Transfer Incentives for High-Performing Teachers: Final Results from a Multisite Randomized Experiment

November 2013

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

U.S. Department of Education

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

Transfer Incentives for High-Performing Teachers: Final Results from a Multisite Randomized Experiment—Executive Summary

November 2013

Steven Glazerman
Ali Protik
Bing-ru Teh
Julie Bruch
Jeffrey Max
Mathematica Policy Research

Elizabeth Warner
Project Officer
Institute of Education Sciences

NCEE 2014-4004
U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

U.S. Department of Education
Arne Duncan
Secretary

Institute of Education Sciences
John Q. Easton
Director

National Center for Education Evaluation and Regional Assistance
Ruth Curran Neild
Commissioner

November 2013

The report was prepared for the Institute of Education Sciences under Contract No. ED-04-CO-0112/007. The project officer is Elizabeth Warner in the National Center for Education Evaluation and Regional Assistance.

IES evaluation reports present objective information on the conditions of implementation and impacts of the programs being evaluated. IES evaluation reports do not include conclusions or recommendations or views with regard to actions policymakers or practitioners should take in light of the findings in the reports.

Glazerman, S., A. Protik, B. Teh, J. Bruch, J. Max. (2013). *Transfer Incentives for High-Performing Teachers: Final Results from a Multisite Experiment—Executive Summary* (NCEE 2014-4004). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

This report also is available on the IES website at http://ies.ed.gov/ncee.

Upon request, this report is available in alternate formats such as Braille, large print, audiotape, or computer diskette. For more information, please contact the Department's Alternate Format Center at 202-260-9895 or 202-205-8113.

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

ACKNOWLEDGMENTS

This study is the product of many people's efforts. We are deeply grateful to the many teachers, principals, district leaders, and central office staff whose hard work and patience made both the intervention and the research possible, although we unfortunately cannot acknowledge them by name. Spearheading the implementation of the intervention were staff from The New Teacher Project (TNTP), including Coral Jenrette, Mónica Vásquez, Emma Cartwright, Latricia Barksdale, and Kristen Rasmussen. At Mathematica, Tim Silva played an important role in overseeing implementation and working closely with TNTP and the districts. Monica Leal Priddy at Optimal Solutions Group led a team, including Kimberly Hahn, Carolina Herrera, Grace Hong, and Mark Partridge, that collected extensive school records data and played an important role in gathering data needed for program implementation.

This report also relies heavily on teacher and principal surveys. At Mathematica, Nancy Carey and Kristina Rall led the survey research effort with invaluable assistance from Theresa Boujada and her team at Mathematica's Survey Operations Center.

The evaluation team at Mathematica benefited from expert programming and research assistance from Alena Davidoff-Gore, Maureen Higgins, and Christopher Jones. John Deke, Duncan Chaplin, and Neil Seftor read and provided helpful comments on earlier versions of the report. A technical working group (TWG) provided useful input on program design and the research. TWG members included Dale Ballou, Lisa Barrow, Jason Kamras, Robert Meyer, Anthony Milanowski, Jeffrey Smith, and Jacob Vigdor. Sharon Peters edited the report, and Jackie McGee prepared it for publication.

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST[1]

The research team for this evaluation consists of a prime contractor, Mathematica Policy Research of Princeton, New Jersey, and a subcontractor, Optimal Solutions Group of College Park, Maryland. Neither of these organizations or their key staff members have financial interests that could be affected by findings from the evaluation. No one on the technical working group, convened by the research team to provide advice and guidance, has financial interests that could be affected by findings from the evaluation.

[1] Contractors carrying out research and evaluation projects for IES frequently need to obtain expert advice and technical assistance from individuals and entities whose other professional work may not be entirely independent of or separable from the tasks they are carrying out for the IES contractor. Contractors endeavor not to put such individuals or entities in positions in which they could bias the analysis and reporting of results, and their potential conflicts of interest are disclosed.

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

EXECUTIVE SUMMARY

One way to improve struggling schools' access to effective teachers is to use selective transfer incentives. Such incentives offer bonuses for the highest-performing teachers to move into schools serving the most disadvantaged students. In this report, we provide evidence from a randomized experiment that tested whether such a policy intervention can improve student test scores and other outcomes in low-achieving schools.

The intervention, known to participants as the Talent Transfer Initiative (TTI), was implemented in 10 school districts in seven states. The highest-performing teachers in each district—those who ranked in roughly the top 20 percent within their subject and grade span in terms of raising student achievement year after year (an approach known as value added)—were identified. These teachers were offered $20,000, paid in installments over a two-year period, if they transferred into and remained in designated schools that had low average test scores. The main findings from the study follow.

Main Findings

- **The transfer incentive successfully attracted high value-added teachers to fill targeted vacancies**. Almost 9 out of 10 targeted vacancies (88 percent) were filled by the high-performing teachers who had been identified as candidates eligible for the transfer intervention. To achieve those results, a large pool of high-performing teachers was identified (1,514) relative to the number of vacancies filled (81). The majority of candidates did not attend an information session (68 percent) or complete an online application to participate in the transfer intervention (78 percent).
- **The transfer incentive had a positive impact on test scores (math and reading) in targeted elementary classrooms**. These impacts were positive in each of the two years after transfer, between 0.10 and 0.25 standard deviations relative to each student's state norms. This is equivalent to moving up each student by 4 to 10 percentile points relative to all students in their state. In middle schools, we did not find evidence of impacts on student achievement. When we combined the elementary and middle school data, the overall impacts were positive and statistically significant for math in year 1 and year 2, and for reading only in year 2. Our calculations suggest that this transfer incentive intervention in elementary schools would save approximately $13,000 per grade per school compared with the cost of class-size reduction aimed at generating the same size impacts. However, overall cost-effectiveness can vary, depending on a number of factors, such as what happens after the last installments of the incentive are paid out after the second year. We also found there was significant variation in impacts across districts.
- **The transfer incentive had a positive impact on teacher-retention rates during the payout period; retention of the high-performing teachers who transferred was similar to their counterparts in the fall immediately after the last payout**. We followed teachers during both the period when they were receiving bonus payments and afterward. Retention rates were significantly higher during the payout period—93 versus 70 percent. After the payments stopped, the difference between cumulative retention of the high-performing teachers who transferred and their counterparts (60 versus 51 percent) was not statistically significant.

