ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 20-190

Does Principal Professional Development Improve Schooling Outcomes? Evidence from Pennsylvania's Inspired Leadership Induction Program

Matthew P. Steinberg
George Mason University

Haisheng Yang
University of Pennsylvania

Principals shape the academic setting of schools. Yet, there is limited evidence on whether principal professional development improves schooling outcomes. In 2008-09, Pennsylvania's Inspired Leadership (PIL) induction program required that newly hired principals complete targeted in-service professional development tied to newly established state leadership standards within five years of employment. Using panel data on all Pennsylvania students, teachers, and principals, we employ difference-in-differences and event study strategies to estimate the impact of PIL induction on teacher and student outcomes. We find that PIL induction improved teacher effectiveness (in math) and student math achievement, and that the effects of PIL induction on teacher effectiveness were concentrated among the most economically and academically disadvantaged schools in Pennsylvania. Principal professional development had the greatest impact on teacher effectiveness when principals completed PIL induction during their first two years in the principalship. We also find evidence that teacher turnover declined in the years following the completion of PIL induction. We discuss the implications of our findings for principal induction efforts.

VERSION: April 2020

Suggested citation: Steinberg, Matthew P., and Haisheng Yang. (2020). Does Principal Professional Development Improve Schooling Outcomes? Evidence from Pennsylvania's Inspired Leadership Induction Program. (EdWorkingPaper: 20-190). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/z46p-cq19

Does Principal Professional Development Improve Schooling Outcomes? Evidence from Pennsylvania's Inspired Leadership Induction Program

Matthew P. Steinberg
George Mason University

Haisheng Yang
University of Pennsylvania

April 17, 2020

The authors thank the Pennsylvania Department of Education for providing data access and, in particular, David Volkman, Carla Wilson, and Thresa Smith for their guidance and insights about the Pennsylvania Inspired Leadership (PIL) induction program. Matthew Steinberg (corresponding author) is an associate professor at George Mason University (msteinb6@gmu.edu). Haisheng Yang (yanghais@upenn.edu) is a doctoral candidate at the University of Pennsylvania's Graduate School of Education.

Abstract

Principals shape the academic setting of schools. Yet, there is limited evidence on whether principal professional development improves schooling outcomes. In 2008-09, Pennsylvania's Inspired Leadership (PIL) induction program required that newly hired principals complete targeted in-service professional development tied to newly established state leadership standards within five years of employment. Using panel data on all Pennsylvania students, teachers, and principals, we employ difference-in-differences and event study strategies to estimate the impact of PIL induction on teacher and student outcomes. We find that PIL induction improved teacher effectiveness (in math) and student math achievement, and that the effects of PIL induction on teacher effectiveness were concentrated among the most economically and academically disadvantaged schools in Pennsylvania. Principal professional development had the greatest impact on teacher effectiveness when principals completed PIL induction during their first two years in the principalship. We also find evidence that teacher turnover declined in the years following the completion of PIL induction. We discuss the implications of our findings for principal induction efforts.

Keywords: Principal induction; professional development; educator mobility; student achievement; teacher effectiveness

JEL Codes: I20, I21, I28

Introduction

Effective school leadership is vital to school improvement efforts. Effective principals attract and retain more effective teachers (Ladd, 2011; Loeb et al., 2012), promote teacher learning and instructional development (Robinson et al., 2008; Steinberg & Sartain, 2015), and improve staff motivation, commitment, and working conditions (Leithwood et al., 2008). Further, student achievement improves when principals have greater decision-making autonomy in their schools (Clark, 2009; Steinberg, 2014; Steinberg & Cox, 2017). Indeed, among all school-level factors, principals account for nearly a quarter of the variation in student achievement, second only to teacher quality in terms of the influence on student learning (Leithwood et al., 2004). Yet, despite the widely acknowledged importance of principal human capital (Ladd, 2011; Loeb et al., 2012; Leithwood et al., 2004), little empirical evidence exists linking principal professional development to teacher and student outcomes.

In 2007, the Pennsylvania Public School Code was amended to provide principals with targeted professional development designed to place an effective school leader in all Pennsylvania schools. The policy reform, known as *Act 45*, directed the Pennsylvania Department of Education (PDE) to establish a principal induction program which focused on developing the capacity of school leaders to improve student achievement. Beginning in January 2008, all school principals employed for the first time were required to complete the Pennsylvania Inspired Leadership (PIL) induction program within the first five years of employment. According to PDE leadership, the PIL induction program, which was tied to newly

established state leadership standards, was designed to endow new principals with the skills necessary to be effective school leaders.[1]

In this paper, we rely on administrative data on all students, teachers, and principals in Pennsylvania public schools during the 2008-09 through 2015-16 school years to address the following questions: (1) Does principal professional development affect teacher retention? (2) Does principal professional development improve teacher effectiveness? (3) Does principal professional development improve student achievement? Though principal professional development via PIL induction was not randomly assigned across schools in Pennsylvania, we employ event study and difference-in-differences strategies to estimate the effect of principal professional development on teacher and student outcomes. Our empirical strategy has three features: first, treatment is defined at the school-year level by a principal's completion of PIL induction; second, identification is restricted to principal-school cells, mitigating concern that non-random principal transitions might bias estimates of PIL induction while enabling us to explicitly model multiple treatment events at the school level; and third, the construction of multiple comparison groups allows us to approximate the counterfactual for what would have happened in treatment schools in the absence of principal professional development.

We find that PIL induction improved student achievement in math through improvements in teacher human capital. Specifically, PIL induction increased student math achievement by 0.02-0.03 student-level standard deviations, and these increases are linked to approximately a 0.10 standard deviation improvement in teacher math effectiveness. Improvements in teacher effectiveness were concentrated among Pennsylvania's most economically and academically

[1] Author's communication with David Volkman, Pennsylvania Department of Education Executive Deputy Secretary and PIL Program Leader (June 16, 2016).

disadvantaged schools located in urban and rural communities. This pattern of findings suggests that the benefits of principal professional development were not only distributed unevenly across the state's schools but that the returns to principal professional development were greatest in schools serving the state's most disadvantaged students. We further find that principal professional development had the greatest impact on improvements in teacher math effectiveness when principals completed PIL induction during their first two years in the principalship. In contrast, PIL induction had no discernible effect on teacher effectiveness and student achievement in English language arts (ELA). Finally, we find evidence that, in the years following PIL induction (though, not in the year of PIL completion), teacher turnover declined by approximately 2 percentage points (or 18 percent).

This paper contributes needed empirical evidence on the efficacy of state-level principal induction efforts designed to provide in-service professional development to novice and early-career principals. Though education scholars and school practitioners widely agree on the importance of effective principal leadership, rigorous empirical evidence examining the effect of in-service principal professional development on teacher and student outcomes is limited (Murphy & Vriesenga, 2006). This limited evidence base reflects state and local policy efforts which typically focus on teacher, rather than principal, professional development despite the acknowledged importance of developing and fostering effective principals (Manna, 2015). Further, principal professional development has, historically, been subsumed under teacher professional development in terms of both content and funding. In fact, most professional development provided to principals is similar to that which is provided to teachers, reflecting state and local education agencies' failure to meaningfully distinguish between principal and teacher professional development (Haller et al., 2016; Manna, 2015; Rowland, 2017).

The recent reauthorization, in 2015, of the Elementary and Secondary Education Act (ESEA) – known as the Every Student Succeeds Act (ESSA) – incentivizes states and local school districts to invest in principal professional development. ESSA authorizes approximately $2.3 billion annually to states to improve teacher and principal human capital. State education agencies can reserve up to three percent of these ESSA funds to improve aspects of principal professional development through preservice programs, differential pay scales, and induction for early career school leaders (Haller et al., 2016; Herman et al., 2017). Yet, in order to maximize the efficacy of ESSA funds so that induction efforts can improve principal human capital, state and local education agencies must implement principal training and induction programs with an established evidence base.[2]

Evidence from this paper should inform nascent policy efforts designed to improve principal human capital through early-career induction and professional development. By examining the consequences of a statewide policy reform in Pennsylvania which codified principal induction for all novice principals statewide, we provide rigorous quasi-experimental evidence on the efficacy of principal induction to improve teacher and student outcomes. Evidence from this study provides needed empirical insight into the effect of targeted, in-service principal professional development for early career school leaders. Findings from this study should inform both state and local policymakers as well as school leaders on the efficacy of targeted principal induction for improving student and teacher performance.

[2] ESSA stipulates three tiers of evidence on educational interventions. Tier I (strong evidence) is evidence derived from a well-implemented randomized control trial. Tier II (moderate evidence) is evidence derived from a single well-designed and implemented quasi-experimental study. Tier III (promising evidence) is evidence derived from at least a single well-designed and implemented study that controls for selection bias. Our study would likely be characterized as Tier II evidence based on ESSA's designations for evidentiary rigor. The only Tier I evidence (to our knowledge) comes from Herrmann et al. (2019), which we discuss below.

Related Literature and Policy Context

To situate our study within the broader policy and research landscape, we organize our discussion of principal professional development as follows. First, we describe how principal quality matters to a range of student and schooling outcomes. Next, we describe the extant evidence on in-service principal professional development; particular attention is given to the National Institute of School Leadership (NISL) Executive Development Program, a principal professional development program adopted by many states and from which Pennsylvania drew select coursework for inclusion in its PIL induction program. We then describe the scope and nature of state-level policy efforts around principal induction and early career professional development, situating Pennsylvania's PIL induction effort in this national policy climate.

Principal Effectiveness

Principal effectiveness and school quality are inextricably linked. Effective principals improve student achievement, develop teacher talent, and manage the organization and mission of schools (Coelli & Green, 2012; Branch et al., 2012; Leithwood et al., 2008). Compared to the average principal, a principal who is one standard deviation above average improves average student achievement from the 50th to the 58th percentile in one academic year (Branch et al., 2012). Similarly, Coelli and Green (2012) estimate that a principal who is one standard deviation above mean principal quality can improve graduation rates by 2.6 percentage points and English standardized tests scores by 2.5 percentage points. Beyond student achievement, principals are instrumental in retaining teachers (Branch et al., 2012; Miller, 2013). In New York City, researchers found that a one standard deviation increase in perceived administrator quality decreases a teacher's likelihood of exit by 44 percent (Boyd et al., 2011).

Principal Professional Development in Pennsylvania

Effective principals create a shared vision for schools, promote school-wide goals, and set high-performance expectations (Leithwood et al., 2008). They develop educator human capital and provide individualized support for staff development. Effective principals manage the instructional mission of the school through targeted instructional support to teachers (Leithwood et al., 2008). Indeed, instructional leadership has been identified as the most direct influence principals have on student achievement (Robinson et al., 2008). For example, intensive instructional coaching – such as pre- and post-observation conferences where principals provide detailed feedback to teachers about their instructional performance – has been found to be positively associated with student achievement (Grissom et al., 2013; Steinberg & Sartain, 2015).

Principal Professional Development

Two avenues of professional development attempt to develop and improve principal human capital. First, principals may participate in *pre-service* training prior to starting the principalship. Second, principals may participate in *in-service* training once they assume the principalship; we differentiate between two dimensions of in-service training: *induction* and *ongoing professional development*. Induction is in-service professional development targeted to novice and early-career principals; ongoing professional development is in-service professional development targeted to more experienced principals (i.e., those not in the early stages of their careers as principals) and is typically a requirement for principals to maintain their active principal licensure. We focus on in-service professional development as the current study examines principal induction for novice and early-career principals in Pennsylvania.[3]

[3] In their review of pre-service training since 2007, Ni et al. (2017) identified 52 published articles across several education journals and research/advocacy organizations. Of these articles, 38 are implementation studies that focus on understanding the programmatic elements of pre-service programs. In the 14 studies that focused on outcomes, only two studies examine the association between pre-service principal training programs and student achievement (Corcoran et al., 2012; Gates et al., 2014).

Though there is little systematic evidence on the efficacy of principal professional development via in-service induction for early career principals (Manna, 2015; Rowland, 2017), recent evidence finds that in-service professional development had no significant impact on teacher retention or student achievement (Herrmann et al., 2019). Among a sample of 100 elementary schools from eight school districts in five states, a professional development program was randomly assigned to school dyads (within districts); the professional development program was focused on developing principal capacity to conduct structured observations of teachers' classroom instruction and provide targeted feedback to teachers (Herrmann et al., 2019). Though the professional development program provided nearly 200 hours of professional development over two years, half of it through individualized coaching, Herrmann et al. (2019) find that principals' practices changed little as a result of professional development, with no discernible effect on either teacher retention or student achievement. In their investigation of the New Leaders Program which recruits, trains, and provides induction support for novice principals, Gates et al. (2014) find that spending three years or more with a New Leaders principal increases student achievement by 0.7 to 1.3 percentile points; notably, however, principal induction was a component of the New Leaders program, limiting insight into the unique effect of in-service principal induction on student achievement.

To the best of our knowledge, ours is the first study to focus exclusively on the effects of a statewide principal induction program on a range of schooling outcomes, although other work has studied principal preparation programs in which induction was a component (Gates et al., 2014). Indeed, state-level policy efforts which have been focused on principal induction and in-service training have received scant attention in the research literature on school leaders. This

paper aims to contribute rigorous evidence on one state's efforts to improve principal human capital via a statewide principal induction policy.

National Institute of School Leadership

The National Institute of School Leadership (NISL) Executive Development Program (EDP) is a widely used principal professional development program.[4] The primary goal of NISL EDP is to provide principals the skills and knowledge to create and implement instructional coaching to promote a high-performing school environment. Three NISL courses – World Class Schooling; Focus on Teaching and Learning; and Sustaining Transformation through Capacity and Commitment – emphasize blended learning and require principals to participate in online professional learning communities, prepare for in-person sessions by completing readings and pre-work, and design an Action Learning Project (Corcoran, 2017).[5] Course delivery for NISL EDP can vary, based on whether district or state policymakers opt for NISL staff to facilitate principal trainings or for district/state staff to be trained to facilitate the program themselves. Further, NISL can be implemented at different policy levels. 10 states have adopted EDP statewide for the purpose of principal professional development.[6] Pennsylvania is the only state to use select NISL coursework for its principal induction program, which we detail in the next section.

[4] NISL EDP includes 24 days of instruction, consisting of 12 two-day units organized into three courses: World Class Schooling; Focus on Teaching and Learning; and Sustaining Transformation through Capacity and Commitment. Historically, NISL included a fourth course – Driving for Results – which was designed to improve data-driven decision-making within schools. However, this course has since been removed from the EDP coursework.

[5] The Action Learning Project asks principals to apply lessons learned in EDP to an issue specific to their school environments. This practice is supported by recent research that suggests that "application-oriented" activities, in which principals apply lessons from coursework to their own school environments, are highly effective for principal professional development (Korach & Cosner, 2017).

[6] The 10 states include Arizona, Kentucky, Louisiana, Massachusetts, Minnesota, Missouri, Mississippi, Oklahoma, Rhode Island, and Virginia. Districts in another 14 states – Alabama, California, Colorado, Florida, Georgia, Illinois, Maryland, Nevada, New Hampshire, New Mexico, North Carolina, Tennessee, Texas, and Wisconsin – have adopted EDP (or a subset of NISL courses) for the purpose of principal professional development (National Institute for School Leadership, 2017).

Evidence on the relationship between NISL EDP and student achievement, though suggestive, is positive. In Wisconsin, Corcoran (2017) finds that students experienced greater achievement growth in schools with NISL EDP trained principals compared to students in schools without NISL EDP trained principals. In Massachusetts, Nunnery et al. (2011A) similarly find that student achievement is greater in schools with NISL EDP trained principals. Nunnery et al. (2011B) find that, in Pennsylvania, schools with EDP trained principals had associated gains in annual student proficiency rates of 0.48 percentage points in mathematics and 0.54 percentage points in English Language Arts, relative to comparison schools without EDP trained principals.

