Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 26-1405

Is Teacher Effectiveness Fully Portable? Evidence from the Random Assignment of Transfer Incentives

Matthew A. Kraft
Brown University

John P. Papay
Brown University

Jessalynn James
TNTP

Manuel Monti-Nussbaum
The Behaviouralist

We examine how performance changes when teachers transfer across very different school contexts. The Talent Transfer Initiative program created a rare natural experiment to study such transfers by randomly assigning low-achieving schools the ability to offer high-performing teachers at higher-achieving schools a $20,000 transfer stipend. Forecast tests show that these high-performing teachers' prior value added is only moderately predictive of their effectiveness in low-achieving schools. Using a difference-in-differences framework, we estimate that incentivized-transfer teachers' value added dropped by 0.12 student standard deviations. This decline appears to be driven by lower match quality, negative indirect school effects, and the loss of student-specific human capital.

VERSION: February 2026

Suggested citation: Kraft, Matthew A., John P. Papay, Jessalynn James, and Manuel Monti-Nussbaum. (2026). Is Teacher Effectiveness Fully Portable? Evidence from the Random Assignment of Transfer Incentives. (EdWorkingPaper: 26-1405). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/9zjf-j191

Is Teacher Effectiveness Fully Portable?
Evidence from the Random Assignment of Transfer Incentives

Matthew A. Kraft
Brown University

John P. Papay
Brown University

Jessalynn K. James
TNTP

Manuel Monti-Nussbaum
The Behaviouralist

February 2026

Abstract

We examine how performance changes when teachers transfer across very different school contexts. The Talent Transfer Initiative program created a rare natural experiment to study such transfers by randomly assigning low-achieving schools the ability to offer high-performing teachers at higher-achieving schools a $20,000 transfer stipend. Forecast tests show that these high-performing teachers' prior value added is only moderately predictive of their effectiveness in low-achieving schools. Using a difference-in-differences framework, we estimate that incentivized-transfer teachers' value added dropped by 0.12 student standard deviations. This decline appears to be driven by lower match quality, negative indirect school effects, and the loss of student-specific human capital.

JEL: I2, I21, J24

Correspondence can be sent to Matthew Kraft at mkraft@brown.edu. We thank Steve Glazerman, Ali Protik, and Chris Jones for their support and guidance in working with the TTI data as well as Jesse Bruhn, Grace Falken, Samantha Kane, Maxwell Robinson, and seminar participants at the University of Virginia, Princeton, and the Inter-American Development Bank for their comments and feedback. All errors and omissions are our own.

1. Introduction

Understanding the degree to which worker productivity is portable across firms has important implications for organizations' efforts to maximize the human capital of their workforce. This is particularly true in the public education sector where teachers are the most important school-based determinant of student success (Rockoff 2004; Aaronson et al. 2007; Chetty et al. 2014b) and over 300,000 teachers change schools annually. Academic research and policy discussions often assume, explicitly or implicitly, that teacher effectiveness is fixed and thus essentially portable across schools. This static characterization of teacher effectiveness has motivated the growing use of targeted bonuses to attract high-performing teachers to low-achieving and hard-to-staff schools.

In this study, we test the assumption that teacher effectiveness is fully portable by examining whether teachers' productivity changes when they transfer across starkly different school environments. We leverage data from the Talent Transfer Initiative (TTI), a novel field experiment in U.S. K-12 public schools designed to evaluate the effect on student achievement of randomly assigning subject-grade levels with vacancies (hereafter "teams") in low-achieving schools the ability to offer financial incentives for high-performing teachers at higher-achieving schools to fill open positions. Findings from the experimental trial show that the incentive program raised student achievement 0.07-0.10 standard deviations [σ] in the first year of the program, and 0.12-0.13σ in the second year (Glazerman et al. 2013). While the positive effects of the incentive program are consistent with the portability of teacher effectiveness, these estimates are not a direct test. Instead, they compare the effectiveness of high-performing teachers induced to transfer to low-achieving schools relative to the counterfactual of the teachers these schools would have hired in the absence of the incentive program.

Three features of the TTI program create an advantageous setting to conduct a direct test of the portability hypothesis. First, the program requires low-achieving "receiving" schools to select from a pre-specified pool of high-performing teachers working at higher-achieving schools, exogenously inducing a stark contrast in school environments. Students in the receiving schools were substantially lower-achieving, much more likely to be from low-income backgrounds, and more likely to be Black or Hispanic than the students these teachers had taught in their higher-achieving schools. Working conditions also differed considerably across the schools, with incentivized-transfer teachers reporting that the low-achieving schools they transferred to provided less autonomy, fewer resources and materials, and less support for students with special needs than their prior schools.

Second, the $20,000 transfer stipend, equivalent to $30,000 in 2025 dollars, created a large compensating differential to induce transfer patterns from higher-achieving to low-achieving schools that we very rarely observe in observational data. Third, the experimental design ensures that the observed and unobserved characteristics of the low-achieving schools with vacancies that were randomized to the incentive or control condition are equal in expectation. This allows us to directly compare the performance of higher-performing teachers who filled vacancies in incentive schools to teachers who filled vacancies in the control schools net of any school-level factors that affect teacher effectiveness and student achievement.

We first explore the predictive validity of incentivized-transfer teachers' prior value-added estimates for their future performance in low-achieving schools. Past studies suggest that value-added estimates have minimal forecast bias with coefficients close to 1 (Kane et al. 2013; Chetty et al. 2014a; Bacher-Hicks et al. 2014; Kane and Staiger 2008). We replicate these findings for teacher teams in the control-group that filled vacancies through naturally occurring processes. We find much weaker predictive validity among teachers who transferred due to the incentive and who thus moved to a very different school context. These findings are consistent with evidence from Bacher-Hicks et al. (2014), who find that prior value-added from a different school is a weaker predictor than prior value-added from the same school. It also aligns with work by Jackson (2013) who finds clear evidence of teacher-school match effects.

Prior research largely interprets imperfect predictive validity through the lens of forecast bias. We explore a second interpretation in our setting — that teachers' effectiveness can change when they move across different school settings. Using time series, difference-in-differences (DiD), and event study approaches, we show that transferring from a higher-achieving to a low-achieving school decreases incentivized-transfer teachers' contributions to student achievement by 0.12σ in the first year, on average. Estimates using a sub-sample of teachers for whom a second year of post-transfer data is available suggests that incentivized-transfer teachers improved in the second year at their new school, but that a sizable portion of the initial decline in effectiveness persisted. These results are robust to precision-based weighting and several alternative constructions of the comparison group, each of which accounts for a different source of potential bias.

We conduct a range of analyses to explore the potential causes of this immediate decline and partial recovery in effectiveness. We find conceptual and empirical support for the loss of student-specific human capital as well as negative match effects and negative indirect school effects as primary drivers. Teachers appear to be less effective in their new schools because they teach different types of students (loss of student-specific capital) and/or because their new schools are less conducive to supporting teacher effectiveness overall (indirect school effects) or for these specific teachers (match).

Our paper contributes to several areas of active research on teacher labor markets. We build on a long literature that identifies the effect of teachers on student performance (Hanushek 1971; Murnane 1975; Rockoff 2004) by conducting the first direct and strong test of the portability hypothesis. Pioneering work by Kane and Staiger (2008) and Kane et al. (2013) leverages classroom randomization to evaluate the predictive validity of value-added estimates. As the authors explain, this within-school randomization design does not allow them to investigate the validity of value-added measures for teachers who move across schools. Empirical tests that exploit changes in the composition of teachers in school-grades (e.g., Chetty et al., 2014) provide an important, but indirect test of the portability hypothesis because vacancies are often filled by within-school churn across grades and subjects (Atteberry et al. 2017). Studies that focus on cross-school transfers exclusively provide a direct, but weaker test because they rely on naturally occurring teacher mobility (Xu et al. 2012). Such moves are overwhelmingly to schools that serve higher-achieving and more affluent student populations (Hanushek et al. 2004; Clotfelter et al. 2023; Grissom et al. 2014). As Koedel et al. (2015) note in their review of the value-added literature, we have a very limited understanding of the cross-school portability of teacher effectiveness.

Second, our conceptual framework and findings help to explain the mixed results of financial incentive programs used to address hiring challenges in hard-to-staff schools (Cowan and Goldhaber 2018; Steele et al. 2010; Castro and Esposito 2022; Cabrera and Webbink 2020; Pugatch and Schroeder 2018; Elacqua et al. 2022; Morgan et al. 2023). We also contribute to an emerging body of research that models the equity-efficiency tradeoffs of alternative allocations of teachers across schools and students (Biasi et al. 2021; Aucejo et al. 2022; Bates et al. 2025; Bobba et al. 2024; Graham et al. 2023; Laverde et al. 2025; Tincani 2021). These studies typically relax strong assumptions about the uniform nature of teacher effectiveness by allowing individual teacher effectiveness to differ across student characteristics (typically a binary measure such as high vs. low achievement or economically advantaged vs. disadvantaged). Finally, our findings contribute to the match effects literature in the context of K-12 education (Jackson 2013; Aucejo et al. 2022) and shed new light on the mechanisms that likely underlie these effects.

In what follows, we develop a conceptual framework for thinking about teacher effectiveness to motivate several possible hypotheses for the dynamic patterns in teacher effectiveness we find. We then review the relevant empirical literature and describe the data and TTI study. Next, we present our econometric approach and describe our primary findings. We follow these with a range of robustness tests, exploratory analyses of alternative hypotheses, and predictive validity tests. We conclude by discussing the implications of our study for research, policy, and practice.

2. The Dynamics Teacher Effectiveness

2.1 Conceptual Framework

Here we present a stylized model of teacher effectiveness as a framework for our analyses. Empirical papers often make strong simplifying assumptions that teacher effectiveness is fixed over time (or varies stochastically around a fixed mean) and homogenous across all students. Our conceptual framework relaxes these assumptions by incorporating theories of specific human capital and employee-firm match into the education production function.

$$E_{icgst} = f\left(HC\left\{G_{it}, F_{ist}, T_{igt}, S_{ict}\right\}, M\left\{F_{is}, T_{ig}, S_{ic}\right\}, F_s\right) \qquad (1)$$

Equation (1) models true teacher effectiveness, E_{icgst}, for teacher i, working with student cohort *c,* in grade/subject g, in school s, at time t. True teacher effectiveness is a function of three broad inputs: a teacher's human capital in that year, HC; a match effect between a teacher and the setting in which they work, M; and an indirect school effect experienced by all teachers in school s, F_s. Teachers' human capital at time t is a function of their accumulated general (industry-specific) human capital (G_{it}), firm-specific human capital at school s (F_{ist}), their task-specific human capital in grade/subject g (T_{igt}), and student-specific human capital they have gained (S_{ict}). Note that we explicitly conceptualize student-specific human capital as including knowledge and skills gained from working with both specific students and with students of particular backgrounds. All types of human capital can vary over time.

Match effects, distinct from time-varying human capital gained through experience, are time-invariant advantages or disadvantages individual teachers have working in certain schools, teaching specific subjects/grades, and working with students of various backgrounds. For example, when a school introduces a new curriculum or instructional technology, some teachers may thrive and others may struggle because of how the curriculum or technology aligns with their instructional strengths. The curriculum or technology may not improve teacher effectiveness on average, but it might be particularly beneficial for certain teachers while limiting the effectiveness of others (Jackson and Makarin 2018; Taylor 2018). Some teachers may also simply have a higher latent ability to connect with certain students due to shared backgrounds, affinities, and/or their teaching styles, even conditional on the student-specific human capital they have gained with experience.

Lastly, indirect school effects are school organizational features that impact teacher effectiveness. For example, schools that provide more structured time for teachers to collaborate and better support for addressing student behavioral challenges provide conditions in which teachers can be more effective with their students (Kraft and Papay 2014). These indirect school effects would improve the effectiveness of all teachers in the school but do not affect student outcomes other than through supporting teacher effectiveness. In other words, they are not direct effects that other non-teacher school inputs might have on student achievement (e.g., tutoring programs).

