NCEE 2009-4073

U.S. DEPARTMENT OF EDUCATION

Impacts of Comprehensive Teacher Induction

Results from the Second Year of a Randomized Controlled Study

Executive Summary

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE

Institute of Education Sciences

Impacts of Comprehensive Teacher Induction

Results from the Second Year of a Randomized Controlled Study

Executive Summary

August 2009

Eric Isenberg
Steven Glazerman
Martha Bleeker
Amy Johnson
Julieta Lugo-Gil
Mary Grider
Sarah Dolfin
Mathematica Policy Research

Edward Britton
WestEd

Melanie Ali
Project Officer
Institute of Education Sciences

NCEE 2009-4073
U.S. DEPARTMENT OF EDUCATION

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

U.S. Department of Education
Arne Duncan
Secretary

Institute of Education Sciences
John Q. Easton
Director

National Center for Education Evaluation and Regional Assistance
Phoebe Cottingham
Commissioner

August 2009

The report was prepared for the Institute of Education Sciences under Contract No. ED-04-CO-0112/001. The project officer is Melanie Ali 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.

To order copies of this report,

- Write to ED Pubs, Education Publications Center, U.S. Department of Education, P.O. Box 1398, Jessup, MD 20794-1398.
- Call in your request toll free to 1-877-4ED-Pubs. If 877 service is not yet available in your area, call 800-872-5327 (800-USA-LEARN). Those who use a telecommunications device for the deaf (TDD) or a teletypewriter (TTY) should call 800-437-0833.
- Fax your request to 301-470-1244.
- Order online at www.edpubs.org.

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.

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST[1]

The research team for this evaluation consists of a prime contractor, Mathematica Policy Research of Princeton, NJ, and one subcontractor, WestEd of San Francisco, CA. Neither of these organizations nor their key staff members have financial interests that could be affected by findings from the evaluation of the two comprehensive induction programs considered in this report. 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.

EXECUTIVE SUMMARY

One of the main policy responses to the problems of turnover and inadequate preparation among beginning teachers is to support them with a formal, comprehensive induction program. Such a program might include a combination of school and district orientation sessions, special in-service training (professional development), mentoring by an experienced teacher, classroom observation, and formative assessment (Berry et al. 2002).

In practice, teacher induction is common, but induction that is intensive, comprehensive, structured, and sequentially delivered in response to teachers' emerging pedagogical needs is not (Berry et al. 2002; Smith and Ingersoll 2004). An example of informal or low intensity teacher induction includes pairing each new teacher with another full-time teacher without providing any training, supplemental materials, or release time for the induction to occur.

There is little empirical evidence on whether investing resources in a more comprehensive, and hence more expensive, induction program would help districts attract, develop, and retain beginning teachers. According to several research reviews (Ingersoll and Kralik 2004; Totterdell et al. 2004; Lopez et al. 2004), little of the research on teacher induction to date has been conclusive or rigorous. Research based on federal statistics (for example, Smith and Ingersoll 2004; Henke et al. 2000; Alt and Henke 2007) can provide a useful, nationally representative perspective on the issue, but it is limited to the extent it can capture the intensity of induction supports and in the range of outcomes that can be examined. Research at the local level (for example, Youngs 2002; Fuller 2003; Rockoff 2008) has relied on non-experimental approaches that do not necessarily provide unbiased estimates of the causal impacts of interest: the retention rate for participants or test scores of participants' students compared to what they would have been in the absence of the program.

Congressional interest in formal, comprehensive teacher induction has grown in recent years. The No Child Left Behind Act of 2001 (NCLB), which reauthorized the Elementary and Secondary Education Act of 1965 (ESEA), emphasizes the importance of teacher quality in student improvement. Title II, Part A of ESEA—the Improving Teacher Quality State Grants program—provides nearly $3 billion a year to states to train, recruit, and prepare high quality teachers. The implementation of teacher induction programs is one allowable use of

these funds. Current discussions on the reauthorization of NCLB argue for a continued focus on supporting teachers through professional development opportunities and teacher mentoring programs, with a call to fund "proven models" to meet these objectives. In addition, the Higher Education Opportunity Act of 2008 authorizes grants that include teacher induction or mentoring programs for new teachers. These initiatives highlight the need to conduct rigorous research to determine whether comprehensive teacher induction programs produce a measurable impact on teacher retention and other positive outcomes for teachers and students.

The National Center for Education Evaluation and Regional Assistance within the U.S. Department of Education's Institute of Education Sciences (IES) contracted with Mathematica Policy Research (MPR) to address this issue by evaluating the impact of structured and intensive teacher induction programs over a three year time period, beginning when teachers first enter the teaching profession. An earlier report (Glazerman et al. 2008) presented results from the first year of the evaluation. The current report presents findings from the second year of the evaluation and a future report will present findings from the third and final year.

Throughout the report, we refer to the more formal, structured programs as "comprehensive" induction. The study examines whether comprehensive teacher induction programs lead to higher teacher retention rates and other positive teacher and student outcomes as compared to prevailing, generally less comprehensive approaches to supporting new teachers. More specifically, the study is designed to address five research questions on the impacts of comprehensive teacher induction:

1. What is the effect of comprehensive teacher induction on the types and intensity of induction services teachers receive compared to the services they receive from the districts' current induction programs?

2. What are the impacts on teachers' classroom practices?[2]

3. What are the impacts on student achievement?

4. What are the impacts on teacher retention?

5. What is the impact on the composition of the district's teaching workforce?

To operationalize the concept of comprehensive teacher induction, we issued a Request for Proposals (RFP) in 2004 to select a comprehensive induction program and program provider for the study. The RFP specified that the induction program should include several

[2] As Glazerman et al. (2008) reports, there was no impact of comprehensive teacher induction on classroom practices in the first year of implementation. Because we did not return to observe classrooms during the second year of the evaluation, we do not re-visit the question about classroom practices in the current report.

components that earlier research and professional wisdom gleaned from practice had suggested were important features of successful teacher induction programs (Alliance for Excellent Education 2004; Ingersoll and Smith 2004; Smith and Ingersoll 2004; Kelly 2004; Serpell and Bozeman 2000). A group of outside expert reviewers ranked the proposals submitted by Educational Testing Service of Princeton, New Jersey (ETS) and the New Teacher Center at the University of California-Santa Cruz (NTC) as most closely meeting the study's specified requirements. The two programs were roughly comparable in structure and included the required components:

- Carefully selected and trained full-time mentors;
- A curriculum of intensive and structured support for beginning teachers that includes an orientation, professional development opportunities, and weekly meetings with mentors;
- A focus on instruction, with opportunities for novice teachers to observe experienced teachers;
- Formative assessment tools that permit evaluation of practice on an ongoing basis and require observations and constructive feedback; and
- Outreach to district and school-based administrators to educate them about program goals and to garner their systemic support for the program.

MPR contracted with both providers to deliver comprehensive induction services to the districts in the study, with one-half of the districts assigned to ETS, the remaining half to NTC. Researchers from WestEd, a subcontractor to MPR, monitored the implementation of the comprehensive induction services to help the providers ensure there was fidelity to the core service model and to identify and help address any implementation challenges that arose.