While these studies inform our understanding of how PIL induction in Pennsylvania may affect student achievement, key differences exist between PIL induction and studies of NISL EDP. First, PIL induction relies on two courses from the EDP curriculum – World Class Schooling and Driving for Results – while existing evidence examines the efficacy of NISL EDP (Corcoran, 2017; Nunnery et al., 2010; Nunnery et al. 2011A; Nunnery, 2011B). Second, existing evidence on NISL EDP relies on study designs that limit the generalizability of study findings, relying on a single-cohort design (Nunnery et al., 2010; Nunnery et al., 2011B), a single school district (Corcoran, 2017), or a select number of schools (Nunnery et al., 2011A). Third, existing studies of NISL EDP rely on school-level aggregates, rather than student-level data, to measure student achievement, masking potentially important variation that not only can be leveraged to control for unit-level heterogeneity, but which likely introduces bias into estimates of principal professional development on student outcomes.

State-Level Principal Induction Policies

Even in the absence of a rigorous evidence-base on in-service principal professional development, many states have enacted principal induction policies. As of 2016, 20 states had introduced principal induction requirements via state-level policy reforms; an additional two states – Illinois and Kentucky – had policy stipulations for induction programs but were unfunded mandates (Goldrick, 2016). Further, three states—Alabama, Connecticut, and New Mexico—have some form of principal induction, but unlike the 20 states with formal induction programs, these programs are not a required component of the principalship (Goldrick, 2016). A key component of many state-level principal induction efforts is that new principals are assigned a principal mentor who provides feedback on a new principal's practice. Seventeen states include mentorship as part of their induction process; of the 15 state policies that include coursework, only three – Hawaii, Pennsylvania, and South Carolina – require specific coursework. The duration of the induction period also varies, although most states require that principals complete induction within 2 years (see Table 1).[7]

<Table 1 about here>

While principal professional development has, historically, received less policy attention than teacher professional development (Manna, 2015; Rowland, 2017), many states have begun to take advantage of ESSA's dedicated school leadership funding to create new principal professional development opportunities. For example, North Dakota is using its ESSA funding from Title II to create a Leadership Academy to provide principal professional development for employed principals and a new mentorship program for novice principals (Espinoza &

[7] Alongside state-level initiatives to improve principal professional development, school districts and private entities also provide opportunities for professional learning (Herman et al., 2017).

Cardichon, 2017).[8] The policy expansion of principal professional development and, specifically, induction programs, motivate the current study's efforts to understand one state's efforts – Pennsylvania – to improve schooling outcomes through targeted principal induction.

Principal Induction in Pennsylvania

The Pennsylvania Inspired Leadership (PIL) induction program was introduced through Pennsylvania's *Act 45* of 2007, which dramatically changed the professional development requirements for newly hired principals in Pennsylvania. Prior to *Act 45*, *Act 48* of 1999 granted principals – both novice and more experienced principals – a variety of professional development options to maintain their active certification status.[9] Under *Act 45*, principal professional development requirements were revised to include more formal coursework – via the National Institute for School Leadership (NISL) – to focus on newly established leadership standards and to establish the PIL induction program for newly hired principals. Beginning in January 2008, all school principals employed for the first time (on or after January 1, 2008) were required to complete the PIL induction program within their first five years of employment (see Table A1 for a comparison of *Act 45* and *Act 48* requirements).

The PIL induction program requires principals to complete two NISL courses designed to meet the three core leadership standards established by *Act 45* (Table A2 summarizes Pennsylvania's core and corollary leadership standards under *Act 45*). Pennsylvania adopted

[8] For more detail on state spending of ESSA funds to improve principal training, see: https://learningpolicyinstitute.org/sites/default/files/product-files/Investing_Effective_School_Leadership_BRIEF.pdf

[9] Under *Act 48*, principals could choose from the following professional development options, which they were required to complete within every five year period in order to maintain their active certification status: (i) earn six credits of collegiate study; (ii) earn six credits of PDE-approved continuing professional education courses; (iii) complete 180 hours of continuing professional education programs, activities or learning experiences through a PDE approved provider; or (iv) any combination of the above (see Table A1).

NISL coursework for its PIL induction program to support the state's newly implemented standards-based approach to school leadership. In the 2004-05 academic year, Pennsylvania's governor and Secretary of Education tasked a group of educators, policymakers, and researchers to establish leadership performance standards (Pennsylvania Department of Education, 2016). Together, this group determined a set of leadership standards deemed necessary for school leaders to improve student achievement and matched these standards to NISL courses constituting the PIL induction requirements.

The first NISL course, World-Class Schooling, is designed to provide principals with the strategic planning tools to implement a vision of high-quality teaching and student achievement. The course curriculum for World-Class Schooling is aligned with the first two core leadership standards: the school leader has the knowledge and skills to think and plan strategically, creating an organizational vision around personalized student success; and the school leader has an understanding of standards-based systems theory and design and the ability to transfer that knowledge to the leader's job as the architect of standards-based reform in the school. The second course, Driving for Results, is designed to provide principals with training to examine school data, including student achievement data, to identify school, teacher, and individual student needs. The course curriculum for Driving for Results is aligned with the third core leadership standard: the school leader has the ability to access and use appropriate data to inform decision-making at all levels of the system. Each of the two courses count for 60 professional development hours, last for six days (36 total hours per course) and have 6 hours of pre-course assignments (Table A3 provides more detail on the NISL courses that comprise both PIL induction and NISL EDP).

Completion of the PIL induction program is tied to principals' administrative licenses; if newly hired principals fail to complete PIL, they are unable to renew their licenses and can no longer continue employment as principals. Moreover, the two NISL courses that constitute PIL induction emphasize the skills that Pennsylvania and PDE policymakers believed early career principals needed to succeed. Indeed, *Act* 45, with its newly implemented leadership standards coupled with PIL induction (via NISL coursework) tied to those standards, represented Pennsylvania's effort to inject greater rigor and accountability into in-service principal professional development.

Consistent with *Act 45*, principals hired after January 1, 2008 are defined as PIL-eligible principals. Among PIL-eligible principals, those who complete two NISL courses – World Class Schooling and Driving for Results – are categorized as having completed PIL induction. Thus, Pennsylvania principals fall into two distinct groups: (i) principals hired on or after January 1, 2008 who are required to complete the PIL Induction Program within their first five years of employment;[10] and (ii) principals employed prior to January 1, 2008 who must complete their continuing professional development requirements established by *Act 48* proportional to their employment post-January 1, 2008 (e.g., if a principal was employed for only two years prior to January 1, 2008, then he/she must complete 60 percent of the remaining professional development hours in the PIL program, unless he/she completed more than 40 percent of the required hours in the first 2 years).[11]

[10] After the initial five years of employment, these principals continue to fulfill their 180 hours of professional development requirement in PIL-approved courses.

[11] PIL induction and continuing professional development is administered within one of eight administrative regions in Pennsylvania (a map of the administrative regions can be found at: https://www.education.pa.gov/Teachers%20-%20Administrators/PA%20Inspired%20Leaders/Pages/default.aspx). The course offerings can be broadly defined as: (i) PIL induction coursework; (ii) NISL, non-PIL induction coursework; and (iii) non-NISL professional development coursework. PIL induction coursework includes World-Class Schooling and Driving for Results.

Ours is the first empirical investigation of PIL induction. As previously discussed, existing evidence has examined the implementation of NISL EDP in Pennsylvania (Nunnery et al., 2011B). We improve upon this prior work in two important ways. First, Nunnery et al. (2011B) relied on school-level proficiency rates as the outcome measure; we employ rich microdata on students, teachers, and principals to estimate the effect of PIL induction on teacher and student outcomes. Second, the current study spans a longer time period and includes the population of Pennsylvania schools, as opposed to a select sample of Pennsylvania schools, as in Nunnery et. al (2011B). The use of individual-level data for the population of Pennsylvania principals enables a more rigorous quasi-experimental approach to uncovering the effect of principal professional development under *Act 45* on student and teacher outcomes. We describe the data and empirical strategy below.

Data & Sample

We construct a panel dataset for all students, teachers, and principals in all traditional and charter public schools in the state of Pennsylvania for the 2008-09 through 2015-16 school years. For each student, we observe a unique student identifier, allowing us to follow students across time; a unique school identifier; teacher identifiers; birth date, which allows us to construct student age; demographic information (race, gender); grade level; free/reduced-price lunch status (eligibility and receipt); poverty status (whether a student receives supplementary government

NISL, non-PIL induction coursework includes Focus on Teaching and Learning and Sustaining Transformation through Capacity and Commitment. Non-NISL professional development coursework includes coursework in school leadership (e.g., effective communication, setting goals and expectations for a school), data use within schools, understanding early childhood education, and emphasizing the need for student equity and career readiness (for more information on the timing or offering of courses, see: https://www.education.pa.gov/Documents/Teachers-Administrators/PA%20Inspired%20Leaders-PIL/PDE%20PIL%20Blended%20Course%20Booklet.pdf). Notably, these professional development courses are not developed by NISL, and can be developed and delivered by any entity that receives approval (e.g., universities) from PDE to provide principal professional development.

services, such as TANF or SNAP); English language learner (ELL) status; special education status; and gifted status. We observe student achievement outcomes – both scaled scores and proficiency levels for math and English language arts (ELA) – for all students in grades 3-8. For all analyses of student achievement and teacher effectiveness, we rely on math and ELA test scores from the Pennsylvania System of School Assessment (PSSA) for students in grades 3-8 (which we standardize at the subject*grade*year level).

For teachers, we observe a unique teacher identifier; a school identifier; demographic information (race, gender); date of birth; educational attainment (i.e., highest degree completed); experience (total years of educational experience in Pennsylvania); and courses taught, allowing us to link individual teachers to individual students for the purposes of constructing teacher effectiveness measures. For principals, we observe a unique principal identifier; a unique school identifier; demographics (race, gender); date of birth; educational experience (total years of educational experience in Pennsylvania); and educational attainment. Importantly, we observe PIL induction coursework taken by PIL eligible principals, as well as ongoing professional development coursework taken by principals hired before January 2008. Specifically, coursework data includes a unique principal identifier, course numbers, course names, course start dates, course end dates, and the credit hours a course is worth. The following NISL courses are included among the professional development coursework data: World-Class Schooling; Focusing on Teaching and Learning; Driving for Results; and Sustaining Transformation through Capacity and Commitment (see Table A3).

Sample

We construct our school-year sample from all traditional and charter public schools in Pennsylvania during the 2008-09 through 2015-16 school years. We limit the sample to school-

year observations with one principal in a given school year (schools with multiple assistant principals are retained in the sample) and drop those school-year observations where the principal is employed at multiple schools in a given school year. Thus, our main analytic sample ("All Schools") consists of 20,689 school-year observations across 3,187 unique public schools and 4,893 unique principals.[12] We also construct a sub-sample of our main analytic sample ("PIL Eligible") which includes school-year observations where the principal is PIL Eligible (i.e., a school's principal became a principal for the first time in Pennsylvania after January 2008 and did not complete a PIL course as an assistant principal); the PIL Eligible sample consists of 7,083 school-year observations across 1,727 unique public schools and 1,879 unique principals.

Table 2 (Panel A) summarizes the characteristics of a school's principal by PIL eligibility and PIL treatment status (i.e., PIL or No PIL). PIL treatment status is defined as the completion of the PIL induction program (i.e., completion of two NISL courses – World Class Schooling and Driving for Results) at the school-year level within principal-school cells.[13] Among the main sample ("All Schools"), treatment schools contain principals who are, on average, younger, more likely to be female, more racially diverse, less experienced in Pennsylvania public education and less likely to hold an advanced degree than comparison schools. Among the PIL Eligible sample, principals are, as expected, much more similar across treatment and comparison schools, differing (statistically) only in that treatment schools are more likely to have a female principal.[14]

[12] We start with 22,805 school-year observations and drop 51 observations due to restricting school-year observations to those where the principal was a principal in no more than one school in a given year. We drop an additional 2,065 school-year observations by restricting school-year observations to one principal per year per school. This gives us 20,689 school-year observations.

[13] Specifically, treatment status (i.e., PIL) is defined at the school-year level and indicates the year (and all subsequent years) in which a principal has completed PIL induction in the same school (i.e., principal-school cells).

[14] Although Act 45 required all principals hired after the 2008-09 school year to complete PIL induction within five years of employment, evidence indicates that compliance was imperfect among PIL eligible principals. Among the 1,879 PIL eligible principals, 600 completed PIL induction (32 percent). Of the 696 PIL eligible principals

Table 2 (Panel B) summarizes the characteristics of a school's students by PIL eligibility and treatment status. Among the main sample ("All Schools"), treatment schools serve, on average, a larger share of students who are economically disadvantaged (based on eligibility for free/reduced-price lunch) and lower- achieving (as measured by proficiency on annual state math and ELA exams) than comparison schools. Among the PIL Eligible sample, treatment schools serve, on average, a larger share of lower-achieving students than comparison schools.

<Table 2 about here>

Empirical Approach

Difference-in-Differences

To estimate the effect of PIL induction on teacher turnover and teacher effectiveness (Appendix B details the construction of teacher-level value added estimates (VAM) of teacher effectiveness), we estimate the following difference-in-differences specification:

$$(1)\quad Y_{jspt} = \beta_1 PIL_{pst} + \boldsymbol{X}_{jt}\boldsymbol{\Gamma} + \boldsymbol{Z}_{st}\boldsymbol{\gamma} + \phi_s + \delta_{dt} + \mu_{jspt}$$

where Y is an outcome (either teacher turnover or teacher effectiveness) for teacher *j* in school *s* with principal *p* in school year *t*. For estimates of teacher turnover (where Y equals 1 if teacher *j* exits school *s* at the end of school year *t*, and 0 otherwise), we include all full-time classroom teachers. For subject-specific estimates of teacher effectiveness (where Y equals teacher *j*'s math (ELA) VAM score from school year *t*), we include teachers in grades 4-8 with available math and/or ELA VAM.[15] **X** is vector of time-varying teacher characteristics, including

employed for five or more years, 348 (50 percent) completed PIL induction (319 completed PIL induction prior to or in their fifth year as principal).

[15] Grade 6-8 teachers with available math and/or ELA VAM are included for the entire study period (i.e., 2008-09 through 2015-16 school years); grade 4-5 teachers with available math and /or ELA VAM are included from the 2013-14 through 2015-16 school years due to data limitations prohibiting the linking of grade 4-5 teachers to students prior to the 2013-14 school year.

age, gender, race, years of experience (in Pennsylvania) and educational attainment. **Z** is a vector of time-varying school characteristics, including the percent of economically disadvantaged students (i.e., the share of a school's students who are free or reduced-price lunch eligible), the percent of racial/ethnic minority students, the percent of students receiving specialized services (i.e., ELL, IEP, gifted), and school size (enrollment).

PIL_{pst} is the treatment indicator of interest and equals 1 in the school year *t* (and in all subsequent years) that principal *p* completed PIL induction at school *s*. Since treatment is defined at the school-year level and because we restrict identification to principal-school cells, β_1 estimates the effect of principal *p* completing PIL induction within the same school *s*. This approach to defining treatment mitigates concern that non-random principal transitions might bias estimates of PIL induction at a given school *s* and, following Bartanen et al. (2019), enables us to explicitly model multiple treatment events at the school level. The variable ϕ_s is a school fixed effect which controls for all time-invariant (and unobserved) school-level heterogeneity; δ_{dt} is a district*year fixed effect which controls for year-specific idiosyncratic shocks common to all schools within a school district; and μ_{jspt} is a random error term. To account for serial correlation across teachers in the same schools, we cluster standard errors at the school level.