2.2 Model Predictions

When a teacher transfers across schools their effectiveness may change due to a range of inputs that impact their overall performance. Teachers' abilities when they enter the profession matter greatly, and teachers continue to acquire skills as they learn on the job. Teachers' general human capital is fully portable across schools by definition. It also grows over time, at a declining rate on average, and so will increase to some degree after a transfer because a teacher starts the year at a new school with an additional year of experience in the profession. However, moving schools causes teachers to lose the firm-specific human capital they had developed with experience in their prior school. Teachers may retain or even increase their task-specific human capital in their new schools if they teach the same grade and subject, but if they change position types it will decline. Similarly, student-specific human capital can either increase, stay the same, or decline depending on the similarity of the students that teachers work with across schools.

Transferring may affect individual teachers' match effects because the new school, task, and students may or may not be well matched to a teacher's individual style and strengths. Thus, individual teachers' match effects could favor or disadvantage their productivity depending on how match quality changes after transfering schools. Indirect school effects on teacher productivity might enhance or undercut teacher effectiveness depending on the relative organizational effectiveness of the schools that teachers leave and enter.

In the context of the TTI study, our conceptual model combined with compensating differentials in the labor market (Rosen 1986; Smith 1979) suggest that incentivized-transfer teachers are likely to experience both temporary and permanent negative declines in their overall effectiveness. The only input likely to drive increased teacher effectiveness post-transfer in the TTI context is the gain of an additional year of general experience in the profession ($\frac{\partial HC(G)}{\partial t} \geq 0$), especially for earlier career teachers for whom the returns to experience are steepest. The effect on task-specific human capital and student-specific human capital is ambiguous, but likely negative in aggregate given that some teachers will switch to teaching new subjects, grades and student groups ($\frac{\partial HC[T]}{\partial t}, \frac{\partial HC[S]}{\partial t} >=< 0$).

The remaining inputs are all likely to decline due to the transfer process in this context. All teachers who transfer lose their firm-specific capital ($\frac{\partial HC(G)}{\partial t} < 0$) and may also experience a loss in relevant task- and student-specific human capital if their teaching assignments and the profile of their students change. These losses of firm-, task-, and school-specific human capital would result in an immediate but more temporary decline in teacher effectiveness that would rebound as teachers gained experience in their new settings and positions. We also expect a permanent decline in the incentivized-transfer teachers' effectiveness because receiving schools are likely to be less supportive of their effectiveness, and transferring teachers are likely to be less well-matched at their new schools. Receiving schools are likely have less favorable teaching conditions given their records of low-performance, also making the change in school indirect effects negative ($\frac{\partial F}{\partial t} < 0$). The powerful incentive required to induce teachers to transfer, combined with the constrained choices afforded to incentivized-transfer teachers, suggests teachers moved to schools in which they were less favorably matched to the school ($\frac{\partial M[F]}{\partial t} < 0$) and potentially to the students it serves ($\frac{\partial M[S]}{\partial t} < 0$).

2.3 Empirical Evidence

A growing body of empirical papers provides evidence for the dynamic nature of teachers' overall human capital. Studies of the productivity returns to experience among teachers capture rapid increases in effectiveness early in their careers followed by more limited growth (Rockoff 2004; Papay and Kraft 2015; Ladd and Sorensen 2017), with corresponding depreciation when teachers have a gap in employment (Dinerstein et al. 2022). Several papers also find evidence of task-specific human capital. Ost (2015) and Blazar (2015) document the importance of grade-specific human capital and the negative effects of grade switching on productivity growth. Cook and Mansfield (2016) examine course-specific returns to experience among high school teachers and find that as much as a quarter of teachers' overall productivity is not portable across subjects. There is also evidence that teachers accumulate student-specific human capital with individual students. Several studies document the positive effects of repeat student-teacher matches across years (Albornoz et al. 2023; Hill and Jones 2018; Hwang et al. 2021; Wedenoja et al. 2022).

A large body of evidence consistently finds that teachers' effects on student achievement are not uniform across all students. This evidence is consistent with both the existence of student-specific human capital that teachers gain working with particular subgroups of students as well teachers' latent comparative advantage in teaching certain students reflected in teacher-student match effects. Teacher effectiveness can differ based on students' prior level of academic achievement (Aaronson et al. 2007; Biasi et al. 2021; Lockwood and McCaffrey 2009; Condie et al. 2014), socio-economic status (Bates et al. 2025), gender (Dee 2007), English language ability (Loeb et al. 2014; Master et al. 2017), and their racial and ethnic backgrounds (Dee 2004; Egalite et al. 2015; Gershenson et al. 2022; Delgado 2022). Research also documents how elementary school teachers' effectiveness can differ meaningfully across subjects (Fox 2016; Goldhaber et al. 2013).

School culture, organizational practices, and colleagues can also affect teacher productivity (indirect school effects). For example, Bryk et al. (2010) show convincingly how specific organizational approaches can create the conditions for teachers to succeed with their students. Kraft et al. (2016) find that improvements in school contexts that support teachers — such as the quality of professional development, teacher collaboration, and teacher relationships — reduce turnover and raise student achievement. Ronfeldt et al. (2013) find that high levels of teacher turnover reduce the effectiveness of teachers who remain in the school. Several studies find that teachers improve when a higher-performing peer enters the school in the same grade and subject area (Jackson and Bruegmann 2009; Sun et al. 2017). Overall, this body of research suggests that an individual teacher may be more effective in some schools than in others.

2.4 The Portability of Teacher Effectiveness

Far fewer studies have examined the portability of teacher effectiveness across schools. Several studies explore this question indirectly through tests designed to evaluate the validity of value-added estimates of teacher effectiveness. One such test leverages naturally occurring churning of teachers across school-grade-subject cells to assess the predictive validity of value-added scores (Bacher-Hicks et al. 2014; Chetty et al. 2014a; Petek and Pope 2023). These teacher switching quasi-experiments provide compelling evidence of the validity of value-added estimates but are not constructed to provide a direct test of the dynamics of how teacher effectiveness changes across transfers for two primary reasons. First, a large portion of changes in the teaching composition of teachers in school-grade-subject-year cells is likely due to teachers switching across grades and subjects within the same school. For example, Atteberry et al. (2017) find that only 25% of teachers filling open positions in New York City Public Schools transferred across schools, while 54% came from within-school teacher churning across grades and subjects. Second, these tests commonly use estimates of teachers' effectiveness in their post-transfer school settings to construct the value-added measure used in the prediction test. Bacher-Hicks and his colleagues (2014) find that value-added measures constructed based on teachers' contributions to test scores from years when teachers were at different schools are less predictive of teachers' impacts in their new school than value-added constructed using data from years they were at their current school.

Three prior studies directly examine the within-teacher portability of teacher effectiveness. Jackson (2013) applies a within-teacher differences-in-differences (DiD) model to estimate changes in teacher effectiveness among teachers who are observed transferring schools relative to those who do not, while removing school effects through the inclusion of school-by-year fixed effects. He finds that, on average, teachers experience a small positive increase in their effects on student achievement when they transfer, which he attributes to a positive match effect. These teacher-school match effects explain between 10% and 40% of the variation in estimates of teacher quality. Xu et al. (2012) specify a triple-difference model comparing teachers who transfer to schools with very different contexts to those who transfer to schools with similar contexts, and then to teachers who remain in the same schools over time. They find evidence of a very small positive increase, on average, in teacher effects on student achievement among these naturally occurring transfers. When they disaggregate within-teacher changes in effects among high- and low-value-added teachers, they find both groups appear to converge towards the population mean. They interpret this evidence to suggest that teachers' measured effectiveness regresses to the mean. Pham (2022) estimates a DiD model among a sample of teachers transferring to the lowest-performing 5% of schools in Tennessee and finds differential changes in teacher effectiveness across school types.

Together, this literature illustrates the multiple ways in which teacher effectiveness might change after transferring schools. Existing evidence suggests that, on average, teachers maintain or very slightly increase their effectiveness when they seek out a transfer. However, whether teacher effectiveness is fully portable when programs incentivize them to transfer to very different school contexts remains an open question, as do the underlying dynamics at play given the intersections of match effects and different types of human capital.

3. The Talent Transfer Initiative Study

3.1 Setting & Sample

The TTI study was commissioned by the Institute of Education Sciences and took place in ten large U.S. school districts across seven states. The TTI research team led by Mathematica Policy Research first targeted large and economically diverse districts with at least 40 elementary schools and a minimum of 10 low-poverty elementary schools (≤40% of students eligible for free or reduced-price lunch [FRPL]) and 15 high-poverty elementary schools (≥70% of students eligible for FRPL). Of the 51 districts that met these criteria, the research team recruited 10 districts based on factors including the availability and quality of administrative data, hiring and transfer practices, and local political conditions that could be favorable for the feasibility of the study. Seven of the districts participated in the first cohort, with incentivized transfers beginning in the 2009-10 school year, and three participated in the second cohort, which began in 2010-11.

Within each district, elementary and middle schools were ranked by average student achievement in levels or based on school accountability ratings and split into a higher-achieving group (the top 77%), from which eligible transfer candidates would be selected, and low-achieving schools (the bottom 23%) that would be randomly assigned the opportunity to offer high-performing teachers the transfer bonus. A total of roughly 110 elementary and 40 middle low-achieving schools opted into the study and were eligible to receive an incentivized-transfer teacher. By design, student performance at these low-achieving receiving schools was substantially lower than the state average on standardized tests in math (-0.43σ) and English Language Arts (ELA) (-0.52σ). Receiving schools served a student population that was 43% Hispanic, 37% Black, and 6% White, with 80% of students from low-income backgrounds (based on FRPL).

Math and ELA teachers in elementary and middle schools were identified as high performing if their value-added scores, averaged across three prior years, placed them in the top 20% of similar subject-grade teachers in their district. Importantly, these average prior value-added estimates did not include the year directly prior to the first transfer year because student test scores from that year were not yet available at the time of the hiring process. Thus, teachers in the first cohort hired in 2009-10 were selected based on data from 2005-06 to 2007-08. Teachers in the second cohort hired in 2010-11 were selected based on data from 2006-07 to 2008-09. The use of up to three years of prior value-added scores (t-4, t-3, t-2), but excluding the year just prior to transfer (t-1) serves to guard against changes in teachers' value-added caused by mean reversion or an Ashenfelter-like dip where teachers who know they will be leaving their school exert less effort prior to transfer.

Of the roughly 1,520 transfer candidates identified, 330 applied, 170 interviewed for at least one position, 100 received offers, and 80 ultimately accepted offers and transferred.[1] Table 1 presents the characteristics of eligible high-performing transfer candidates in higher-achieving schools. Incentive-eligible teachers who ultimately accepted offers for a position in low-achieving receiving schools were equally as effective as those that did not apply, but were more likely to be Black, younger, and less experienced, on average.

3.2 Data

The Institute for Education Sciences (IES) makes available a secondary dataset from the TTI study that includes data from two cohorts of districts spanning from 2005-06 through 2010-11. This time span encompasses data from three years prior to implementation, two years of post-randomization for the initial cohort, and one year of post-randomization for the second cohort. Seven districts provided raw student data for prior years which the original research team used to calculate empirical Bayes shrunken value-added estimates using a standard lagged-dependent outcomes model that corrects for measurement error in prior test scores.[2] Two districts provided value-added estimates produced by outside vendors, which were combined across three prior years. The remaining district shared only a list of the top performing teachers, which was based on value-added estimates conducted by an outside vendor. The secondary dataset contains student-level administrative data for the post-randomization years that includes students' current and lagged test scores, student demographics, course scheduling, and student-teacher links. All

[1] Surveys administered to candidates suggest that of those who interviewed for teaching positions, roughly 100 teachers interviewed at one school, 40 interviewed at two schools, and the remaining 30 interviewed at three or more schools.

[2] Appendix B provides a detailed explanation of the models used for value-added estimates.

test scores were converted into z-scores that express student achievement relative to the average statewide performance in a given grade and subject. It also includes school-level average student characteristics for all schools in eight of the 10 participating districts.

The secondary dataset also makes available survey data collected from incentive-eligible high-performing teachers in the 10 participating districts. These data include information about the teachers' experiences in their current schools and factors affecting their willingness to apply, interview, and transfer to low-achieving schools. The research team also administered a teacher background survey to all teachers in participating subject/grade teams to collect information on teachers' experiences, satisfaction, and challenges at their schools, along with information on their demographics and their educational and professional experience.