STUDY DESIGN

The centerpiece of the study design is the use of random assignment to create a group of teachers exposed to comprehensive teacher induction (treatment) and an equivalent group exposed to the district's usual set of induction services (control). The study design allows us to measure and compare outcomes for these two groups to estimate the impacts of comprehensive induction relative to the services teachers receive from their district's prevailing induction program. We used surveys and school records to measure the background of the study teachers, their receipt of induction services and alternative support services, their attitudes, and the key outcomes of student achievement and teacher mobility.

We selected 17 school districts to participate in the study. District selection was based upon factors such as district size and poverty, whether the district was already implementing a comprehensive teacher induction program, and district willingness to participate in the evaluation. The selected districts, which were spread across 13 states, served low-income students, with every district in the study having more than 50 percent of its students

qualifying for the federal School Lunch Program. We then assigned each district to one of the two providers of comprehensive induction, either ETS or NTC, based primarily on district preferences. Nine districts participated in the ETS program; eight districts participated in the NTC program. The preference-based method of assigning districts to providers does not allow for and should not be used to make direct comparisons of one provider to the other.

IES later expanded the treatment to include a second year of services for a subsample of the districts, in effect creating two studies: one for districts that received one year of services (during the 2005-2006 school year), and the other for districts that received two years (during the 2005-2006 and 2006-2007 school years). In the two-year districts, teachers who had been assigned to the treatment group were offered continued services for a second year. The goal of this expansion was to enable the study to address its main research questions separately for one-year and two-year comprehensive induction programs. Policymakers are interested in both models of service delivery because they are both viable policy options for future implementation.

We used convenience sampling to select the districts to receive a second year of the treatment; we selected the districts based upon factors such as whether the mentors who had been trained within the district by ETS or NTC were available for a second year and whether the group of districts selected for a second year would include approximately one-half of the total number of teachers participating in the evaluation. Dividing the sample in this way does not allow for and should not be used to make direct comparisons between the districts that received one year of treatment and districts that received two years of treatment, but instead allows us to investigate the effectiveness of one-year programs separately from that of two-year programs.

In this Year 2 impact report, unlike the Year 1 impact report (Glazerman et al. 2008), we present findings separately for the set of 10 districts that received one year of treatment ("one-year districts") and the other set of 7 districts that received two years of treatment ("two-year districts"). Both sets of findings are based on data collected through two years of the study. When appropriate, however, we compare outcomes from the first year of the study to outcomes from the second year of the study within the one-year districts and within the two-year districts.

Within each district, a subset of elementary schools participated in the study. As noted above, we randomly assigned these elementary schools to either a treatment group, which was offered comprehensive teacher induction, or a control group, which took part in the district's usual teacher induction program. The final sample size included 418 schools across the 17 districts.

Within each study school, we selected all eligible teachers, defined as beginning teachers who met certain criteria: taught in an elementary grade (K-6); were new to the profession; and were not already receiving induction support from a teacher preparation or certification program. Under these criteria, the 252 schools in the one-year districts contained 561 eligible teachers, and the 166 schools in the two-year districts contained 448 eligible teachers. For the student achievement analysis, we limited the collection of student test score data to

teachers meeting another set of eligibility criteria, including teaching a self-contained classroom in a tested grade and subject. This resulted in the collection of reading test scores for 139 teachers and math scores for 123 teachers in the one-year districts, and of reading scores for 96 teachers and math scores for 95 teachers in the two-year districts.[3]

Eligible teachers in a school were either all exposed or all not exposed to treatment, a method known as cluster random assignment. Cluster random assignment was necessary because varying the types of induction services available in the same school building could result in contamination of the control group. Therefore, we assigned all eligible teachers to treatment or control status based on the school where they were expected to teach at the point of random assignment.

METHODS AND DATA

We used a model-based approach to estimate program impacts. The statistical model explicitly acknowledges the hierarchical structure of the data—for example, the nesting of teachers within schools—an approach that is sometimes referred to as a hierarchical linear model (HLM). Accordingly, we can properly specify the units of analysis (teachers and schools) and devise unbiased estimates of the standard errors that we used to conduct hypothesis tests. The model also allows us to control for the effects of a range of teacher and school characteristics on the outcomes of interest to increase the precision of the estimates of treatment effects.

For each outcome, we use a different set of control variables (covariates), described in the discussion of key study findings. The control variables used in the body of the report are called the benchmark control variables; in sensitivity analyses presented in appendices to the report, we alter the control variables to test the robustness of the results. These sensitivity tests included re-estimation of the study's main impacts with different sets of covariates, using different samples or sample weights, and different statistical model assumptions.

Data for the study were collected from a variety of sources. In fall 2005 we surveyed mentors participating in the comprehensive induction programs on their background characteristics and reviewed program documents from ETS and NTC. We administered a baseline survey of beginning teachers in fall 2005, at which time we also requested teachers' permission to obtain their college entrance examination scores (SAT or ACT). The baseline survey asked teachers about their formal education, professional training, current teaching assignment, and personal background. We surveyed teachers twice during the 2005-2006 school year on the induction activities in which they participated, including questions about duration and intensity of mentoring and professional development as well as questions about satisfaction with different aspects of their current teaching position. During the 2006-2007 school year, we surveyed teachers in the two-year districts twice and teachers in the one-year

[3] The standard errors of test score impact estimates were in the range of 0.05 to 0.08, meaning that an impact in effect size units of 0.10 to 0.16 would be statistically significant. The study was originally designed to detect test score impacts of 0.10 to 0.22 (Glazerman et al. 2005).

districts once on the induction activities in which they participated and on their job satisfaction.

For the report's core outcomes measuring the impacts of comprehensive teacher induction, we collected districts' student records data at the end of the 2006-2007 school year and conducted the second of three mobility surveys in fall 2007 to learn about teacher retention. We measured student achievement outcomes using district-administered test score data from the spring 2007 (posttest) for students taught by study teachers in the 2006-2007 school year and students' linked scores from the prior grade in spring 2006 (pretest).[4] We conducted all treatment-control comparisons within grade and within district to ensure that treatment status was not confounded with properties of the test. Response rates on teacher surveys ranged from 88 percent to 97 percent for the treatment group and 78 percent to 92 percent for the control group. We used nonresponse adjustment weights and sensitivity analyses to address the differential response rates in the analysis of teacher mobility.

THE TREATMENT: COMPREHENSIVE INDUCTION SERVICES

Treatment teachers in each district were given the opportunity (but were not required) to participate in the comprehensive induction program implemented there. The comprehensive induction program components included carefully selected and trained full-time mentors; a curriculum of intensive and structured support for beginning teachers; a focus on instruction, with opportunities for novice teachers to observe experienced teachers; formative assessment tools that permit evaluation of practice on an ongoing basis and require observations and constructive feedback; and outreach to district and school-based administrators to educate them about program goals and to garner their systemic support for the program.

Both the ETS and NTC programs are based on a curriculum expected to promote effective teaching. The ETS program defines effective teaching in terms of 22 components organized into four domains of professional practice. The components are aligned with the Interstate New Teacher Assessment and Support Consortium (INTASC 1992) principles. The NTC induction model defines effective teaching in terms of six Professional Teaching Standards. Each standard, or domain, is broken into a succession of more discretely defined categories of teaching behaviors.