To estimate the effect of PIL induction on student achievement, we estimate the following difference-in-differences specification:

$$(2)\ Achievement_{ispt} = \beta_1 PIL_{pst} + \boldsymbol{V}_{it}\boldsymbol{\Gamma} + \boldsymbol{Z}_{st}\boldsymbol{\gamma} + \phi_s + \delta_{dt} + \mu_{ispt}$$

where *Achievement* is the academic achievement, in either math or English language arts (ELA), of student *i* attending school *s* led by principal *p* in school year *t*. Students' scaled scores are standardized within year-subject-grade in our sample to account for test differences across years. For student math and ELA achievement outcomes, estimates are based on students in

grades 3 through 8 with available math and/or ELA test scores from the Pennsylvania System of School Assessment (PSSA). **V** is a vector of time-varying student characteristics, including: age, race, gender, grade level, free/reduced price-lunch eligibility status, poverty status (whether a student receives supplementary governmental services, such as TANF or SNAP), special education status, English language learner (ELL) status, and gifted status. All other variables are defined as in equation (1), and we cluster the standard errors at the school level to account for serial correlation across students within the same school.

Event Study

To assess the identifying assumption of the difference-in-differences approach (i.e., parallel pre-trends) and to estimate the potentially dynamic effects of PIL induction, we disaggregate the treatment indicator (PIL_{pst}) from equations (1) and (2) in the following event study framework. For teacher-level outcomes, we estimate the following specification:

$$(3)\ Y_{jspt} = \sum_{r=-4+,r!=-1}^{r=4+} \beta_r \left(PIL_{ps,t+r}\right) + \boldsymbol{X}_{jt}\boldsymbol{\Gamma} + \boldsymbol{Z}_{st}\boldsymbol{\gamma} + \phi_s + \delta_{dt} + \mu_{jspt}$$

where Y is an outcome (either teacher turnover or teacher effectiveness) for teacher *j* in school *s* with principal *p* in school year *t*. In equation (3), β_r estimates the year-specific effects of PIL induction in the *r* years before and after the completion of PIL induction. All other variables are defined as in equation (1) and standard errors are clustered at the school level.

For student-achievement outcomes, we estimate the following specification:

$$(4)\ Achievement_{ispt} = \sum_{r=-4+,r!=-1}^{r=4+} \beta_r \left(PIL_{ps,t+r}\right) + \boldsymbol{V}_{it}\boldsymbol{\Gamma} + \boldsymbol{Z}_{st}\boldsymbol{\gamma} + \phi_s + \delta_{dt} + \mu_{ispt}$$

where *Achievement* is the academic achievement, in either math or English language arts (ELA), of student *i* attending school *s* led by principal *p* in school year *t*. In equation (4), β_r estimates the year-specific effects of PIL induction in the *r* years before and after the completion

of PIL induction. All other variables are defined as in equation (2) and standard errors are clustered at the school level. For estimates of equations (3) and (4), we assess the parallel trends assumption underlying the difference-in-differences approach by testing the joint significance of the pre-treatment year-specific effects (i.e., $\beta_{r=-2} = \beta_{r=-3} = \beta_{r=-4+}$), the reference category for which is the year prior to the completion of PIL induction (i.e., $\beta_{r=-1}$).

Multiple Comparison Groups

To examine the robustness of our results, we estimate equations (1) through (4) on three different school-year samples. The first sample is the main sample (*All Schools*), which includes all school-year observations in the 2008-09 through 2015-16 school years. The second sample is the PIL Eligible Sample, a subsample of the main sample, which includes school-year observations where the school principal is PIL Eligible (i.e., a school's principal became a principal for the first time in Pennsylvania after January 2008 and did not complete a PIL course as an assistant principal). The third sample is based on a matching procedure in which we construct a matched comparison group by matching, among the main sample, treated school-year observations – those in which a school's principal completed PIL induction – to untreated school-year observations.

To generate the matched samples, we follow Bartanen et al. (2019) and implement the following matching procedure. First, we estimate the probability that a given unit (i.e., teachers or students) was in a treated school-year cell. For each of three teacher-level outcomes (i.e., teacher turnover, math VAM and ELA VAM), we estimate logit models of the following form:

$$(5)\ Pr\left(PIL_{jspt} = 1\right) = f\left(\boldsymbol{X_{jt}}, \boldsymbol{Z_{st}}, \boldsymbol{\theta_t}, \mu_{jspt}\right)$$

where PIL_{jspt} is an indicator for whether teacher *j* is in a school-year cell (i.e., school *s* in year *t*) in which principal *p* completed PIL induction. **X** is a vector of teacher level covariates, including age, gender, race/ethnicity, whether a teacher holds an advanced degree, and years of educational experience in Pennsylvania. Further, **X** also includes contemporaneous teacher-level outcomes (an indicator for turnover, math VAM or ELA VAM) and a one-period lag for a given outcome. **Z** is a vector of school-level covariates, including the percent of economically disadvantaged students (i.e., the share of a school's students who are free or reduced-price lunch eligible), the percent of racial/ethnic minority students, the percent of students receiving specialized services (i.e., ELL, IEP, gifted), school size (enrollment), urbanicity indicators (whether a school is in a rural, town, suburb, or city), an indicator for charter school status and school-level proficiency rates separately for math and ELA achievement which have been standardized at the year*subject level for years *t* and *t-1*. **θ** is a vector of school year dummies and μ_{jspt} is a random error term.

For each of two student-level outcomes (i.e., math and ELA achievement), we estimate logit models of the following form:

$$(6)\ Pr\left(PIL_{ispt} = 1\right) = f\left(\boldsymbol{V_{it}}, \boldsymbol{Z_{st}}, \boldsymbol{\theta_t}, \mu_{jspt}\right)$$

where PIL_{ispt} is an indicator for whether student *i* is in a school-year cell (i.e., school *s* in year *t*) in which principal *p* completed PIL induction. **V** is a vector of student characteristics, including age, race, gender, grade level, free/reduced price-lunch eligibility status, poverty status, special education status, English language learner (ELL) status, and gifted status. Further, **V** also includes student achievement in year *t*, in either mathematics or ELA, standardized at the year*subject*grade level. All other variables are defined as in equation (5).

From equations (5) and (6), we recover five sets of propensity scores for each of the five outcomes (teacher math and ELA VAM, teacher mobility, and student math and ELA achievement). Using these propensity scores, we employ a kernel matching algorithm to construct comparison groups of teachers and students across treated and untreated school-year cells. Then, we constrain our matching procedure to units that share a common support, thereby restricting comparisons to units with similar likelihoods of treatment. Table A4 shows the number of matched units for each outcome; for any given outcome, we are able to match at least 93 percent of teachers or students in treated school-year cells. Tables A5 through A11 show the differences in observable characteristics before and after the matching procedure. The standardized differences in observable characteristics for units across treated school-year and untreated school-year cells are generally within a tenth of standard deviation, demonstrating that units in treated and untreated school-year cells are similar on observable characteristics.

Results

Did the completion of PIL induction improve teacher effectiveness in math and ELA instruction? Difference-in-differences estimates indicate that PIL induction improved teacher effectiveness in math by, on average, 0.01-0.02, which corresponds to an effect size of approximately 0.10 standard deviations of teacher math VAM. In contrast, we find no evidence of any significant improvements in teacher ELA effectiveness (see Table 3).

Though teacher math effectiveness improved, on average, following the completion of PIL induction, improvements in teacher effectiveness may not emerge until years after the completion of principal professional development. Alternatively, if the human capital benefits of principal professional development are short-lived and fade over time, the average effect of PIL induction may mask potentially important dynamic effects of principal professional development

in the years after PIL induction. Event study results indicate that improvements in teacher math effectiveness, which first occur in the year of PIL completion, persist in the years following the year of PIL completion. This represents an upward shift in teacher math effectiveness in principal-school cells where a principal completed PIL induction, and evidence that principal human capital acquired under PIL induction does not fade over time. These results are robust across comparison groups.[16]

<Table 3 about here>

Given that effective principals improve teacher retention (Ladd, 2011; Loeb et al., 2012), we next examined the impact of PIL induction on teacher turnover. Table 4 summarizes these results. Though we find no effect of PIL induction, on average, on teacher turnover, teacher turnover declines significantly in the years after PIL completion. Indeed, in the second and third years after PIL completion, teacher turnover declines by 0.02 percentage points, representing an approximately 18 percent decline relative to the mean teacher turnover rate (of 11 percent) among Pennsylvania teachers during the study period. The difference-in-differences and event study estimates of PIL induction on teacher turnover are robust across the *All Schools* and *PIL Eligible* comparison groups, with similar findings based on the *Matched* comparison group.

<Table 4 about here>

Teacher effectiveness and teacher turnover are two important channels through which improvements in principal human capital, via professional development, might operate to improve student achievement. Indeed, prior evidence suggests that teacher churn reduces student

[16] One potential is that results for teacher effectiveness may reflect survivor bias. To address this, we re-estimate models on a restricted sample that includes principals who were at the same school for a minimum of four consecutive years. The results (available from the authors upon request) are robust to restricting the school-year sample in this way.

achievement (Atteberry et al., 2017), while evidence herein indicates that PIL induction improved teacher effectiveness in math. Therefore, to what extent did PIL induction generate improvements in student achievement? Table 5 summarizes these results. On average, student math achievement improved by approximately 0.01-0.02 student-level standard deviations across the three comparison groups, though only the estimate from the matched comparison group is statistically different from zero.

<Table 5 about here>

Notably, the student achievement estimates summarized in Table 5 are based on all grade 3-8 students, independent of whether a school-grade-year observation contains teachers for whom a teacher-level VAM score can be calculated. To provide additional insight into the link between improvements in principal human capital (via PIL induction), changes in teacher effectiveness and, ultimately, changes in student achievement, we estimate the returns to PIL induction on student achievement among the same school-grade-year observations for which teacher VAM can be calculated (i.e., *VAM Teacher Sample*). These results again indicate that PIL induction generated improvements in student math achievement, on the order of 0.02-0.03 student-level standard deviations and are robust across comparison groups (see Table A12). The impact of PIL induction on student achievement first appears in the year in which a school's principal completed PIL induction and persist in the years after PIL completion. These results not only provide additional support for the positive impact of PIL induction on student achievement, but also highlight a key mechanism through which principal professional development improves student achievement; namely, through improvements in teachers' instructional effectiveness.

Heterogeneous Effects of PIL Induction

Schools in Pennsylvania vary in the characteristics of the students that they serve and the geographic settings in which its schools are located (see Table 2). The effect of PIL induction might therefore also vary across schools that serve different student populations. Indeed, evidence suggests that principals adjust their leadership behaviors based on the characteristics of their school settings (Goldring et al., 2008; May et al., 2012). We examine five dimensions of school settings for which the consequences of principal professional development might differentially impact teacher and student outcomes: (i) poverty; (ii) racial/ethnic minority; (iii) achievement; (iv) geographic location (i.e., urbanicity); and (v) sector (charter or traditional public schools). Tables 6-8 summarize the heterogeneous effects of PIL induction as a function of school poverty, racial/ethnic minority, and achievement; Table 9 summarizes the heterogeneous effects of PIL induction as a function of school geographic location and sector.

We find that the impact of PIL induction on teacher (math) effectiveness is concentrated among those schools serving the most academically struggling and economically disadvantaged students. Indeed, PIL induction improves teacher (math) effectiveness more in schools with the highest share of students eligible for free/reduced-price lunch, schools with the highest share of racial/ethnic minority students, and in schools with the lowest average student achievement. In contrast, PIL induction had no impact on teacher effectiveness in schools serving the most economically and academically advantaged students and in schools serving the lowest share of minority students (see Table 6).

Consistent with evidence of greater returns to PIL induction in the most disadvantaged Pennsylvania schools, we further find that PIL induction generated greater improvements in urban and rural schools than in schools located in suburban communities (Table 9). And, while

we find little to no evidence of heterogeneous effects of PIL induction on teacher turnover (Table 7), we find that, among the *VAM Teacher Sample*, student math achievement increases by 0.03 standard deviations in the most economically disadvantaged schools. This result provides further evidence that improvements to student achievement and teacher effectiveness driven by PIL induction are concentrated among the most disadvantaged schools (Table A13).

<Tables 6-9 about here>

The effects of professional development might also depend on the timing of when early-career principals complete PIL induction. Specifically, principals who complete PIL induction earlier in their tenure as principal may benefit differently than principals who complete PIL induction later in their (early-career) tenure. For example, principals who delay the completion of PIL induction until later in their (early-career) tenure may be more resistant to incorporating the skills taught via PIL induction than those who are in their first or second year in the principalship. Table 10 summarizes evidence on the potentially differential returns to PIL induction by the year in which a principal completed PIL.

We find that principal professional development has the greatest impact on teacher effectiveness when principals complete PIL induction in their first two years as principal. Indeed, among school-year observations in which the principal is in his/her first two years as school leader, PIL induction improves teacher math effectiveness by 0.04 (or nearly 0.40 standard deviations); in contrast, PIL induction has no significant impact on teacher effectiveness among school-year observations with principals in at least their 3rd year as school leader. We find no differential effect of PIL induction on teacher turnover by the year of PIL completion.

<Table 10 about here>

Discussion

Principals are among the most important inputs to the operation and performance of schools. Yet, little work has examined whether efforts to improve principal human capital via in-service induction and professional development can positively affect a range of schooling outcomes. In this paper, we examine a statewide policy reform in Pennsylvania aimed to improve principal human capital through targeted professional development for novice principals tied to the state's leadership standards – the PIL induction program. Relying on difference-in-differences and event study strategies, we estimate the effect of PIL induction on teacher effectiveness, teacher turnover and student achievement. We find that PIL induction improved teacher effectiveness (in math) and student math achievement, and that the effects of PIL induction on teacher effectiveness were concentrated among the most economically and academically disadvantaged schools in Pennsylvania. In contrast, we find no impact of PIL induction on ELA achievement or teacher effectiveness (in ELA). We do find evidence that teacher turnover declined by approximately 2 percentage points in the years following PIL induction, but find no evidence that PIL induction decreased teacher turnover in the year of PIL completion. Finally, principal professional development had the greatest impact on teacher effectiveness when principals completed PIL induction during their first two years as principal.

The heterogeneous effects of PIL induction – which are concentrated among schools serving the lowest-achieving and most economically disadvantaged students in the state of Pennsylvania – speak to the need to tailor principal induction to school contexts. Not all schools are created equal, and prior research shows that principal behavior varies with school contexts (Goldring et al., 2008; May et al., 2012). Our investigation of PIL induction reveals heterogeneous returns to teacher effectiveness based on school characteristics and the timing of

PIL completion in the early part of a principal's career. Moreover, we find suggestive evidence that PIL induction improves student math achievement in Pennsylvania's most disadvantaged schools. Together, these findings reveal the importance of considering school context and the local human capital needs of principals in very different school settings. Indeed, policymakers should recognize that providing targeted professional development to principals who lead the lowest performing, most disadvantaged schools, has the potential to generate the greatest improvements in teacher effectiveness and student achievement.