The distinct advantages of these secondary data come with important limitations. The data only include student-level information for students in subject-grade teams that participated in the study during the two post-transfer years rather than for the full panel of pre- and post-randomization years or for all students in participating districts. This requires us to rely on the value-added estimates provided in the secondary dataset rather than constructing alternative estimates and constrains the range of analyses we can conduct. Furthermore, three districts in the study only provided a single pooled pre-period value-added estimate for teachers.[3]

3.3 Experimental Design

The TTI experiment was designed to test the effect of offering a $20,000 salary bonus to high-performing teachers to transfer and remain (for two years) in low-achieving schools.[4]

[3] We also must round sample sizes to the nearest tens place per IES reporting guidelines.
[4] Installments of $10,000, contingent on remaining in the school, were paid out to teachers at the beginning of each of the two school years of the program. A second feature of the study was to offer high-performing teachers already working in low-achieving receiving schools a $10,000 retention bonus to stay in their school for the two-year period. We exclude these teachers from our analyses.

Randomization was conducted across subject-grade teams, given that schools were allowed to submit more than one open vacancy. Prior to random assignment, teams were grouped into blocks that were matched according to grade and subject and, when possible, on similar size and student characteristics. The program included a total of roughly 170 subject-grade teams that were randomly assigned eligibility to offer the talent transfer incentive. Teams were composed of teachers who filled the vacancies as well as incumbent teachers already in the same subject-grade team. High-performing teachers from low-achieving schools filled 100% of the vacancies in incentive teams for which we have data[5]; we refer to vacancy-filling teachers in the treatment group specifically as "incentivized-transfer teachers." In contrast, 62% of teachers who filled vacancies in control teams for which we have data were from within the school, 24% were teachers transferring from other schools, and 14% were novice teachers; we refer to these teachers in the control group who filled vacancies as "control vacancy-filling teachers." These patterns further suggest that tests of value-added based on changes in school-grade-subject teams rely heavily on the within-school churn of teachers across grades and subjects.

Table 2 shows the characteristics of vacancy-filling and incumbent teachers across treatment and control teams. As expected, incentivized-transfer teachers had substantially higher average pre-transfer value-added estimates in both math (0.21σ) and ELA (0.12σ) relative to teachers who filled the vacancies in teams assigned to the control group. Incentivized-transfer teachers were also more experienced overall and had taught for three more years in the district, on average, than control vacancy-filling teachers. Compared to their team-level peers in treatment teams, incentivized-transfer teachers were also more effective and experienced.

[5] The very few incentive vacancies not filled by incentive-eligible teachers went unfilled due to reductions in staffing or were filled by other teachers.

Comparing the characteristics of the teachers, students and working conditions across treatment and control teams confirms that the randomization process was implemented successfully. Column 8 of Table 2 illustrates that 28 out of 29 measures are balanced across incumbent teachers in treatment and control teams, with the lone exception of the probability of being married. Appendix Table A1 compares prior student achievement and demographics across incentive and control teams for all teachers and for only vacancy-filling teachers. Across both groups, we find no significant differences in prior achievement or student demographics, with the exception that incentive teams overall and incentivized-transfer teachers specifically taught more Hispanic students and fewer Black students, on average. Glazerman et al. (2013) contains further details about the setting, sample, data and design of the TTI study.

3.4 Incentivized Teacher Transfer Process

The nature of the matching procedures between high-performing teachers eligible for the incentive and receiving schools with treatment teams is particularly salient for our analyses. The research team managed this process in ways that meaningfully constrained the choices teachers and schools had. In contrast to teacher transfers in the open market, high-performing teachers eligible for the incentive were recruited to transfer and then guided by research managers to apply to individual vacancies among randomly selected low-achieving schools based on position type and geographic proximity. While all transfers were based on mutual consent, the TTI program limited the market to a narrow set of teachers and schools, playing a match-making role to expedite the hiring process during a compressed two-month window. Most incentivized-transfer teachers applied to a single school (60%) and nearly all received a single job offer (90%). The time from random assignment to incentive eligibility and filling a vacancy identified for treatment teams was generally short; site managers reported that vacancies were filled in as few as two days of being assigned. Most vacancies were assigned and filled in May and June, prior to the start of the school year.

4. Econometric Methods

4.1 The Predictive Validity of Teacher Effectiveness across School Contexts

We begin by testing the predictive validity of incentivized-transfer teachers' prior-value-added using the random assignment of incentives across teacher teams as an instrument that induced exogenous increases in the prior value-added among vacancy-filling teachers in treated teams. Glazerman and Protik (2015) conducted parallel 2SLS analyses in an unpublished manuscript and report coefficient estimates from prior value-added scores that are highly variable across grades and subjects (ranging from -1.18 to 1.18σ; see their Table 10) and very imprecisely estimated. We replicate these 2SLS analyses and estimate a similar coefficient on prior value added of -0.57σ with an extremely large 95% confidence interval of [-2.06, 0.92], considerably limiting the utility of this 2SLS approach (see Appendix C).

As an alternative, we adapt the approach used by Bacher-Hicks et al. (2014) to the experimental structure of our data. Specifically, we model the conditional relationship between teachers' average prior value-added scores from *t*-4 to *t*-2 (VA_j^{pre}) and their students' achievement (Y_{ijst}) in first year of the study. We then test whether this relationship differs, in aggregate, across all teachers in treatment teams and those in control teams. We include both vacancy-filling and incumbent teachers on teams rather than only vacancy-filling teachers to guard against any potential for dynamic within-team sorting of students to incentivized-transfer teachers to drive our estimates. In practice, we find little evidence for this type of sorting and the ex-ante direction of such sorting is unclear (see Appendix Table A2). We fit models of the following form:

$$Y_{ijst} = \varphi Y_{ijs,t-1} + \alpha X_{ijst} + \beta_1 (VA_j^{pre} * Incentive\ Team_s) + \beta_2 (VA_j^{pre} * Control\ \ Team_s) + \beta_3 (Incentive\ Team_s) + \varepsilon_{ijst} \quad (2)$$

Here we model student achievement for student i with teacher j in grade-level team s in year t as a function of a student's lagged achievement score ($Y_{ijs,t-1}$) as well as a vector of student demographic characteristics (X_{ijst}) including gender, race, eligibility for free or reduced price lunch, English language learner status, special education services, and having been retained in grade. β_1 captures the association between incentive-group teachers' prior effectiveness and their students' achievement, while β_2 captures the association between control-group teachers' prior effectiveness and their students' achievement. If teacher effectiveness is fully portable, we would expect teachers on teams that were randomized to the incentive condition to have prior value-added scores that were similarly predictive of post-transfer student achievement to that of their control-group peers; more precisely we would expect to fail to reject the null hypothesis that both β_1 and β_2 are different than one. As above, we cluster our standard errors at the teacher level.

Prior research by Jackson and Bruegmann (2009) documents positive peer spillovers when effective teachers transfer to new school contexts. Such dynamics create the potential that the high-performing incentivized-transfer teachers helped their team members improve their performance, weakening the predictive validity of prior value-added for incumbent teachers in treatment teams. To guard against this potential threat to our inferences, we also remove incumbent teachers on treatment teams from the sample and re-estimate equation (3) to examine the predictive validity of prior value-added scores for incentivized-transfer teachers relative to teachers on control teams.

4.2 Testing the Stability of Teacher Effectiveness across Transfers

We then explore directly whether teacher effectiveness changes when teachers transfer across very different school settings using a canonical difference-in-differences (DiD) model adapted to our context. We first estimate a 2x2 DiD model with the two cohorts of participating districts where changes in the effectiveness of incentivized-transfer teachers serve as the first difference and changes in the effectiveness of control vacancy-filling teachers serve as the second difference. Our single pre-period measure of value added is an average of up to three years of estimates from the pre-transfer period. Our single post-period measure is teachers' value added in their first year in the new position. We fit the following model:

$$VA_{jt} = \beta(Incentivized\,Transfers * Post)_{jt} + \alpha * f(Exp)_{jt} + \pi_j + \lambda_t + \varepsilon_{jt} \qquad (3)$$

We proxy for true teacher effectiveness by modeling estimated teacher effectiveness (i.e. value added) for teacher j in year t as the outcome. As we discuss below, this has important implications for the potential interpretation of our results given that year-specific value-added estimates are the product of a teacher-school pairing and may not fully remove direct school effects. We model VA_{jt} as a function of experience[6], $f(Exp)_{jt}$, and fixed effects for teachers, π_j,

[6] It is important to control for experience because changes in value-added scores differ on average at different stages of teachers' careers (Harris and Sass 2011; Papay and Kraft 2015; Rockoff 2004). We parameterize experience as a set of indicator variables (2, 3, 4, 5, 6, 7, 8, 9, 10-15, 16-20, and 21 and above with 1 year as the reference group). We impute the average level of experience in the first program year for teachers with missing data and project experience in pre- and post-transfer years.

and calendar years, λ_t. The coefficient β associated with the interaction term of *Incentivized Transfer* * *Post* provides an estimate of the differential change in value added among incentivized-transfer teachers relative to control new hires. We pool our analytic sample across the seven districts in the first cohort and three districts in the second cohort to maximize our statistical power. In our preferred models, we present results from estimates stacked across subjects to maximize power and include teacher-by-subject fixed effects and a fixed effect for subject. We also present subject-specific estimates in math and ELA.

Pooling our sample across both cohorts maximizes our precision but requires that we limit our primary analysis in equation (1) to only one year post transfer. We next restrict our sample to the first cohort of districts which allows us to extend the canonical DiD approach to model treatment effects in the first year and second year of transferring separately as follows:

$$VA_{jt} = \sum_{t=2010}^{2011} \beta_t D_{jt} + \alpha * f(Exp)_{jt} + \pi_j + \lambda_t + \varepsilon_{jt} \qquad (4)$$

The simple 2x2 design of equation (1) and the restriction of our sample to a single cohort in equation (2) eliminate any concerns about potential biases that may arise in settings with multiple cohorts, staggered treatments, and multiple pre/post periods due to heterogeneous effects (Baker et al. 2025). However, the validity of these estimates still rests on the parallel trends assumption. We examine this assumption by presenting estimates from a fully saturated event-study model where we replace our pooled average estimates of teachers' value added in the pre-period with year-specific estimates. We also employ a number of alternative specifications to test the robustness of our findings, which we describe and report below.

5. Findings

5.1 Differences in School Contexts

The TTI program introduced a strong incentive for high-performing teachers in higher-achieving schools to transfer to low-achieving schools with substantially different school environments. As seen in Table 3, incentivized-transfer teachers experienced a large change in the characteristics of the students in their classrooms. In incentivized-transfer teachers' classrooms, students' prior achievement was 0.40σ lower in math and 0.29σ lower in ELA than in their previous schools. The percentage of students eligible for FRPL increased from 68% to 92%, and the percentage of Black and Hispanic students increased by a combined total of 15 percentage points. In comparison, control vacancy-filling teachers experienced more modest changes in the types of students they taught in their new positions. There was no change in the prior math achievement or racial composition among the students they taught, although they did teach students with substantially lower ELA achievement (0.25σ) and that were more likely to be from low-income backgrounds (8 percentage points).

5.2 Predictive Tests

If measured teacher effectiveness is fully portable across contexts, we would expect prior value-added estimates to predict student achievement gains with a coefficient close to 1. Similar to prior studies (Kane & Staiger, 2008; Kane et al. 2013; Chetty et al., 2014; Bacher-Hicks et al., 2014), our estimated coefficient for prior value-added among control teams reported in Table 4 Column 1 is 0.92 and we cannot reject that this estimate is statistically different from a coefficient of 1 ($H_0: \beta_2 = 1$). We find a much weaker relationship between prior value-added and student achievement gains among incentive teams who filled vacancies with high-performing teachers who transferred from higher-achieving schools. We estimate a coefficient of 0.42 and can reject the null hypothesis that this estimate is equal to 1 (H_0: $\beta_1 = 1$: p=0.003).

