total value-added in the position and the mean preference percentile of the principal for the teacher in the principal preference model. The middle set of points (red circle) is the mean percentile, while the top (orange cross) and bottom (blue x) sets of points are the 10th and 90th percentiles, respectively.

Figure A7: Economic disadvantage by teacher value-added in a school-year

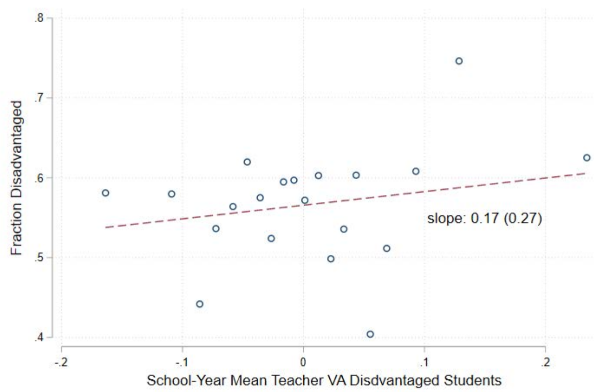
(a) Average VA



(b) VA for economically advantaged



(c) VA for economically disadvantaged



These figures show binscatters relating how a school's mean teacher value-added varies with the fraction of students at the school that are disadvantaged. An observation is a school-year. The x-axis is mean teacher value-added (panel a), mean teacher value-added for advantaged students (panel b), or mean teacher value-added for disadvantaged students (panel c). Binscatters weight observations by the number of students for each school-year. We list estimated slopes from an OLS regression with a constant, with standard errors clustered by school.

Table A1: Applications: sample and summary statistics

	Full Sample	Elementary Sample	Value-Added Sample	2016 Sample
Applications				
N	2,163,711	337,754	13,797	2,702
On-Cycle	0.68	1.00	1.00	1.00
Instructional	0.70	1.00	1.00	1.00
Elementary	0.39	1.00	1.00	1.00
Applicants				
N	104,795	14,864	866	178
Female		0.92	0.87	0.89
Black		0.24	0.30	0.25
Hispanic		0.03	0.01	0.03
In-District		0.12	0.43	0.44
Choice Set Size		159.10	151.19	151.38
Application Rate		0.18	0.11	0.10
Transferred		0.23	0.43	0.50
Mean Commute Time		14.52	18.46	18.32
Experience		5.81	9.23	9.89
VA Econ Adv		-0.03	-0.03	-0.03
VA Econ Disadv		-0.02	-0.02	-0.03
Abs Adv		-0.03	-0.03	-0.03
Comp Adv in Econ Disadv		0.01	0.01	0.00
Positions				
N	38,921	1,824	1,784	296
Choice Set Size		1,293.54	71.83	88.65
Application Rate		0.14	0.11	0.10
Mean Class Size		26.41	26.41	25.70
Frac Econ Disadv		0.65	0.65	0.68
Frac Black		0.43	0.43	0.45
Frac Hispanic		0.24	0.24	0.25

The table shows count or mean statistics across different samples. The “Full Sample” includes all of the raw data, the “Elementary Sample” restricts to on-cycle elementary school instructional positions without specialization, the “Value-Added Sample” further restricts to teachers with value-added forecasts based on prior years, and the “2016 Sample” further restricts to the 2015 application cycle for positions in the 2016 school year. We use the “Elementary Sample” for estimating principal preferences, the “Value-Added Sample” for estimating teacher preferences, and the “2016 Sample” for estimating counterfactual allocations. We do not include mean statistics for applicants and positions for the complete sample because we built the data on the subsample. Commute time is measured in minutes, absolute advantage is value-added at the representative school in the district, and choice set size is the number of positions in a teacher’s choice set (Applicants panel) or the number of teachers with the position in their choice set (Positions panel).

Table A2: Match rate to administrative data

	Obs	Match Rate
Hired Teachers, 2011	573	0.97
Hired Teachers, 2012	372	0.98
Hired Teachers, 2013	550	0.94
Hired Teachers, 2014	495	0.99
Hired Teachers, 2015	535	1.00
Hired Teachers, 2016	638	1.00

The table shows the match rate of our labor market (applications and vacancies) data to administrative data. We show the fraction of teachers hired in each application cycle in our focal district that match to the administrative data. Hired teachers are designated in the labor market data. The administrative data consists of teacher records for public school teachers who have worked in the state of North Carolina. A non-match would indicate a teacher whose application was recorded as leading to a hire but who does not have a teacher record in the administrative data.

Table A3: Segregation by school and classroom

	Adj R2, Sch	Adj R2, Cls	Sch Peers — Type 0	Sch Peers — Type 1	Cls Peers — Type 0	Cls Peers — Type 1
Economic disadvantage	0.401	0.404	0.33	0.73	0.33	0.74
Non-white	0.333	0.337	0.45	0.78	0.45	0.78
Low lagged achievement	0.163	0.191	0.38	0.54	0.36	0.55

The table shows measures of segregation by school and classroom in our focal district. For each characteristic, we divide the students into binary types. Type 0 refers to students who do not have the characteristic in each row (i.e., not economically disadvantaged, white, high lagged achievement) and type 1 refers to students who have the characteristic. The first two columns report adjusted R-squareds from a regressions of a student's type on school fixed effects (column 1) and class fixed effects (column 2). Classes are specific to each school, so class fixed effects are nested within school fixed effects. The third column calculates the mean fraction of a student's school peers who are type 1, conditional on the student being type 0. For example, the top row's number corresponds to the average fraction of an economically advantaged student whose school peers are economically disadvantaged. The fourth column calculates the mean fraction of a student's school peers who are type 1, conditional on the student being type 1. The last two columns are parallel to columns 3 and 4 except a student's peers are defined as those in the same classroom as the student.

Table A4: Distribution of interviews

	Obs	Mean	10th	25th	50th	75th	90th	Std. dev.
Interviews per Position	1,824	1.9	0.0	1.0	1.0	2.0	4.0	2.93
Internal Interviews per Position	1,824	0.3	0.0	0.0	0.0	0.0	1.0	0.68
Interviews per Teacher	14,864	0.2	0.0	0.0	0.0	0.0	1.0	0.46
Interviews per Internal Teacher	1,746	0.3	0.0	0.0	0.0	1.0	1.0	0.51
Interviews per Teacher ($\geq$1)	3,223	1.1	1.0	1.0	1.0	1.0	1.0	0.29

The table shows statistics from the distribution of the interviews. The top two rows are the number of interviews (or internal interviews) per vacancy, where internal interviews are interviews of candidates already working in the focal district. The last three rows are the number of interviews per candidate. The last row conditions on teachers who receive at least one interview in the cycle. The sample consists of interviews in our focal district in the 2011-2016 application cycles.