The curriculum that formed the foundation of both programs included a number of activities. Mentors were asked to meet weekly with treatment teachers for approximately two hours. Conversation was expected to center around the induction programs' teacher learning activities, but mentors also exercised professional judgment in selecting additional activities to meet beginning teachers' needs, including observing instruction or providing a demonstration lesson; reviewing lesson plans, instructional materials, or student work; or interacting with students to gain an additional perspective on teachers' instructional

[4] For three districts that tested at least some students in the fall, we used a fall 2006 test as a pretest and/or a fall 2007 test as a posttest.

practices. Treatment teachers were provided monthly professional development sessions to complement their interactions with mentors, and the ETS districts also offered monthly study groups—mentor-facilitated peer support meetings for treatment teachers during which beginning teachers met monthly to discuss their local needs and practices. Treatment teachers also observed veteran teachers once or twice during the year. At the end of each school year, treatment teachers in both ETS and NTC districts participated in a colloquium celebrating the year's successes and teachers' professional growth.

The providers adapted the curricula of the second year of their usual induction programs for the second year of induction services in the two-year districts. While programs provided induction activities to these districts' treatment teachers during the second year that were similar to those in the first year, the content was designed to reflect the growth of mentors and beginning teachers and the evolution of their circumstances and needs. In two-year districts served by ETS, mentors led Teacher Learning Communities, an adaptation of the first year's study groups that included specific content for each session and a formal structure for teachers to try out approaches to instruction. During second year professional development sessions in the two-year districts served by NTC, mentors elaborated on standardized topics and designed activities to reflect local needs.

At the heart of the comprehensive induction services was the support provided by a full-time mentor trained by the program providers. The goal of the study was to assign each mentor to 12 beginning teachers. At the outset of the study, the program providers sought mentor candidates with a minimum of five years of teaching experience in elementary school, recognition as an exemplary teacher, and experience in providing professional development or mentoring other teachers (particularly beginning teachers).

In Year 1, the providers brought their respective mentors together for 10 to 12 days of training. The training was spread across four sessions of 2 to 3 days, with the first session held during the summer of 2005 and the rest taking place throughout the school year. Trainings previewed the content of upcoming professional development sessions and gradually introduced processes of mentor/mentee work in such areas as reflecting on instructional practices and analyzing student work. During Year 2, ETS and NTC continued intensive training of their respective mentors in the seven districts that were selected to continue program implementation. ETS brought mentors together for a total of 8 days over 3 sessions. NTC did so over 10 days and 4 sessions. The providers devoted 1.5 to 2.5 days per session. All mentors participated in the trainings, which reflected a focus similar to Year 1. In sum, in two-year districts ETS mentors participated in 18 days of training; NTC mentors participated in 22 days.

Practitioners and policymakers should be aware of two issues related to program implementation. The first is the voluntary nature of teachers' participation in the treatment services. The program models that were implemented did not necessarily require teachers to participate but rather made services available to them, so not all teachers attended every professional development session provided.

The second issue for practitioners and policymakers to be aware of is that the programs implemented in this study by ETS and NTC were not necessarily the same models that would be delivered outside the study context. First, for study purposes, we aimed for consistent implementation of each program, with a high level of fidelity to the program design and a quick response to any implementation issues. Second, the providers adapted their program for the study to ensure that the required components were included in a one-year curriculum. Once it was decided to add a second year, the programs made additional modifications and adaptations to extend the curriculum another year. Finally, each provider organized off-site mentor training sessions, bringing together the mentors from all of the provider's study districts. For district-wide implementation with a larger number of mentors, training typically occurs within the district, rather than off-site together with mentors from other districts.

THE COUNTERFACTUAL: PREVAILING INDUCTION SERVICES

We designed the study to compare teachers who were exposed to comprehensive teacher induction services (treatment) to an equivalent group that was exposed to the induction services normally offered by the districts (control). We purposefully selected districts whose schools were not already working with ETS or NTC on induction projects, were not using the providers' induction materials, were not spending more than $1,000 per teacher on induction, and did not assign full-time release mentors to work with beginning teachers.

SUMMARY OF FINDINGS AFTER ONE YEAR: ONE-YEAR AND TWO-YEAR DISTRICTS COMBINED

An earlier report (Glazerman et al. 2008) presented findings after the first year of implementation of the comprehensive induction program within study districts. That report showed that teachers assigned to the treatment group reported significantly more induction support, but also that the additional support did not translate into positive impacts on key outcomes after one year.[5] The additional induction support amounted to a greater likelihood of having a mentor formally assigned to beginning teachers (93 versus 75 percent), more time spent in meetings with the mentor (95 versus 74 minutes per week), and greater likelihood of receiving "a moderate amount" or "a lot" of assistance from mentors in areas such as classroom management (65 versus 40 percent), reviewing student work (55 versus 30 percent), and communicating with parents (38 versus 31 percent). There were no positive impacts on classroom practices, student achievement, teacher retention, or the composition of the district's teaching workforce after one year. Nor did we find any evidence of positive impacts on teachers' satisfaction or feelings of preparedness.

[5] All references to "significance" in this report refer to statistical significance. A difference is deemed statistically significant in this report if the probability that it was observed by chance is less than 5 percent. The term "statistically insignificant" does not imply irrelevance for policymakers and similarly the term "statistically significant" does not necessarily mean "large" or meaningful for policy.

SUMMARY OF FINDINGS AFTER TWO YEARS: TREATMENT-CONTROL DIFFERENCES IN ONE-YEAR DISTRICTS

Induction Services Received

Within one-year districts, during Year 1—the year in which comprehensive teacher induction was implemented—we found statistically significant differences between the treatment and control group; the treatment group reported receiving more induction support than the control group across a broad range of measures of the amount, types, and content of supports.

In Year 2—the year in which treatment teachers no longer received comprehensive teacher induction supports—the percentage of teachers with an assigned mentor and the weekly minutes spent with that mentor declined from Year 1 to Year 2 (differences with a p-value of 0.000) for both the treatment and control groups. During this second year, we found statistically significant negative impacts on these and other measures of support, as described below.

Because teachers in one-year districts were not surveyed in the spring of Year 2, we focus the discussion on findings for the fall of each year.[6] Estimates were computed using an ordinary least squares model with district and grade assignment fixed effects that accounted for clustering of teachers within schools; weights were applied to adjust for survey nonresponse and the study design.[7]

Amount of Mentoring. In Year 1, we found statistically significant differences in the likelihood of teachers reporting having a mentor assigned to them and having a full-time mentor. As part of the intervention, every treatment teacher was assigned a mentor by ETS or NTC, but that did not guarantee that all teachers would work with their mentor or acknowledge having had one assigned to them. Still, treatment teachers were more likely than control teachers to report having a mentor assigned to them (90 versus 70 percent) and to report having a full-time mentor (74 versus 8 percent). We found statistically significant differences in teachers' likelihood of having a mentor who was another teacher and in the amount of time teachers reported spending with a mentor during the most recent full week of teaching. Treatment teachers were less likely than control teachers to report having a mentor who was another teacher (25 versus 64 percent). In addition, treatment teachers reported spending an average of 87 minutes per week in mentor meetings compared to 67 minutes for control teachers, with the 20-minute difference attributable entirely to differences in the duration of scheduled meetings, as opposed to informal meetings.