We test the robustness of these results to the possibility of positive peer spillover effects among incentive teams, such as those documented in Jackson and Bruegmann (2009), by isolating the predictive validity of incentive new hires. In Column 3, we report coefficients associated with teacher average value-added in the pre-transfer period for incentivized-transfer teachers alone and control teams. The magnitude of our estimates remains essentially unchanged, with a control-team prior value-added coefficient of 0.92 and incentivized-transfer teacher coefficient of 0.41. This suggests that changes in performance of incumbent teachers on incentive teams are not the primary driver of our results.

5.3 The Portability of Teacher Effectiveness Across Different School Contexts

We now examine the possibility that changes in teacher effectiveness explain the weak predictive validity of value-added measures for incentivized-transfer teachers. Before discussing our DiD model estimates, we present times series estimates in Figure 1 Panel A and B to provide visual intuition of the analysis. These figures present trends in year-specific value-added estimates conditional on teaching experience for incentivized-transfer teachers and control vacancy-filling teachers. Three important patterns emerge from this comparison.

First, as described above, incentivized-transfer teachers were substantially more effective in the pre-transfer period than control vacancy-filling teachers. This suggests that the incentive worked as designed to attract teachers who had higher value-added scores than these schools would have been able to recruit otherwise. Second, despite the drop in value added among incentivized-transfer teachers, they remained more effective in their new schools relative to the counterfactual group of vacancy-filling teachers that low-achieving schools were able to hire through standard practices. This is consistent with findings that the transfer incentive program was somewhat effective in improving student achievement in these schools (Glazerman et al., 2013).

Our focus is on the third pattern that emerges: the average value added of incentivized-transfer teachers dropped substantially in the year in which they moved to low-achieving schools, while the average value-added of vacancy-filling teachers on control teams declined a small and not statistically significant amount. The relatively stable pattern for the control vacancy-filling teachers that occurred through a traditional labor-market process is consistent with findings from the existing literature (Chetty et al., 2014; Bacher-Hicks et al., 2014; Xu et al., 2012), but the divergent pattern for incentivized-transfer teachers suggests teacher effectiveness is not fully portable across all contexts. These figures also suggest this decline is not simply a reversion to the mean given the stability of pre-transfer value-added estimates among incentivized-transfer teachers. Figure 1 Panel B focuses on the first cohort only, for whom we can extend our time series to include a second post-transfer year. We again see a similar pattern, but where incentivized-transfer teachers' effectiveness appears to recover somewhat in the second year.

We next present our DiD estimates where we formally model the difference in the change in value-added among incentivized-transfer teachers pre- and post-transfer, removing any secular trends using control vacancy-filling teachers as the comparison group. As shown in Panel A of Table 5, we estimate that the value-added of incentivized-transfer teachers dropped by 0.12σ relative to control vacancy-filling teachers in the first year of their new position. This decline represents moving the typical high-performing incentivized-transfer teacher who was at the 85th percentile of effectiveness down to the 66th percentile of performance and is equivalent to a half of a standard deviation in teacher effectiveness in our sample. Results suggest that the drop in value added was larger for math (-0.15σ) than ELA (-0.09σ).

In Panel B (Table 5), we present estimates using teachers in the first cohort, which we can follow for two years post-transfer. Among these teachers, average value-added scores in the year of transfer, pooled across subjects, were -0.11σ lower in the first year but rose to -0.06σ in the second. Disaggregating our results by subject suggests that value-added scores for incentivized-transfer teachers in their second year recovered very little in math but almost entirely in ELA. However, these disaggregated estimates for a single cohort are somewhat imprecise due to the small sample size.

5.4 Robustness Tests

We test the robustness of our findings by examining the parallel trends assumption, exploring student sorting patterns, and applying alternative value-added estimators, weighting approaches, and comparison groups. Our results are broadly consistent across all these tests. First, to explore the parallel trends assumption, we plot estimates from event-study models for a subsample of teachers who have year-specific pre-period estimates of value-added in Figure 2 Panels A and B. These event-study analyses suggest no discernable pattern of differential trends in prior performance, supporting the validity of our primary DiD design.

A second potential concern is that our findings are driven by a dynamic pattern of student sorting where incentivized-transfer teachers were assigned students with unobservable characteristics related to lower academic growth given that they were viewed as expert teachers. We examine this possibility by testing for differential assignment patterns based on observable student characteristics among teachers in treatment teams. Appendix Table A2 reports coefficients from regressions of observable student characteristics on an indicator for being an incentivized-transfer teacher relative to an incumbent teacher in an incentive team. Consistent with Glazerman et al. (2013), we find little evidence of differential teacher sorting. Average prior student test scores differ by only 0.009σ in math and 0.018σ in reading and are not statistically significant. The only marginally significant difference suggests incentive incumbent teachers were 3.8 percentage points more likely to be assigned a White student.

As a further test of potential bias due to student sorting, we replicate our analyses using value-added estimates that include average student peer characteristics in the model and report the results in Column 1 of Table 6. The results from these models are nearly identical to our primary findings and add further evidence that dynamic student sorting across classrooms does not explain our results. Next, we re-estimate our primary models using inverse variance weights based on the standard error associated with each teacher's individual value-added estimates. This approach places greater weight on those estimates that are more precise and down-weights less precise estimates. As shown in Table 6 Column 2, we find similar estimates for the decline in effectiveness in the first year. However, results from these weighted regressions suggest that incentivized-transfer teachers' effectiveness did not rebound in the second year. Finally, we re-estimate our results using only teachers who are incumbents in a school (i.e. stayers) as the comparison groups following Jackson (2013), rather than control vacancy-filling teachers that are a mix of within school moves, transfers, and teachers new to the profession. Here again, our primary results remain consistent (Table 6 Column 3).

6. Potential Explanations for Dynamic Teacher Effectiveness

We explore the conceptual alignment and empirical evidence for the range of potential mechanisms that might explain our findings. The analyses below rule out changes in general, firm-specific, and task-specific human capital as possible explanations. We find supporting evidence that a combination of changes to student-specific human capital, negative match effects and negative indirect school effects explain our findings. We also discuss the potential that our value-added estimates are conflated with direct school effects, accounting for the patterns we find.

6.1 Human Capital Effects

The drop in estimated teacher effectiveness we observe among the high-performing teachers who transfer to low-achieving schools and the subsequent partial recovery is consistent with the loss and then new accumulation of human capital. Our DiD model includes teacher fixed effects and controls for teacher experience, thus accounting for average changes in general human capital.

Dynamic changes in firm-specific human capital could contribute to this pattern. We explore the role of changes in firm-specific human capital directly by constructing a new comparison group restricted to non-incentivized teachers who transferred into receiving schools.[7] This allows us to more directly remove the loss of firm-specific capital through our second difference but also limits the precision of our estimates given that it narrows the comparison group substantially. As shown in Table 6 Column 4, these results are similar to those from the primary analytic sample, if not slightly more negative, albeit less precise. This suggests that the loss of firm-specific human capital is not likely to be a primary explanation for the pattern of results we find. Teachers in this comparison group experienced a corresponding loss of human capital specific to their prior school, but one that did not result in the same large drop in estimated effectiveness.

[7] This consists of both teachers who joined incentive teams but were not high-performing teachers who were eligible for the transfer bonus as well as control transfer teachers who were new to their school.

A third possible explanation is the loss of task-specific human capital. Our context rules out this possibility. The loss of grade-level human capital is unlikely to be a primary factor given that similar percentages of incentivized-transfer teachers and control vacancy-filling teachers switched grades in our analytic sample (42% vs. 36% , respectively).[8] Subject-specific human capital is also unlikely to be a primary explanation given that by design, incentivized-transfer teachers were hired to teach the same subjects they were previously teaching.

Dynamic changes in student-specific human capital is a final possible explanation. As shown in Table 3 and discussed above, the change in the characteristics and performance of the students that incentivized-transfer teachers taught pre and post transfer was much larger than that of control vacancy-filling teachers. In particular, incentivized-transfer teachers taught students whose prior-year math test scores were 0.40 standard deviations lower in their new school than their old school. This is a substantial shift in the types of students taught. Results from Table 7 comparing incentivized-transfer teachers' perceptions of their students also illustrate a large change in their experiences. Teachers' satisfaction with student motivation dropped from 86% to 39%. These teachers likely had less experience working with the profile of students at the low-achieving schools than the control, vacancy-filling teachers who filled these roles in the absence of the incentive program. We see this as one likely explanation for the patterns we find.

6.2 Match Effects and Indirect School Effects

We now turn to explanations other than a loss of human capital – match effects and indirect school effects. These are conceptually quite similar as both imply that a teacher may be more (or less) effective in one school environment than another. While indirect school effects would shift the effectiveness of all teachers, match effects explicitly suggest that specific

---

[8] Of those, we have data for roughly 30 incentive teachers and 20 control teachers who switched grades in the first program year.

environments may be more or less effective for individual teachers. We cannot disentangle these empirically in our data.

Given our context, we expect that any match effect (at the school, task, or student level) for incentivized-transfer teachers is likely to be negative given that they only accepted to transfer to a low-achieving school in exchange for a large compensating differential. Their choice of what school they transferred to was also substantially constrained, further limiting the likelihood of finding a more positive match.

Descriptive data also suggest that high-performing incentivized-transfer teachers appear to have experienced a large decline in the quality of their teaching environment that are suggestive of a negative indirect school effect. For example, incentivized-transfer teachers' satisfaction with student discipline dropped from 81% in their sending school to 52% in their receiving schools, suggesting their new schools were far less conducive to teaching and learning. The randomization process ensured that vacancy-filling teachers in the incentive and control teams were exposed to school environments in their receiving schools that were equal in expectation. However, the much higher academic performance of students in sending schools for incentivized-transfer teachers compared to sending schools for control vacancy-filling teachers (see Table 3) is suggestive of school environments that might have been more conducive to teaching. Thus, incentivized-transfer teachers may have experienced a larger decline in benefits of indirect school effects compared to their control new hire peers. In other words, the school environment in their pre-transfer schools may have supported their ability to teach more effectively than in their post-transfer schools.

Consistent with both negative match effects and negative indirect school effects, incentivized-transfer teachers reported that they were much less satisfied with the school environments in their new schools and that the students were much more challenging to teach. Table 7 shows that satisfaction with student testing policies and the availability of resources and classroom materials both dropped by 16 percentage points, while satisfaction with autonomy over the classroom dropped by 12 percentage points. Satisfaction with parental involvement dropped by 27 percentage points. These differences illustrate large changes in the teaching and learning environments of their old higher-achieving schools and new low-achieving schools. This descriptive evidence is consistent with a more pronounced decline in match effects and indirect school effects for incentivized-transfer teachers.

6.3 Direct School Effects

Exploring our conceptual model with data requires us to use an empirical estimate of teacher effectiveness in place of true teacher effectiveness. This introduces the potential for another possible explanation for the results we find – direct school effects – given that researchers only observe the combination of a teacher-school pairing in a single year (Raudenbush 2004). Here we characterize direct school effects as the impact that schools have on students directly that are not mediated through teachers (e.g. extended school days, effective tutoring supports, and so on). Some studies explicitly include school fixed effects in their value-added models as an approach to disentangle the potential confounding of teacher and school effects, with the drawback that they remove any true average differences in teacher effectiveness across schools (Bacher-Hicks and Koedel 2023). The value-add models used in the TTI study do not.

The randomized design of the TTI study makes school direct effects equal in expectation across teachers that filled vacancies in incentive and control teams. Thus, the only avenue for school direct effects to explain our results is if the higher-achieving schools where incentivized-transfer teachers worked had larger positive direct school effects than the sending schools where control vacancy-filling teachers previously worked. Although we cannot rule out this potential explanation, it would require quite large school direct effects — which are only one component of school-value added estimates — to fully account for our findings (Angrist et al. 2017; Carrell et al. 2023).

7. Conclusion

Understanding the nature of teacher productivity across schools is essential for informing policies designed to improve educator effectiveness and make access to high-quality teachers more equitable. Teachers differ substantially in their demonstrated practice and their ability to support student learning. They also face strong incentives to work in some schools rather than others. Some of these incentives are pecuniary, with schools offering different salaries even in a single labor market. Many others are non-pecuniary, including working conditions, the professional environment, and the quality of the school leader which all influence teachers' career decisions (Johnson 2020; Kraft et al. 2015). Some policies designed to improve teacher effectiveness rest on the assumption that teacher performance is largely invariant across work environments. Our study illustrates that this assumption does not hold in all contexts.