Background

There is growing concern that the nation's most effective teachers are not working in the schools with the most disadvantaged students (Goldhaber 2008; Peske and Haycock 2006; Tennessee Department of Education 2007; Sass et al. 2012; Glazerman and Max 2011). One strategy to remedy this situation is the use of monetary incentives to recruit teachers who have demonstrated success in raising student test scores ("value added") to teach in low-achieving schools. This strategy has been tried in some fashion in several places, such as Mobile, Alabama; Chattannooga, Tennessee; Palm Beach, Florida; and the states of California and Virginia (Max et al. 2007). It has the potential to redistribute some of the highest-performing teachers in a district from higher-achieving schools to lower-achieving schools. However, there is a need for research that addresses several questions related to such a policy, including whether teachers would be willing to transfer if offered incentives, and, if they do transfer, how their presence will change the dynamics in their new schools, how long they would stay in those schools, and whether they would improve student achievement in those schools.

The U.S. Department of Education's Institute of Education Sciences (IES) contracted with Mathematica Policy Research to study the effectiveness of an intervention that is based on this strategy. The intervention, known to participating districts as the TTI (and described in Box ES.1), offers $20,000 to the highest-performing teachers in tested grades and subjects within each district who agree to transfer into one of the lowest-achieving schools in their district and stay for at least two years. Highest-performing teachers were identified based on their value-added scores[2] because some of the study districts were already using value added as one of the measures of teacher performance and because pay-for-performance policies like TTI are likely to use value-added scores as performance measures. We used whatever value-added measure the district was using because that is what would have been used in the absence of the study. In cases where such a measure was not in use, we calculated it ourselves. Teachers were eligible to transfer through TTI if, based on value-added measures, they were among the top 20 percent of teachers in the district and were not currently teaching in the lowest-achieving schools. Teachers who were in the top 20 percent but were already teaching in the lowest-achieving schools were offered $10,000 in retention stipends to continue teaching at those schools for two years. In this final report, we cover implementation and impacts for 10 districts that agreed to participate in the study. Seven began implementation in 2009 (cohort 1); an additional three began implementation in 2010 (cohort 2). In an earlier report (Glazerman et al. 2012), we presented early implementation and intermediate impacts for the first 7 districts.

Research Questions and Study Design

In this study, we address a set of specific research questions related to this transfer incentive policy, including both implementation and impact questions:

- What can we learn from the **implementation** of TTI? Specifically, what can we learn about the timing and scale of implementation, who transfers, and from where they transfer?

[2] Value-added measures seek to describe the contribution that teachers make (the value that they add) to student achievement growth, holding constant factors outside the teacher's control, such as student background and prior learning (McCaffrey et al. 2004; Lipscomb et al. 2010).

- What were the **intermediate impacts** in schools receiving the transfer teachers (referred to as receiving schools)? Specifically, how did TTI affect the dynamics within those schools, such as the allocation of resources, staffing patterns, assignment of students to teachers and courses, and school climate?
- What was TTI's **impact on student test scores** in receiving schools?
- What was TTI's **impact on teacher retention** in receiving schools?

The impact questions relate to the effect of the transfer incentive policy relative to the absence of such a policy. In other words, we sought to measure effects relative to the outcomes that would have been realized had the school not had the opportunity to use the $20,000 incentive to fill its vacancy with a teacher designated as highest performing.

Box ES.1. How the Talent Transfer Initiative Works

The intervention is designed to proceed within each district according to the following steps. The first step is to conduct a value-added analysis of student test scores to identify the highest-performing teachers, defined as the top 20 percent based on a value-added measure of teachers in tested grades and subjects in each district. The second step is to classify schools as "potential receiving" or "potential sending" schools. Potential receiving schools are those with the lowest achievement in the district, based on school-average test scores in the most recent year, and, in some cases, rankings on school accountability. The rare exceptions that are already participating in a comparable intervention are exempted. The rest of the schools in the district are potential sending schools.

The third step is recruitment of (1) eligible high-performing teachers in sending schools, whom we refer to as "transfer candidates," and, simultaneously, (2) principals of receiving schools. The highest-performing teachers (identified in the first step) in potential sending schools are offered a series of transfer incentive payments, totaling $20,000 over two years, to transfer into and remain in one of the receiving schools in their district. The offer is made to these teachers, known as "transfer candidates," in the spring, at which point they are invited to apply to the program.

At the same time, principals of potential receiving schools are invited to an information session and asked to identify likely teaching vacancies in targeted grades and subjects. To be considered for inclusion in TTI, principals must volunteer a vacancy. Eligibility is based on grade level and subject of the vacancy. A site manager in each district helps principals fill the targeted vacancies by providing information about transfer candidates and arranging and encouraging interviews. This extra hiring support is in addition to the TTI transfer incentive.

Next, applicants must interview with and be offered a position by the receiving-school principal and then voluntarily transfer to qualify for the transfer incentive. To improve the probability of matching high-performing teachers with low-achieving schools, the implementation team works with each district to finalize offers and acceptances by early summer.

Finally, the transfer teachers participate in a half-day orientation just before the start of the school year. Because they are selected on the basis of their performance in the classroom, it is assumed that they do not require additional formal support beyond what teachers normally receive. To facilitate the transition, however, the site manager provides informal support and answers any questions throughout the two school years of the intervention period. TTI teachers who fill study-assigned vacancies receive their first incentive payment after the orientation, and those who remain during the intervention period in the positions into which they transferred receive incentive payments in December and June, for a total of $20,000.

Teachers who are identified as highest-performing but who are already teaching in low-achieving (potential receiving) schools are not eligible to transfer, but they are offered a retention stipend of $10,000 for staying at their schools over the same two-year period as transfer teachers.