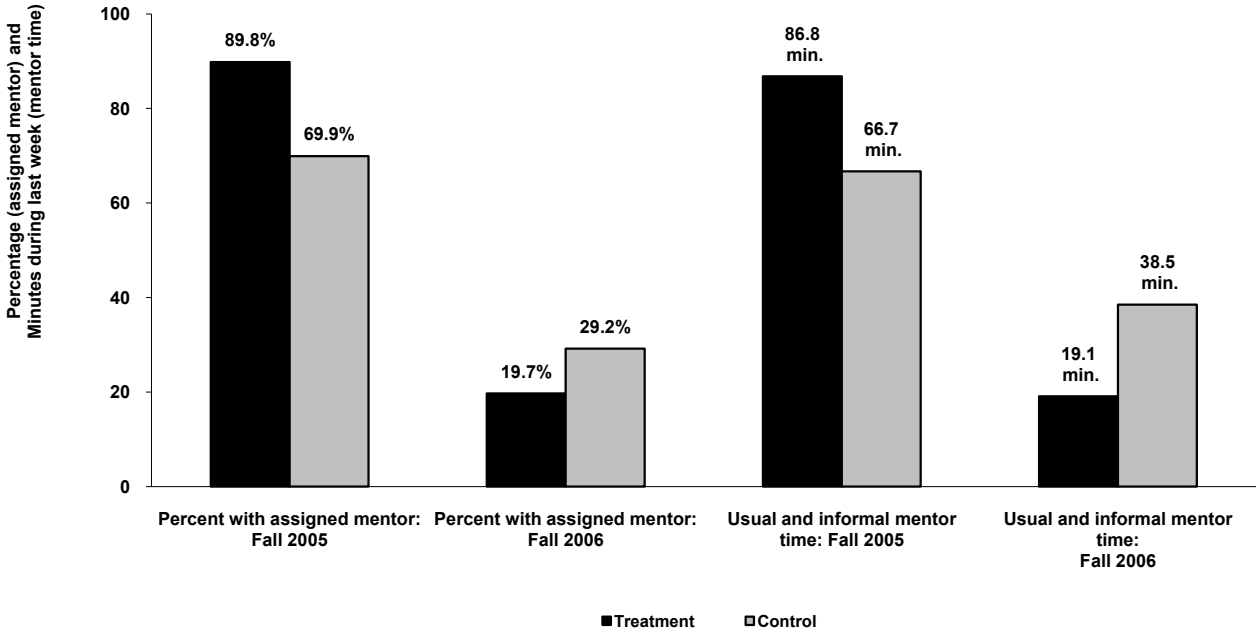
In Year 2, we found statistically significant differences in the prevalence of and time spent in mentoring. Treatment teachers were *less* likely than control teachers to report having

[6] Findings from the fall of Year 1 can be compared to findings from the spring of Year 1, which are shown in Appendix C.

[7] Across all outcomes, the same methods were used in the analysis of two-year districts.

a mentor assigned to them (20 versus 29 percent). Treatment teachers were also less likely than control teachers to report having a mentor who was another teacher (21 versus 31 percent). Treatment teachers spent less time in mentor meetings than control teachers (19 versus 39 minutes per week). Figure ES.1 shows treatment-control differences for having an assigned mentor and time in mentor meetings in Year 1 and Year 2.

Figure ES.1. Treatment-Control Differences in Percent Assigned a Mentor and Total Minutes Spent in Mentoring Per Week: One-Year Districts, Fall 2005 and Fall 2006



Note: All treatment-control differences are significantly different from zero at the 0.05 level, two-tailed test (N=503 teachers in fall 2005 and 472 teachers in fall 2006).

Mentor Activities and Assistance. In Year 1, treatment and control teachers' reports showed statistically significant differences in the amounts of time in various mentor activities and the kinds of assistance received from their mentors. Treatment teachers reported spending more time during the most recent full week of teaching being observed by mentors (34 versus 10 minutes), meeting one-on-one with mentors (34 versus 23 minutes), meeting with mentors together with other first-year teachers (29 versus 9 minutes), and having mentors model lessons (9 versus 6 minutes). During the most recent full week of teaching, treatment teachers were 14 to 27 percentage points more likely than control teachers to report having received mentors' assistance in a variety of topic areas, such as receiving suggestions to improve practice (77 versus 53 percent) and discussing instructional goals (73 versus 48 percent).

By Year 2, we found statistically significant differences in the amount of time teachers reported being observed by mentors during the most recent full week of teaching in fall 2006. Treatment teachers reported less time in a list of six common mentoring activities (22 versus 36 minutes per week) including less time being observed by mentors than control

teachers (2 versus 6 minutes). No statistically significant differences were found between treatment and control group teachers on their reported time spent in any of the other five activities covered by the survey. During the most recent full week of teaching in fall 2006, treatment and control teachers' reports showed statistically significant differences in the likelihood of receiving mentors' assistance in each of the topic areas covered by the survey. Treatment teachers were less likely than control teachers to report receiving mentors' assistance in each topic area, with effects ranging from 8 to 14 percentage points, including, for example, impacts on receiving suggestions to improve practice (15 versus 27 percent) and discussing instructional goals (14 versus 24 percent).

Professional Development. We did not find statistically significant differences between treatment and control teachers in their reported attendance in professional development, except in certain areas. Of the 12 professional development topics covered by the survey, treatment teachers were less likely than control teachers to report having attended professional development sessions in two areas in fall 2005 (Year 1): content area knowledge (61 versus 72 percent) and preparing students for standardized testing (30 versus 41 percent). We did not find statistically significant differences between treatment and control teachers in their reported attendance in any of the 12 professional development activities in fall 2006 (Year 2).

Student Achievement

In Year 2 (school year 2006-2007), we found no statistically significant impacts on reading or math scores in the one-year districts. We compared the test scores for students of treatment teachers to those of control teachers using post-test scores measured in 2007 adjusted for pre-test scores measured in 2006. The test score analysis was based on standardized achievement tests that the district normally conducts.[8] Though district-administered test scores do not cover every domain of student achievement that induction might affect, they do capture the content that school districts or states deem most important and worthy of assessing. We aggregated test scores across districts and grades by standardizing each test to a common metric called a z-score, which has a mean of zero and a standard deviation of one. We kept two broad subject areas, math and reading, distinct. The benchmark model accounts for the nesting of students within schools, using the normalized student pretest score and district-by-grade fixed effects as covariates.

The benchmark impacts on math and reading scores in Year 2 were not significantly different from zero (see Table ES.1). We confirmed that the impact on math and reading in the second year was not statistically significant when the impacts were re-estimated using different samples, sets of covariates, or estimation techniques.

[8] The specific test differs from district to district, and in some cases by grade within district. However, all treatment-control comparisons were made using a common set of tests within grade within district.

Table ES.1. Impacts on Test Scores: One-Year Districts, 2006-2007 School Year

	Adjusted Mean Test Scores					Unweighted Sample Sizes		
Subject	Treatment	Control	Difference	Effect Size	P-value	Students	Teachers	Districts
Reading	0.05	0.01	0.04	0.04	0.380	2,245	135	9
Math	0.05	-0.02	0.08	0.08	0.367	1,995	117	9

Source: MPR analysis of data from 2005-2006 and 2006-2007 school years provided by participating school districts.

Notes: Data are regression-adjusted to account for pretest, district-by-grade fixed effects, and clustering of students within schools. For Reading, there were 1,193 students and 72 teachers in the treatment group, and 1,052 students and 63 teachers in the control group. For Math, there were 994 students and 57 teachers in the treatment group, and 1,001 students and 60 teachers in the control group.

None of the differences is statistically significant at the 0.05 level, two-tailed test.