Table A5: Time between posting and interview, by period

	Obs	Mean	10th	25th	50th	75th	90th	Std. dev.
April	893	26.9	0.0	3.0	14.0	41.0	77.0	31.32
May/June	1,608	13.2	0.0	1.0	6.0	17.0	39.0	18.87
July/August	952	4.8	0.0	0.0	1.0	6.0	13.0	12.99

The table shows statistics of the difference (in days) of when an interviewed candidate applied for a vacancy and when the vacancy was posted. The difference is weakly positive by definition. We split the sample separately by the months of the cycle when the vacancy was posted. The sample consists of applications sent to vacancies in our focal district in the 2011-2016 application cycles.

Table A6: Rates of applying to multiple school vacancies

	N	Mean	Unconditional
Applied to Second Vacancy, Given Applied to First	39,532	0.48	0.16
Applied to First Vacancy, Given Applied to Second	37,380	0.51	0.17

The table shows application patterns across vacancies in the same cycle that are posted by the same school. The sample consists of vacancies at schools in our focal district that posted exactly two vacancies in a given cycle. We include the 2011-2016 application cycles. "Applied to Second Vacancy, Given Applied to First" is the conditional probability a teacher applies to the later vacancy given she applied to the earlier vacancy, among teachers with both vacancies in their choice set (i.e., active while both vacancies are available). "Applied to First Vacancy, Given Applied to Second" is the conditional probability a teacher applies to the earlier vacancy given she applied to the later vacancy, among teachers with both vacancies in their choice set. The "Unconditional" column shows the probability a teacher applied to the later vacancy (not conditional on applying to the earlier one) for the top row and the probability a teacher applied to the earlier vacancy (not conditional on applying to the later one) for the bottom row. These unconditional application rates are calculated among teachers with both vacancies in their choice set.

Table A7: Application timing

(a) Wait times until applying

	Obs	Mean days	Median days	Share 0 days
Stock	196,779	3.6	0	0.72
Flow	146,382	2.1	0	0.75

(b) First day versus subsequent days

	Obs	Mean fraction of days	Mean fraction of applications	Mean days since posting
First day	14,864	0.61	0.65	23.47
Subsequent days	40,850	0.14	0.13	11.55

(c) Timing of first and last days

	Obs	April or before	May	June	July	August
First day (all teachers)	14,864	0.20	0.25	0.22	0.18	0.15
Last day (all teachers)	14,864	0.09	0.15	0.21	0.26	0.29
First day (transfers)	2,547	0.27	0.30	0.24	0.14	0.05
Last day (transfers)	2,547	0.10	0.17	0.25	0.29	0.19

The tables show statistics related to application timing. Panel (a) shows how long it took an applicant to apply to positions that were in "stock" (already posted) on the day the teacher first applied on the platform or in "flow" (posted after the day the teacher first applied on the platform). Panel (b) shows application statistics for the first day a teacher applied on the platform in a cycle versus subsequent days. "Mean days since posting" is the mean number of days a vacancy had been posted at the time the teacher applied. Panel (c) shows the (monthly) timing of when an applicant's first and last application days of the cycle occurred. "All teachers" includes all applicants while "transfers" includes just teachers who ended up in new schools.

Table A8: Time between posting and application

	Obs	Mean	10th	25th	50th	75th	90th	Std. dev.
All applications	343,161	15.4	0.0	2.0	7.0	20.0	43.0	20.81
Hired candidates	3,163	14.1	0.0	1.0	5.0	16.0	43.0	22.87
Interviewed candidates	3,462	14.4	0.0	1.0	5.0	17.0	43.0	23.05
Positive assessment candidates	6,349	19.5	0.0	0.0	7.0	30.0	61.0	27.07

The table shows statistics of the difference (in days) of when a vacancy receives applications and when the vacancy was posted. The difference is weakly positive by definition. "Positive Assessment" is an offer, interview, and/or a positive rating. The sample consists of applications sent to vacancies in our focal district in the 2011-2016 application cycles.

Table A9: Timing of posting, applying, and hiring

(a) Monthly shares by position

	Posting			***Applying***			***Hiring***		
	Vacs	Share	Share TI	Apps	Share	Share TI	Apps	Share	Share TI
April	295	16.24	0.62	24799	7.13	0.50	393	13.23	0.69
May	392	21.57	0.52	70248	20.21	0.50	585	19.70	0.63
June	502	27.63	0.52	108776	31.29	0.51	827	27.85	0.60
July	451	24.82	0.42	94171	27.09	0.50	755	25.42	0.50
August	167	9.19	0.46	44673	12.85	0.51	358	12.05	0.57
Total	1807	100		342667	100		2918	2918	

(b) Monthly shares by teacher value-added

	Has VA			***Above median VA***			***Top decile VA***		
	Apps	Share	Share TI	Apps	Share	Share TI	Apps	Share	Share TI
April	3050	6.23	0.44	1552	7.16	0.42	373	9.15	0.41
May	9662	19.75	0.44	4218	19.46	0.44	918	22.53	0.45
June	16832	34.40	0.46	8035	37.08	0.45	1396	34.26	0.47
July	13673	27.95	0.47	5600	25.84	0.46	944	23.17	0.46
August	5522	11.29	0.48	2189	10.10	0.47	434	10.65	0.52
Total	48739	100		21594	100		4065	100	

(c) Early vs. late posting times by school

	Posts in July		
Posts in April	No	Yes	Total
No	8	15	23
Yes	10	88	98
Total	18	103	121

This table shows the timing of posting, applying, and hiring during a cycle. Panel (a) shows the distribution of vacancy postings, applications, and hires by month, where hires correspond to the timing of the applicant who was hired to the position. For each type of action, we show the share that corresponds to Title I positions. Some of the vacancies produce multiple hires. In Panel (b) we show the distribution of applications by month, where we split the sample of applicants into those with a value-added forecast (i.e., had taught in tested grades and subjects in North Carolina prior to applying), those with above median value-added, and those in the top decile. Panel (c) shows the cross-tabulation of whether a school posts a vacancy in April and whether that school posts a vacancy in July (in the same cycle).

Table A10: Application evaluations, outcomes, and timing

(a) Outcomes at the application level

	Hired successfully	Hired but taught elsewhere	Hired but not in district	Declined offer	Interview	Positive	Middle	Negative	Withdrew	No comment
mean	0.00051	0.00003	0.00016	0.00006	0.00000	0.00064	0.00029	0.00037	0.00002	0.07367
count	2,291	125	747	292	7	2,887	1,300	1,655	74	333,780

(b) Outcomes at the position level

	Hired	Declined offer	Interview	Positive	Middle	Negative	Withdrew	No comment	Any Non-Hire Action
mean	0.799	0.117	0.001	0.101	0.023	0.075	0.037	0.985	0.179
count	1,457	213	2	184	42	136	67	1,797	327

(c) Timing relative to hired applicant

	Obs	Mean	10th	25th	50th	75th	90th	Std. dev.
All applications	292,410	2.2	-17.3	-4.9	1.0	8.0	23.0	19.32
No recorded evaluation	283,706	2.2	-17.0	-4.5	1.0	8.0	22.3	19.04
Recorded evaluation	8,704	1.3	-30.8	-11.8	0.0	11.7	38.2	26.97