The methods for answering these questions included descriptive tabulations (for implementation questions) and causal analysis (for impact questions). The causal analysis relied on an experimental design in which we used random assignment to form equivalent groups of classrooms with and without the intervention to compare outcomes after one and two years. The units we assigned were teacher teams, which we defined as all teachers in a specific grade and subject. In elementary schools, the teacher team consisted of all classroom teachers in the grade. In middle schools, the teacher team consisted of all the math or English/language arts (ELA) teachers who taught at least one class in the grade level of interest. For example, all teachers responsible for teaching 7th-grade math in the same school made up one team. All teachers in the school who were responsible for 8th-grade ELA were considered another team. Thus, teacher teams were based on the grade span (elementary and middle school) as well as the subject (math or reading) the teachers were teaching. Because our unit of assignment was teacher teams, we present impact estimates by grade span and subject.

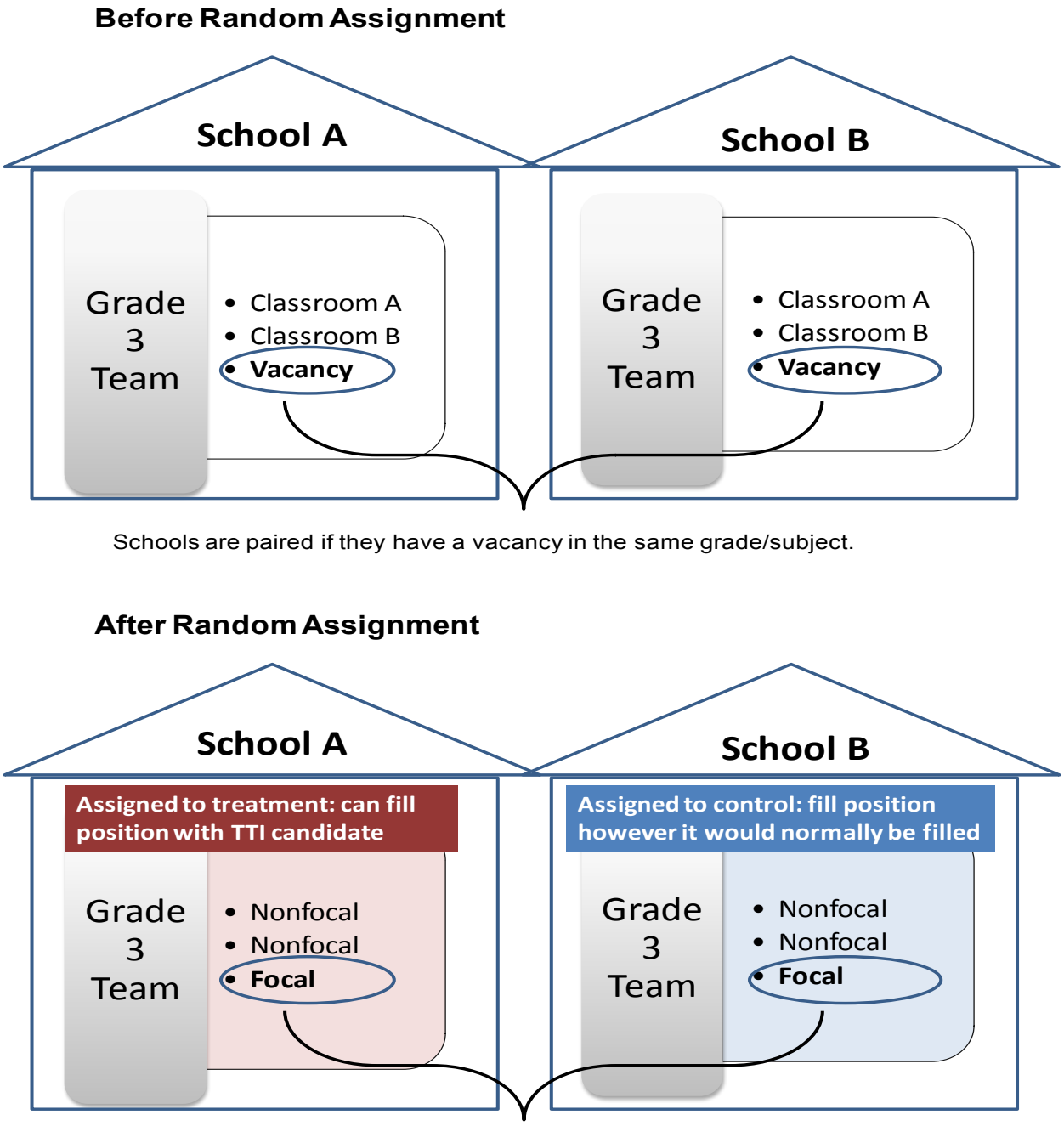
Random Assignment

The research team randomly assigned study subjects to a treatment or control group in the following way. First, we identified low-achieving schools that had a vacancy within a teaching team. If we learned of multiple eligible teacher teams in the same district at approximately the same time, we matched schools with vacancies in the same grade (and subject, in the case of middle school teams) within the same district. When possible, we also matched schools with vacancies based on their student achievement ranking and the percentage of students eligible for free or reduced-priced lunch (FRL). These matched schools formed blocks, and we then randomly assigned teacher teams within each block to either treatment status (with the opportunity to fill the team's vacancy with a TTI teacher) or control status (in which vacancies were filled through whatever process the school would normally use). For example, consider two schools, A and B, each of which had one vacancy in the grade 3 teacher team (see Figure ES.1). The grade 3 teacher team in school A was randomly assigned to treatment and is eligible to fill its vacancy through TTI; the grade 3 teacher team in school B was consequently assigned to control status, so normal hiring practices were to be followed. Teams that could not be assigned in pairs were assigned in blocks containing an odd number of teams. This situation arose either because pairs of vacancies were not available at about the same time or because a close match—in terms of student achievement ranking and/or percentage of students eligible for FRL—could not be found among schools with vacancies.

Repeated for many blocks, this process created two groups of teacher teams that were, on average, similar in terms of student characteristics and school context. The only systematic difference between the two groups was whether, in hiring for the vacancy, there was the opportunity to use the TTI policy with the associated $20,000 transfer incentive. Comparing outcomes for these groups will generate unbiased estimates of the impact of TTI on student achievement and other outcomes.