Instead, we find that the estimated effectiveness of high-performing teachers working at higher-achieving schools declines substantially when they transfer to low-achieving schools. This suggests, at the least, that estimated teacher effectiveness is not always fully portable. The portability of teacher effectiveness is also highly relevant in contexts with centralized assignment mechanisms such as those that arise due to reductions-in-force or school turnaround processes in the United States, or among more centralized assignment schemes used by many national ministries of education abroad. The teachers we observe moving did so in response to a high-powered financial incentive, but for many more teachers that incentive was still not enough of a compensating differential. Systems that simply require teachers to move rather than using high-powered incentives may be contexts where teacher portability is particularly weak.

The specific dynamics underlying the lack of full portability are also critical for structuring and supporting teachers' work in different ways. First, student-specific human capital appears to be important. When teachers are asked to teach very different type of students, they lose valuable human capital and may require additional support. Second, the role of teacher-school match suggests the value of robust hiring practices designed to improve match quality. Third, indirect school effects appear to play an important role. Schools are critical in supporting or constraining teacher effectiveness, and teachers are substantially less effective in less supportive schools. These patterns suggest that improving teacher effectiveness requires attention not only to identifying teachers with the skills and capacities to be effective, but also to how teacher effectiveness can differ across settings and how organizational conditions affect teachers' work.

References

Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. "Teachers and Student Achievement in the Chicago Public High Schools." *Journal of Labor Economics* 25 (1): 95–135. https://doi.org/10.1086/508733.

Albornoz, Facundo, David Contreras, and Richard Upward. 2023. "Let's Stay Together: The Effects of Repeat Student-Teacher Matches on Academic Achievement." *Economics of Education Review* 94 (June): 102375. https://doi.org/10.1016/j.econedurev.2023.102375.

Angrist, Joshua D., Peter D. Hull, Parag A. Pathak, and Christopher R. Walters. 2017. "Leveraging Lotteries for School Value-Added: Testing and Estimation." *The Quarterly Journal of Economics* 132 (2): 871–919. https://doi.org/10.1093/qje/qjx001.

Atteberry, Allison, Susanna Loeb, and James Wyckoff. 2017. "Teacher Churning: Reassignment Rates and Implications for Student Achievement." *Educational Evaluation and Policy Analysis* 39 (1): 3–30. https://doi.org/10.3102/0162373716659929.

Aucejo, Esteban, Patrick Coate, Jane Cooley Fruehwirth, Sean Kelly, and Zachary Mozenter. 2022. "Teacher Effectiveness and Classroom Composition: Understanding Match Effects in the Classroom." *The Economic Journal* 132 (648): 3047–64.

Bacher-Hicks, Andrew, Thomas Kane, and Douglas Staiger. 2014. *Validating Teacher Effect Estimates Using Changes in Teacher Assignments in Los Angeles*. No. W20657. National Bureau of Economic Research. https://doi.org/10.3386/w20657.

Bacher-Hicks, Andrew, and Cory Koedel. 2023. "Estimation and Interpretation of Teacher Value Added in Research Applications." In *Handbook of the Economics of Education*, vol. 6. Elsevier. https://doi.org/10.1016/bs.hesedu.2022.11.002.

Baker, Andrew, Brantly Callaway, Scott Cunningham, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2025. "Difference-in-Differences Designs: A Practitioner's Guide." arXiv:2503.13323. Preprint, arXiv, June 17. https://doi.org/10.48550/arXiv.2503.13323.

Bates, Michael, Michael Dinerstein, Andrew C Johnston, and Isaac Sorkin. 2025. "Teacher Labor Market Policy and the Theory of the Second Best." *The Quarterly Journal of Economics* 140 (2): 1417–69. https://doi.org/10.1093/qje/qjae042.

Biasi, Barbara, Chao Fu, and John Stromme. 2021. "Equilibrium in the Market for Public School Teachers: District Wage Strategies and Teacher Comparative Advantage." *NBER*.

Bobba, Matteo, Tim Ederer, Gianmarco Leon-Ciliotta, Christopher Neilson, and Marco G Nieddu. 2024. "Teacher Compensation and Structural Inequality: Evidence from Centralized Teacher School Choice in Peru." *NBER*.

Bryk, Anthony S., Penny Bender Sebring, Elaine Allensworth, Stuart Luppescu, and John Q. Easton. 2010. *Organizing Schools for Improvement: Lessons from Chicago*. University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/O/bo8212979.html.

Cabrera, José María, and Dinand Webbink. 2020. “Do Higher Salaries Yield Better Teachers and Better Student Outcomes?” *Journal of Human Resources* 55 (4): 1222–57. https://doi.org/10.3368/jhr.55.4.0717-8911R3.

Carrell, Scott, Michal Kurlaender, Paco Martorell, Matthew Naven, and Christina Sun. 2023. “Do Schools Matter? Measuring the Impact of California High Schools on Test Scores and Postsecondary Enrollment.” https://conference.nber.org/conf_papers/f191771.pdf.

Castro, Juan F., and Bruno Esposito. 2022. “The Effect of Bonuses on Teacher Retention and Student Learning in Rural Schools: A Story of Spillovers.” *Education Finance and Policy* 17 (4): 693–718. https://doi.org/10.1162/edfp_a_00348.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014a. “Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates.” *The American Economic Review* 104 (9): 2593–632.

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. https://doi.org/10.1257/aer.104.9.2633.

Clotfelter, Charles T., Helen F. Ladd, and Calen R. Clifton. 2023. “Racial Differences in Student Access to High-Quality Teachers.” *Education Finance and Policy* 18 (4): 738–52. https://doi.org/10.1162/edfp_a_00402.

Condie, Scott, Lars Lefgren, and David Sims. 2014. “Teacher Heterogeneity, Value-Added and Education Policy.” *Economics of Education Review* 40 (June): 76–92. https://doi.org/10.1016/j.econedurev.2013.11.009.

Cowan, James, and Dan Goldhaber. 2018. “Do Bonuses Affect Teacher Staffing and Student Achievement in High Poverty Schools? Evidence from an Incentive for National Board Certified Teachers in Washington State.” *Economics of Education Review* 65 (August): 138–52. https://doi.org/10.1016/j.econedurev.2018.06.010.

Dee, Thomas S. 2004. “Teachers, Race, and Student Achievement in a Randomized Experiment.” *The Review of Economics and Statistics* 86 (1): 195–210.

Dee, Thomas S. 2007. “Teachers and the Gender Gaps in Student Achievement.” *Journal of Human Resources* 42 (3): 528–54.

Delgado, William. 2022. “Heterogeneous Teacher Effects, Comparative Advantage, and Match Quality.” Unpublished manuscript.

Dinerstein, Michael, Rigissa Megalokonomou, and Constantine Yannelis. 2022. “Human Capital Depreciation and Returns to Experience.” *American Economic Review* 112 (11): 3725–62. https://doi.org/10.1257/aer.20201571.

Egalite, Anna J., Brian Kisida, and Marcus A. Winters. 2015. “Representation in the Classroom: The Effect of Own-Race Teachers on Student Achievement.” *Economics of Education Review* 45 (April): 44–52. https://doi.org/10.1016/j.econedurev.2015.01.007.

Elacqua, Gregory, Diana Hincapie, Isabel Hincapie, and Veronica Montalva. 2022. “Can Financial Incentives Help Disadvantaged Schools to Attract and Retain High-Performing Teachers? Evidence from Chile.” *Journal of Policy Analysis and Management* 41 (2): 603–31. https://doi.org/10.1002/pam.22375.

Fox, Lindsay. 2016. “Seeing Potential: The Effects of Student–Teacher Demographic Congruence on Teacher Expectations and Recommendations.” *AERA Open* 2 (1): 2332858415623758. https://doi.org/10.1177/2332858415623758.

Gershenson, Seth, Cassandra M. D. Hart, Joshua Hyman, Constance A. Lindsay, and Nicholas W. Papageorge. 2022. “The Long-Run Impacts of Same-Race Teachers.” *American Economic Journal: Economic Policy* 14 (4): 300–342. https://doi.org/10.1257/pol.20190573.

Glazerman, Steven, and Ali Protik. 2015. “Validating Value-Added Measures of Teacher Performance.” Unpublished manuscript.

Glazerman, Steven, Ali Protik, Bing-ru The, Julie Bruch, and Jeffrey Max. 2013. *Transfer Incentives for High-Performing Teachers: Final Results from a Multisite Randomized Experiment*.

Goldhaber, Dan, James Cowan, and Joe Walch. 2013. “Is a Good Elementary Teacher Always Good? Assessing Teacher Performance Estimates across Subjects.” *Economics of Education Review* 36 (October): 216–28. https://doi.org/10.1016/j.econedurev.2013.06.010.

Graham, Bryan S., Geert Ridder, Petra Thiemann, and Gema Zamarro. 2023. “Teacher-to-Classroom Assignment and Student Achievement.” *Journal of Business & Economic Statistics* 41 (4): 1328–40. https://doi.org/10.1080/07350015.2022.2126480.

Grissom, Jason A., Susanna Loeb, and Nathaniel A. Nakashima. 2014. “Strategic Involuntary Teacher Transfers and Teacher Performance: Examining Equity and Efficiency.” *Journal of Policy Analysis and Management* 33 (1): 112–40. https://doi.org/10.1002/pam.21732.

Hanushek, Eric. 1971. “Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data.” *American Economic Review* 62 (2): 280–88.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2004. “Why Public Schools Lose Teachers.” *The Journal of Human Resources* 39 (2): 326–54. https://doi.org/10.2307/3559017.

Harris, Douglas N., and Tim R. Sass. 2011. “Teacher Training, Teacher Quality and Student Achievement.” *Journal of Public Economics* 95 (7–8): 798–812. https://doi.org/10.1016/j.jpubeco.2010.11.009.

Hill, Andrew J., and Daniel B. Jones. 2018. "A Teacher Who Knows Me: The Academic Benefits of Repeat Student-Teacher Matches." *Economics of Education Review* 64 (June): 1–12. https://doi.org/10.1016/j.econedurev.2018.03.004.

Hwang, NaYoung, Brian Kisida, and Cory Koedel. 2021. "A Familiar Face: Student-Teacher Rematches and Student Achievement." *Economics of Education Review* 85 (December): 102194. https://doi.org/10.1016/j.econedurev.2021.102194.

Jackson, C. Kirabo. 2013. "Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers." *The Review of Economics and Statistics* 95 (4): 1096–116. https://doi.org/10.1162/REST_a_00339.

Jackson, C. Kirabo, and Elias Bruegmann. 2009. "Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers." *American Economic Journal: Applied Economics* 1 (4): 85–108. https://doi.org/10.1257/app.1.4.85.

Jackson, Kirabo, and Alexey Makarin. 2018. "Can Online Off-the-Shelf Lessons Improve Student Outcomes? Evidence from a Field Experiment." *American Economic Journal: Economic Policy* 10 (3): 226–54. https://doi.org/10.1257/pol.20170211.

Johnson, Susan Moore. 2020. *Where Teachers Thrive: Organizing Schools for Success*. Harvard Education Press.

Kane, Thomas J, Daniel F McCaffrey, Trey Miller, and Douglas O Staiger. 2013. "Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment." *Research Paper. MET Project. Bill & Melinda Gates Foundation*.

Kane, Thomas, and Douglas Staiger. 2008. *Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation*. No. W14607. National Bureau of Economic Research. https://doi.org/10.3386/w14607.

Koedel, Cory, Kata Mihaly, and Jonah E. Rockoff. 2015. "Value-Added Modeling: A Review." *Economics of Education Review* 47 (August): 180–95. https://doi.org/10.1016/j.econedurev.2015.01.006.

Kraft, Matthew A., William H. Marinell, and Darrick Shen-Wei Yee. 2016. "School Organizational Contexts, Teacher Turnover, and Student Achievement: Evidence From Panel Data." *American Educational Research Journal* 53 (5): 1411–49. https://doi.org/10.3102/0002831216667478.

Kraft, Matthew A., and John P. Papay. 2014. "Can Professional Environments in Schools Promote Teacher Development? Explaining Heterogeneity in Returns to Teaching Experience." *Educational Evaluation and Policy Analysis* 36 (4): 476–500.