Notably, our study offers a methodological advancement over prior studies of NISL-specific coursework and principal induction. Prior studies of NISL relied on school-level measures instead of individual-level data (Nunnery et al., 2011A; Nunnery et al. 2011B), studied principal induction in individual school districts (Corcoran, 2017), or relied on just one cohort of principals participating in professional development (Nunnery et al., 2011A; Nunnery et al. 2011B). In comparison, we leverage individual-level panel data for the population of Pennsylvania principals in both traditional and charter public schools. This data allows us to estimate the effect of PIL induction among all Pennsylvania principals between the 2008-09 and 2015-16 school years. Further, our econometric approach, which relies on quasi-experimental strategies and multiple comparison groups, mitigates concern that the timing of a principal's completion of PIL induction at the principal-school level is correlated with the teacher and student outcomes that we study. However, given the programmatic differences between PIL induction in Pennsylvania and prior studies of NISL, we cannot conclusively say that the differences between our findings are driven solely by methodological differences.

Though this study offers a rigorous empirical investigation of a statewide principal induction program in Pennsylvania, a deeper understanding of the implementation and effects of

principal induction would likely benefit from additional analyses. For example, prior evidence finds that when principals participate in effective preparation programs, they gain an understanding of effective leadership practices and later use those leadership practices in their schools (Orr & Orphanos, 2011). By interviewing principals who completed PIL induction, we could learn more about the aspects of principal induction that supported (and constrained) the development of principals' leadership skills. Such interviews could shed additional light on principal perceptions about why the completion of PIL induction had a disproportionate effect across schools in Pennsylvania. Indeed, we do not observe important aspects of PIL implementation and the extent of principals' experiences in PIL coursework, including the quality of PIL instructors who deliver the course content, the extent of principal engagement in the PIL coursework, nor the quality of course instruction or specific course materials. Both interviews with and surveys of principals could provide important insights into the extent to which principals were engaged with PIL training and employed specific aspects of their PIL induction training in their schools. For example, were principals at high-poverty schools more (or less) likely to apply certain leadership practices than principals in schools serving more economically advantaged students? These, among other questions, could support state and local policymakers in their efforts to refine and improve principal induction programs targeted at improving the academic settings of schools, the effectiveness of teachers and, ultimately, student academic achievement. Despite these limitations, this paper is the first to study a state-wide principal induction reform and contributes much needed evidence on how principal professional development shapes teacher and student outcomes.

References

Atteberry, A., Loeb, S., & Wyckoff, J. (2015). Do first impressions matter? Predicting early career teacher effectiveness. *AERA Open*, *1*(4), 1-23.

Atteberry, A., Loeb, S., & Wyckoff, J. (2017). Teacher churning: Reassignment rates and implications for student achievement. *Educational Evaluation and Policy Analysis*, *39*(1), 3-30.

Bartanen, B., Grissom, J. A., & Rogers, L. K. (2019). The Impacts of Principal Turnover. *Educational Evaluation and Policy Analysis*, *41*(3), 350-374.

Boyd, D., Grossman, G., Ing, M., Lankford, H., Loeb, S. & Wyckoff, J. (2011). The influence of school administrators on teacher retention decisions. *American Educational Research Journal, 48*(2) 303-333.

Branch, G. F., Hanushek, E. A., & Rivkin, S. G. (2012). Estimating the effect of leaders on public sector productivity: The case of school principals (No. w17803). *National Bureau of Economic Research.*

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, *104*(9), 2593-2632.

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

Chetty, R., Friedman, J., & Rockoff, J. (2014c). Discussion of the American Statistical Association's Statement (2014) on using value-added models for educational assessment. *Statistics and Public Policy*, *1*(1), 111-113.

Clark, D. (2009). The performance and competitive effects of school autonomy. *Journal of Political Economy 117*(4)*,* 745–783.

Coelli, M., & Green, D. A. (2012). Leadership effects: School principals and student outcomes. *Economics of Education Review*, *31*(1), 92-109.

Corcoran, R. P. (2017). Preparing principals to improve student achievement. *Child & Youth Care Forum, 46*(5), 769-781.

Corcoran, S. P., Schwartz, A. E., & Weinstein, M. (2012). Training your own: The impact of New York City's aspiring principals program on student achievement. *Educational Evaluation and Policy Analysis*, *34*(2), 232-253.

Espinoza, D., & Cardichon, J. (2017). Investing in effective school leadership: How states are taking advantage of opportunities under ESSA. *Palo Alto, CA: Learning Policy Institute*.

Gates, S. M., Hamilton, L. S., Martorell, P., Burkhauser, S., Heaton, P., Pierson, A., Baird, M., Vuollo, M., Li, J., Lavery, D., Harvey, M., & Gu, K. (2014). Preparing Principals to Raise Student Achievement: Implementation and Effects of the New Leaders Program in Ten Districts. *RAND Corporation*.

Goldrick, L. (2016). Support from the start: A 50-state review of policies on new educator induction and mentoring. *New Teacher Center*.

Goldring, E., Huff, J., May, H., & Camburn, E. (2008). School context and individual characteristics: what influences principal practice?. *Journal of Educational Administration*, *46*(3), 332-352.

Grissom, J. A., & Bartanen, B. (2019). Principal effectiveness and principal turnover. *Education Finance and Policy*, *14*(3), 355-382.

Grissom, J. A., & Loeb, S. (2017). Assessing principals' assessments: Subjective evaluations of teacher effectiveness in low-and high-stakes environments. *Education Finance and Policy*, *12*(3), 369-395.

Grissom, J. A., Loeb, S., & Master, B. (2013). Effective instructional time use for school leaders: Longitudinal evidence from observations of principals. *Educational Researcher*, *42*(8), 433-444.

Grossman, P., Loeb, S., Cohen, J., & Wyckoff, J. (2013). Measure for measure: The relationship between measures of instructional practice in middle school English language arts and teachers' value-added scores. *American Journal of Education*, *119*(3), 445-470.

Guarino, C. M., Maxfield, M., Reckase, M. D., Thompson, P. N., & Wooldridge, J. M. (2015). An evaluation of empirical Bayes's estimation of value-added teacher performance measures. *Journal of Educational and Behavioral Statistics*, *40*(2), 190-222.

Haller, A., Hunt, E., Pacha, J., & Fazekas, A. (2016). Lessons for States: The" Every Student Succeeds Act" (ESSA) Increases Focus on and Investment in Supporting Principal Preparation and Development. *Center for the Study of Education Policy*.

Heck, R. H., & Hallinger, P. (2009). Assessing the contribution of distributed leadership to school improvement and growth in math achievement. *American Educational Research Journal*, *46*(3), 659-689.

Herman, R., Gates, S.M., Arifkhanova, A., Bega, A., Chavez-Herreias, E.R., Han, E., Harris, M., Tamargo, J. and Wrabel, S.L. (2017). School Leadership Interventions Under the Every Student Succeeds Act: Evidence Review. *RAND Corporation.*

Herrmann, M., Clark, M., James-Burdumy, S., Tuttle, C., Kautz, T., Knechtel, V., Dotter, D., Wulsin, C. S., and Deke, J. (2019). *The Effects of a Principal Professional Development Program Focused on Instructional Leadership: Appendices* (NCEE 2020-0002). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Jacob, R., Goddard, R., Kim, M., Miller, R., & Goddard, Y. (2015). Exploring the causal impact of the McREL Balanced Leadership Program on leadership, principal efficacy, instructional climate, educator turnover, and student achievement. *Educational Evaluation and Policy Analysis*, *37*(3), 314-332.

Kane, T. J., & Staiger, D. O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation (No. w14607). *National Bureau of Economic Research.*

Korach, S., & Cosner, S. (2017). Developing the leadership pipeline: Comprehensive leadership development. *Handbook of Research on the Education of School Leaders*, 262-282.

Kraft, M. A. (2019). Teacher effects on complex cognitive skills and social-emotional competencies. *Journal of Human Resources*, *54*(1), 1-36.

Ladd, H. F. (2011). Teachers' perceptions of their working conditions: How predictive of planned and actual teacher movement?. *Educational Evaluation and Policy Analysis*, *33*(2), 235-261.

Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership. *School Leadership & Management, 28*(1), 27–42.

Leithwood, K., Seashore Louis, K., Anderson, S., & Wahlstrom, K. (2004). Review of research: How Leadership Influences Student Learning. *The Wallace Foundation.*

Loeb, S., Kalogrides, D., & Béteille, T. (2012). Effective schools: Teacher hiring, assignment, development, and retention. *Education Finance and Policy*, *7*(3), 269-304.

Manna, P. (2015). Developing Excellent School Principals to Advance Teaching and Learning: Considerations for State Policy. *The Wallace Foundation*.

May, H., Huff, J., & Goldring, E. (2012). A longitudinal study of principals' activities and student performance. *School Effectiveness and School Improvement*, *23*(4), 417-439.

Miller, A. (2013). Principal turnover and student achievement. *Economics of Education Review*, *36*, 60-72.

Morganstein, D., & Wasserstein, R. (2014). ASA statement on value-added models. *Statistics and Public Policy*, *1*(1), 108-110.

Murphy, J., & Vriesenga, M. (2006). Research on school leadership preparation in the United States: An analysis. *School Leadership and Management*, *26*(2), 183-195.

National Institute of School Leadership (2016). NISL Executive Development Program. Retrieved from http://www.education.pa.gov/Documents/Teachers-Administrators/PA%20Inspired%20Leaders-PIL/NISL%20Overview%20Description.pdf

National Institute of School Leadership (2017). Retrieved from http://www.nisl.org/

Ni, Y., Hollingworth, L., Rorrer, A., & Pounder, D. (2017). The evaluation of educational leadership preparation programs. *Handbook of Research on the Education of School Leaders*, 285-308.

Nunnery, J. A., Ross, S. M., & Yen, C. J. (2010). The Effect of the National Institute for School Leadership's Executive Development Program on School Performance Trends in Pennsylvania. *Online Submission*.

Nunnery, J. A., Ross, S. M., Chappell, S., Pribesh, S., & Hoag-Carhart, E. (2011A). The Impact of the NISL Executive Development Program on School Performance in Massachusetts: Cohort 2 Results. *Online Submission*.

Nunnery, J. A., Yen, C. J., & Ross, S. M. (2011B). Effects of the National Institute for School Leadership's Executive Development Program on School Performance in Pennsylvania: 2006-2010 Pilot Cohort Results. *Online Submission*.

Orr, M. T., & Orphanos, S. (2011). How graduate-level preparation influences the effectiveness of school leaders: A comparison of the outcomes of exemplary and conventional leadership preparation programs for principals. *Educational Administration Quarterly*, *47*(1), 18-70.

Pennsylvania Department of Education (2016). PA Inspired Leadership Legislation: Frequently Asked Questions. Retrieved from: http://www.education.pa.gov/Teachers%20-%20Administrators/PA%20Inspired%20Leaders/Pages/PIL-Legislation-FAQ.aspx#tab-1

Pennsylvania Department of Education (2016). Principal Induction Program. Retrieved from: http://www.education.pa.gov/Teachers%20-%20Administrators/PA%20Inspired%20Leaders/Pages/Principals'-Induction-Program.aspx#tab-1

Pennsylvania Department of Education (2018). Invitation to Qualify: Information and Instructions. Retrieved from: https://www.education.pa.gov/Documents/Teachers-Administrators/PA%20Inspired%20Leaders-PIL/PIL%20-%20Invitation%20to%20Qualify.pdf

Robinson, V. M. J., Lloyd, C. A., & Rowe, K. J. (2008). The impact of leadership on school outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly, 44*(5), 635–674.

Rowland, C. (2017). Principal Professional Development: New Opportunities for a Renewed State Focus. *Education Policy Center at American Institutes for Research*.

Steinberg, M. P. (2014). Does greater autonomy improve school performance? Evidence from a regression discontinuity analysis in Chicago. *Education Finance and Policy, 9*(1), 1-35.

Steinberg, M. P., & Cox, A. B. (2017). School autonomy and district support: how principals respond to a tiered autonomy initiative in Philadelphia public schools. *Leadership and Policy in Schools, 16*(1), 130-165.

Steinberg, M. P., & Kraft, M. A. (2017). The sensitivity of teacher performance ratings to the design of teacher evaluation systems. *Educational Researcher, 46*(7), 378-396.

Steinberg, M. P., & Sartain, L. (2015). Does teacher evaluation improve school performance? Experimental evidence from Chicago's Excellence in Teaching project. *Education Finance and Policy, 10*(4), 535-572.

Tables & Figures

Table 1. State-Level Principal Induction Policies

State	Coursework	Mentorship	Tied to Licensure	Dosage/Duration of Induction
Arkansas	Yes	Yes	Yes	New principals must complete mentorship within first 3 years
California	Yes	Yes	Yes	New principals must complete 2 years of induction
Colorado	Yes	Yes	Yes	New principals must complete induction within first 3 years
Delaware	Yes	Yes	Yes	New principals receive 30 hours of mentorship in the first year; additional induction may be completed within first 3 years
Hawaii	Yes	Yes	n/a	New principals must participate in Hawaii's New Principal Academy within the first 2 years of principalship
Iowa	Yes	Yes	n/a	New principals must complete mentorship, which is required for the first year and may last upwards of 3 years
Kansas	Yes	Yes	Yes	New principals must complete one year of mentorship
Maryland	Yes	Yes	n/a	New principals must complete one year of induction
Massachusetts	Yes	Yes	Yes	New principals must complete one year of induction
Missouri	Yes	No	n/a	New principals must complete two years of induction
New Jersey	Yes	Yes	Yes	New principals must complete a one-year residency program
New York	No	Yes	n/a	New principals must complete one year of mentorship
Pennsylvania	Yes	No	Yes	New principals have 5 years to complete the PIL Induction program
South Carolina	Yes	No	n/a	New principals must complete one year of induction
Texas	Yes	Yes	n/a	New principals must complete one year of induction
Utah	No	Yes	n/a	New principals must complete one year of mentorship
Vermont	No	Yes	n/a	New principals must complete two years of mentorship
Virginia	No	Yes	n/a	New principals must complete one year of mentorship
West Virginia	Yes	Yes	n/a	New principals must complete one year of induction
Wisconsin	No	Yes	n/a	New principals must complete up to five years of mentorship

Notes. State-level policy summary of principal induction derived from Goldrick (2016). *Coursework* indicates whether a principal is required to complete formal coursework as part of principal induction. *Mentorship* indicates whether a principal is required to receive mentorship as part of principal induction. *Induction Tied to Licensure* indicates whether completion of induction is a requirement of principals to obtain and maintain their principal license. Cells with "n/a" indicate that information on principal licensure was not available. *Dosage/Duration of Induction* indicates the type and length of induction required of new principals. Goldrick (2016) can be downloaded from: https://newteachercenter.org/wp-content/uploads/2016CompleteReportStatePolicies.pdf; individual state summaries can be downloaded from: https://newteachercenter.org/policy/state-policy-reviews/.