Each teacher team included the teacher who filled the vacancy as well as his or her grade-level colleagues at the same school. We expect much of the impact of TTI to operate through the teachers who filled the vacancies in the treatment and control teacher teams. We refer to them as "focal" teachers. Therefore, in addition to the team-level analysis, we are interested in the comparison between focal treatment and focal control teachers. We refer to the other teachers on a given grade-level team as "nonfocal" teachers.

Figure ES.1. Random Assignment Study Design



Data Collection

To gather data for the study, we administered surveys of teachers and principals and collected administrative records from schools and districts. We also gathered information from the program implementation process. For cohort 1, the report covers two program years, 2009–10 and 2010–11; for cohort 2, the report includes information from the first program year only, 2010–11. Data for the study are summarized below.

Candidate survey. High-performing teachers who were eligible to apply to TTI were designated as TTI transfer candidates. We surveyed candidates during the first program year. The survey asked about their background; the factors affecting their decision to apply for, interview for, and transfer into TTI positions; and their experiences in the hiring process, if applicable. The response rate was 81 percent.

Teacher background survey. All teachers on study teacher teams—focal and nonfocal—were surveyed during the first program year. The survey asked about their background, their experiences at study schools, and other factors that might affect their students' achievement.[3] The response rate was 77 percent.

Principal survey. Principals of the receiving schools were surveyed in the spring of both program years for cohort 1, and spring of the first program year only for cohort 2. The survey asked about teacher recruitment and hiring, principals' assessments of newly hired teachers, redistribution of resources across classrooms, and the school environment. The response rate was 90 percent in the first program year (cohorts 1 and 2) and 82 percent in the follow-up year (cohort 1 only).

Teacher rosters. Study schools provided teacher rosters in the fall of program years 1 and 2, as well as in the fall after the program ended for cohort 1. The rosters included information about each teacher's school and teaching assignment, and they were used to estimate teachers' retention rates. We obtained rosters from 100 percent of the schools.

Student achievement record. Districts provided student test scores linked to teachers and student demographic data in the fall after program year 1. Cohort 1 districts provided similar data in the fall after program year 2. These were used to estimate impacts on student achievement. We obtained achievement data on 100 percent of the schools in the study.

TTI program implementation records. Districts provided information related to teachers' value-added performance that was used to identify transfer candidates and retention bonuses. Some districts provided more information than others; the research team used all available data. In cases where Mathematica conducted the value-added analysis, we had detailed information on student-teacher links and student background characteristics to estimate value-added scores. In cases where value-added analysis was conducted by a third-party vendor hired by the district, we were given value-added scores directly, or, in one case, just names of highest-scoring teachers. Districts also provided data on school-level student achievement that was used to determine which schools were eligible to participate as potential receiving schools. The TTI site managers provided principal consent forms and information on the timing of teaching vacancies and when they were filled.

Study Sample

We selected school districts that were large and economically diverse. They had to have no fewer than 40 elementary schools, at least 10 of which had to be low-poverty schools and at least 15 of which had to be high-poverty schools. Low- and high-poverty schools were defined as

[3] Survey questions on teacher background information were the same for the candidate and the teacher background survey. Transfer teachers who responded to the candidate survey were not asked about their background in the teacher background survey.

having less than 40 percent or more than 70 percent of students eligible for FRL, respectively. In addition to the quantitative criteria, we selected districts according to a variety of qualitative factors related to the feasibility of implementation, including availability of test scores, data quality, hiring/transfer practices, and the local political environment. The resulting set of districts was not a random sample of a well-defined population of districts, so findings from this study cannot necessarily be generalized to other districts.

We excluded school districts in which existing or planned teacher-incentive programs would have duplicated the intervention under study, but we did come across some existing performance-incentive initiatives in some of the 10 participating school districts. In each case, we determined that the existing programs were different enough, isolated to a few schools that could be excluded from our study, or involved small enough dollar amounts that they would not interfere with the study design. Teachers and schools receiving more than $5,000 were excluded so as to avoid complicating the study by changing the effective differential in the TTI transfer incentives relative to the counterfactual. The $5,000 threshold we established was based on information in the literature on teacher responsiveness to pay (Max et al. 2007) that suggests this amount would plausibly influence teacher behavior.

Working with each district, the implementation team divided the elementary and middle schools into potential sending or potential receiving schools according to academic ranking. Schools were ranked by their students' average prior achievement level, which was determined by the previous three years of achievement data or by the past year's achievement data, depending on the district leaders' preferences.[4] The lowest-ranking schools were designated as potential receiving schools that could benefit from the intervention, and the rest were potential sending schools. We removed some schools from both pools and referred to them as exempt schools because they served a special population of students or were already implementing a program that was meant to address the problem that TTI aims to address. In the end, 21 percent of the schools were classified as potential receiving schools, 72 percent were potential sending schools, and 7 percent were exempt.

The study focused on teachers in a subset of these potential sending and receiving schools. From the potential sending schools, we surveyed the teachers who were identified as highest-performing in the district and eligible for TTI. From the potential receiving schools, we collected data on "study schools," those with teaching teams that had been randomly assigned to treatment or control status. There were 114 study schools, which represents 56 percent of the potential receiving schools.

The final sample had the following features:

Study districts. Ten school districts participated in the study, contributing both elementary and middle schools except for 3 that contributed only elementary schools or only middle schools. Six of the 10 districts were countywide, encompassing urban and nonurban areas. The districts ranged in size from just under 100 square miles to more than 1,200 square miles, which is larger than the state of Rhode Island.

[4] Achievement data from the year before the implementation of TTI were used for all but two districts, where three prior years of achievement data were used.

Study schools. Across the 10 districts, 114 of the potential receiving schools had teams that were randomly assigned to treatment or control status. The average study school was 80 percent low income (FRL).

Teacher teams. Teacher teams in the study ranged from 3rd grade through 8th grade. Some teams included more than one vacancy, and some schools included more than one team. Eighty-five teams were assigned to participate in the intervention and 80 teams were assigned to the control group. Note that we randomly assigned teacher teams within blocks and because we had some blocks with an odd number of teacher teams, the process of random assignment generated by chance an unequal number of treatment and control teams.