Kraft, Matthew A., John P. Papay, Susan Moore Johnson, Megin Charner-Laird, Monica Ng, and Stefanie Reinhorn. 2015. "Educating Amid Uncertainty: The Organizational Supports Teachers Need to Serve Students in High-Poverty, Urban Schools." *Educational Administration Quarterly* 51 (5): 753–90. https://doi.org/10.1177/0013161X15607617.

Ladd, Helen F., and Lucy C. Sorensen. 2017. "Returns to Teacher Experience: Student Achievement and Motivation in Middle School." *Education Finance and Policy* 12 (2): 241–79. https://doi.org/10.1162/EDFP_a_00194.

Laverde, Mariana, Elton Mykerezi, Aaron Sojourner, and Aradhya Sood. 2025. *Gains from Alternative Assignment? Evidence from a Two-Sided Teacher Market*.

Lockwood, J. R., and Daniel F. McCaffrey. 2009. "Exploring Student-Teacher Interactions in Longitudinal Achievement Data." *Education Finance and Policy* 4 (4): 439–67. https://doi.org/10.1162/edfp.2009.4.4.439.

Loeb, Susanna, James Soland, and Lindsay Fox. 2014. "Is a Good Teacher a Good Teacher for All? Comparing Value-Added of Teachers With Their English Learners and Non-English Learners." *Educational Evaluation and Policy Analysis* 36 (4): 457–75. https://doi.org/10.3102/0162373714527788.

Master, Benjamin, Susanna Loeb, and James Wyckoff. 2017. "More Than Content: The Persistent Cross-Subject Effects of English Language Arts Teachers' Instruction." *Educational Evaluation and Policy Analysis* 39 (3): 429–47. https://doi.org/10.3102/0162373717691611.

Morgan, Andrew, Minh Nguyen, Eric Hanushek, Ben Ost, and Steven Rivkin. 2023. *Attracting and Retaining Highly Effective Educators in Hard-to-Staff Schools*. No. W31051. National Bureau of Economic Research. https://doi.org/10.3386/w31051.

Murnane, Richard J. 1975. *The Impact of School Resources on the Learning of Inner City Children*. Balinger Publishing Company.

Papay, John P., and Matthew A. Kraft. 2015. "Productivity Returns to Experience in the Teacher Labor Market: Methodological Challenges and New Evidence on Long-Term Career Improvement." *Journal of Public Economics* 130 (October): 105–19. https://doi.org/10.1016/j.jpubeco.2015.02.008.

Petek, Nathan, and Nolan G. Pope. 2023. "The Multidimensional Impact of Teachers on Students." *Journal of Political Economy* 131 (4): 1057–107. https://doi.org/10.1086/722227.

Pham, Lam D. 2022. "Is Teacher Effectiveness Stable Across School Contexts? An Examination of Teachers Who Transfer Into Turnaround Schools." *AERA Open* 8 (January): 23328584221139763. https://doi.org/10.1177/23328584221139763.

Pugatch, Todd, and Elizabeth Schroeder. 2018. "Teacher Pay and Student Performance: Evidence from the Gambian Hardship Allowance." *Journal of Development Effectiveness* 10 (2): 249–76. https://doi.org/10.1080/19439342.2018.1452778.

Raudenbush, Stephen W. 2004. "What Are Value-Added Models Estimating and What Does This Imply for Statistical Practice?" *Journal of Educational and Behavioral Statistics* 29 (1): 121–29. https://doi.org/10.3102/10769986029001121.

Rockoff, Jonah E. 2004. “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data.” *American Economic Review* 94 (2): 247–52. https://doi.org/10.1257/0002828041302244.

Ronfeldt, Matthew, Susanna Loeb, and James Wyckoff. 2013. “How Teacher Turnover Harms Student Achievement.” *American Educational Research Journal* 50 (1): 4–36. https://doi.org/10.3102/0002831212463813.

Rosen, Sherwin. 1986. “Chapter 12 The Theory of Equalizing Differences.” In *Handbook of Labor Economics*, vol. 1. Elsevier. https://doi.org/10.1016/S1573-4463(86)01015-5.

Smith, Robert S. 1979. “Compensating Wage Differentials and Public Policy: A Review.” *ILR Review* 32 (3): 339–52. https://doi.org/10.1177/001979397903200304.

Steele, Jennifer L., Richard J. Murnane, and John B. Willett. 2010. “Do Financial Incentives Help Low-Performing Schools Attract and Keep Academically Talented Teachers? Evidence from California.” *Journal of Policy Analysis and Management* 29 (3): 451–78.

Sun, Min, Susanna Loeb, and Jason A. Grissom. 2017. “Building Teacher Teams: Evidence of Positive Spillovers From More Effective Colleagues.” *Educational Evaluation and Policy Analysis* 39 (1): 104–25. https://doi.org/10.3102/0162373716665698.

Taylor, Eric S. 2018. “Skills, Job Tasks, and Productivity in Teaching: Evidence from a Randomized Trial of Instruction Practices.” *Journal of Labor Economics* 36 (3): 711–42. https://doi.org/10.1086/696144.

Tincani, Michela M. 2021. “Teacher Labor Markets, School Vouchers, and Student Cognitive Achievement: Evidence from Chile.” *Quantitative Economics* 12 (1): 173–216. https://doi.org/10.3982/QE1057.

Wedenoja, Leigh, Papay, John, and Kraft, Matthew A. 2022. *Second Time’s the Charm? How Repeat Student-Teacher Matches Build Academic and Behavioral Skills*. https://doi.org/10.26300/SDDW-AG22.

Xu, Zeyu, Umut Özek, and Matthew Corritore. 2012. “Portability of Teacher Effectiveness Across School Settings.” *CALDER Working Paper* No. 77-0612.

Tables

Table 1. Characteristics of High-Performing Teachers Eligible for Transfer Incentive Bonus

	Did Not Apply	Applied	
		All	Transferred
Math VA (avg across years)	0.21	0.22	0.21
ELA VA (avg across years)	0.12	0.13	0.12
Female	0.82	0.81	0.82
Hispanic	0.18	0.16	0.17
Black	0.15	0.27	0.33
White	0.79	0.68	0.62
Age	46.3	42.8	42.0
Years Experience as Classroom Teacher	16.9	12.7	12.5
Barron's Selectivity Rank of Bachelor's Degree	3.83	3.97	4.01
Graduate Degree	0.49	0.54	0.53
National Board Certification	0.16	0.14	0.15
Married	0.71	0.61	0.62
Home Owner	0.87	0.80	0.81
Living within School District	0.66	0.64	0.62
Distance to School (minutes)	19.8	20.1	21.6
Total N (teachers)	1190	330	80

Notes: Columns represent teachers that made it to each level (mutually exclusive). Math VA and ELA (English language arts) VA are the teacher's value-added measure averaged across years prior to the start of the transfer incentive program. Barron's rating is a measure of the selectivity of the institution where the teacher obtained their Bachelor's degree; values range from 1 "Most Competitive" to 6 "Non-Competitive". Missingness varies across variables. Sample sizes rounded to the nearest ten.

Source: Moving High-Performing Teachers restricted dataset, U.S. Department of Education, National Center for Education Statistics.

Table 2. Descriptive Characteristics of Teachers in Program Schools

	Incentive Teams		Control Teams		Significance Level			
	Vacancy-Filling Teachers	Incumbent Teachers	Vacancy-Filling Teachers	Incumbent Teachers	*p-value (Incentive: V-F v. incumbent)*	*p-value (V-F: incentive v. control)*	*p-value (Control: VF v. incumbent)*	*p-value (Incumbents: incentive v. control)*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Math VA (avg across years)	0.21	0.00	-0.04	0.01	0.00	0.00	0.09	0.62
ELA VA (avg across years)	0.12	-0.01	0.00	-0.02	0.00	0.00	0.34	0.25
Female	0.82	0.73	0.82	0.77	0.12	0.97	0.35	0.41
Hispanic	0.17	0.27	0.18	0.25	0.10	0.88	0.18	0.75
Black	0.33	0.34	0.39	0.39	0.88	0.43	0.99	0.38
White	0.62	0.60	0.53	0.58	0.81	0.27	0.52	0.70
Age	41.99	39.46	37.03	39.70	0.10	0.00	0.10	0.86
Years Experience as Classroom Teacher	12.49	10.37	9.20	11.26	0.07	0.01	0.09	0.35
Years Experience in District	10.19	8.48	6.77	9.69	0.10	0.00	0.01	0.15
Years Experience in School	1.00	6.28	4.39	6.21	0.00	0.00	0.01	0.91
Rating for Selectivity of Bachelor's Degree	4.01	3.91	3.81	4.02	0.58	0.25	0.22	0.43
Graduate Degree	0.53	0.47	0.43	0.53	0.32	0.20	0.16	0.31
Certifications								
National Board	0.15	0.10	0.09	0.16	0.28	0.17	0.12	0.11
Pre-School	0.11	0.08	0.18	0.10	0.58	0.23	0.10	0.61
Elementary (K-5)	0.90	0.75	0.81	0.67	0.00	0.10	0.01	0.11
Middle (6-8)	0.63	0.54	0.57	0.54	0.14	0.45	0.61	0.96
Secondary (9-12)	0.20	0.26	0.21	0.26	0.34	0.86	0.36	0.90
Special Subject Areas	0.42	0.20	0.21	0.22	0.00	0.00	0.90	0.70
Exceptional Children	0.15	0.04	0.06	0.06	0.00	0.06	0.98	0.42
Other Area	0.26	0.16	0.11	0.16	0.06	0.01	0.30	0.98
ELL	0.14	0.11	0.14	0.12	0.39	0.97	0.65	0.87
Math	0.17	0.10	0.09	0.09	0.07	0.11	0.97	0.84
ELA	0.21	0.11	0.09	0.12	0.02	0.03	0.60	0.98
Special Ed	0.16	0.07	0.07	0.07	0.03	0.07	0.92	0.96
Married	0.62	0.47	0.55	0.63	0.02	0.34	0.21	0.00
Home Owner	0.81	0.62	0.54	0.69	0.00	0.00	0.02	0.17
Living within School District	0.62	0.56	0.59	0.51	0.40	0.69	0.25	0.39

Distance to School in Previous Year (minutes)	21.61	23.92	21.86	23.47	0.23	0.90	0.43	0.80
Distance to School in Current Year (minutes)	25.35	23.13	22.23	23.37	0.26	0.14	0.56	0.88
N (teachers)	80	270	120	180				

Notes: Math VA and ELA (English language arts) VA are the teacher's value-added score averaged across years prior to the start of the transfer incentive program. Missingness varies across variables. Last three columns show *p*-values from t-tests between groups of teachers. Sample sizes rounded to the nearest ten.

Source: Moving High-Performing Teachers restricted dataset, U.S. Department of Education, National Center for Education Statistics.

Table 3. Average Student Characteristics in Incentivized-Transfer and Control Vacancy-Filling Teachers' Classrooms in Sending and Receiving Schools

	Incentivized-Transfer Teachers				Control Vacancy-Filling Teachers			
	Sending School	Receiving School	Difference		Sending School	Receiving School	Difference	
Prior Achievement								
Math	-0.11	-0.51	-0.40	***	-0.47	-0.51	-0.04	
ELA	-0.10	-0.39	-0.29	**	-0.40	-0.65	-0.25	**
Student Characteristics								
White	0.23	0.09	-0.14	***	0.04	0.04	-0.01	
Black	0.28	0.37	0.08	**	0.38	0.39	0.02	
Hispanic	0.37	0.43	0.06		0.50	0.49	-0.02	
FRPL	0.68	0.92	0.24	***	0.82	0.91	0.08	***
ELL	0.16	0.15	-0.01		0.28	0.20	-0.08	**
Special Ed	0.13	0.11	-0.02		0.10	0.11	0.01	
N (classrooms)	50	50			40	40		

Notes: Table shows average characteristics of students taught by teachers in pre-transfer years and in the first year post-transfer. Math and ELA (English language arts) achievement values are average scores from the previous years, reported in standard deviations relative to state averages. Missingness varies across variables. All differences are estimated within teacher. Sample sizes rounded to the nearest ten. *p<0.1, **p<0.05, ***p<0.01.