Table 2. Principal and School Characteristics, by PIL Treatment Status

	All Schools			PIL Eligible		
	All	**No PIL**	**PIL**	**All**	**No PIL**	**PIL**
Panel A: Principal Characteristics						
Age	46.66 (8.62)	46.91 (8.67)	43.36*** (7.13)	43.61 (8.16)	43.68 (8.42)	43.36 (7.13)
Female	0.45	0.44	0.52***	0.47	0.45	0.52***
White	0.88	0.88	0.86**	0.85	0.85	0.86
Black	0.10	0.10	0.11*	0.12	0.12	0.11
Hispanic	0.01	0.01	0.01	0.02	0.02	0.01***
Other Race	0.01	0.01	0.02***	0.01	0.01	0.02**
Experience	19.06 (9.17)	19.32 (9.25)	15.66*** (7.39)	15.33 (8.23)	15.24 (8.44)	15.66* (7.39)
Bachelor's Degree	0.13	0.13	0.18***	0.18	0.17	0.18
Advanced Degree	0.86	0.87	0.82***	0.82	0.82	0.82
Panel B: School Characteristics						
Enrollment	610.25 (368.3)	615.07 (372.06)	548.65*** (309.93)	590.11 (367.89)	601.26 (381.24)	548.65*** (309.93)
Female	0.48 (0.04)	0.48 (0.04)	0.48 (0.04)	0.49 (0.05)	0.49 (0.05)	0.48 (0.04)
Age	10.83 (3.32)	10.87 (3.32)	10.38*** (3.17)	10.87 (3.34)	11.00 (3.37)	10.38*** (3.17)
White	0.71 (0.31)	0.71 (0.31)	0.69** (0.32)	0.67 (0.34)	0.67 (0.34)	0.69*** (0.32)
Minority	0.23 (0.30)	0.23 (0.30)	0.24 (0.30)	0.27 (0.33)	0.28 (0.33)	0.24*** (0.30)
FRPL	0.44 (0.27)	0.43 (0.27)	0.49*** (0.28)	0.48 (0.28)	0.48 (0.28)	0.49* (0.28)
IEP	0.16 (0.06)	0.16 (0.07)	0.16 (0.05)	0.16 (0.07)	0.17 (0.07)	0.16 (0.05)
ELL	0.03 (0.05)	0.03 (0.05)	0.03** (0.05)	0.03 (0.06)	0.03 (0.06)	0.03 (0.05)
Gifted	0.04 (0.05)	0.04 (0.05)	0.03*** (0.03)	0.03 (0.04)	0.03 (0.04)	0.03* (0.03)
Math Proficiency	0.65 (0.24)	0.66 (0.23)	0.56*** (0.26)	0.6 (0.25)	0.61 (0.25)	0.56*** (0.26)
ELA Proficiency	0.67 (0.19)	0.68 (0.18)	0.62*** (0.20)	0.64 (0.20)	0.64 (0.20)	0.62*** (0.20)
Charter	0.04	0.04	0.03*	0.08	0.09	0.03***

City	0.19	0.19	0.21**	0.23	0.23	0.21
Suburban	0.45	0.45	0.42**	0.43	0.43	0.42
Rural	0.25	0.24	0.28***	0.25	0.24	0.28***
Town	0.11	0.12	0.08***	0.09	0.10	0.08
Principals	4,893	4,855	571	1,879	1,841	571
Schools	3,187	3,180	551	1,727	1,678	551
School*Years	20,689	19,188	1,501	7,083	5,582	1,501

Notes. In Panel A, proportions are reported, except for age and experience, which report means (standard deviation). In Panel B, school-level means (standard deviation) reported, except for charter status and urbanicity, which report proportions. The sample includes K-12 traditional and charter public schools present in any school year during the 2008-09 through 2015-16 school years. *PIL Eligible* includes schools with principals hired during or after the 2008-09 school year who did not complete World Class Schooling or Driving for Results as assistant principals. Treatment status (i.e., *PIL*) is defined at the school-year level and indicates the year (and all subsequent years) in which a principal has completed PIL induction in the same school (i.e., principal-school cells). Differences between *PIL* and *No PIL*, within a sample, are statistically significant at *10%, **5%, and ***1%.

Table 3. Effect of PIL Induction on Teacher Effectiveness

	Math VAM						ELA VAM					
	All Schools		PIL Eligible		Matched		All Schools		PIL Eligible		Matched	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PIL	0.01***		0.02***		0.01*		0.00		-0.00		-0.00	
	(0.00)		(0.01)		(0.00)		(0.00)		(0.00)		(0.00)	
4+ Years Before		-0.01		-0.07		-0.17***		-0.01		-0.04		-0.05
		(0.04)		(0.05)		(0.06)		(0.02)		(0.03)		(0.04)
3 Years Before		0.01		-0.03		-0.02		-0.00		-0.01		0.01
		(0.02)		(0.02)		(0.02)		(0.01)		(0.02)		(0.01)
2 Years Before		0.01*		-0.00		-0.01		0.01*		-0.00		0.01
		(0.01)		(0.01)		(0.02)		(0.00)		(0.01)		(0.01)
Year of PIL		0.02***		0.02***		0.01**		0.00		0.00		0.00
		(0.00)		(0.01)		(0.01)		(0.00)		(0.00)		(0.00)
1 Year After		0.01**		0.01**		-0.00		-0.00		-0.00		-0.01
		(0.00)		(0.01)		(0.01)		(0.00)		(0.00)		(0.00)
2 Years After		0.01**		0.01**		-0.00		0.00		-0.00		-0.00
		(0.01)		(0.01)		(0.01)		(0.00)		(0.01)		(0.01)
3 Years after		0.01**		0.02**		-0.00		0.00		0.00		-0.01
		(0.01)		(0.01)		(0.01)		(0.00)		(0.01)		(0.01)
4+ Years After		0.02**		0.02*		-0.00		0.01		0.01		-0.00
		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)
P-Value from F-Statistic:												
Before PIL		0.87		0.06		0.02		0.45		0.34		0.31
Year of/After PIL		0.61		0.72		0.52		0.48		0.33		0.40
Outcome Mean (SD)	0.00 (0.09)		0.00 (0.09)		0.00 (0.09)		0.00 (0.06)		0.00 (0.06)		0.00 (0.06)	

Teacher* Years	60,053	60,053	23,220	23,220	30,763	30,763	71,365	71,365	27,027	27,027	36,772	36,772
Teachers	23,542	23,542	11,263	11,263	14,014	14,014	27,682	27,682	12,931	12,931	16,639	16,639
School* Years	9,462	9,462	3,706	3,706	6,176	6,176	9,439	9,439	3,701	3,701	6,277	6,277
Schools	2,256	2,256	1,199	1,199	2,006	2,006	2,258	2,258	1,198	1,198	2,019	2,019
R^2	0.23	0.23	0.25	0.25	0.30	0.30	0.25	0.25	0.28	0.28	0.32	0.32

Notes. Each column within a panel represents a separate regression. Coefficients are reported with robust standard errors (clustered at the school level). All regressions include school and district-year fixed effects. Teacher math (ELA) VAM constructed for grades 6-8 teachers in the 2008-09 through 2015-16 school years and grades 4-5 teachers in the 2013-14 through 2015-16 school years. *All Schools* is the main analytic sample; *PIL Eligible* includes schools with principals hired during or after the 2008-09 school year who did not complete World Class Schooling or Driving for Results as assistant principals. The *Matched* sample is constructed by matching, among the *All Schools* sample, teachers in treated school-year observations – those in which a school's principal completed PIL induction – to teachers in untreated school-year observations. *PIL* is an indicator variable, defined at the school-year level, which equals one in the year (and all subsequent years) in which a principal has completed PIL induction in the same school (i.e., principal-school cells). *P-Value from F-Statistic* displays the p-values of F-tests of the joint significance of the pre-treatment effects (i.e., *Before PIL*) and of the post-treatment effects (i.e., *Year of/After PIL*). Coefficients statistically significant at *10% **5% and ***1% levels.

Table 4. Effect of PIL Induction on Teacher Turnover

	All Schools		PIL Eligible		Matched	
	(1)	(2)	(3)	(4)	(5)	(6)
PIL	-0.00 (0.00)		0.00 (0.00)		0.01* (0.00)	
4+ Years Before		0.00 (0.01)		0.00 (0.02)		-0.02 (0.02)
3 Years Before		-0.00 (0.01)		-0.01 (0.01)		-0.02* (0.01)
2 Years Before		-0.00 (0.00)		0.00 (0.01)		-0.01 (0.01)
Year of PIL		0.00 (0.00)		0.01 (0.01)		0.01 (0.01)
1 Year After		-0.00 (0.00)		-0.00 (0.01)		0.00 (0.01)
2 Years After		-0.02*** (0.01)		-0.02*** (0.01)		-0.01* (0.01)
3 Years after		-0.02*** (0.01)		-0.02* (0.01)		-0.00 (0.01)
4+ Years After		-0.01 (0.01)		-0.00 (0.01)		0.00 (0.01)
P-Value from F-Statistic:						
Before PIL		0.93		0.54		0.69
Year of/After PIL		0.00		0.00		0.02
Outcome Mean	0.11		0.12		0.11	
Teacher*Years	786,840	786,840	258,421	258,421	522,331	522,331
Teachers	154,654	154,654	78,837	78,837	127,294	127,294
School*Years	20,669	20,669	7,073	7,073	15,779	15,779
Schools	3,181	3,181	1,723	1,723	2,914	2,914
R^2	0.06	0.06	0.08	0.08	0.08	0.08

Notes. Each column within a panel represents a separate regression. Coefficients are reported with robust standard errors (clustered at the school level). All regressions include school and district-year fixed effects. Sample includes all K-12 traditional and charter public schools present in any school year during the 2009-09 through 2015-16 school years. *All Schools* is the main analytic sample; *PIL Eligible* includes schools with principals hired during or after the 2008-09 school year who did not complete World Class Schooling or Driving for Results as assistant principals. The *Matched* sample is constructed by matching, among the *All Schools* sample, teachers in treated school-year observations – those in which a school's principal completed PIL induction – to teachers in untreated school-year observations. *PIL* is an indicator variable, defined at the school-year level, which equals one in the year (and all subsequent years) in which a principal has completed PIL induction in the same school (i.e., principal-school cells). *P-Value from F-Statistic* displays the p-values of F-tests of the joint significance of the pre-treatment effects (i.e., *Before PIL*) and of the post-treatment effects (i.e., *Year of/After PIL*). Coefficients statistically significant at *10% **5% and ***1% levels.

Table 5. Effect of PIL Induction on Student Achievement

	Mathematics						ELA					
	All Schools		PIL Eligible		Matched		All Schools		PIL Eligible		Matched	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PIL	0.01 (0.01)		0.02 (0.01)		0.02*** (0.01)		0.00 (0.01)		0.00 (0.01)		0.02*** (0.01)	
4+ Years Before		-0.01 (0.02)		-0.02 (0.03)		-0.02 (0.04)		-0.01 (0.02)		-0.03 (0.03)		0.00 (0.03)
3 Years Before		-0.00 (0.01)		-0.01 (0.02)		0.00 (0.02)		0.00 (0.01)		0.00 (0.02)		0.02 (0.01)
2 Years Before		-0.01 (0.01)		0.00 (0.02)		-0.01 (0.01)		-0.01* (0.01)		0.00 (0.01)		-0.01 (0.01)
Year of PIL		0.01 (0.01)		0.02 (0.01)		0.02*** (0.01)		0.00 (0.01)		0.00 (0.01)		0.02*** (0.01)
1 Year After		0.00 (0.01)		0.01 (0.01)		0.02* (0.01)		-0.01 (0.01)		-0.01 (0.01)		0.01 (0.01)
2 Years After		0.01 (0.01)		0.02 (0.02)		0.02 (0.01)		-0.00 (0.01)		-0.01 (0.02)		0.01 (0.01)
3 Years after		0.02 (0.01)		0.03 (0.02)		0.04*** (0.01)		0.01 (0.01)		0.01 (0.02)		0.02* (0.01)
4+ Years After		0.02 (0.02)		0.01 (0.02)		0.04** (0.02)		-0.00 (0.02)		-0.01 (0.02)		0.03 (0.02)
<u>P-Values from F-Statistic:</u>												
Before PIL		0.62		0.70		0.54		0.21		0.62		0.20
Year of/After PIL		0.62		0.43		0.16		0.20		0.14		0.34
Outcome Mean (SD)	0.00 (1.00)		-0.06 (1.00)		0.01 (1.00)		0.00 (1.00)		-0.06 (1.01)		0.04 (0.99)	

Student* Years	5,064,932	5,064,932	1,713,460	1,713,460	4,033,845	4,033,845	5,073,554	5,073,554	1,712,089	1,712,089	4,020,425	4,020,425
Students	1,673,099	1,673,099	822,728	822,728	1,429,136	1,429,136	1,661,463	1,661,463	819,659	819,659	1,422,954	1,422,954
School* Years	16,486	16,486	5,653	5,653	13,695	13,695	16,486	16,486	5,653	5,653	13,701	13,701
Schools	2,559	2,559	1,387	1,387	2,464	2,464	2,559	2,559	1,387	1,387	2,464	2,464
R^2	0.40	0.40	0.41	0.41	0.39	0.39	0.41	0.41	0.43	0.43	0.40	0.40

Notes. Each column within a panel represents a separate regression. Coefficients are reported with robust standard errors (clustered at the school level). All regressions include school and district-year fixed effects. The sample includes K-12 traditional and charter public schools with tested grades (3-8) in any school year during the 2008-09 through 2015-16 school years. *All Schools* is the main analytic sample; *PIL Eligible* includes schools with principals hired during or after the 2008-09 school year who did not complete World Class Schooling or Driving for Results as assistant principals. The *Matched* sample is constructed by matching, among the *All Schools* sample, teachers in treated school-year observations – those in which a school's principal completed PIL induction – to teachers in untreated school-year observations. *PIL* is an indicator variable, defined at the school-year level, which equals one in the year (and all subsequent years) in which a principal has completed PIL induction in the same school (i.e., principal-school cells). *P-Value from F-Statistic* displays the p-values of F-tests of the joint significance of the pre-treatment effects (i.e., *Before PIL*) and of the post-treatment effects (i.e., *Year of/After PIL*). Coefficients statistically significant at *10% **5% and ***1% levels.

Table 6. Heterogeneous Effects of PIL on Teacher Effectiveness, by School Characteristics

	Math VAM			ELA VAM		
	Low	**Medium**	**High**	**Low**	**Medium**	**High**
Panel A: FRPL						
PIL	0.00 (0.01)	0.02** (0.01)	0.02*** (0.01)	-0.00 (0.01)	0.01 (0.01)	0.00 (0.00)
Tercile Mean (SD)	0.16 (0.08)	0.39 (0.06)	0.77 (0.18)	0.16 (0.08)	0.39 (0.06)	0.77 (0.18)
Teacher*Years	19,729	19,751	20,510	24,631	23,469	23,212
Teachers	7,929	8,746	9,987	9,779	10,374	11,322
School*Year	2,565	3,201	3,633	2,559	3,201	3,626
Schools	769	972	1,046	768	978	1,044
R^2	0.25	0.28	0.21	0.26	0.31	0.23
Panel B: Minority						
PIL	0.01 (0.01)	0.01** (0.01)	0.02*** (0.01)	-0.00 (0.01)	0.00 (0.00)	0.00 (0.00)
Tercile Mean (SD)	0.03 (0.01)	0.09 (0.04)	0.56 (0.28)	0.03 (0.01)	0.09 (0.04)	0.56 (0.28)
Teacher*Years	16,982	21,975	21,053	20,789	26,803	23,738
Teachers	7,052	9,064	9,389	8,423	11,039	10,631
School*Year	2,973	2,940	3,506	2,979	2,942	3,483
Schools	805	885	867	808	886	866
R^2	0.27	0.24	0.21	0.31	0.26	0.22
Panel C: Achievement						
PIL	0.02*** (0.01)	0.01 (0.01)	0.00 (0.01)	0.00 (0.00)	0.00 (0.01)	-0.01 (0.01)
Tercile Mean (SD)	0.38 (0.15)	0.71 (0.06)	0.87 (0.04)	0.38 (0.15)	0.71 (0.06)	0.87 (0.04)
Teacher*Years	24,944	19,789	15,214	28,417	23,465	19,405
Teachers	14,592	12,263	8,414	16,635	14,345	10,431
School*Year	4,000	3,232	2,124	4,003	3,219	2,139
Schools	1,686	1,520	966	1,686	1,521	972
R^2	0.24	0.29	0.27	0.25	0.32	0.30

Notes. Each column within a panel represents a separate regression. Coefficients are reported with robust standard errors (clustered at the school level). All regressions include school and district-year fixed effects. Estimates based on *All Schools* sample. Teacher math (ELA) VAM constructed for grades 6-8 teachers in the 2008-09 through 2015-16 school years and grades 4-5 teachers in the 2013-14 through 2015-16 school years. *FRPL* is the proportion of students in a school*year who are eligible to receive free-or-reduced price lunch. *Minority* is the proportion of students in a school*year who are either Black or Hispanic. *Achievement* is the proportion of students in school*year who tested either proficient or advanced on the mathematics standardized exam. Terciles (*Low*, *Medium*, *High*) are based on

school-by-year characteristics. *PIL* is an indicator variable, defined at the school-year level, which equals one in the year (and all subsequent years) in which a principal has completed PIL induction in the same school (i.e., principal-school cells). Coefficients statistically significant at *10% **5% and ***1% levels.