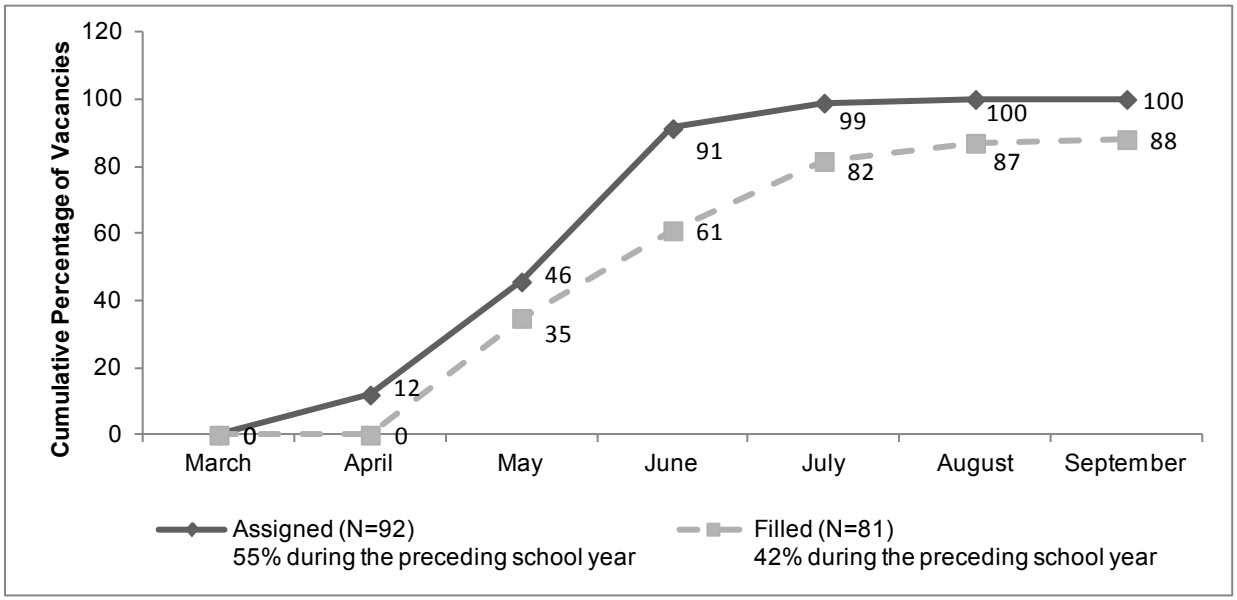
Students. The students of teacher teams in the study were low achieving and disadvantaged. In the year before implementation, the students on study teams performed at approximately the 32nd and the 33rd percentile on state standardized tests in reading and math, respectively, compared with other students in their state.

Implementation Findings: What Can We Learn from the Implementation of TTI?

The implementation of TTI can offer insights regarding the use of transfer incentives to redistribute a district's highest-performing teachers. Here, we summarize the key findings about how vacancies were filled normally and under TTI, how teachers responded to the transfer incentive, where they transferred from, and the resulting treatment contrast in teacher background. We focus especially on the teachers who filled vacancies targeted by TTI, teachers we refer to as "focal," as shown in Figure ES.1. We also consider possible effects on nonfocal teachers, defined as the focal teachers' peers in the same grade and subject.

Almost all vacancies in treatment schools were filled by TTI teachers, although a large pool of candidates was used to yield the desired number of successful TTI transfers. The implementation of TTI demonstrated that it is possible to implement a transfer-incentive program as designed for this study. The highest-performing teachers were identified in approximately the first three months of the calendar year using value-added analysis. Beginning as early as March, the implementation team, consisting of district personnel working with staff from The New Teacher Project (TNTP), conducted several months of intensive recruitment of receiving schools and transfer candidates, resulting in 88 percent of the treatment school vacancies being filled with TTI teachers. In Figure ES.2, we show that most treatment vacancies were assigned and filled in May and June. However, an initial pool of 1,514 candidates was identified to yield the 81 who ultimately transferred. Thus, an average of 5 percent of each district's highest-performing teachers in sending schools ultimately transferred to low-achieving schools. Most did not even attempt to transfer: 32 percent of eligible TTI candidates attended an information session, leaving 68 percent who did not attend; 22 percent completed an application, leaving 78 percent who did not. Fifty-five percent of the applicants interviewed for at least one vacancy, and the other 45 percent either did not follow through or were not given an opportunity to interview. Principals in treatment schools conducted an average of 3.1 interviews per vacancy, and most principals with treatment teams made an offer to only one TTI candidate.

Figure ES.2. Percentage of TTI Vacancies Assigned and Filled, by Month



Source: TTI program records.

Standard practice in the absence of treatment was to fill vacancies through a combination of new hires, transfers in, and within-school reassignments. Nineteen percent of control group vacancies were filled by teachers new to the district, 22 percent by teachers transferring from another school in the district, and 30 percent by teachers reassigned within the school. No TTI transfer candidates transferred to any of these vacancies. The average teaching experience of control focal teachers was eight years, reflecting the fact that many were experienced teachers who simply moved from elsewhere in the school or district and were not hired out of the pool of novice teachers. Only 17 percent reported being new to teaching, whereas 45 percent reported being in at least their sixth year of teaching.

The TTI teachers were more experienced than teachers who would normally fill the vacancies. The average difference in teaching experience between treatment and control focal teachers was about four years. There was also a significant difference in the percentage of teachers with National Board Certification (20 percent of focal treatment teachers compared with 9 percent of focal control teachers).

Intermediate Impacts: How Did TTI Affect School Staffing Assignments and Resource Allocation?

The opportunity for a school to fill a teaching vacancy with one of the district's highest-performing teachers could lead to changes in a range of behaviors, such as student or teaching assignment within the teaching team and the school. These changes can be considered intermediate outcomes, effects on other teachers' behavior and the allocation of resources within the teacher team or school. Understanding these intermediate impacts is important for explaining how TTI influences the internal dynamics of schools as well as for interpreting the impacts on student achievement and teacher retention.