This table shows the frequency and timing of application outcomes. The data record a single outcome per application; as an example, "Interview" implies not hired as otherwise the "Interview" outcome would be replaced by "Hired." The data record "Hired," which we split into "Hired successfully" for teachers who taught in the position's school the following year, "Hired but taught elsewhere" for teachers hired who taught in district but not at that position's school, and "Hired but not in district" for teachers hired who did not appear in the district the following year. "Positive," "Middle," and "Negative" reflect the authors' coding of different text categories. "No comment" includes applications without an updated status. Panel (a) shows frequencies at the application level and panel (b) shows frequencies at the position level for at least one outcome across all applications to that position (i.e., "Hired" indicates at least one application led to a hire). "Any Non-Hire Action" is a positive, middle, or negative assessment or an application withdrawal. In panel (c) we calculate the difference in timing (in days) between when an application was made and when the application that led to a hire was made. A value of 1 would indicate an application made 1 day after the one that led to a hire. In the last two rows, we split the sample into those with no notes ("No comment") and those with an outcome.

Table A11: Position characteristics by posting period

	April	May-June	July-August
Title I	0.62	0.54	0.40
Fraction of Students Economically Disadvantaged	0.71	0.66	0.59
Fraction of Students Non-White	0.78	0.74	0.69
Fraction of Students Below Median Lagged Math Scores	0.52	0.49	0.46
Number of Students per Teacher	24	25	26
Attendance Rate	0.96	0.97	0.97
School Exceeded Expectations	0.28	0.23	0.14
School Met Expectations	0.33	0.24	0.19
Mean Applications in Previous Hiring Cycle	239	264	276

The table shows mean school characteristics for vacancies posted during different periods of the application cycle. For example, "Title I" in April is the fraction of vacancies posted in April that are at Title I schools. "Exceeded Expectations" and "Met Expectations" are designations made by the district of a school's progress on various measures. The excluded category is "Did Not Meet Expectations." "Mean Applications in Previous Hiring Cycle" is the average number of applications the same school received for vacancies it posted in the previous application cycle. The sample consists of vacancies in our focal district in the 2011-2016 application cycles.

Table A12: Differences in position outcomes, for schools posting two vacancies

	First vacancy	Second vacancy	p-value on difference	Num obs
Applicants VA per Student	-0.05	-0.07	0.000	2,239
Applicants non-Missing VA	0.06	0.04	0.000	58,538
Postitive Asssessments VA per Student	-0.01	0.06	0.101	88
Postitive Asssessments non-Missing VA	0.17	0.07	0.000	1,041
Hired Teachers VA per Student	-0.01	0.04	0.403	55
Hired Teachers non-Missing VA	0.18	0.06	0.000	568

The table shows difference in outcomes across vacancies in the same cycle that are posted by the same school. The sample consists of vacancies at schools in our focal district that posted exactly two vacancies in a given cycle. "First Vacancy" refers to the vacancy posted earlier and "Second Vacancy" refers to the vacancy posted later. We include the 2011-2016 application cycles. Each number refers to a mean characteristic for the first or second vacancy or the p-value on the difference across first and second vacancies. Means and p-values come from regressions of the outcome on a constant and an indicator for being the second vacancy, with school-cycle fixed effects. Observations are weighted by the inverse number of observations in each school-cycle-first/second group. The rows include two outcomes: mean per-student value-added (among non-missing value-added applicants) or the fraction applicants with non-missing value-added. We split the rows further by the full set of teachers that applied ("Applicants"), those who received an offer, interview, and/or positive rating ("Positive Assessment"), and hired applicants.

Table A13: Forecast Unbiasedness Tests for Value-Added Predictions

	Mean Res	Mean Diff	Mean Res	Mean Res	Mean Res	Mean Res
VA (Heterog)	1.060			1.068		
	(0.00655)			(0.00686)		
VA Diff		0.870				
		(0.0225)				
Post Transfer			-0.00245	0.00587		
			(0.00367)	(0.00281)		
VA * Post Transfer				-0.0903		
				(0.0214)		
VA – below 10th (disadv)					0.998	
					(0.0226)	
VA – 10th-90th (disadv)					1.066	
					(0.00703)	
VA – above 90th (disadv)					1.068	
					(0.0230)	
VA – below 10th (size)						1.018
						(0.0226)
VA – 10th-90th (size)						1.073
						(0.00718)
VA – above 90th (size)						0.969
						(0.0190)
Constant	0.00864	0.0483	0.00752	0.00796	0.00863	0.00853
	(0.000834)	(0.00100)	(0.00174)	(0.000883)	(0.000834)	(0.000842)
Subject	Math	Math	Math	Math	Math	Math
Mean DV	0.00735	0.0528	0.00727	0.00735	0.00735	0.00735
Clusters	21519	21519	21840	21519	21519	21519
N	74560	74560	75467	74560	74560	74560

The table includes tests of whether a value-added estimate is forecast unbiased. In the first and third through sixth columns, the outcome ("Mean Res") is the mean student math test score, residualized by student demographics including lagged scores, school fixed effects, and teacher experience measures. The mean is taken over all students for a given teacher-year. In the second column, the outcome ("Mean Diff") is the difference in the mean residualized math scores between a teacher's economically disadvantaged and advantaged students. The "VA" measures allow for match effects ("Heterog"). The measures predict mean student residuals using data from all prior years a teacher taught. "VA Diff" is the difference in predicted value-added between a teacher's economically disadvantaged and advantaged students (i.e., the predicted comparative advantage). "Post Transfer" refers to years after a teacher switched schools. The interaction with "VA" multiplies the post-transfer indicator with the heterogeneous value-added measure. Column (5) splits the year t observations into bins as a function of the change in share of disadvantaged students relative to the data observed for the teacher before year t. The split is based on percentiles of the change. Column (6) splits the year t observations into bins as a function of the change in classroom size relative to the data observed for the teacher before year t. The split is based on percentiles of the change. Standard errors are clustered at the teacher level.

Table A14: Pseudo R-squareds for principal rating models

	Non-Title I	Title I
Demographics	0.009	0.004
Teacher Characteristics	0.030	0.019
Value Added	0.006	0.003
EVAAS	0.000	0.000
Demographics + Teacher Characteristics	0.038	0.024
Demographics + Value Added	0.017	0.006
Teacher Characteristics + Value Added	0.033	0.021
EVAAS + Value Added	0.007	0.003
Demographics + Teacher Characteristics + Value Added	0.040	0.026
Demographics + Teacher Characteristics + EVAAS + Value Added	0.041	0.026
Observations	25,834	31,792

The table shows pseudo R-squareds from logit models for whether a principal rates an application highly (a positive rating, an interview, or an offer). Each model includes position fixed effects. The pseudo R-squared is the percentage improvement in the likelihood relative to a model with only the fixed effects. Demographics are measures of the teacher's race and gender, interacted with the school's racial composition. Teacher characteristics are experience, licensing, certification, and Praxis scores. Value Added is our model's forecast of the teacher's causal effect on student test scores from the assignment. EVAAS is the measure of teacher performance that the state uses and released to teachers.