Teacher Retention

We found that comprehensive teacher induction had no statistically significant impact on teacher retention after two years. We measured teacher retention in terms of the percentage of teachers who remained in their originally assigned school, their district, and the teaching profession. Table ES.2 shows the result of the three hypothesis tests specifically focused on retention in the school, in the district, and in the profession as binary outcomes. For each of the outcomes, there was no statistically significant impact. The same result was obtained when we expanded the number of outcomes to differentiate between moving to a school in another public school district and moving to a private, parochial, or other school, and expanded the outcomes for leaving to include leaving to stay at home, leaving to attend school or take a new job, and other reasons for leaving.

Table ES.2. Impacts on Teacher Retention Rates after Two Years (Percentages): One-Year Districts

Outcome	All Teachers	Treatment	Control	Difference	P-value
Retained in the same school	62.5	60.3	64.7	-4.5	0.280
Retained in the same district	79.5	78.6	80.3	-1.7	0.619
Retained in the teaching profession	90.1	90.4	89.8	0.7	0.789
Unweighted Sample Size (Teachers)	476	244	232		
Unweighted Sample Size (Schools)	227	114	113		

Source: MPR Mobility Survey administered in 2007-2008 and Teacher Background Survey administered in 2005-2006 to all study teachers.

Note: Data are regression-adjusted using a logit model with robust standard errors to account for baseline characteristics and clustering of teachers within schools.

None of the differences is statistically significant at the 0.05 level, two-tailed test.

We also examined the reasons that teachers who left their districts (movers) or left the teaching profession (leavers) gave for leaving and found no statistically significant impacts of treatment. When we asked leavers whether they expected to return and if so, when they would do so, we did not find evidence of a treatment-control difference. In addition, we found that treatment teachers did not report feeling more satisfied with their jobs than control teachers.

Composition of District Teaching Force

The last major research question concerned the impact of comprehensive teacher induction on the composition of the teaching workforce in the district. As shown below, we found no statistically significant impacts on the composition of the district teaching force in one-year districts after two years.

For comprehensive teacher induction to affect the composition of the district's teaching workforce, it has to produce a difference in the types of teachers who decide to remain in the district. As teachers leave the district, the average qualifications of the teachers who remain in the district begin to change, perhaps differentially between the treatment and control groups. We tested this hypothesis by comparing the characteristics of district stayers between the treatment and control groups along two dimensions: (1) their impact on student achievement; and (2) their professional characteristics such as SAT/ACT scores and advanced degrees. The student achievement outcome is regression-adjusted using the same model used in the main analysis.

We found that the treatment had no statistically significant impacts on the student achievement or professional background characteristics of district stayers. Table ES.3 presents the impacts on student achievement outcomes for district stayers. Table ES.4 shows the background characteristics of teachers by mobility status.

Table ES.3. Impacts on Test Scores, District Stayers Only: One-Year Districts, 2005-2006 School Year

Outcome	Treatment	Control	Difference	Effect Size	P-value
Reading scores (all grades)	0.02	-0.03	0.05	0.05	0.331
Unweighted Sample Size (Students)	975	942	1,917		
Unweighted Sample Size (Teachers)	53	56	109		
Unweighted Sample Size (Schools)	47	41	88		
Math scores (all grades)	0.01	-0.02	0.03	0.03	0.629
Unweighted Sample Size (Students)	826	857	1,683		
Unweighted Sample Size (Teachers)	47	52	99		
Unweighted Sample Size (Schools)	43	38	81		

Source: MPR analysis of data from 2004-2005 and 2005-2006 school years provided by participating school districts; MPR Second Mobility Survey administered in 2007-2008 to all study teachers.

Notes: Data are regression-adjusted to account for pretest, district-by-grade fixed effects, and clustering of students within schools.

None of the differences is statistically significant at the 0.05 level, two-tailed test.

Table ES.4. Characteristics of District Stayers, Movers, and Leavers after Two Years by Treatment Status (Percentages Except Where Noted): One-Year Districts

	Treatment			Control			Difference		
Teacher characteristic	Stayers	Movers	Leavers	Stayers	Movers	Leavers	Stayers	Movers	Leavers
College entrance exam scores (SAT combined score or equivalent)	1,026	1,029	1,082	1,021	984	1,080	4	45	2
Attended highly selective college	30.3	27.3	46.0	27.2	50.5	33.3	3.1	-23.2	12.7
Major or minor in education	79.8	65.5	76.1	81.1	65.9	67.2	-1.3	-0.4	8.9
Student teaching experience (Weeks)	16.5	13.9	14.2	15.1	13.5	12.4	1.5	0.4	1.8
Entered the profession through traditional four-year program	64.4	61.0	45.8	60.3	58.7	30.8	4.1	2.4	15.0
Unweighted Sample Size (Teachers)	191	29	24	187	23	22			
Unweighted Sample Size (Schools)	100	25	18	104	22	21			

Source: MPR calculations using data from the College Board and ACT, Inc.; MPR Teacher Background Survey administered in 2005-2006, MPR Second Mobility Survey administered in 2007-2008; MPR First and Second Induction Activities Surveys administered in fall/winter 2005-2006 and spring 2006 to all study teachers.

Notes: Data are weighted to account for the study design. Sample sizes vary due to item nonresponse. The analysis of college entrance exam scores relied on a smaller sample of teachers (191/29/24 treatment stayers/movers/leavers and 187/23/22 control stayers/movers/leavers) and schools (100/25/18 treatment and 104/22/21 control).
Stayer: retained in the same school district.
Mover: retained in the teaching profession, but not in the same school district.
Leaver: no longer teaching.

None of the differences between treatment and control stayers, between treatment and control movers, or between treatment and control leavers is statistically significant at the 0.05 level, two-tailed test. P-values are suppressed to make the table easier to read.