The evidence on the strategic assignment of students to teachers as a result of TTI was mixed. We hypothesized that one way principals could react to an intervention like TTI was to

assign more challenging students to the transfer teachers on the assumption that their high performance meant transfer teachers were more capable of teaching struggling students. We did not find evidence of impacts on student assignment when we examined administrative data that described the characteristics of students assigned to treatment and control focal teachers relative to their nonfocal counterparts. We also found no evidence of impacts on student assignment when we examined principals' reports of how they assigned students to classrooms in treatment versus control teams. However, treatment focal teachers in middle schools were more likely than their peers in control teams to say they had more academically challenging students. Control focal teachers did not report differences between their students and their peer teachers' students.

We found evidence of reassignment of teachers across grades due to TTI. Another way that TTI could alter the school's internal dynamics is through resource allocation across grades. Under the status quo, principals might compensate for weak incoming teachers by moving strong peers from elsewhere in the school into their grade team. TTI may have the opposite effect: principals might move weak teachers into the grades with TTI teachers.

Using experience as an indicator of teacher quality, we found support for this hypothesis of pairing high-performing teachers with inexperienced ones. Among nonfocal teachers who had moved within their schools, the control movers were more experienced than treatment movers (by a statistically significant difference of nearly five years).

TTI teachers used less mentoring and provided more mentoring than their counterparts. We found that treatment focal teachers were less likely to receive mentoring (39 versus 59 percent had a mentor) and more likely to provide mentoring than control focal teachers (15 versus 5 percent provided mentor support). We defined a mentor as "someone who provides professional advice and direct assistance to classroom teachers." Because teachers in TTI positions were using fewer mentoring resources than the new hires on control teams, more resources were potentially available in TTI schools for supporting other teachers. However, changes in mentoring services used and provided by focal teachers were not offset by equal and opposite changes for nonfocal teachers. This opens the possibility that resource-allocation effects could spread to the larger school community, beyond just the teacher teams for which we collected data.

Test-Score Impacts: Did TTI Raise Student Achievement?

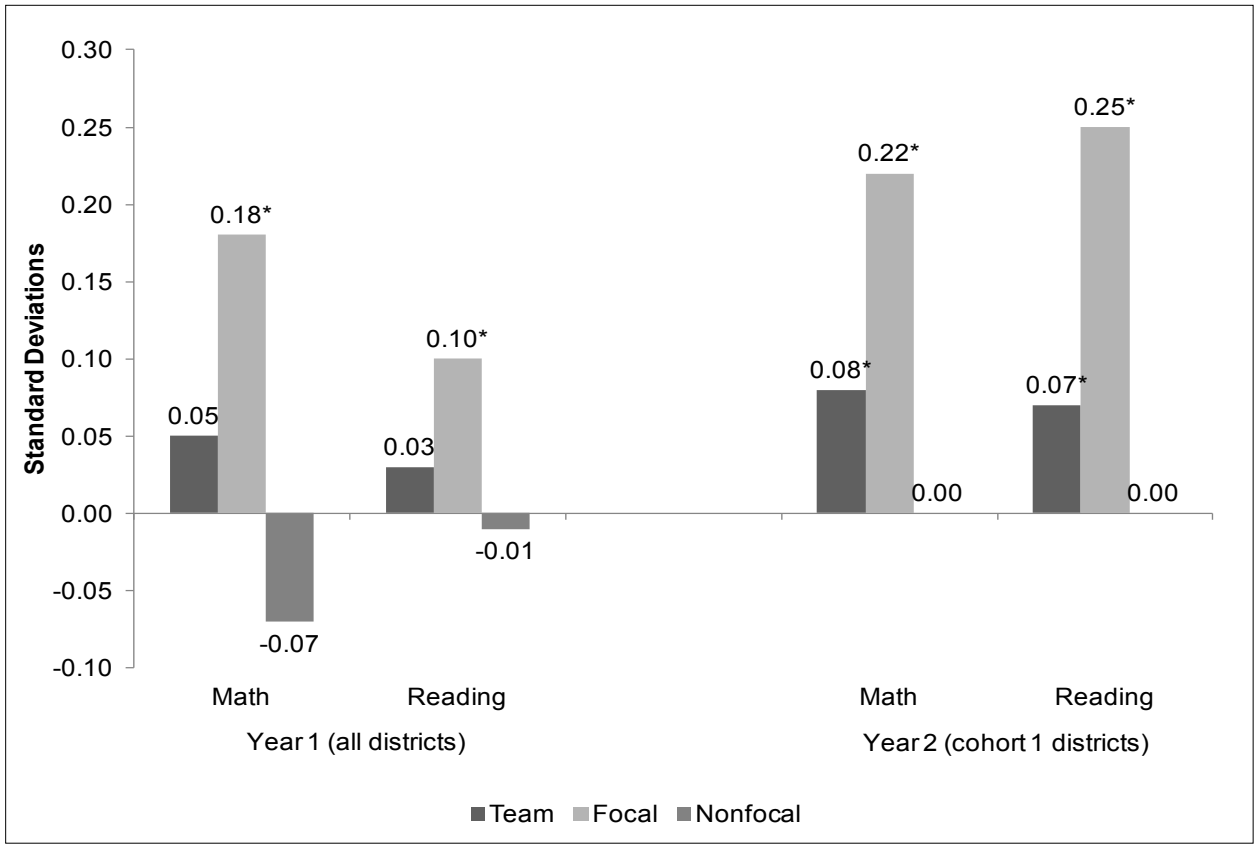
To estimate the impact of TTI on student achievement, we compared the test-score performance of students from treatment teams to the corresponding performance of students from control teams because the teacher team is the unit of random assignment. However, we expected much of the effect to be captured directly by comparing the performance of students who were taught by teachers who filled treatment vacancies (treatment focal teachers) against the performance of students who were taught by teachers who filled control vacancies (control focal teachers).

It is possible that the presence of a TTI teacher could have an effect on the team composition or the performance of other team members (nonfocal teachers). Therefore, we also report both results of the corresponding comparisons between treatment focal and control focal teachers and between treatment nonfocal and control nonfocal teachers within those teams. Program year 1 impacts were estimated using data from all 10 districts; year 2 impacts were estimated based on

cohort 1 districts because second-year follow-up data for the three cohort 2 districts were not collected.