Table A15: Relationship between Teacher Characteristics and Teacher Value-Added

	VA Mean	VA Adv	VA Disadv
Experience 1-2	0.0797	0.0744	0.0816
	(0.0326)	(0.0315)	(0.0334)
Experience 3-5	0.134	0.123	0.138
	(0.0322)	(0.0312)	(0.0331)
Experience 6-12	0.139	0.126	0.144
	(0.0320)	(0.0310)	(0.0329)
Experience 13-20	0.137	0.125	0.142
	(0.0320)	(0.0310)	(0.0329)
Experience 21-27	0.149	0.138	0.155
	(0.0322)	(0.0312)	(0.0331)
Experience 28+	0.132	0.121	0.135
	(0.0324)	(0.0314)	(0.0333)
Graduate degree	0.00263	0.00442	0.000950
	(0.00364)	(0.00352)	(0.00373)
Regular license	0.0531	0.0443	0.0574
	(0.0183)	(0.0177)	(0.0188)
NBPTS certified	0.0303	0.0303	0.0307
	(0.00528)	(0.00511)	(0.00542)
Praxis	0.00414	0.00573	0.00323
	(0.00241)	(0.00233)	(0.00247)
Mean DV	-0.00366	-0.0130	0.000960
R squared	0.0228	0.0219	0.0232
N	7335	7335	7335

The table shows the relationship between teacher characteristics and value added across student types ("Adv" and "Disadv") or mean value added. The omitted experience category is having no experience

Table A16: Summary statistics for 2015-16, by race

	Focal, Non-white	Focal, White	Other, Non-white	Other, White
Students				
White (%)	0.00	100.00	0.00	100.00
Black (%)	55.43	0.00	48.14	0.00
Hispanic (%)	32.92	0.00	34.48	0.00
Student performance (level scores)				
Math	-0.08	0.70	-0.29	0.25
Student performance (gain scores)				
Math	0.06	0.07	0.01	-0.02
Teachers				
Experience (% of teachers)				
0 years	10.16	4.58	4.97	3.56
1-2 years	16.39	10.66	9.85	7.51
3-5 years	18.48	18.72	12.84	11.57
6-12 years	23.57	29.79	26.69	26.16
13 or more years	31.40	36.26	45.65	51.20
Graduate degree (%)	43.67	44.91	37.98	38.67
Regular license (%)	86.46	94.96	92.74	95.82
NBPTS certified (%)	8.13	15.37	9.15	14.04
Praxis score	0.07	0.36	0.12	0.25
Age	37	39	40	41
Mean math value-added				
Baseline, white	0.04	0.03	0.01	0.00
Baseline, non-white	0.05	0.03	0.02	0.01
Homogeneous	0.01	0.02	-0.01	-0.00
CFR	0.07	0.09	0.00	0.01
Using school means	0.13	0.15	0.07	0.08
Imputed, econ disadv	0.03	0.02	0.02	0.01
Imputed, econ adv	0.00	0.01	0.01	0.00
Fraction imputed	0.31	0.15	0.19	0.14
Residual	0.02	0.05	0.01	-0.03
Mean behavioral value-added				
Baseline	-0.00	-0.01		
Sample size				
Number of students	24,930	10,027	157,913	162,018

The table shows mean student and teacher characteristics in our sample for the 2015-16 school year. We split the sample into whether the student is in our focal district ("Focal") or in the rest of North Carolina ("Other") and whether he or she is Non-white or White. Math scores are standardized to have mean 0 and standard deviation 1 at the state-grade-year level. The "Baseline" value-added estimates incorporate comparative advantage according to student's race while the other estimates have no comparative advantage or comparative advantage according to economic disadvantage. The alternate VA estimators are (a) a homogeneous value-added model with constant effects across student types, (b) using the estimator from Chetty, Friedman, and Rockoff (2014a), and (c) a model that uses school mean characteristics rather than school fixed effects.Imputed value-added predicts value-added based on how observable teacher characteristics predict value-added. "Residual" is the unshrunken value-added for 2015-16, which has no missing values.

Table A17: Summary statistics for 2015-16, by lagged achievement

	Focal, Below	Focal, Above	Other, Below	Other, Above
Students				
White (%)	13.44	34.27	37.89	56.15
Black (%)	55.08	33.82	34.40	19.17
Hispanic (%)	26.94	22.21	20.24	15.63
Student performance (level scores)				
Math	-0.67	0.45	-0.74	0.30
Student performance (gain scores)				
Math	0.18	-0.03	0.12	-0.12
Teachers				
Experience (% of teachers)				
0 years	10.98	6.64	4.96	3.59
1-2 years	16.56	13.26	9.82	7.55
3-5 years	18.00	18.95	12.68	11.72
6-12 years	24.26	26.33	26.50	26.34
13 or more years	30.20	34.82	46.04	50.80
Graduate degree (%)	43.31	44.58	37.28	39.28
Regular license (%)	85.15	91.88	93.24	95.35
NBPTS certified (%)	7.25	12.57	9.54	13.64
Praxis score	0.03	0.26	0.13	0.24
Age	37	38	40	41
Mean math value-added				
Baseline, above	0.03	0.01	0.02	-0.00
Baseline, below	0.05	0.04	0.04	0.02
Homogeneous	0.00	0.02	-0.02	-0.00
CFR	0.06	0.09	-0.00	0.01
Using school means	0.12	0.14	0.06	0.08
Imputed, econ disadv	0.03	0.02	0.04	0.02
Imputed, econ adv	0.00	-0.00	0.01	-0.00
Fraction imputed	0.27	0.21	0.18	0.15
Residual	0.02	0.01	0.01	-0.04
Mean behavioral value-added				
Baseline	-0.00	-0.01		
Sample size				
Number of students	9,381	25,576	96,472	223,459

The table shows mean student and teacher characteristics in our sample for the 2015-16 school year. We split the sample into whether the student is in our focal district ("Focal") or in the rest of North Carolina ("Other") and whether the student's lagged math achievement is below or above the state median. Math scores are standardized to have mean 0 and standard deviation 1 at the state-grade-year level. The "Baseline" value-added estimates incorporate comparative advantage according to student's lagged math achievement while the other estimates have no comparative advantage or comparative advantage according to economic disadvantage. The alternate VA estimators are (a) a homogeneous value-added model with constant effects across student types, (b) using the estimator from Chetty, Friedman, and Rockoff (2014a), and (c) a model that uses school mean characteristics rather than school fixed effects. Imputed value-added predicts value-added based on how observable teacher characteristics predict value-added. "Residual" is the unshrunken value-added for 2015-16, which has no missing values.