Table 7. Heterogeneous Effects of PIL on Teacher Turnover, by School Characteristics

	Low	Medium	High
Panel A: FRPL			
PIL	-0.00	0.00	-0.01
	(0.01)	(0.00)	(0.01)
Tercile Mean	0.16	0.39	0.76
(SD)	(0.08)	(0.06)	(0.18)
Teacher*Years	297,691	250,594	238,513
Teachers	71,576	70,472	64,195
School*Years	6,893	6,886	6,888
Schools	1,422	1,604	1,442
R^2	0.04	0.06	0.08
Panel B: Minority			
PIL	-0.00	-0.01	-0.01
	(0.00)	(0.00)	(0.01)
Tercile Mean	0.03	0.09	0.57
(SD)	(0.01)	(0.04)	(0.28)
Teacher*Years	242,627	280,362	263,850
Teachers	55,754	66,459	62,990
School*Years	6,880	6,895	6,893
Schools	1,298	1,370	1,241
R^2	0.05	0.04	0.08
Panel C: Achievement			
PIL	-0.01	0.02***	0.00
	(0.01)	(0.01)	(0.01)
Tercile Mean	0.37	0.71	0.88
(SD)	(0.16)	(0.06)	(0.04)
Teacher*Years	273,500	264,533	227,019
Teachers	103,913	101,020	67,446
School*Years	6,593	6,590	6,589
Schools	2,250	2,298	1,641
R^2	0.09	0.06	0.06

Notes. Each column within a panel represents a separate regression. Coefficients are reported with robust standard errors (clustered at the school level). All regressions include school and district-year fixed effects. Estimates based on *All Schools* sample. *FRPL* is the proportion of students in a school*year who are eligible to receive free-or-reduced price lunch. *Minority* is the proportion of students in a school*year who are either Black or Hispanic. *Achievement* is the proportion of students in school*year who tested either proficient or advanced on the mathematics standardized exam. Terciles (*Low*, *Medium*, *High*) are based on school-by-year characteristics. *PIL* is an indicator variable, defined at the school-year level, which equals one in the year (and all subsequent years) in which a principal has completed PIL induction in the same school (i.e., principal-school cells). Coefficients statistically significant at *10% **5% and ***1% levels.

Table 8. Heterogeneous Effects of PIL on Student Achievement, by School Characteristics

	Math			ELA		
	Low	**Medium**	**High**	**Low**	**Medium**	**High**
Panel A: FRPL						
PIL	0.02	-0.01	0.01	0.01	-0.01	-0.00
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Tercile Mean	0.16	0.39	0.77	0.16	0.39	0.77
(SD)	(0.08)	(0.06)	(0.18)	(0.08)	(0.06)	(0.18)
Student*Years	1,819,785	1,605,509	1,639,637	1,829,922	1,610,677	1,632,954
Students	686,188	672,146	660,189	683,390	671,103	655,471
School*Years	5,243	5,395	5,847	5,243	5,395	5,847
Schools	1,079	1,255	1,205	1,079	1,255	1,205
R^2	0.33	0.31	0.35	0.34	0.33	0.37
Panel B: Minority						
PIL	0.01	0.01	0.02	-0.02	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Tercile Mean	0.03	0.09	0.57	0.03	0.09	0.57
(SD)	(0.01)	(0.04)	(0.28)	(0.01)	(0.04)	(0.28)
Student*Years	1,479,984	1,872,790	1,712,158	1,487,441	1,879,285	1,706,828
Students	563,168	725,646	649,382	561,090	723,190	643,595
School*Years	5,477	5,481	5,528	5,477	5,481	5,528
Schools	1,050	1,096	997	1,050	1,096	997
R^2	0.32	0.35	0.40	0.34	0.35	0.42
Panel C: Achievement						
PIL	0.02	0.00	0.00	0.01	-0.02	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Tercile Mean	0.38	0.72	0.88	0.38	0.72	0.88
(SD)	(0.15)	(0.06)	(0.04)	(0.15)	(0.06)	(0.04)
Student*Years	1,456,203	1,600,675	2,008,053	1,450,507	1,608,847	2,014,199
Students	811,721	899,483	840,188	808,601	898,393	837,543
School*Years	4,632	5,415	6,438	4,632	5,415	6,438
Schools	1,783	1,944	1,572	1,783	1,944	1,572
R^2	0.42	0.35	0.30	0.45	0.37	0.32

Notes. Each column within a panel represents a separate regression. Coefficients are reported with robust standard errors (clustered at the school level). All regressions include school and district-year fixed effects. Estimates based on *All Schools* sample. *FRPL* is the proportion of students in a school*year who are eligible to receive free-or-reduced price lunch. *Minority* is the proportion of students in a school*year who are either Black or Hispanic. *Achievement* is the proportion of students in school*year who tested either proficient or advanced on the mathematics standardized exam. Terciles (*Low*, *Medium*, *High*) are based on school-by-year characteristics. *PIL* is an indicator variable, defined at the school-year level, which equals one in the year (and all subsequent years) in which a principal has completed

PIL induction in the same school (i.e., principal-school cells). Coefficients statistically significant at *10% **5% and ***1% levels.

Table 9. Heterogeneous Effects of PIL, by Urbanicity and Sector

	Urbanicity				School Sector	
	City	**Suburb**	**Town**	**Rural**	**TPS**	**Charter**
Panel A: Math VAM						
PIL	0.02*** (0.01)	0.01* (0.00)	-0.01 (0.02)	0.02** (0.01)	0.01*** (0.00)	0.03 (0.02)
Teacher*Years	11,805	29,881	5,856	12,495	57,985	2,011
Teachers	5,493	11,525	2,330	5,092	22,382	1,209
School*Years	2,272	3,973	947	2,254	8,915	490
Schools	509	1,034	240	583	2,126	126
R^2	0.19	0.23	0.27	0.29	0.24	0.28
Panel B: ELA VAM						
PIL	0.00 (0.00)	-0.00 (0.00)	0.00 (0.01)	0.01 (0.01)	0.00 (0.00)	-0.03* (0.01)
Teacher*Years	12,971	36,189	7,261	14,928	69,199	2,126
Teachers	5,987	13,864	2,844	6,034	26,445	1,299
School*Years	2,251	3,968	951	2,253	8,889	510
Schools	508	1,031	243	584	2,128	129
R^2	0.18	0.25	0.30	0.32	0.26	0.30
Panel C: Teacher Turnover						
PIL	-0.01 (0.01)	-0.00 (0.00)	-0.02* (0.01)	0.00 (0.00)	-0.00 (0.00)	-0.01 (0.03)
Teacher*Years	146,021	379,686	87,805	173,328	761,978	24,862
Teachers	35,796	78,014	21,659	41,198	146,866	9,528
School*Years	4,015	9,226	2,353	5,075	19,807	862
Schools	690	1,489	469	953	3,015	166
R^2	0.07	0.04	0.06	0.05	0.06	0.09
Panel D: Math Achievement						
PIL	0.01 (0.02)	0.01* (0.01)	-0.00 (0.01)	0.03** (0.01)	0.02** (0.01)	-0.01 (0.08)
Student*Years	977,823	2,515,197	522,791	1,049,121	4,896,103	168,829
Students	365,335	895,490	215,006	435,108	1,627,681	77,027
School*Years	3,283	7,391	1,742	4,070	15,765	721
Schools	560	1,190	354	777	2,417	142
R^2	0.36	0.37	0.29	0.31	0.39	0.31
Panel E: ELA Achievement						

PIL	-0.00 (0.02)	0.00 (0.01)	-0.03* (0.02)	0.03** (0.01)	0.01 (0.01)	-0.00 (0.05)
Student*Years	972,141	2,522,447	525,259	1,053,707	4,904,688	168,866
Students	360,891	890,954	214,475	434,349	1,616,185	76,904
School*Years	3,283	7,391	1,742	4,070	15,765	721
Schools	560	1,190	354	777	2,417	142
R^2	0.39	0.39	0.32	0.33	0.41	0.34
Panel F: Math Achievement (VAM Teacher Sample)						
PIL	0.02 0.02)	0.04*** 0.01)	0.02 0.03)	0.01 0.02)	0.03** 0.01)	-0.01 0.08)
Student*Years	770,026	1,722,897	341,151	670,827	3,361,757	143,144
Students	340,974	823,579	176,947	348,953	1,538,174	70,566
School*Years	2,275	3,975	947	2,265	8,920	542
Schools	510	1,036	240	585	2,126	130
R^2	0.38	0.40	0.32	0.34	0.42	0.33
Panel G: ELA Achievement (VAM Teacher Sample)						
PIL	0.01 0.02)	0.01 0.01)	-0.03 0.03)	0.02 0.02)	0.01 0.01)	0.00 0.04)
Student*Years	640,131	1,682,619	331,373	649,217	3,188,608	114,732
Students	325,926	814,638	174,549	344,330	1,512,456	65,386
School*Years	2,257	3,968	951	2,261	8,891	546
Schools	509	1,033	243	586	2,128	130
R^2	0.42	0.41	0.34	0.36	0.44	0.37

Notes. Each column within a panel represents a separate regression. Coefficients are reported with robust standard errors (clustered at the school level). All regressions include school and district-year fixed effects. In Panels A-E, estimates based on *All Schools* sample; in Panels F-G, estimates based on *All Schools* sample that includes school-grade-year observations with teacher VAM (i.e., *VAM Teacher Sample*). *PIL* is an indicator variable, defined at the school-year level, which equals one in the year (and all subsequent years) in which a principal has completed PIL induction in the same school (i.e., principal-school cells). Coefficients statistically significant at *10% **5% and ***1% levels.

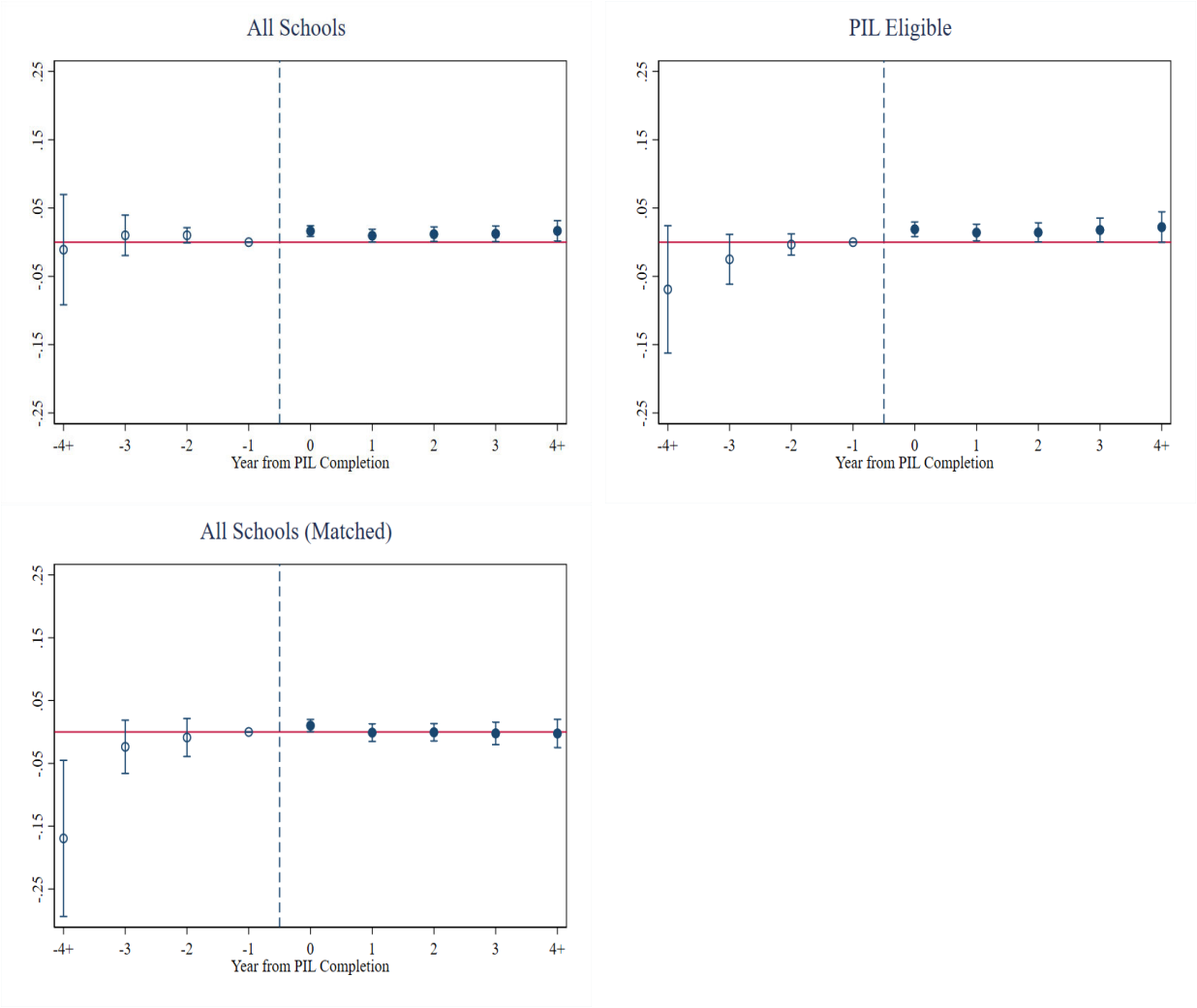
Table 10. Heterogeneous Effects of PIL, by Year Principal Completed PIL

	Year of PIL Completion		
	1-2 Years	**3-4 Years**	**5+ Years**
Panel A: Math VAM			
PIL	0.04** (0.02)	0.02 (0.01)	0.01 (0.02)
Teacher*Years	2,472	5,013	1,476
Teachers	1,291	2,447	725
School*Years	390	798	285
Schools	148	252	82
R^2	0.29	0.27	0.29
Panel B: ELA VAM			
PIL	0.02* (0.01)	0.00 (0.01)	0.01 (0.01)
Teacher*Years	2,699	5,942	1,657
Teachers	1,429	2,832	804
School*Years	394	791	281
Schools	146	251	82
R^2	0.35	0.27	0.33
Panel C: Teacher Turnover			
PIL	-0.02 (0.03)	0.02* (0.01)	0.02 (0.02)
Teacher*Years	21,729	52,395	22,610
Teachers	8,158	16,271	6,926
School*Years	669	1,537	653
Schools	200	360	135
R^2	0.08	0.07	0.08
Panel D: Math Achievement			
PIL	-0.02 (0.03)	0.00 (0.02)	0.03 (0.07)
Student*Years	156,455	370,372	120,804
Students	88,753	198,160	65,032
School*Years	577	1,277	522
Schools	170	305	104
R^2	0.40	0.39	0.43
Panel E: ELA Achievement			