TTI elementary school teachers had positive impacts on test scores. Comparing teams as a whole (see Figure ES.3), there were positive impact estimates for both subjects, but they were significant only in the second year of implementation: 0.08 standard deviations for math and 0.07 standard deviations for reading. The impact estimates for focal teachers ranged from one-tenth to one-quarter of a standard deviation, depending on subject and implementation year, and were positive and statistically significant for both subjects in both implementation years. The impacts on nonfocal teachers were not significantly different from zero in either year or subject. This suggests that the TTI teachers have minimal or no effect on their colleagues' performance.

Figure ES.3. Test-Score Impacts in Elementary Schools



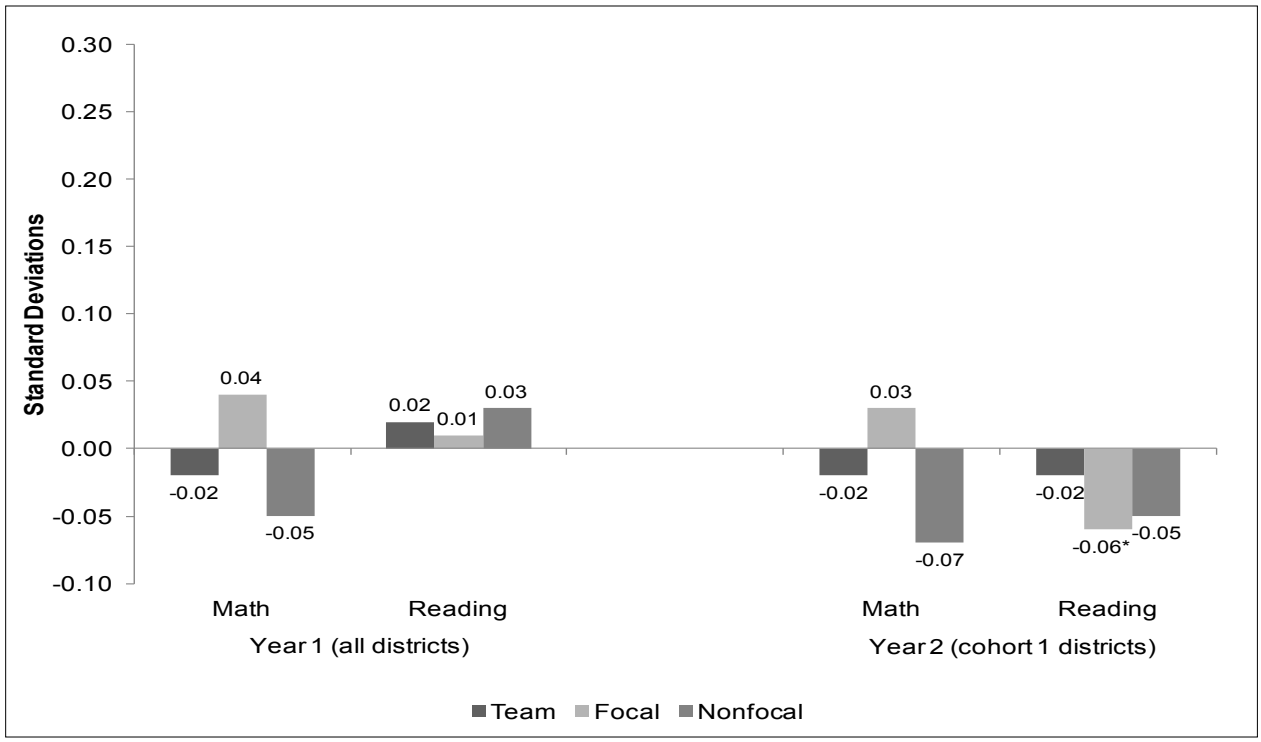
Source: District administrative data.

Note: A team consists of all classroom teachers in the grade and subject for a school. Focal teachers are those who filled study vacancies. Nonfocal teachers are the rest of the teachers on the team.

*Statistically significant at the 0.05 level, two-tailed test.

We did not find evidence that TTI was effective in middle schools. The results are shown in Figure ES.4. The impact estimates were all statistically insignificant for program years 1 and 2 except for the year 2 focal teacher impact on reading, which was negative (impact = -0.06, *p*-value = 0.031). This finding may be a middle school phenomenon or, as will be discussed next, it may be a result of the particular districts where middle schools were most heavily represented.

Figure ES.4. Test-Score Impacts in Middle Schools



Source: District administrative data.

*Statistically significant at the 0.05 level, two-tailed test.

Impacts varied by district. The impact estimates for the team comparisons and the focal teacher comparisons varied across districts more than would be expected with sampling variation if TTI were equally effective in all sites. For example, elementary team-level district-specific impacts on math in program year 1 ranged from -0.25 to 0.48 of a standard deviation; elementary focal teacher district-specific impacts on math in program year 1 ranged from -0.15 to 0.57 of a standard deviation in program year 1. The district-specific impacts are based on small samples and most are not statistically significant, but their distribution suggests that neither the team nor focal teacher impacts are driven by results from one or two outlier districts. Also, the math and reading impacts by district are positively correlated.

These results suggest that although the variation in impacts across different grade spans may be due to real differences between elementary and middle schools, they may also be partially driven by differences in district-specific impacts because the shares of elementary and middle school teams differed across districts. One district contributed only elementary school teams to the study; two districts contributed only middle school teams; the remaining seven districts had different mixes of elementary school and middle school teams.