Table A18: Summary statistics for 2015-16, by persistent economic disadvantage

	Focal, Adv	Focal, Disadv	Other, Adv	Other, Disadv
Students				
White (%)	59.13	8.73	73.49	35.16
Black (%)	20.97	51.69	10.49	32.76
Hispanic (%)	8.68	33.17	7.33	23.59
Student performance (level scores)				
Math	0.66	-0.20	0.41	-0.31
Student performance (gain scores)				
Math	0.07	0.06	-0.01	0.00
Teachers				
Experience (% of teachers)				
0 years	5.16	11.44	3.56	4.94
1-2 years	11.21	17.74	7.28	10.01
3-5 years	18.21	18.86	11.44	12.93
6-12 years	28.58	22.65	26.55	26.26
13 or more years	36.84	29.32	51.17	45.86
Graduate degree (%)	45.13	43.06	39.26	37.38
Regular license (%)	93.61	84.94	95.46	93.21
NBPTS certified (%)	14.62	6.47	13.65	9.71
Praxis score	0.30	0.03	0.26	0.11
Age	39	37	41	40
Mean math value-added				
Baseline, econ disadv	0.02	0.02	-0.01	-0.01
Baseline, econ adv	0.01	0.01	-0.01	-0.01
Homogeneous	0.02	0.01	-0.00	-0.01
CFR	0.08	0.07	0.01	0.01
Using school means	0.15	0.13	0.07	0.08
Imputed, econ disadv	-0.03	-0.01	-0.01	-0.01
Imputed, econ adv	0.01	0.01	-0.01	-0.01
Fraction imputed	0.19	0.43	0.15	0.23
Residual	0.02	0.02	-0.04	0.00
Mean behavioral value-added				
Baseline	-0.01	0.01		
Sample size				
Number of students	13,840	19,448	129,217	163,484

The table shows mean student and teacher characteristics in our sample for the 2015-16 school year. We split the sample into whether the student is in our focal district ("Focal") or in the rest of North Carolina ("Other") and whether he or she is persistently economically disadvantaged ("Disadv") or not ("Adv"). Persistent disadvantaged students are those who are economically disadvantaged for all years in our sample (Michelmore and Dynarski, 2017). Math scores are standardized to have mean 0 and standard deviation 1 at the state-grade-year level. All value-added estimates either incorporate no comparative advantage or comparative advantage according to economic disadvantage. The alternate VA estimators are (a) a homogeneous value-added model with constant effects across student types, (b) using the estimator from Chetty, Friedman, and Rockoff (2014a), and (c) a model that uses school mean characteristics rather than school fixed effects. Imputed value-added predicts value-added based on how observable teacher characteristics predict value-added. "Residual" is the unshrunken value-added for 2015-16, which has no missing values.

Table A19: Persistent economic disadvantage

	Fraction of current disadvantaged students always disadvantaged
Students appearing 1 times	1.00
Students appearing 2 times	0.93
Students appearing 3 times	0.86
Students appearing 4 times	0.81
Students appearing at least 5 times	0.79

The table shows what fraction of disadvantaged students in 2015-16 are persistently economically disadvantaged (Michelmore and Dynarski, 2017). For each student, we calculate the number of years she appears in our sample and the number of years is classified as economically disadvantaged. The rows condition on the number of years a student appears in the sample. For each row, we condition on students that are economically disadvantaged in 2015-16 and calculate the fraction that are economically disadvantaged every year they are in our sample.

Table A20: Summary statistics for 2015-16, by school's level of economic disadvantage

	Focal, Adv	Focal, Disadv	Other, Adv	Other, Disadv
Students				
White (%)	52.09	6.10	63.39	30.11
Black (%)	25.12	53.44	15.08	37.74
Hispanic (%)	12.58	34.00	12.86	23.72
Student performance (level scores)				
Math	0.49	-0.18	0.17	-0.31
Student performance (gain scores)				
Math	0.06	0.07	-0.01	0.01
Teachers				
Experience (% of teachers)				
0 years	4.50	12.14	3.71	5.86
1-2 years	9.89	18.66	7.73	11.54
3-5 years	16.73	20.47	11.75	14.03
6-12 years	30.58	20.47	27.62	26.77
13 or more years	38.31	28.26	49.18	41.79
Graduate degree (%)	47.03	41.89	38.43	37.81
Regular license (%)	93.17	84.06	95.44	91.48
NBPTS certified (%)	15.11	5.43	13.59	7.59
Praxis score	0.32	0.02	0.28	0.05
Age	40	36	41	40
Mean math value-added				
Baseline, econ disadv	0.01	0.01	-0.01	-0.01
Baseline, econ adv	0.01	0.02	-0.01	-0.00
Homogeneous	0.01	0.01	-0.01	-0.02
CFR	0.08	0.08	0.01	0.02
Using school means	0.16	0.16	0.10	0.12
Imputed, econ disadv	-0.03	-0.00	-0.01	-0.00
Imputed, econ adv	-0.00	0.01	-0.01	-0.01
Fraction imputed	0.21	0.53	0.18	0.35
Residual	0.01	0.07	-0.05	0.02
Mean behavioral value-added				
Baseline	-0.01	0.01		
Sample size				
Number of students	12,329	22,628	122,903	197,028

The table shows mean student and teacher characteristics in our sample for the 2015-16 school year. We split the sample into whether the school is in our focal district ("Focal") or in the rest of North Carolina ("Other") and school has fewer than 70% of students as economically disadvantaged ("Adv") or more than 70% ("Disadv"). Math scores are standardized to have mean 0 and standard deviation 1 at the state-grade-year level. The alternate VA estimators are (a) a homogeneous value-added model with constant effects across student types, (b) using the estimator from Chetty, Friedman, and Rockoff (2014a), and (c) a model that uses school mean characteristics rather than school fixed effects. Imputed value-added predicts value-added based on how observable teacher characteristics predict value-added. "Residual" is the unshrunken value-added for 2015-16, which has no missing values.

Table A21: Summary statistics for 2015-16, by teachers at the school level

	Focal, Adv	Focal, Disadv	Other, Adv	Other, Disadv
Students				
White (%)	64.61	9.11	75.58	35.09
Black (%)	17.04	51.78	9.54	32.63
Hispanic (%)	6.77	32.58	6.00	23.90
Student performance (level scores)				
Math	0.70	-0.16	0.43	-0.30
Student performance (gain scores)				
Math	0.07	0.07	-0.01	0.00
Teachers				
Experience (% of teachers)				
0 years	4.78	11.08	3.65	4.87
1-2 years	11.02	17.02	7.22	9.70
3-5 years	16.93	18.44	11.31	12.80
6-12 years	29.42	22.69	27.14	26.31
13 or more years	37.85	30.76	50.69	46.32
Graduate degree (%)	44.78	43.28	39.43	37.79
Regular license (%)	93.69	85.36	95.03	93.24
NBPTS certified (%)	14.80	7.44	13.36	9.87
Praxis score	0.34	0.05	0.29	0.13
Age	39	37	41	41
Mean math value-added				
Baseline, econ disadv	0.02	0.01	-0.01	-0.01
Baseline, econ adv	0.01	0.01	-0.01	-0.01
Homogeneous	0.02	0.01	-0.01	-0.01
CFR	0.09	0.07	0.00	0.01
Using school means	0.15	0.13	0.07	0.08
Imputed, econ disadv	-0.03	-0.01	-0.01	-0.01
Imputed, econ adv	0.00	0.01	-0.01	-0.01
Fraction imputed	0.19	0.44	0.17	0.25
Residual	0.02	0.04	-0.04	0.01
Mean behavioral value-added				
Baseline	-0.01	0.01		
Sample size				
Number of students	12,329	22,628	122,903	197,028