PIL	-0.01 (0.03)	0.01 (0.02)	0.05 (0.06)
Student*Years	156,183	369,875	120,498
Students	88,638	197,360	64,715
School*Years	578	1,277	522
Schools	170	305	104
R^2	0.42	0.41	0.44
Panel F: Math Achievement (VAM Teacher Sample)			
PIL	-0.02 (0.02)	0.01 (0.03)	0.03 (0.08)
Student*Years	118,862	267,823	77,954
Students	71,725	152,712	45,169
School*Years	402	802	287
Schools	148	253	82
R^2	0.42	0.41	0.45
Panel G: ELA Achievement (VAM Teacher Sample)			
PIL	-0.02 (0.02)	0.04** (0.02)	0.02 (0.05)
Student*Years	114,099	253,311	71,754
Students	70,495	148,245	43,515
School*Years	403	804	283
Schools	148	254	82
R^2	0.44	0.43	0.47

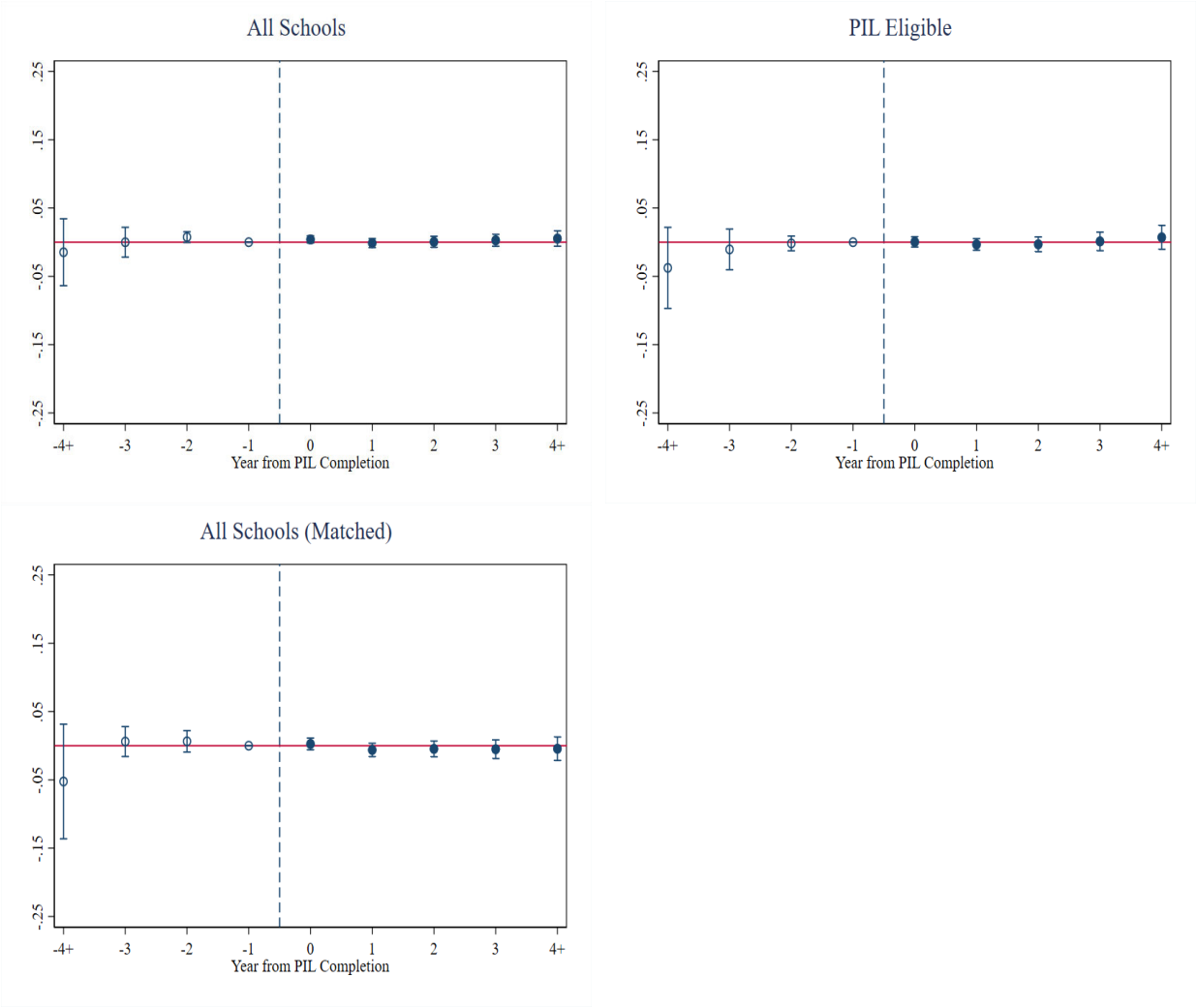
Notes. Each column within a panel represents a separate regression. Coefficients are reported with robust standard errors (clustered at the school level). All regressions include school and district-year fixed effects. In Panels A-E, estimates based on *All Schools* sample; in Panels F-G, estimates based on *All Schools* sample that includes school-grade-year observations with teacher VAM (i.e., *VAM Teacher Sample*). *PIL* is an indicator variable, defined at the school-year level, which equals one in the year (and all subsequent years) in which a principal has completed PIL induction in the same school (i.e., principal-school cells). *Year of PIL Completion* indicates the year during a principal's career he/she completed PIL induction. Coefficients statistically significant at *10% **5% and ***1% levels.

Figure 1. Event Study Estimates of PIL Induction on Teacher Effectiveness (Math VAM)



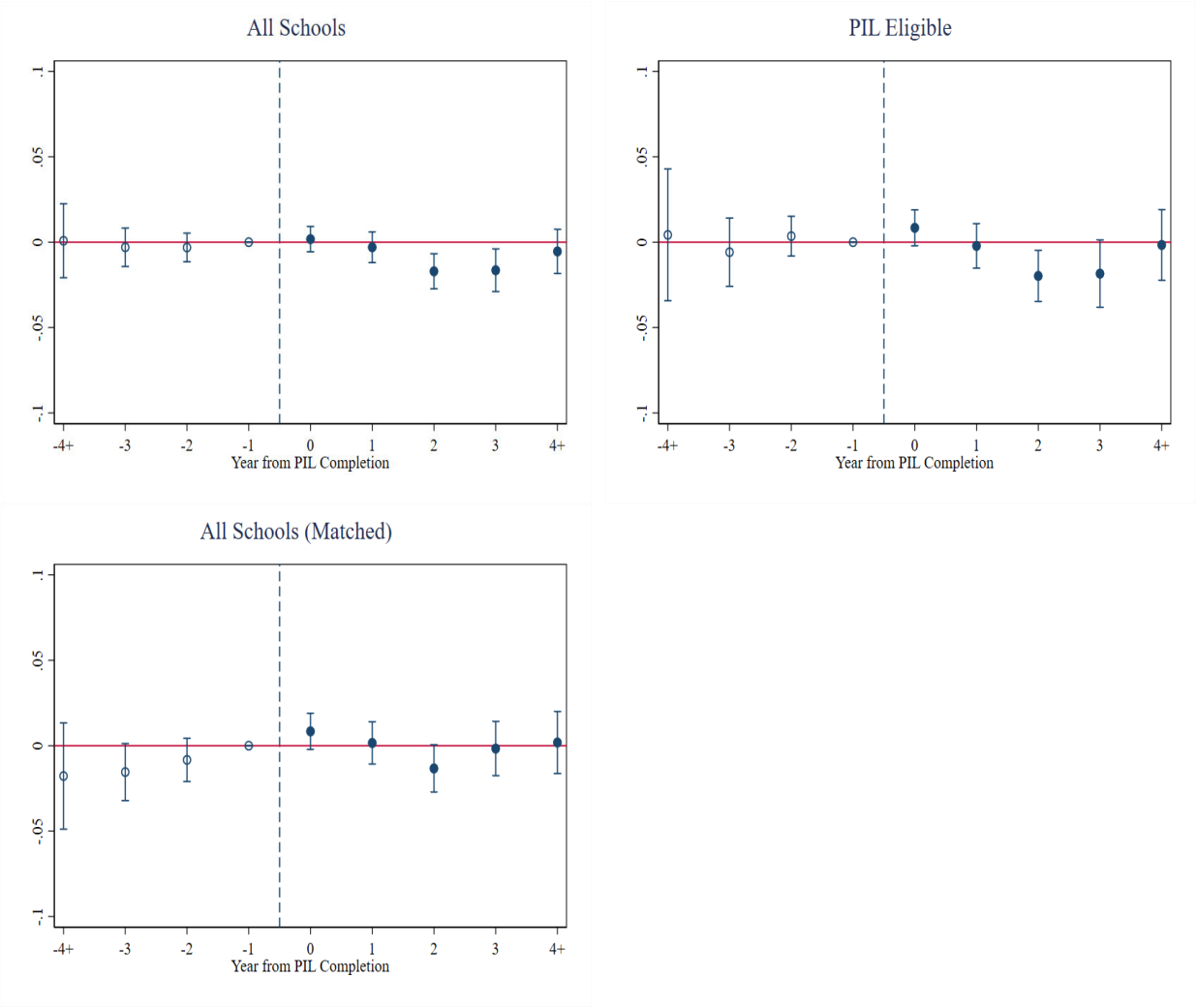
Notes. Figures report coefficient estimates and 95 percent confidence intervals for teacher effectiveness (math VAM). Coefficients reported relative to the year prior to PIL completion. *All Schools* includes all school-year observations in the 2008-09 through 2015-16 school years. *PIL Eligible* is a subset of *All Schools* and includes school-year observations where the school principal is PIL Eligible (i.e., a school's principal became a principal for the first time in Pennsylvania after January 2008 and did not complete a PIL course as an assistant principal). *All Schools (Matched)* is a subsample of *All Schools* and is based on a matching procedure in which we construct a matched comparison group by matching treated school*year observations – those in which a school's principal completed PIL induction – to untreated school*year observations. See Table 3 for event study estimates.

Figure 2. Event Study Estimates of PIL Induction on Teacher Effectiveness (ELA VAM)



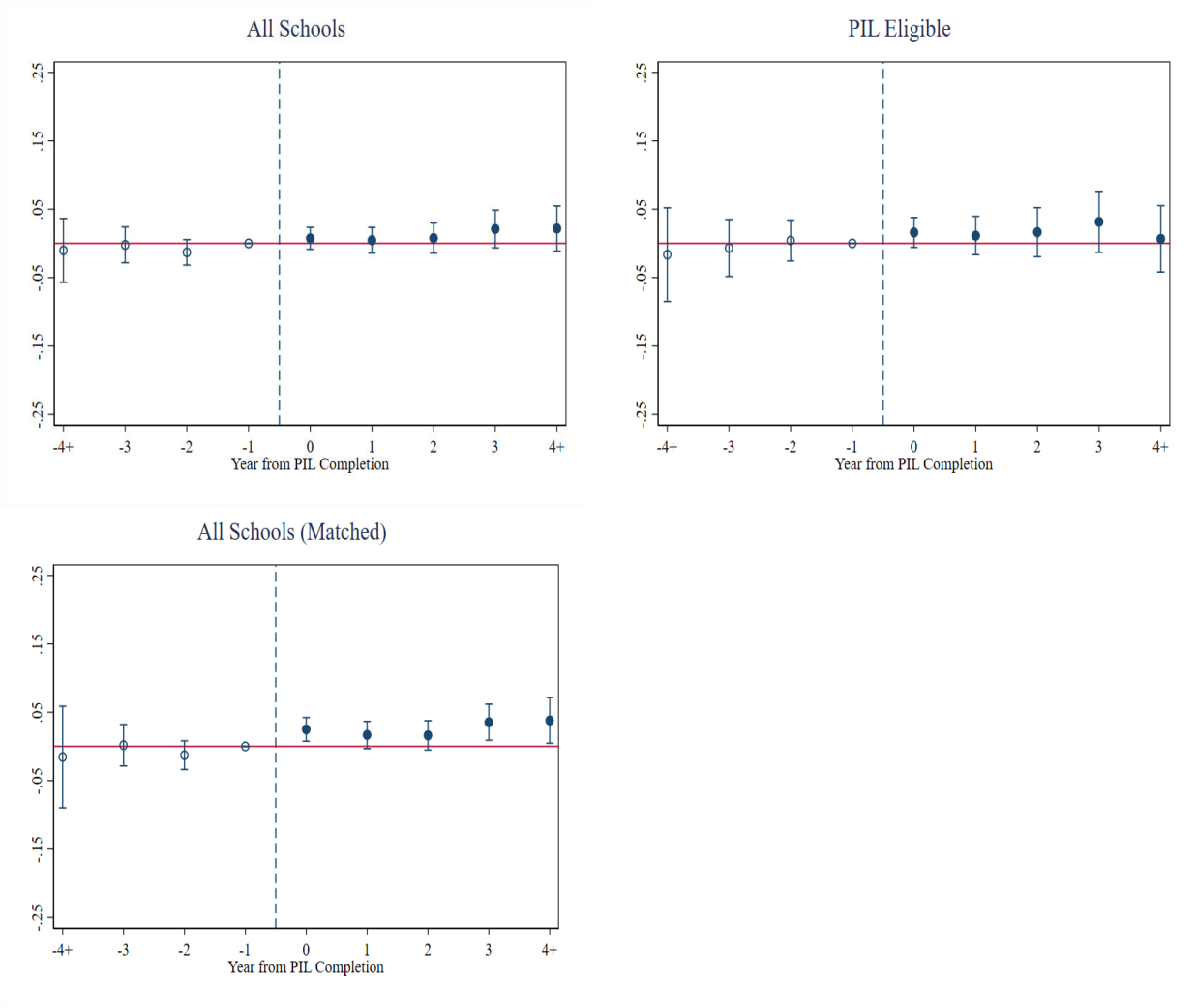
Notes. Figures report coefficient estimates and 95 percent confidence intervals for teacher effectiveness (ELA VAM). Coefficients reported relative to the year prior to PIL completion. *All Schools* includes all school-year observations in the 2008-09 through 2015-16 school years. *PIL Eligible* is a subset of *All Schools* and includes school-year observations where the school principal is PIL Eligible (i.e., a school's principal became a principal for the first time in Pennsylvania after January 2008 and did not complete a PIL course as an assistant principal). *All Schools (Matched)* is a subsample of *All Schools* and is based on a matching procedure in which we construct a matched comparison group by matching treated school*year observations – those in which a school's principal completed PIL induction – to untreated school*year observations. See Table 3 for event study estimates.