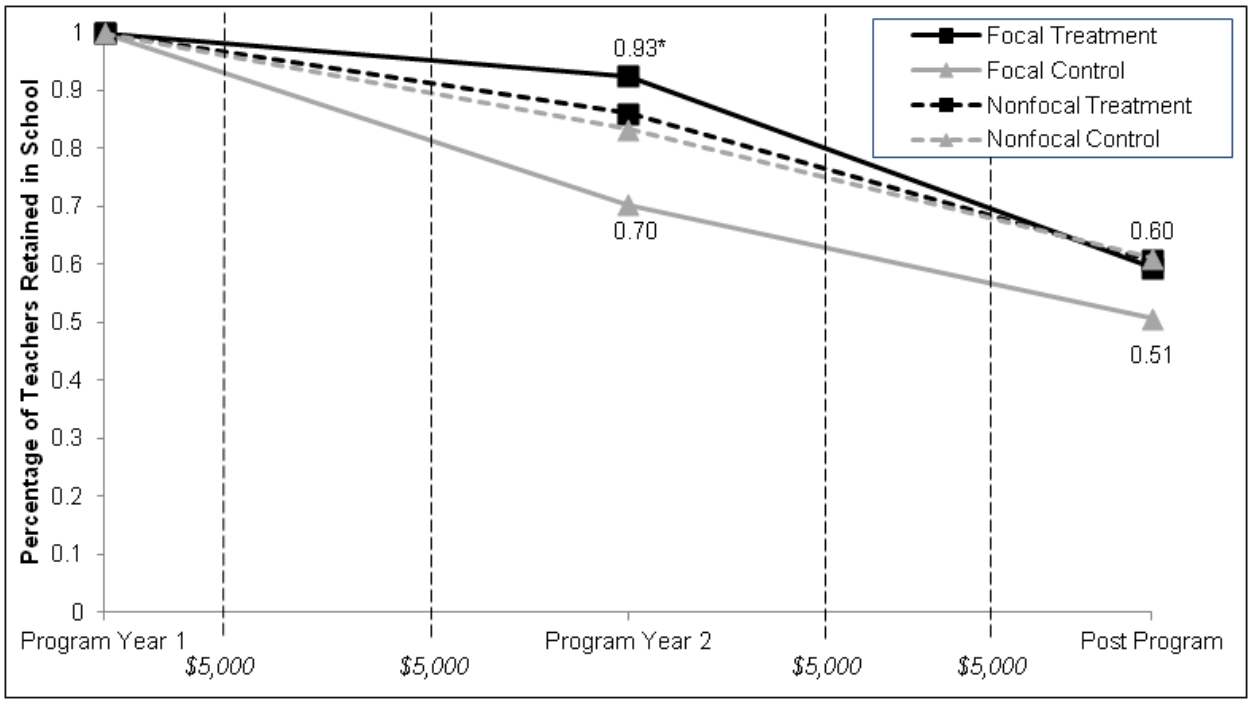
Retention Impacts: Did TTI Keep High-Performing Teachers in Their New School?

TTI teachers were offered $20,000 in five installments over a two-year period to transfer to and continue teaching in a low-achieving school in their district. Above, we discussed teachers' initial responses to the offer of a transfer incentive. Another important question for policymakers is whether the incentive would be sufficient to keep teachers at the schools into which they transferred. We examined the rates of teacher retention on treatment and control teams using teacher rosters collected during the program—while TTI teachers were still receiving incentive payments—and after the incentive payments ended.

Over the two years of the intervention, retention rates were higher for TTI teachers than for their counterparts. This finding is based on cohort 1 districts, in which impacts can be estimated after one and two years of the program (Figure ES.5). After the first year, when TTI teachers were still receiving payments for remaining at their schools, the difference in school retention between treatment and control focal teachers was 22 percentage points.[5] During the same period, the difference in school retention between treatment and control nonfocal teachers was not statistically different from zero. We observed a similar pattern for focal and nonfocal comparisons in the full sample of cohort 1 and 2 districts after the first year.

Retention rates after the TTI intervention ended were not statistically different. We used the cohort 1 districts to measure longer-term retention. After the second (last) program year, in the fall after TTI transfer teachers had received their final stipend payment, about 60 percent of the original treatment focal group returned to the same school, which was statistically indistinguishable from the control focal teachers (Figure ES.5, Post Program). The end result was that TTI transfer teachers did not leave their schools—to return to their original schools or to transfer to any destination—at a rate that was higher than any teacher hired into such a school.

Figure ES.5. Impacts on Retention in School, Cohort 1 Districts Only



Source: School rosters.

Note: Vertical dotted lines represent points at which TTI teachers received payments. Note that the first $5,000 payment was actually paid in two installments of $2,500. One installment was paid before the start of the first school year and the other was paid in the fall of the first school year.

N = 80 focal treatment teachers, 96 focal control teachers, 193 nonfocal treatment teachers, and 183 nonfocal control teachers.

*Statistically significant at the 0.05 level, two-tailed test.

[5] This impact estimate is calculated as the difference between the unrounded treatment and control means. The treatment and control means (0.93 and 0.70, respectively) presented in Figure ES.5 are rounded means.

Cost-Effectiveness

We showed that in at least some settings TTI had positive impacts on test scores, and after the two-year study period when the payments ended, treatment group teachers had not all left, but returned to their schools in year 3 at rates that were similar to their control group counterparts. Thus, the question arises as to whether the impacts are large and meaningful enough to offset the costs of generating them.

To provide a point of comparison, we compared the cost of generating the impacts of TTI with the costs of generating similar impacts if we were to implement an alternative policy, such as reduction in class size. We first estimated the incremental cost per team of implementing TTI. Then we estimated what it would have cost in terms of class-size reduction to generate those same impacts using results from the Tennessee STAR Class Size Reduction (CSR) experiment. We then subtracted the actual costs incurred per team from the estimated costs per team that would have been incurred using CSR to generate the same impacts as TTI. A positive number suggests that TTI was cheaper, and, therefore, more cost-effective.

The results of this cost-effectiveness analysis suggest that the impacts of TTI found for elementary schools would make it the cheaper alternative, compared to CSR, by $13,154 per team. Considering the long-term benefits of having high-performing teachers remain in the school could make TTI the cheaper alternative by an estimated $40,043 per team.

We found less-favorable results, however, in middle schools and in selected districts, so we cannot say that the cost comparison will always favor TTI if replicated. When we repeated the cost-effectiveness analysis based on replicating the full intervention in both elementary and middle schools, we found differences in the cost of TTI and the alternative intervention to be both negative and positive, depending on the assumptions we made. This implies that it is unclear whether TTI is more cost-effective overall for all grades and districts.


DEPARTMENT OF EDUCATION
UNITED STATES OF AMERICA