Source: Moving High-Performing Teachers, U.S. Department of Education, National Center for Education Statistics.

Table 4. Predictive Validity Tests of Prior Value Added Across Teams and Teachers

	Coefficients		P-value from t-test of β=1	Coefficients		P-value from t-test of β=1
	(1)		(2)	(3)		(4)
VA^{pre} Incentive Teams	0.421	**	0.003			
	(0.193)					
VA^{pre} Incentivized-Transfer Teachers				0.405		0.170
				(0.432)		
VA^{pre} Control Teams	0.917	***	0.763	0.919	***	0.768
	(0.274)			(0.274)		
VA^{pre} Difference (Control - Incentive)	-0.496			-0.515		
	(0.334)			(0.509)		
n (students)	18,000			10,250		

Notes: We estimate this model at the student-subject level, with student achievement in math or ELA as the outcome variable. The model includes controls for students' lagged within-subject achievement and fixed or pre-treatment characteristics (race/ethnicity, gender, English language learner status, special education status, free or reduced-price lunch eligibility, gifted status, and age), as well as subject fixed effects. Sample sizes rounded to the nearest ten. *p<0.1, **p<0.05, ***p<0.01.

Source: Moving High-Performing Teachers restricted dataset, U.S. Department of Education, National Center for Education Statistics.

Table 5. Difference-in-Differences Estimates of the Portability of Teacher Effectiveness

	Average Pre-Period VA					
	Stacked		Math		ELA	
	(1)		(2)		(3)	
	Panel A. One Year Post Transfer					
Incentivized Transfer * Post y1	-0.121	***	-0.149	***	-0.087	*
	(0.044)		(0.056)		(0.050)	
N (teacher subject year)	390		190		190	
	Panel B. Two Years Post Transfer					
Incentivized Transfer * Post y1	-0.105	**	-0.132	**	-0.079	
	(0.050)		(0.055)		(0.058)	
Incentivized Transfer * Post y2	-0.059		-0.118	**	-0.004	
	(0.051)		(0.054)		(0.065)	
N (teacher subject year)	600		300		310	

Notes: The estimation sample consists of all incentivized-transfer teachers and control vacancy-filling teachers. Value-added (VA) scores for the average pre-treatment period is a pooled measure for most teachers. Where pooled value-added was not provided, this value is an average of value-added scores across t-4 to t-2 prior to transfer. VA scores are estimated separately for math and English language arts (ELA). All models are estimated with teacher (by-subject) fixed effects and include controls for teacher experience. Standard errors, clustered at the teacher level, are in parentheses. Sample sizes rounded to the nearest ten. *p<0.1, **p<0.05, ***p<0.01.

Source: Moving High-Performing Teachers restricted dataset, U.S. Department of Education, National Center for Education Statistics.

Table 6. Robustness Tests of Difference-in-Differences Estimates of the Portability of Teacher Effectiveness

	Average Pre-Period VA									
	VA estimates from models that include average peer student characteristics		Inverse Variance Weights		All Incumbents Comparison (stayers)		Non-Incentivized Transfers Comparison		Incentive Incumbents Comparison	
	(1)		(2)		(3)		(4)		(5)	
	Panel A. One Year Post									
Incentivized Transfer * Post y1	-0.115	***	-0.131	***	-0.132	***	-0.145	*	-0.133	***
	(0.040)		(0.048)		(0.035)		(0.074)		(0.037)	
N (teacher subject year)	280		280		960		290		634	
	Panel B. Two Years Post									
Incentivized Transfer * Post y1	-0.116	***	-0.104	**	-0.122	***	-0.098		-0.124	***
	(0.040)		(0.041)		(0.040)		(0.083)		(0.044)	
Incentivized Transfer * Post y2	-0.073	*	-0.104	***	-0.079	**	-0.122		-0.093	*
	(0.039)		(0.038)		(0.040)		(0.087)		(0.050)	
N (teacher subject year)	540		540		1450		430		940	

Notes: Value-added (VA) scores for the average pre-treatment period is a pooled measure for most teachers. Where pooled value-added was not provided, this value is an average of value-added scores across t-4 to t-2 prior to transfer. VA scores are estimated separately for math and English language arts (ELA). All models are estimated with teacher (by-subject) fixed effects and include controls for teacher experience. Standard errors, clustered at the teacher level, are in parentheses. Sample sizes rounded to the nearest ten. *p<0.1, **p<0.05, ***p<0.01.

Source: Moving High-Performing Teachers restricted dataset, U.S. Department of Education, National Center for Education Statistics.

Table 7. Self-Reported Working Conditions for Incentivized-Transfer Teachers in Sending and Receiving Schools

	Sending School	Receiving School	Difference	
Principal's leadership and vision	0.68	0.69	0.01	
Recognition for positive accomplishments	0.73	0.73	0.00	
Student testing policies	0.81	0.65	-0.16	**
Other school policies	0.69	0.57	-0.12	
Salary	0.70	0.78	0.08	
Benefits	0.77	0.81	0.04	
Caliber of colleagues	0.81	0.81	0.00	
Opportunities for professional dev.	0.84	0.79	-0.05	
Opportunities to provide input into school policies	0.68	0.60	-0.08	
Autonomy over classroom	0.87	0.75	-0.12	**
Workload	0.65	0.64	-0.01	
Teacher support from administration	0.62	0.65	0.03	
Support/Collaboration from faculty	0.87	0.81	-0.06	
Support for teachers with students with special needs	0.74	0.61	-0.13	*
Availability of resources and materials for classroom	0.81	0.65	-0.16	**
School Facilities	0.79	0.78	-0.01	
Safety on school grounds	0.90	0.88	-0.01	
Safety in school neighborhood	0.84	0.82	-0.03	
Student motivation	0.86	0.39	-0.47	***
Student discipline	0.81	0.52	-0.29	***
Student academic performance	0.82	0.40	-0.42	***
Parental involvement in the school	0.60	0.32	-0.27	***
N (teachers)	80	80		

Notes: Table shows average response scores to questions about satisfaction in current school. Responses, answered on a Likert scale ranging from 1 = "Very Dissatisfied" to 4 = "Very Satisfied", are collapsed into a score of 1 for Satisfied and 0 for Dissatisfied. Sample sizes rounded to the nearest ten. *p<0.1, **p<0.05, ***p<0.01.

Source: Moving High-Performing Teachers restricted dataset, U.S. Department of Education, National Center for Education Statistics.

Figures

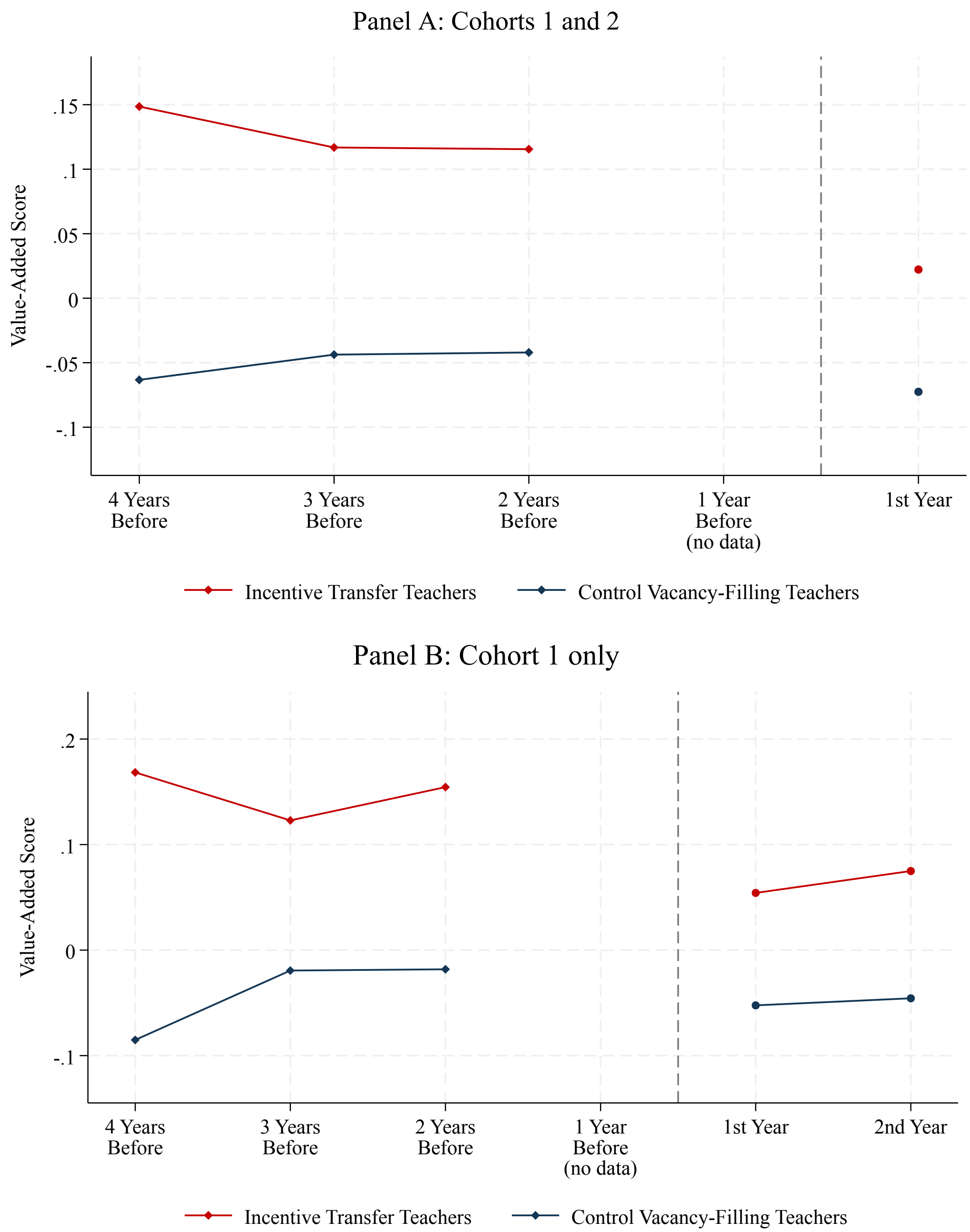

Figure 1. Trends in average value-added scores for incentivized-transfer teachers and control vacancy-filling teachers prior to and one year after filling open positions in low-achieving receiving schools.
Notes: Estimates from a linear regression of yearly value-added scores on interactions between incentive status and year, where year is centered on the transfer year. The regression model includes controls for ranges of teacher experience. Observations are teacher-by-subject-by-year in an unbalanced panel. Standard errors are clustered within teacher.
Source: Moving High-Performing Teachers restricted dataset, U.S. Department of Education, National Center for Education Statistics.