SUMMARY OF FINDINGS AFTER TWO YEARS: TREATMENT-CONTROL DIFFERENCES IN TWO-YEAR DISTRICTS

Induction Services Received

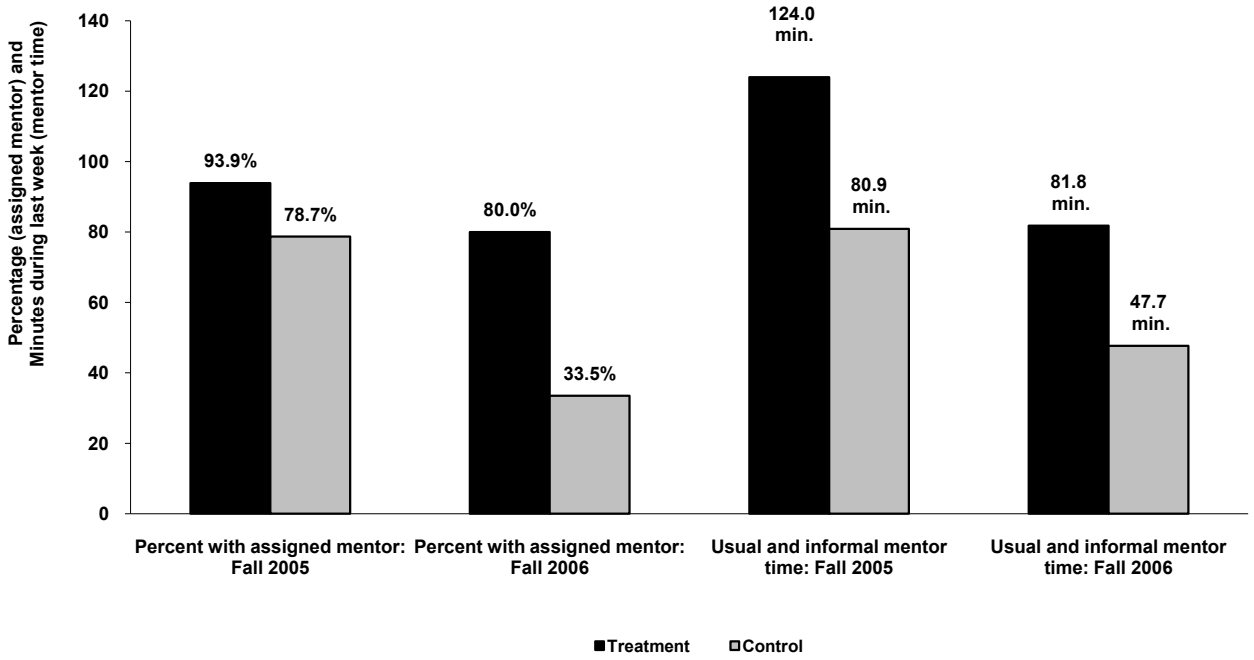
During Year 1 and Year 2, both years in which comprehensive teacher induction services were offered to the treatment group in the two-year districts, treatment and control teachers' reports showed statistically significant differences favoring the treatment group on many measures of the amount, types, or content of supports. For consistency with the way in which results are reported for one-year districts, we report on findings for the fall of each year.[9]

Amount of Mentoring. We found statistically significant differences between the treatment and control teachers with regard to the likelihood of teachers reporting having a mentor assigned to them, having a full-time mentor, and having a mentor who was another teacher. Treatment teachers were more likely than control teachers to report having a mentor assigned to them (94 versus 79 percent in Year 1; 80 versus 34 percent in Year 2), and to report having a full-time mentor (72 versus 16 percent in Year 1; 64 versus 7 percent in Year 2). Treatment teachers were less likely than control teachers to report having a mentor who was another teacher (38 versus 62 percent in Year 1; 12 versus 27 percent in Year 2). We also found statistically significant differences in the amount of time teachers reported spending with their mentors. Treatment teachers reported spending more time working with their mentors than control teachers did during the most recent full week of teaching. Treatment teachers reported spending more time on average in mentor meetings (124 minutes per week versus 81 minutes in Year 1; 82 minutes versus 48 minutes in Year 2). In both years, the differences were attributable primarily to differences in the duration of scheduled meetings. Figure ES.2 shows treatment-control differences for having an assigned mentor and time in mentor meetings in Year 1 and Year 2.

Mentor Activities and Assistance. Treatment and control teachers' reports showed statistically significant differences in the amount of time in various mentor activities and in the kinds of assistance teachers reported receiving from their mentors. Treatment teachers reported spending more time being observed by mentors (38 versus 17 minutes in Year 1; 22 versus 7 minutes in Year 2), meeting one-on-one with mentors (43 versus 23 minutes in Year 1; 25 versus 12 minutes in Year 2), meeting together with mentors and other first-year teachers (38 versus 11 minutes in Year 1; 25 versus 6 minutes in Year 2), and having mentors model lessons (16 versus 10 minutes in Year 1; 12 versus 5 minutes in Year 2). During the most recent full week of teaching, treatment teachers were more likely than control teachers to report receiving mentors' assistance in each of the topic areas covered by the survey: effects ranged from 14 to 28 percentage points in Year 1 and 28 to 44 percent in Year 2.

[9] For two-year districts, findings from spring of Year 1 were consistent with the findings from fall of Year 1. Likewise, findings from spring of Year 2 were consistent with the findings from fall of Year 2.

Figure ES.2. Treatment-Control Differences in Percent Assigned a Mentor and Total Minutes Spent in Mentoring Per Week: Two-Year Districts, Fall 2005 and Fall 2006



Note: All treatment-control differences are significantly different from zero at the 0.05 level, two-tailed test (N=395 teachers in fall 2005 and 360 teachers in fall 2006).

Professional Development. We did not find statistically significant differences between treatment and control teachers' reported attendance in professional development, except that treatment teachers were more likely than control teachers to report having attended sessions focused on classroom management techniques (61 versus 48 percent) in fall 2005 (Year 1).

Student Achievement

We found no evidence of statistically significant impacts on student test scores in two-year districts. The benchmark impacts on math and reading scores in the second year of the study were not significantly different from zero (Table ES.5). The data confirm that the impacts on reading and math in the second year were not statistically significant when we re-estimated the impacts using different samples, different sets of covariates, or different estimation techniques.

Table ES.5. Impacts on Test Scores: Two-Year Districts, 2006-2007 School Year

	Adjusted Mean Test Scores					Unweighted Sample Sizes		
Subject	Treatment	Control	Difference	Effect Size	P-value	Students	Teachers	Districts
Reading	0.00	0.00	0.00	0.00	0.967	1,732	100	7
Math	-0.03	-0.01	-0.02	-0.02	0.746	1,736	99	7

Source: MPR analysis of data from 2005-2006 and 2006-2007 school years provided by participating school districts.

Notes: Data are regression-adjusted to account for pretest, district-by-grade fixed effects, and clustering of students within schools. For Reading, there were 856 students and 52 teachers in the treatment group, and 876 students and 48 teachers in the control group. For Math, there were 780 students and 50 teachers in the treatment group, and 956 students and 49 teachers in the control group.

None of the differences is statistically significant at the 0.05 level, two-tailed test.

Teacher Retention

We found that comprehensive teacher induction had no statistically significant impact on teacher retention after two years. Table ES.6 shows the result of the three hypothesis tests specifically focused on retention in the school, in the district, and in the profession as binary outcomes. For each of the outcomes, there was no statistically significant impact. The same result was obtained when we expanded the number of outcomes to differentiate between moving to a school in another public school district and moving to a private, parochial, or other school, and expanded the outcomes for leaving to include leaving to stay at home, leaving to attend school or take a new job, and other reasons for leaving.

We also examined the reasons that teachers who left their districts (movers) or left the teaching profession (leavers) gave for leaving and found no statistically significant impacts of treatment. When we asked leavers whether they expected to return and if so, when they would do so, we did not find evidence of a treatment-control difference. In addition, we found that treatment teachers did not report feeling more satisfied with or prepared for their jobs than control teachers.

Table ES.6. Impacts on Teacher Retention Rates after Two Years (Percentages): Two-Year Districts

Outcome	All Teachers	Treatment	Control	Difference	P-value
Retained in the same school	64.1	62.2	66.2	-4.0	0.386
Retained in the same district	72.3	69.6	75.3	-5.7	0.208
Retained in the teaching profession	88.8	86.9	90.8	-3.9	0.241
Unweighted Sample Size (Teachers)	364	203	161		
Unweighted Sample Size (Schools)	151	81	70		

Source: MPR Second Mobility Survey administered in 2007-2008 and Teacher Background Survey administered in 2005-2006 to all study teachers.

Note: Data are regression-adjusted using a logit model with robust standard errors to account for baseline characteristics and clustering of teachers within schools.

None of the differences is statistically significant at the 0.05 level, two-tailed test.