The table shows mean student and teacher characteristics in our sample for the 2015-16 school year. Teacher characteristics and value-added are means based on the teachers in the student's school, not necessarily in the same classroom. We split the sample into whether the school is in our focal district ("Focal") or in the rest of North Carolina ("Other") and school has fewer than 70% of students as economically disadvantaged ("Adv") or more than 70% ("Disadv"). Math scores are standardized to have mean 0 and standard deviation 1 at the state-grade-year level. The alternate VA estimators are (a) a homogeneous value-added model with constant effects across student types, (b) using the estimator from Chetty, Friedman, and Rockoff (2014a), and (c) a model that uses school mean characteristics rather than school fixed effects. Imputed value-added predicts value-added based on how observable teacher characteristics predict value-added. "Residual" is the unshrunken value-added for 2015-16, which has no missing values.

Table A22: Potential Gains from Reassignment

	Adv	Disadv	Mean
Targeting Disadvantaged Students			
Max Disadvantaged VA	-0.050	0.075	0.016

The table shows the potential gains from reassignments of teachers to different schools. The sample is all teachers with non-missing value-added forecasts in 2016 (based on prior data), along with their corresponding 2016 assignments. Gains come from matching better teachers to disadvantaged students. Gains are measured in student standard deviations (σ). The first and second columns show the per-student gains, relative to the actual allocation, for non-disadvantaged and disadvantaged students while the third column shows the per-student gains across all students.

Table A23: Reflecting teacher preferences

Description	VA disadv.	VA adv.	Dis. share among matched	Dis share among unmatched	Share matched
Teacher most preferred	-0.025	-0.037	0.509	0.786	0.470
Serial dictatorship (VA Dis.)	-0.061	-0.004			

This table shows allocations that reflect teacher preferences under two market clearing protocols. In the first row, we assign teachers to their most preferred school according to our estimated preferences. As a result, multiple teachers can be assigned to each position. In the second row, we clear the market by a serial dictatorship where we order teachers by their value added with disadvantaged students (starting with the best teachers). The first two columns show the average value-added of the teachers assigned to disadvantaged and advantaged students. Note that in the first row this average is taken over all of the teachers assigned to the student. The third and fourth columns show the share of disadvantaged students among those who are matched with a teacher (third column) and those who are not matched with a teacher (fourth column). The fifth column shows the overall share of students who are matched to a teacher.

Table A24: Difference between value-added among all applicants and applicants with specific outcomes

	Mean	Mean, non-TI	Mean, TI	non-TI vs TI p value	25th Perc.	50th Perc.	75th Perc.
Applicants vs. Positive Assessment	0.13	0.12	0.13	0.22	0.01	0.10	0.20
Applicants vs. Interview	0.14	0.13	0.14	0.21	0.01	0.12	0.21
Applicants vs. Offer	0.13	0.13	0.14	0.26	0.00	0.12	0.21

The table shows statistics of the difference in the highest value-added among all applicants to a vacancy and the highest value-added among applicants who received a specific outcome. "Positive Assessment" is an offer, interview, and/or a positive rating. For instance, "Applicants vs. Positive Assessment" is the difference between the highest value-added among applicants to a vacancy and the highest value-added among applicants who received a positive assessment from the hiring principal of that vacancy. All statistics are weakly positive by definition. "TI" and "non-TI" refer to Title I and non-Title I schools, respectively, and "p value" is the p-value from a t-test on the difference in means across Title I and non-Title I schools. The sample consists of applications sent to vacancies in our focal district in the 2011-2016 application cycles.

Table A25: Including imputed teachers: allocations and market balance

Description	VA disadv.	VA adv.	Teacher-weighted VA by student type			
			Imputed		Nonimputed	
			Disadv.	Adv.	Disadv.	Adv.
A. Balanced market						
Status quo	-0.003	-0.035	0.018	-0.030	-0.030	-0.034
All options (timing)	-0.005	-0.033	0.017	-0.030	-0.030	-0.034
Teachers N Dis	-0.018	-0.034	0.018	-0.030	-0.030	-0.034
Principals VA	-0.015	-0.016	0.017	-0.030	-0.030	-0.034
Previous two	0.008	-0.036	0.018	-0.030	-0.030	-0.034
First best	0.042	-0.078	0.018	-0.030	-0.030	-0.034
B. Unbalanced market: schools short						
Status quo	0.010	-0.027	0.043	-0.005	-0.030	-0.034
All options (timing)	0.014	-0.025	0.049	0.001	-0.030	-0.034
Teachers N Dis	-0.012	-0.005	0.052	0.006	-0.030	-0.034
Principals VA	0.018	-0.002	0.077	0.029	-0.030	-0.034
Previous two	0.018	0.016	0.073	0.034	-0.030	-0.034
First best	0.068	-0.069	0.078	0.024	-0.030	-0.034

This table shows the effect of including imputed teachers (teachers for whom we cannot compute value-added) on the allocation. In this table we include all 296 positions. First, we run deferred acceptance with all 296 positions and the 178 nonimputed teachers. Second, we clear the market with the unmatched positions and the imputed teachers. In Panel A, we randomly select a subset of the imputed teachers so that the market is balanced. In Panel B, we include all imputed teachers so that schools are short. Columns 1 and 2 show the student weighted value added for disadvantaged and advantaged students. Columns 3 through 6 show teacher weighted value-added for the matched teachers with disadvantaged and advantaged students. The status quo uses teacher and principal estimated preferences and restricted choice sets, and solves for the teacher proposing stable allocation. All options relaxes the timing constraint in the status quo. Teachers rank by N disadvantage changes the teacher preferences in the status quo. Principals rank by VA changes the principal preferences in the status quo. Previous two changes takes the status quo and replaces the teacher preferences with teachers ranking on the number of disadvantaged students and principals ranking by value added. The first-best is the allocation where the planner maximizes the achievement of disadvantaged students. We report averages over 400 simulations.