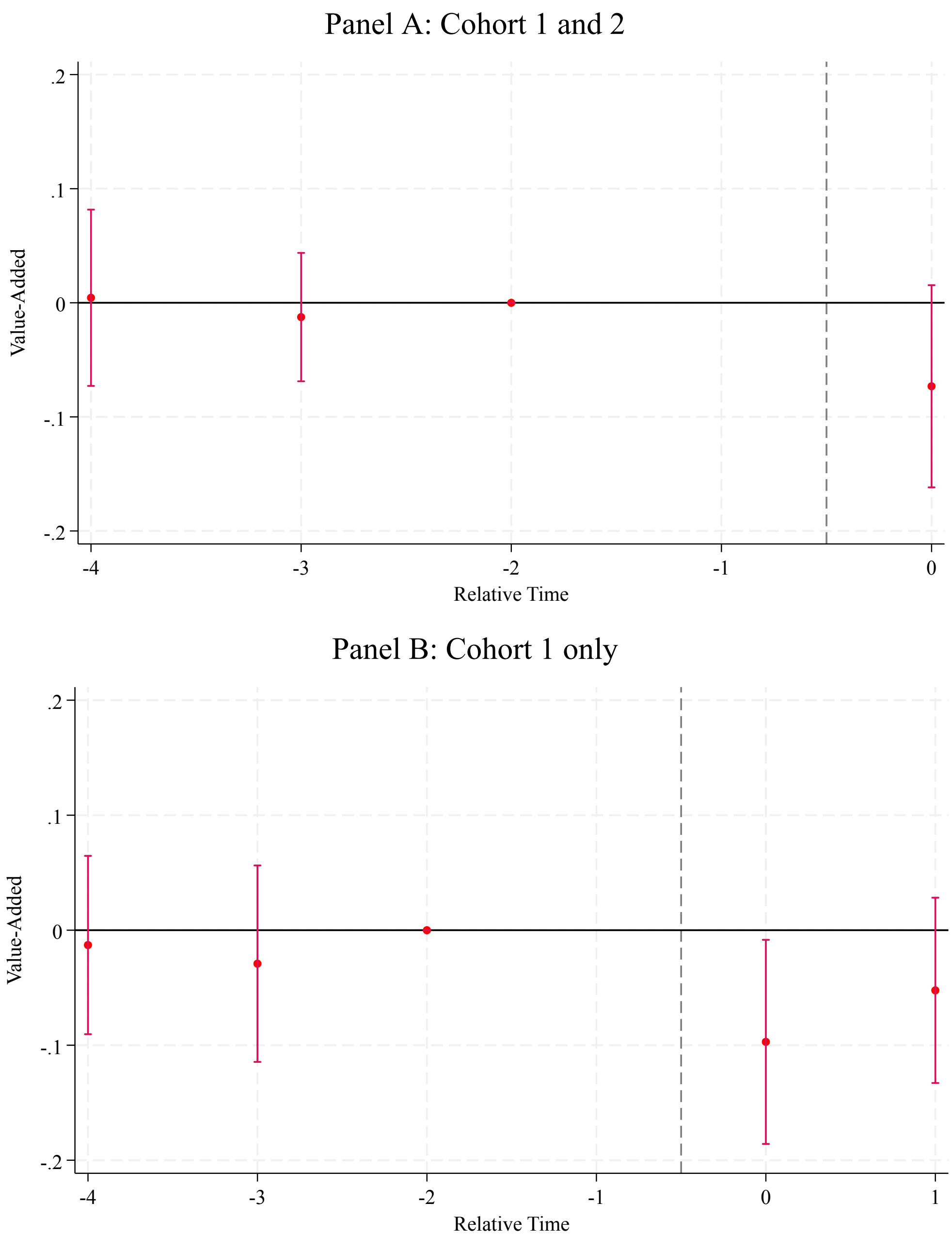

Figure 2. Event-study estimates depicting the relative change in value-added for incentivized-transfer teachers compared to control vacancy-filling teachers prior to and one year after filling open positions in low-achieving receiving schools.
Notes: Observations are teacher-by-subject-by-year in an unbalanced panel. Standard errors are clustered within teacher.
Source: Moving High-Performing Teachers restricted dataset, U.S. Department of Education, National Center for Education Statistics.

Appendix A: Additional Tables

Table A1. Mean Characteristics of Students in Study

	All Teachers			Vacancy-Filling Teachers		
	Incentive	Control		Incentive	Control	
Prior Achievement						
Math pre-test score	-0.44	-0.47		-0.45	-0.49	
ELA pre-test score	-0.55	-0.54		-0.56	-0.55	
Student Characteristics						
Male	0.51	0.51		0.51	0.50	
White	0.05	0.05		0.08	0.05	
Black	0.28	0.45	**	0.30	0.47	**
Hispanic	0.64	0.45	**	0.57	0.42	*
FRPL	0.84	0.81		0.80	0.82	
ELL	0.26	0.21		0.23	0.18	
Special Education	0.08	0.09		0.09	0.10	
N (students)	21,160	16,430		6,140	6,680	

Notes: Characteristics of elementary and middle school students in program year 1, by teachers' treatment designation. Missingness varies across variables. Standard errors are clustered at the teacher team level. Sample sizes rounded to the nearest ten. ELA = English language arts.

Source: Moving High-Performing Teachers restricted dataset, U.S. Department of Education, National Center for Education Statistics.

Table A2: Tests of Differential Student Sorting Across Incentivized-Transfer Teachers vs. Incentive Incumbent Teachers

	Prior Math	Prior ELA	Male	White	Black	Hispanic	FRPL	ELL	SPED
Incentive New Hire	-0.009	0.018	0.010	0.038*	0.018	-0.081	-0.035	-0.062	0.015
	(0.058)	(0.062)	(0.008)	(0.021)	(0.053)	(0.057)	(0.028)	(0.039)	(0.010)
N	18020	17960	20880	20760	20760	20760	20720	20510	20990

Notes: Standard errors in parentheses clustered at the teacher level. Sample sizes rounded to the nearest ten. ELA = English language arts. *p<0.1, **p<0.05, ***p<0.01.

Source: Moving High-Performing Teachers restricted dataset, U.S. Department of Education, National Center for Education Statistics.

Appendix B

Value-Added Model[9]

As a first step of their study, the TTI team identified the highest-performing teachers in each district by calculating value added to student achievement using up to three years of test-score data from state assessments. Value added represents the amount of learning growth that can be attributed to the teacher, holding constant the factors outside the teacher's control.

Value-added estimates were calculated either by the participating districts, with the assistance of outside vendors, or by the study team. Of the ten participating districts, seven provided raw data on student achievement, demographics, and enrollment to link students to teachers, with which the study team estimated teachers' value added. The remaining three districts used outside vendors for the value-added estimation. One of these districts supplied a list of the top-performing teachers instead of value-added estimates. This district is thus dropped from our sample.

The study team's method, described below, was not a duplicate of the methods used by districts with outside vendors. Instead, their goal was to estimate a model that could possibly have been adopted by districts in future implementations of interventions such as TTI. While this approach poses a threat to the internal validity of our research, the study team examined the main impacts separately for the districts that used different estimates[10] and did not find differences in the results.

Estimation Equation

The value-added model employed for the study took the following form:[11]

$$Y_{ijt} = \lambda_{t-1} Y_{ij,t-1} + \alpha_1 X_{ijt} + \alpha_2 Z_{jt} + \beta_j D_{ijt} + e_{ijt} \ , \tag{1}$$

where, Y_{ijt} is the post-test score for student i with teacher j in year t; $Y_{ij,t-1}$ is the pre-test score for that same student, which captures previous inputs into student achievement; X_{ijt} is a vector of student-level control variables that includes gender, race/ethnicity, disability type, and FRLP, ELL, special education, grade repetition, and over-age-for-grade status; Z_{jt} is a vector of teacher-level variables that includes the percentage of a teacher's students who were mobile, who were grade repeaters, class size, and grade-by-year dummies; D_{ijt} is a vector of dosage (the percentage of the year student i in year t was taught by teacher j) that includes separate values for each teacher-year; and e_{ijt} is the error term. The performance measures are contained in the vector β_j, which is the set of coefficients of D_{ijt}.

---

[9] The explanation of the value-added analysis conducted by the TTI study team in order to select top-performing teachers from districts can be found in Appendix B, page 237, of the Glazerman et al. (2013) report.

[10] Main outside vendor used by districts was SAS Institute

[11] Value-added scores were estimated individually for elementary school teachers, middle school math teachers, and middle school ELA teachers. Observations of teachers in a given year linked to fewer than five students' test scores were dropped from the estimation sample. Students that spent less than 20% of the school year with a teacher were also excluded from the estimation sample for that teacher.

After initial estimation of equation (1), subject-specific performance measures were standardized within each grade level.

Controlling for Measurement Error in the Pre-Test

A two-stage procedure to correct for measurement error in the pre-test was conducted before estimating equation (1). The first stage is the estimation of an errors-in-variables regression model by using the average reliability of the test across grades and years to remove bias caused by the measurement error in the pre-test. The errors-in-variables regression[12], estimated with the reliability for each test, when available, is:

$$Y_{ijt} = \lambda_{t-1} Y_{ij,t-1} + \alpha_1 X_{ijt} + \beta_j D_{ijt} + e_{ijt} \tag{2}$$

where, the student-level control variables are the same as in equation (1). With $\hat{\lambda}_{t-1}$, the estimated value for the coefficient of the pre-test, the estimated adjusted gain for each student in each year is calculated by:

$$\hat{G}_{ijt} = Y_{ijt} - \hat{\lambda}_{t-1} Y_{ij,t-1} \tag{3}$$

The second-stage regression model pools data from all years and uses the adjusted gain as the dependent variable[13]:

$$\hat{G}_{ijt} = \alpha_1 X_{ijt} + \alpha_2 Z_{jt} + \beta_j D_{ijt} + e_{ijt} \tag{4}$$

Equation (1) is arrived at by substituting equation (3) into (4), rearranging terms, and treating $\hat{\lambda}_{t-1}$as λ_{t-1}. This method underestimates the standard errors of β_j because it treats $\hat{\lambda}_{t-1}$as λ_{t-1}. If λ_{t-1} is estimated precisely, the understatement in the standard errors is negligible.

Shrinkage Estimator

After estimating equation (1) a shrinkage procedure to calculate empirical Bayes performance measures and standard errors is applied. With this procedure, the empirical Bayes estimate for each performance measure is approximately the precision-weighted average of the original measure and the mean of all point estimates. The empirical Bayes shrinkage adjusts estimates by placing relatively more weight on the mean of all point estimates when the individual estimate has a high standard error. This approach corrects for potential variance in precision of value added across teachers.

---

[12] The model is implemented in Stata using the *eivreg* command.

[13] The correlation in outcomes for students in different years is accounted by using robust standard errors.

Appendix C

We estimate the predictive validity of value-added in the context of the TTI experiment in an approach that adapts Kane et al. (2013) and replicates Glazerman and Protik (2015). The TTI experiment created an exogenous shock to the value-added measures of teachers who filled vacancies on teams that were assigned to the incentive condition. We exploit the between-school variation in value-added induced by the randomization process to address the selection challenge when estimating the predictive power of value-added for the performance of teachers' current students in a 2SLS framework. Teams assigned to the incentive condition had an increased probability of filling a vacancy with a high-performing transfer teacher.

In our first stage, we estimate:

$$VA_j^{preavg} = \varphi Y_{ijs,t-1} + \alpha X_{ijst} + \beta_1(Incentive\ Team_s) + \tau_b + \varepsilon_{ijst} \quad (1)$$

where we model the average pre-period value-added for all teachers on incentive and control teams as a function of prior student achievement for student i with teacher j in grade-level team s in year t ($Y_{ijs,t-1}$), as well as a vector of student demographic characteristics (X_{ijst}) including gender, race, eligibility for free or reduced-price lunch, English language learner status, special education services, and having been retained in grade. $Incentive\ Team_s$, an indictor for whether a student is on a team randomly assigned the opportunity to hire high-performing teachers with incentive pay, serves as our exogenous instrument. τ_b represent fixed effects for randomization blocks.

In our second stage, we estimate:

$$Y_{ijst} = \beta_2 \widehat{VA_j^{preavg}} + \delta Y_{ijs,t-1} + \theta X_{ijst} + \tau_b + \epsilon_{ijst} \quad (2)$$

Where student achievement (Y_{ijst}) for student i with teacher j in grade-level team s in year t is a function of the same covariates as in equation (1) and predicted average pre-period value-added for all teachers generated in first stage model (VA_j^{preavg}). We stack subjects to maximize power and include a control fixed effect. We cluster our standard errors at the teacher level.

We report first-stage results for β_1 and second-stage results for β_2 Appendix Table C1.

Table C1. 2SLS Predictive Validity Test of Prior Value Added Stacked Across Subjects

	First Stage	Second Stage	P-value from t-test of β=1
	(1)	(2)	(3)
Incentive	0.054 ***		
	(0.016)		
Average Prior Value Added		-0.572	0.039
		(0.762)	
n (students)	18,000	18,000	
F-statistic			

Notes: We estimate this model at the student-subject level, with student achievement in math or English language arts as the outcome variable in the second stage. The model includes controls for students' lagged within-subject achievement and fixed or pre-treatment characteristics (race/ethnicity, gender, English language learner status, special education status, free or reduced-price lunch eligibility, gifted status, and age), as well as subject and randomization block fixed effects. Sample sizes rounded to the nearest ten. *p<0.1, **p<0.05, ***p<0.01.

Source: Moving High-Performing Teachers restricted dataset, U.S. Department of Education, National Center for Education Statistics.