Composition of the District Teaching Force

We found that the treatment had no statistically significant impacts on the student achievement outcomes or professional background characteristics of district stayers. Table ES.7 presents the impacts on student achievement outcomes for district stayers. Table ES.8 shows the background characteristics of teachers by mobility status.

Table ES.7. Impacts on Test Scores, District Stayers Only: Two-Year Districts, 2005-2006 School Year

Outcome	Treatment	Control	Difference	Effect Size	P-value
Reading scores (all grades)	0.03	-0.03	0.06	0.06	0.591
Unweighted Sample Size (Students)	745	558	1,303		
Unweighted Sample Size (Teachers)	45	30	75		
Unweighted Sample Size (Schools)	31	24	55		
Math scores (all grades)	-0.04	0.07	-0.11	-0.11	0.162
Unweighted Sample Size (Students)	693	549	1,242		
Unweighted Sample Size (Teachers)	43	30	73		
Unweighted Sample Size (Schools)	29	24	53		

Source: MPR analysis of data from 2004-2005 and 2005-2006 school years provided by participating school districts; MPR Second Mobility Survey administered in 2007-2008 to all study teachers.

Notes: Data are regression-adjusted to account for pretest, district-by-grade fixed effects and clustering of students within schools.

None of the differences is statistically significant at the 0.05 level, two-tailed test.

Table ES.8. Characteristics of District Stayers, Movers, and Leavers after Two Years by Treatment Status (Percentages Except Where Noted): Two-Year Districts

	Treatment			Control			Difference		
Teacher Characteristic	Stayers	Movers	Leavers	Stayers	Movers	Leavers	Stayers	Movers	Leavers
College entrance exam scores (SAT combined score or equivalent)	916	1,006	1,095	967	1,040	1,081	-51	-34	14
Attended highly selective college	23.4	28.6	59.9	25.1	37.1	52.4	-1.7	-8.5	7.5
Major or minor in education	67.0	70.9	38.9	66.6	70.8	74.7	0.4	0.0	-35.8
Student teaching experience (weeks)	12.2	14.1	6.2	11.9	11.7	9.3	0.3	2.4	-3.1
Entered the profession through traditional four-year program	61.5	76.8	25.2	66.0	61.3	56.1	-4.5	15.5	-30.9
Unweighted Sample Size (Teachers)	143	35	25	121	25	15			
Unweighted Sample Size (Schools)	71	28	20	62	21	13			

Source: MPR calculations using data from the College Board and ACT, Inc.; MPR Teacher Background Survey administered in 2005-2006, MPR Second Mobility Survey administered in 2007-2008; MPR First and Second Induction Activities Surveys administered in fall/winter 2005-2006 and spring 2006 to all study teachers.

Notes: Data are weighted to account for the study design. Sample sizes vary due to item nonresponse. The analysis of college entrance exam scores relied on a smaller sample of teachers (143/35/25 treatment stayers/movers/leavers and 121/25/15 control stayers/movers/leavers) and schools (71/28/20 treatment and 62/21/13 control).
Stayer: retained in the same school district.
Mover: retained in the teaching profession, but not in the same school district.
Leaver: no longer teaching.

None of the differences between treatment and control stayers, between treatment and control movers, or between treatment and control leavers is statistically different from zero. P-values are suppressed to make the table easier to read.

CORRELATIONAL ANALYSES

Given the prevalence of supports reported by control teachers, we explored the relationship between induction supports and outcomes independent of group assignment (treatment or control) and district type (one-year or two-year). Using data from the first three Induction Activities surveys, we created a variable that reflects the number of years (0, 1, or 2) the beginning teacher had an assigned mentor and constructed three other new measures[10]:

- The Induction Services Index measuring breadth of services received by the beginning teacher,
- The Instructional Support Index measuring suggestions, guidance, and feedback on teaching, and
- The Induction Intensity Index measuring program duration and intensity.

The analyses use the same methods as the experimental analyses, but instead of assignment to treatment status, which was randomly determined, the key explanatory variables are the number of years the beginning teacher had an assigned mentor and the three indices, included jointly in a regression model. The results should be interpreted with caution because the analyses are correlational and not causal. In particular, a nonexperimental estimate of the relationship of induction services with outcomes may be spurious, as it will confound the true (causal) impact of mentoring with the effect of the teacher's own ability or motivation.

Overall, we found that induction measures were not significantly related to math test scores (p-value of F-test = 0.068) or reading scores (p-value of F-test = 0.651). However, we found that the association between the years the beginning teacher had a mentor and math test scores was statistically significant (regression coefficient = 0.12, p-value = 0.015). For measures of teacher retention, there was a statistically significant relationship between the induction activities variables and retention (p-value of F-test = 0.016 for remaining in the

[10] The variable that reflects the number of years the beginning teacher had an assigned mentor is constructed using three items: the indicator variables at fall 2005, spring 2006, and fall 2006, on whether the beginning teacher had an assigned mentor. This variable has the values 0, 1, and 2 years. The Induction Services Index is the sum of nine indicator variables at fall 2005, spring 2006, and fall 2006, on whether the beginning teacher: (1) met with a literacy or math coach, (2) met with a study group, and (3) observed others teaching. The Induction Services Index has values in the range 0 to 9. The Instructional Support Index is constructed similarly using eight indicator variables on whether the beginning teacher received: (1) suggestions from a mentor to improve his/her teaching, (2) at least a moderate amount of guidance in subject area content, and (3) feedback on teaching. The Instructional Support Index has values in the range 0 to 8. The Induction Intensity Index is the sum of the average number of hours per week at fall 2005, spring 2006, and fall 2006 (3 items) that beginning teachers reported spending: (1) in mentoring sessions, (2) being observed teaching by mentor, (3) in professional development learning instructional techniques and strategies, and (4) in professional development learning content area knowledge, specifically language arts, math, and science. The Induction Intensity Index has values in the range 0 to 20.8.

district; p-value of F-test = 0.001 for remaining in teaching). One measure—the Induction Services Index—was positively related and no measures were negatively related to teacher mobility for both remaining in the district and remaining in teaching. The estimate of the regression coefficient on the Induction Services Index for remaining in the district was 0.02; for remaining in teaching, it was 0.01. This implies that, for example, if the retention rate in a district were 80 percent, then an additional induction service, such as meeting with a study group in one semester, would be associated with a district retention rate of 82 percent, all else equal. All results were robust to alternate methods of constructing the indices and alternate model specifications.

SUMMARY OF FINDINGS

The report presents findings from an experimental test of the impact of comprehensive teacher induction on student achievement in beginning teachers' classrooms and on the teachers' retention rates in urban elementary schools. In ten of the study districts, a comprehensive induction program was implemented during beginning teachers' first year in the classroom. In the remaining seven study districts, comprehensive induction was implemented during beginning teachers' first two years in the classroom. This design does not allow for and should not be used to make direct comparisons between the districts that received one year of treatment and districts that received two years of treatment, but instead allows us to investigate the effectiveness of one-year programs separately from that of two-year programs. The main findings are summarized below.

- During their first year in the classroom, in both one- and two-year districts, treatment and control teachers' reports showed statistically significant differences in the amount and types of support received. Treatment teachers were more likely than control teachers to report having an assigned mentor (90 versus 70 percent of teachers reported having an assigned mentor in one-year districts; 94 versus 79 percent in two-year districts) and reported spending more time per week with a mentor (87 versus 67 minutes in one-year districts; 124 versus 81 minutes in two-year districts). Treatment teachers reported spending more time being observed by mentors (34 versus 10 minutes during the most recent full week of teaching in one-year districts; 38 versus 17 minutes in two-year districts) and meeting with mentors together with other first-year teachers (29 versus 9 minutes in one-year districts; 38 versus 11 minutes in two-year districts).