Table A26: Outcomes by vacancy posting period

Description	VA disadv.	VA adv.	VA disadv.: By period			VA adv.: By period		
A. Clearing the market in one period								
			1	2	3	1	2	3
Status quo	-0.027	-0.039	-0.008	-0.018	-0.066	-0.019	-0.028	-0.073
Title I	-0.025	-0.032	-0.005	-0.018	-0.066	-0.015	-0.020	-0.067
Non Title I	-0.032	-0.041	-0.020	-0.020	-0.067	-0.020	-0.029	-0.074
All options (timing)	-0.029	-0.034	-0.030	-0.032	-0.026	-0.033	-0.035	-0.033
B. Clearing the market in three periods								
			1	2	3	1	2	3
Status quo	-0.027	-0.041	-0.007	-0.024	-0.062	-0.023	-0.031	-0.069
C. Share of vacancy posting and teacher entry by sub-period								
	1	2	3					
Vacancy posting	0.35	0.38	0.27					
Title I	0.43	0.31	0.26					
Non-Title I	0.24	0.48	0.28					
First application	0.53	0.32	0.15					

This table shows the effect of timing. In Panel A we implement our baseline timing and clear the market in one period. In Panel B we clear the market in three periods: April, May and June, and July and August. Panel C shows the timing of initial vacancy posting and teacher applications, where we split the vacancy posting between Title I (high-poverty) and non-Title I schools. For Panels A and B, columns 1 and 2 show the student weighted value added for disadvantaged and advantaged students. The remaining columns show these outcomes by subperiod of when the vacancy was posted. The status quo uses teacher and principal estimated preferences and restricted choice sets, and solves for the teacher proposing stable allocation. In Panel A we split the status quo between Title I (high-poverty) and non-Title I schools. In Panel A, we show all options, which relaxes the timing constraint in the status quo. We report averages over 400 simulations.

Table A27: Rank in teachers preferences by entry date (lower is better)

Description	Overall	By period		
Clearing the market in one period				
		1	2	3
Status quo	15.4	18.4	14.6	6.7
All options (timing)	20.5	19.1	21.2	23.8

This table shows the rank of the position to which the teacher is matched in the teacher's preferences. Lower is better so a school ranked 1 is the teachers most preferred school. The overall column shows the average outcome, while the by period columns show how the teacher's outcomes vary by subperiod in which they initially apply where 1 is April, 2 is May and June, and 3 is July and August. The status quo uses teacher and principal estimated preferences and restricted choice sets, and solves for the teacher proposing stable allocation. The all options row relaxes the timing constraints. We report averages over 400 simulations.

Table A28: Transferring and non-transferring teachers' value added

	Did not transfer			Applied to transfer		
	mean	sd	count	mean	sd	count
Comparative advantage	0.0023	0.0191	530	0.0024	0.0170	507
Absolute advantage	0.0053	0.1401	530	0.0058	0.1248	507

The table shows the means and standard deviations of absolute and comparative advantage for teaching economically advantaged students by whether the teacher ever submits an application to transfer. An observation is a teacher with a value-added forecast. These are pooled over years 2010 through 2018.

Table A29: Multi-classroom teacher prevalence

Year	All	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
2012	0.264	0.109	0.187	0.618	0.621	0.631
2013	0.287	0.124	0.210	0.636	0.631	0.649
2014	0.300	0.152	0.227	0.633	0.625	0.644
2015	0.363	0.256	0.345	0.615	0.598	0.602
2016	0.391	0.305	0.392	0.595	0.591	0.595
2017	0.385	0.291	0.399	0.612	0.569	0.596
2018	0.393	0.307	0.425	0.596	0.586	0.578
Estimation sample	0.417					

The table shows the prevalence of teachers having multiple classrooms, separately by teacher's grade and year. The sample includes teachers for whom we can calculate math value-added. Our estimation sample consists of teachers, with value-added forecasts, who applied to elementary school positions.

Table A30: Teacher Value-Added Structural Parameters

	Estimates	Standard Errors	95% CI Lower Bound	95% CI Upper Bound
$\sigma_{\varepsilon 0}$	0.450	0.000	0.456	0.457
$\sigma_{\varepsilon 1}$	0.470	0.000	0.477	0.479
$\sigma_{\theta 0}$	0.110	0.007	0.108	0.137
$\sigma_{\theta 1}$	0.088	0.015	0.089	0.143
correlation$(\theta_{c0t}, \theta_{c1t})$	0.657	0.162	0.126	0.844
$\sigma_{\mu 0}$	0.249	0.007	0.262	0.284
$\sigma_{\mu 1}$	0.243	0.015	0.254	0.316
correlation(μ_{j0t}, μ_{j1t})	0.859	0.035	0.729	0.872

	Race	Achievement
$\sigma_{\varepsilon 0}$	0.465	0.481
$\sigma_{\varepsilon 1}$	0.457	0.439
$\sigma_{\theta 0}$	0.091	0.099
$\sigma_{\theta 1}$	0.110	0.102
correlation$(\theta_{c0t}, \theta_{c1t})$	0.637	0.628
$\sigma_{\mu 0}$	0.233	0.240
$\sigma_{\mu 1}$	0.261	0.282
correlation(μ_{j0t}, μ_{j1t})	0.900	0.844

	1	2	3	4	5	6	7+
Estimate	0.056	0.077	0.083	0.088	0.088	0.091	0.070
Standard Error	0.004	0.004	0.005	0.005	0.005	0.005	0.005

In Panel A, the table shows the estimates of a subset of the structural parameters of the production model – specifically the parameters corresponding to contemporaneous output. Non-disadvantaged students have index 1 while disadvantaged students have index 2. ε is the student-year idiosyncratic component, θ captures classroom effects, and μ describes a teacher's value-added. The remaining structural parameters describe the drift process of teacher value-added over time. Standard errors and confidence intervals are estimated with 100 bootstrap iterations. In Panel B, The table shows the estimates of a subset of the structural parameters of production models with alternate forms of heterogeneous teacher effects – specifically by race and prior achievement. In the first column, non-white students have index 1 while White students have index 2. In the second column, students with below median prior math achievement have index 1 while students with above median prior math achievement have index 2. ε is the student-year idiosyncratic component, θ captures classroom effects, and μ describes a teacher's value-added. The remaining structural parameters describe the drift process of teacher value-added over time. The table shows the estimated experience returns for math test scores, where the scores have been normalized to have mean 0 and standard deviation 1 for students in a given grade-year. Columns designate the number of prior years of experience. The omitted category is teachers with no prior experience.

Table A31: Predicting Student Residuals by Teacher Type

	Student res	Student res
Share disadvantaged	-0.0549	-0.0409
	(0.0251)	(0.0202)
Share disadvantaged x CA in disadvantaged	0.0820	0.0697
	(0.0356)	(0.0283)
Num teachers	3214	3214
Num students	157671	157671
Mean CA	-0.00805	-0.00805
SD CA	0.0624	0.0624
Controls	No	Yes

The table assesses whether changes in the share of economically disadvantaged students predict changes in student test score residuals differently by teacher comparative advantage in pre-transfer schools. The outcome is changes in average teacher-by-school student residuals across transfers. "Share disadvantaged" is the change in the average share of economically disadvantaged students teacher j taught when moving from one school to another. Controls include a cubic in average experience in the school, an indicator for experience missingness, and transfer year indicators. Standard errors are clustered at the teacher level.