- During their second year in the classroom, treatment teachers in one-year districts received less support than did control teachers. During Year 2, we found a statistically significant difference favoring the control group in teachers' likelihood of having an assigned mentor and in the amount of time teachers spent per week with a mentor. Treatment teachers were less likely than control teachers to report having an assigned mentor (20 versus 29 percent) and reported spending less time per week with a mentor (19 versus 39 minutes).

- During their second year in the classroom, treatment teachers in two-year districts received more support than did control teachers. During Year 2, we found a statistically significant difference favoring the treatment group in teachers' likelihood of having an assigned mentor and in the amount of time teachers spent per week with a mentor. Treatment teachers were more likely than control teachers to report having an assigned mentor (80 versus 34 percent) and reported spending more time per week with a mentor (82 versus 48 minutes).

- No impacts of comprehensive teacher induction were found on student achievement during teachers' second year in the classroom. In both one- and two-year districts, we did not find statistically significant impacts on student achievement across all elementary grade levels in reading or math during the teachers' second year.

- No impacts of comprehensive teacher induction were found on teacher retention rates after two years. There was also no evidence that comprehensive teacher induction induced a change in the kind of teachers retained within the district. In both one- and two-year districts, we did not find statistically significant impacts of comprehensive teacher induction on teacher retention rates in the school, district or profession after two years. In both one- and two-year districts, we did not find statistically significant impacts on the composition of the district teaching workforce after two years, whether measured by district stayers' impacts on student achievement or by their professional background characteristics (for example, SAT/ACT scores or whether the teacher attended a highly selective college).

- In a correlational (nonexperimental) analysis of induction and student test scores, the relationship between four composite induction measures (considered jointly) and test scores was statistically insignificant for both math and reading. When we tested the variables individually, one of the four measures of beginning teacher support (years had a mentor) was positively related to math scores (coefficient = 0.12, p-value = 0.015) and none were related to student achievement in reading. The significant result can be interpreted as a student scoring 12 percent of a standard deviation higher on the math test for each year the beginning teacher had a mentor. The nonexperimental results should be interpreted with caution because the analyses are correlational and not causal.

- In the correlational analysis of induction and teacher mobility, there was a positive relationship between the four composite induction measures and retention that was statistically significant for both retention in the district (p-value=0.016) and retention in the profession (p-value=0.001). When we tested the induction indices one at a time, one of the four explanatory variables was positively related to retention in the district, none were positively related to retention in the profession, and none were negatively related to either type of teacher retention. The estimate of the regression coefficient on the Induction

Services Index for remaining in the district was 0.02. This implies that, for example, if the retention rate in a district were 80 percent, then an additional induction service, such as meeting with a study group in one semester, would be associated with a district retention rate of 82 percent, all else equal. As mentioned above, the nonexperimental results should always be interpreted with caution because the analyses are correlational and not causal.

FUTURE RESEARCH

This report focused on the second year of findings, updating an earlier report (Glazerman et al. 2008) that presented results after one year of implementation for one-year and two-year districts combined. The research team is conducting a follow-up analysis that will include a third and final year of test score and teacher mobility data in one-year and two-year districts.

REFERENCES

Alliance for Excellent Education. *Tapping the Potential: Retaining and Developing High-Quality New Teachers*. New York: Carnegie Corporation, 2004.

Alt, Martha Naomi, and Robin R. Henke. "To Teach or Not to Teach: Teaching Experience and Preparation Among 1992-93 Bachelor's Degree Recipients 10 Years After College." (NCES 2007-163). Washington, DC: U.S. Department of Education, National Center for Education Statistics. July 2007.

Berry, Barnett, Peggy Hopkins-Thompson, and Mandy Hoke. "Assessing and Supporting New Teachers: Lessons from the Southeast." Chapel Hill, NC: Southeast Center for Teaching Quality, December 2002.

Fuller, Edward J. "Beginning Teacher Retention Rates for TxBESS and Non-TexBESS Teachers." Austin, TX: State Board for Educator Certification, 2003.

Glazerman, Steven, Sarah Dolfin, Martha Bleeker, Amy Johnson, Eric Isenberg, Julieta Lugo-Gil, Mary Grider, and Edward Britton. "Impacts of Comprehensive Teacher Induction: Results from the First Year of a Randomized Controlled Study." (NCEE 2009-4034). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Educational Evaluation and Regional Assistance, 2008.

Glazerman, Steven M., Sarah Senesky, Neil Seftor, and Amy Johnson. "Design of an Impact Evaluation of Teacher Induction Programs." Princeton, NJ: Mathematica Policy Research, Inc., September 2005.

Henke, Robin R., Xianglei Chen, and Sonya Geis. "Progress Through the Teacher Pipeline: 1992–93 College Graduate and Elementary/Secondary School Teaching as of 1997." Washington, DC: U.S. Department of Education, National Center for Education Statistics, 2000.

Ingersoll, Richard M., and Jeffrey M. Kralik. "The Impact of Mentoring on Teacher Retention: What the Research Says." Denver, CO: Education Commission of the States, 2004, pp. 1–23.

Ingersoll, Richard M., and Thomas M. Smith. "Do Teacher Induction and Mentoring Matter?" *NASSP Bulletin*, vol. 88, no. 638, 2004, pp. 28–40.

Interstate New Teacher Assessment and Support Consortium. "Model Standards for Beginning Teacher Licensing, Assessment and Development: A Resource for State Dialogue." Washington, DC: Council of Chief State School Officers, 1992.

Kelly, Linda Molner. "Why Induction Matters." *Journal of Teacher Education*, vol. 55, no. 5, winter 2004, pp. 438–448.

Lopez, Alejandra, Andrea Lash, Monika Schaffner, Patrick Shields, and Mary Wagner. "Review of Research on the Impact of Beginning Teacher Induction on Teacher Quality and Retention." Menlo Park, CA: SRI International, 2004.

Rockoff, Jonah. "Does Mentoring Reduce Turnover and Improve Skills of New Employees? Evidence from Teachers in New York City." NBER Working Paper 13868. Available at [www.nber.org/papers/w13868]. 2008.

Serpell, Zewelanji, and L.A. Bozeman. "Beginning Teacher Induction: A Report on Beginning Teacher Effectiveness and Retention." Washington, DC: American Association of Colleges for Teacher Education, 2000.

Smith, Thomas M., and Richard M. Ingersoll. "What Are the Effects of Induction and Mentoring on Beginning Teacher Turnover?" *American Educational Research Journal*, vol. 41, no. 3, 2004, pp. 681–714.

Totterdell, Michael, Lynda Woodroffe, Sara Bubb, and Karen Hanrahan. "The Impact of NQT Induction Programmes on the Enhancement of Teacher Expertise, Professional Development, Job Satisfaction or Retention Rates: A Systematic Review of Research Literature on Induction." London: University of London Institute of Education, EPPI-Centre, 2004.

Youngs, Peter A. "State and District Policy Related to Mentoring and New Teacher Induction in Connecticut." New York, NY: National Commission on Teaching and America's Future, 2002.