Table A32: Imputing value-added for new teachers

	VA Advantaged	VA Disadvantaged	CA
Praxis	0.0295	0.0171	-0.00889
	(0.00447)	(0.00398)	(0.00431)
Praxis missing	-0.000658	0.0150	0.0101
	(0.00823)	(0.00700)	(0.00795)
Graduate degree	0.00811	0.0122	0.00541
	(0.00645)	(0.00574)	(0.00624)
Graduate degree missing	0.0464	-0.0847	-0.0747
	(0.0698)	(0.0534)	(0.0665)
NBPTS certified	0.0107	0.0326	0.0245
	(0.0173)	(0.0160)	(0.0171)
NBPTS certified missing	-0.0450	-0.0217	0.0141
	(0.00961)	(0.00930)	(0.00940)
Regular license	-0.00785	0.0403	0.0549
	(0.0365)	(0.0277)	(0.0348)
Regular license missing	-0.0173	0.0391	0.0350
	(0.0402)	(0.0306)	(0.0383)
Constant	0.0346	0.00821	-0.0368
	(0.0378)	(0.0292)	(0.0360)
Mean DV	-0.00122	0.0420	0.0328
R squared	0.00894	0.00449	0.00166
N	10102	12454	9801

The table shows estimated coefficients from our model of value-added for new teachers. The sample consists of teacher-years from 2008-2018 without a value-added score. These teachers are largely novice teachers or those for whom prior years of teaching came without classes that mixed advantaged and disadvantaged students. “VA Advantaged” is the unshrunken teacher output for advantaged students from a given year and similarly for “VA Disadvantaged.” Both measures exist, even for teachers without shrunken value-added scores. “CA” is a teacher’s comparative advantage with disadvantaged students. The sample is larger for “VA Disadvantaged” because more teachers have only disadvantaged students than only advantaged students. “Praxis” is a score from a teacher test standardized to have mean 0 and standard deviation 1. Graduate degree, NBPTS certification, and regular license are indicator variables, filled in to be 0 when missing.

Table A33: Advantaged VA: Sources of imputed value-added differences

	β VA New	β VA Experienced	Mean New	Mean Experienced	Diff. from X	Diff. from β
Constant	0.03	0.04	1.00	1.00	0.00	-0.01
Praxis	0.02	-0.01	0.13	0.08	0.00	0.00
Praxis missing	0.01	-0.06	0.21	0.21	-0.00	0.02
Graduate degree	0.01	0.01	0.49	0.55	-0.00	0.00
Graduate degree missing	-0.08	-0.08	0.00	0.00	-0.00	0.00
NBPTS certified	0.03	0.04	0.04	0.07	-0.00	-0.00
NBPTS certified missing	-0.02	-0.02	0.85	0.93	0.00	0.00
Regular license	0.04	0.04	0.94	1.00	-0.00	0.00
Regular license missing	0.04	0.04	0.05	0.00	0.00	0.00
Total					0.00	0.01

The table decomposes the sources of variation in imputed value-added (VA) for advantaged students. The first column shows the estimated regression coefficients from the model we use for imputation – i.e., estimated on teachers without VA scores ("New"). The second column is for comparison and shows the estimated regression coefficients from the same model, but estimated on teachers with VA scores ("Experienced"). The third and fourth columns show the mean of the characteristic in each row for the teacher sample, separately for "New" and "Experienced" teachers, respectively. The last two columns show the components of a Oaxaca-Blinder Decomposition explaining the differences in unshrunken value-added across "New" and "Experienced" teachers. The fifth column is the difference in mean characteristics, multiplied by the coefficients from the first column. The last column is the difference in coefficients multiplied by the mean characteristic for "Experienced" teachers. The "Total" row shows the sum of the decomposition components.

Table A34: Disadvantaged VA: Sources of imputed value-added differences

	β VA New	β VA Experienced	Mean New	Mean Experienced	Diff. from X	Diff. from β
Constant	0.03	0.04	1.00	1.00	0.00	-0.10
Praxis	-0.01	0.05	0.16	0.15	-0.00	-0.01
Praxis missing	0.01	0.00	0.19	0.21	-0.00	0.00
Graduate degree	0.01	-0.07	0.51	0.56	-0.00	0.04
Graduate degree missing	-0.07	-0.07	0.00	0.00	-0.00	0.00
NBPTS certified	0.02	-0.01	0.04	0.09	-0.00	0.00
NBPTS certified missing	0.01	0.01	0.83	0.91	-0.00	0.00
Regular license	0.06	0.06	0.96	0.99	-0.00	0.00
Regular license missing	0.04	0.04	0.04	0.00	0.00	0.00
Total					-0.00	-0.06

The table decomposes the sources of variation in imputed value-added (VA) for disadvantaged students. The first column shows the estimated regression coefficients from the model we use for imputation – i.e., estimated on teachers without VA scores ("New"). The second column is for comparison and shows the estimated regression coefficients from the same model, but estimated on teachers with VA scores ("Experienced"). The third and fourth columns show the mean of the characteristic in each row for the teacher sample, separately for "New" and "Experienced" teachers, respectively. The last two columns show the components of a Oaxaca-Blinder Decomposition explaining the differences in unshrunken value-added across "New" and "Experienced" teachers. The fifth column is the difference in mean characteristics, multiplied by the coefficients from the first column. The last column is the difference in coefficients multiplied by the mean characteristic for "Experienced" teachers. The "Total" row shows the sum of the decomposition components.

Table A35: Comparative Advantage: Sources of imputed value-added differences

	β VA New	β VA Experienced	Mean New	Mean Experienced	Diff. from X	Diff. from β
Constant	0.03	0.04	1.00	1.00	0.00	-0.01
Praxis	0.03	-0.08	0.17	0.16	0.00	0.02
Praxis missing	-0.00	-0.05	0.18	0.21	0.00	0.01
Graduate degree	0.01	0.05	0.51	0.56	-0.00	-0.03
Graduate degree missing	0.05	0.05	0.00	0.00	0.00	0.00
NBPTS certified	0.01	0.01	0.05	0.10	-0.00	0.00
NBPTS certified missing	-0.04	-0.13	0.83	0.90	0.00	0.08
Regular license	-0.01	-0.01	0.96	0.99	0.00	0.00
Regular license missing	-0.02	-0.02	0.04	0.00	-0.00	0.00
Total					0.00	0.08

The table decomposes the sources of variation in imputed comparative advantage: VA for disadvantaged students minus VA for advantaged students. The first column shows the estimated regression coefficients with the sample of teachers without VA scores ("New"). The second column is for comparison and shows the estimated regression coefficients from the same model, but estimated on teachers with VA scores ("Experienced"). The third and fourth columns show the mean of the characteristic in each row for the teacher sample, separately for "New" and "Experienced" teachers, respectively. The last two columns show the components of a Oaxaca-Blinder Decomposition explaining the differences in unshrunken value-added across "New" and "Experienced" teachers. The fifth column is the difference in mean characteristics, multiplied by the coefficients from the first column. The last column is the difference in coefficients multiplied by the mean characteristic for "Experienced" teachers. The "Total" row shows the sum of the decomposition components.