Figure 3. Event Study Estimates of PIL Induction on Teacher Turnover



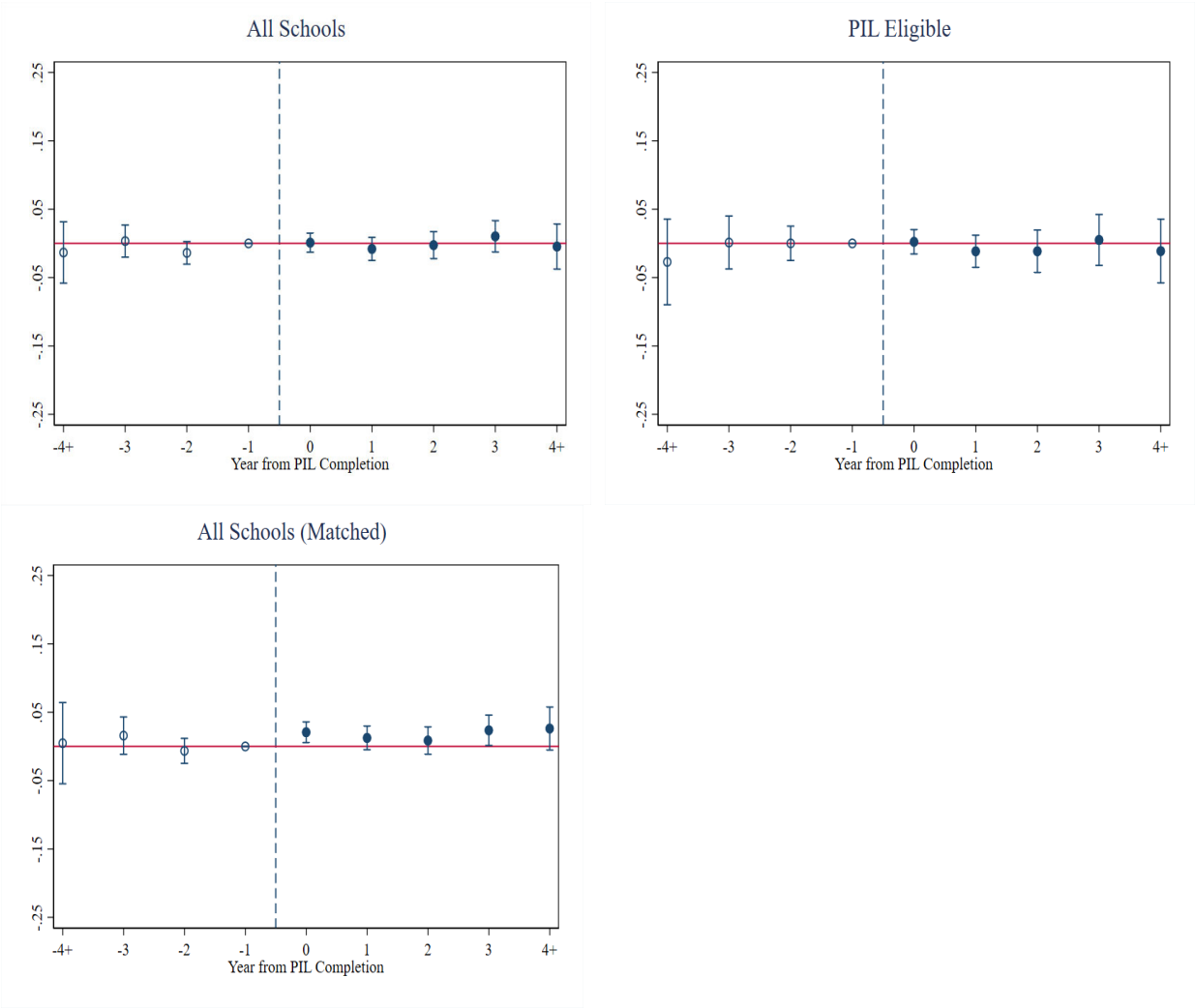
Notes. Figures report coefficient estimates and 95 percent confidence intervals for teacher mobility. Coefficients reported relative to the year prior to PIL completion. *All Schools* includes all school-year observations in the 2008-09 through 2015-16 school years. *PIL Eligible* is a subset of *All Schools* and includes school-year observations where the school principal is PIL Eligible (i.e., a school's principal became a principal for the first time in Pennsylvania after January 2008 and did not complete a PIL course as an assistant principal). *All Schools (Matched)* is a subsample of *All Schools* and is based on a matching procedure in which we construct a matched comparison group by matching treated school*year observations – those in which a school's principal completed PIL induction – to untreated school*year observations. See Table 4 for event study estimates.

Figure 4. Event Study Estimates of PIL Induction on Student Math Achievement



Notes. Figures report coefficient estimates and 95 percent confidence intervals for student math achievement. Coefficients reported relative to the year prior to PIL completion. *All Schools* includes all school-year observations in the 2008-09 through 2015-16 school years. *PIL Eligible* is a subset of *All Schools* and includes school-year observations where the school principal is PIL Eligible (i.e., a school's principal became a principal for the first time in Pennsylvania after January 2008 and did not complete a PIL course as an assistant principal). *All Schools (Matched)* is a subsample of *All Schools* and is based on a matching procedure in which we construct a matched comparison group by matching treated school*year observations – those in which a school's principal completed PIL induction – to untreated school*year observations. See Table 5 for event study estimates.

Figure 5. Event Study Estimates of PIL Induction on Student ELA Achievement



Notes. Figures report coefficient estimates and 95 percent confidence intervals for student ELA achievement. Coefficients reported relative to the year prior to PIL completion. *All Schools* includes all school-year observations in the 2008-09 through 2015-16 school years. *PIL Eligible* is a subset of *All Schools* and includes school-year observations where the school principal is PIL Eligible (i.e., a school's principal became a principal for the first time in Pennsylvania after January 2008 and did not complete a PIL course as an assistant principal). *All Schools (Matched)* is a subsample of *All Schools* and is based on a matching procedure in which we construct a matched comparison group by matching treated school*year observations – those in which a school's principal completed PIL induction – to untreated school*year observations. See Table 5 for event study estimates.

Appendix A. Supplemental Tables & Figures

Table A1. Professional Development and Induction Requirements under Pennsylvania *Act 48* and *Act 45*

	Act 48	Act 45
Year Enacted	1999	2007
Effective Date	July 1, 2000	January 1, 2008
Professionals Affected	All educators in Pennsylvania that hold the following certificates: Instructional I and II, Educational Specialist I and II, Administrative, Supervisory, Letters of Eligibility, and all vocational certificates	Principals, Assistant/Vice Principals, Superintendents, Assistant Superintendents, Intermediate Unit Executive Director, Intermediate Unit Assistant Executive Director, Director of an Area Vocational-Technical School
Continuing Professional Development Requirements	Every five years, educators must either earn six hours of college credits, six credits of PDE approved professional development courses, 180 hours of professional development programs approved by PDE, or any combination of the three	Principals employed before January 1, 2008 must complete their professional development requirements proportional to their employment period (e.g., if a principal has worked was employed for one year before January 1, 2008, she must complete 80% of her professional development requirements in a PIL course)
Principal Induction Requirements	N/A	All principals employed for the first time on or after January 1, 2008 must complete the Pennsylvania Inspired Leadership Induction Program within five years of employment
Alignment to Pennsylvania Standards	N/A	Aligned to 3 core leadership standards and 6 corollary standards
Consequence of Not Meeting Requirements	Suspension of license, resulting in suspension of employment	Suspension of license, resulting in suspension of employment

Notes. Source: Pennsylvania Department of Education (www.education.pa.gov). Although Act 48 affects all Pennsylvania educators, Act 45 only affects those employed as principals, assistant/vice principals, superintendents, assistant superintendents, intermediate unit executive directors, intermediate unit assistant executive directors, or directors of an area vocational-technical schools (i.e., school or district leaders).

Table A2. Pennsylvania Leadership Standards

Core Leadership Standards	Corollary Leadership Standards
• The leader has the knowledge and skills to think and plan strategically, creating an organizational vision around personalized student success	• The leader knows how to create a culture of teaching and learning with an emphasis on learning
• The leader has an understanding of standards-based systems theory and design and the ability to transfer that knowledge to the leader's job as the architect of standards-based reform in the school	• The leader knows how to manage resources for effective results
• The leader has the ability to access and use appropriate data to inform decision-making at all levels of the system	• The leader knows how to collaborate, communicate, engage and empower others inside and outside of the organization to pursue excellence in learning • The leader knows how to operate in a fair and equitable manner with personal and professional integrity • The leader knows how to advocate for children and public education in the larger political, social, economic, legal and cultural context • The leader knows how to support professional growth of self and others through practice and inquiry

Notes. Source: Pennsylvania Department of Education (www.education.pa.gov). The Pennsylvania Inspired Leadership (PIL) induction program focuses on the 3 core leadership standards.

Table A3. National Institute of School Leadership (NISL) Coursework

Course Title	Course Description	Course Units
Panel A: PIL Courses		
World Class Schooling: Vision and Goals	This course focuses on providing principals the strategic planning tools to implement a vision of high-quality teaching and student achievement.	• The Educational Challenge: This unit emphasizes the need for all students to be college and career ready. • Principal as Strategic Thinker: This unit gives principals the tools to be strategic thinkers and effective decision-makers. • Elements of Standards-Aligned Instructional Systems: This unit emphasizes an understanding of standards, assessments, and how to align instruction to standards.
Driving for Results	Principals are trained to examine student achievement and many other types of data to identify school, teacher and individual student needs.	• Driving for Change • Leading for Results • Culminating Simulation
Panel B: Non-PIL Courses		
Focus on Teaching and Learning	This course focuses on the principal as an instructional leader; participants learn to "integrate curriculum, instruction, and assessment within the instructional core."	• Foundations of Effective Learning: This unit is designed to help participants understand "the relationship between ideas about learning, the alignment of standards, curriculum, instruction, and assessment." • Leadership in the Instructional core: Part 1: This unit provides participants the tools to implement and support effective English Language Arts and History instruction. • Leadership in the Instructional Core: Part 2: This unit provides participants the tools to implement and support effective Mathematics and Science instruction. • Coaching for High Quality Teaching: This unit gives participants practice in coaching and developing human capital within a school.
Sustaining Transformation through Capacity and Commitment	This course focuses on principals as organizational leaders of schools.	• Promoting the Learning Organization: This unit teaches principals to view schools as learning organizations and apply teacher accountability to improve instruction.

- Teams for Instructional Leadership: This unit teaches principals the importance of distributing leadership throughout the school and how to create and foster leadership teams.
- Ethical Leadership for Equity: This unit teaches principals how to make moral and ethical decisions despite facing external and operational pressures.
- Driving and Sustaining Transformation: This unit gives principals the skills to maintain changes over time.

Notes. Source: Pennsylvania Department of Education (www.education.pa.gov). The Pennsylvania Inspired Leadership (PIL) program consists of World-Class Schooling and Driving for Results from the 2008-09 academic year through the 2015-16 year. In 2016-17, PDE replaced Driving for Results with Focusing on Teaching and Learning as a PIL course.

Table A4. Propensity Score Matching, All Outcomes

	PIL		No PIL		All		
	Matched	Unmatched	Matched	Unmatched	Matched	Unmatched	Bandwidth
Panel A: Teacher Effectiveness							
Math VAM	3,381	206	27,478	3,971	30,859	4,177	0.00
ELA VAM	3,883	271	32,887	4,906	36,770	5,177	0.00
Panel B: Teacher Turnover							
Teacher Turnover	42,916	2,686	476,850	85,082	519,766	87,768	0.00
Panel C: Student Achievement							
Math	334,761	22,843	3,677,871	875,616	4,012,632	898,459	0.00
ELA	333,837	22,496	3,665,714	895,959	3,999,551	918,455	0.00
Panel D: Student Achievement (VAM Sample)							
Math	273,464	18,191	2,489,409	636,500	2,762,873	654,691	0.00
ELA	264,697	17,964	2,356,628	584,108	2,621,325	602,072	0.00

Notes. Table reports the count of observations that were matched using propensity score matching (for a full set of matching variables, see Tables A5-A11). Treatment is defined as being a school-year observation in the year (and subsequent years) with a principal that completed PIL induction in that school.

Table A5. Propensity Score Matching (Math VAM): Covariate Balance

	Unmatched			Matched		
	Treated	**Untreated**	**Std. Diff**	**Treated**	**Untreated**	**Std. Diff**
Outcome Variable and Lag						
VAM	0.01	0.00	0.06	0.00	0.00	0.01
VAM (t-1)	0.01	0.01	0.04	0.01	0.01	0.00
School Characteristics						
Enrollment	618.48	694.07	-0.27	675.24	646.23	0.10
% Female	0.49	0.49	-0.01	0.49	0.49	-0.02
Age	10.41	11.28	-0.38	11.03	10.73	0.13
% Minority	0.24	0.21	0.09	0.21	0.21	0.00
% FRPL	0.48	0.42	0.23	0.43	0.44	-0.04
% IEP	0.16	0.16	-0.02	0.16	0.16	-0.04
% ELL Participation	0.03	0.03	0.11	0.03	0.03	-0.02
% Gifted	0.04	0.05	-0.24	0.05	0.04	0.06
Math Achievement	0.13	0.20	-0.08	0.21	0.20	0.01
Reading/ELA Achievement	0.11	0.25	-0.16	0.24	0.22	0.02
Math Achievement (t-1)	0.17	0.23	-0.06	0.24	0.25	-0.01
Reading/ELA Achievement (t-1)	0.09	0.23	-0.16	0.22	0.20	0.02
Charter	0.02	0.02	-0.06	0.02	0.02	0.03
City	0.21	0.16	0.13	0.15	0.16	-0.01
Suburb	0.49	0.53	-0.09	0.55	0.53	0.04
Town	0.08	0.11	-0.11	0.10	0.10	0.00
Rural	0.23	0.21	0.05	0.20	0.22	-0.06
Teacher Characteristics						
Age	41.30	41.01	0.03	41.08	41.34	-0.03
Female	0.70	0.69	0.03	0.69	0.70	-0.02
White	0.93	0.95	-0.08	0.95	0.95	0.00
Black	0.05	0.04	0.08	0.03	0.04	0.00
Hispanic	0.01	0.01	0.05	0.00	0.00	0.01
Asian	0.00	0.01	-0.02	0.00	0.00	0.00
Other	0.00	0.00	0.00	0.00	0.00	-0.01
Advanced Degree	0.53	0.56	-0.07	0.58	0.56	0.02
Years Experience	13.73	13.78	-0.01	13.89	13.99	-0.01
Years						
2011	0.02	0.09	-0.34	0.06	0.02	0.18
2012	0.05	0.09	-0.14	0.08	0.06	0.10
2013	0.06	0.09	-0.10	0.10	0.08	0.09

2014	0.10	0.10	0.00	0.11	0.11	0.00
2015	0.37	0.28	0.19	0.32	0.37	-0.11
2016	0.41	0.27	0.28	0.32	0.36	-0.09

Notes. Treated units are teachers in school-year observations with a principal who completed PIL induction in that school year (and all subsequent years) in the same school. Matching procedure matches teacher-year observations across treated and untreated school-year cells using the variables reported above. School level achievement is standardized within subject-year at the school level. *Std. Diff* is the standardized difference between treated and untreated units.

Table A6. Propensity Score Matching (ELA VAM): Covariate Balance

	Unmatched			Matched		
	Treated	**Untreated**	**Std. Diff**	**Treated**	**Untreated**	**Std. Diff**
Outcome Variable and Lag						
VAM	0.00	0.00	0.01	0.00	0.00	-0.01
VAM (t-1)	0.00	0.00	0.03	0.00	0.00	-0.01
School Characteristics						
Enrollment	0.49	0.49	-0.01	0.48	0.49	-0.03
% Female	10.53	11.43	-0.41	11.11	10.99	0.06
Age	0.22	0.20	0.09	0.22	0.20	0.07
% Minority	0.47	0.41	0.21	0.43	0.43	-0.01
% FRPL	0.16	0.16	-0.03	0.16	0.16	-0.06
% IEP	0.03	0.03	0.11	0.03	0.03	0.00
% ELL Participation	0.04	0.05	-0.24	0.05	0.04	0.04
% Gifted	0.49	0.49	-0.01	0.48	0.49	-0.03
Math Achievement	0.15	0.21	-0.06	0.19	0.20	-0.01
Reading/ELA Achievement	0.16	0.29	-0.15	0.24	0.25	-0.02
Math Achievement (t-1)	0.20	0.24	-0.04	0.23	0.26	-0.04
Reading/ELA (t-1)	0.14	0.27	-0.15	0.23	0.24	-0.02
Charter	0.01	0.02	-0.05	0.02	0.02	0.00
City	0.20	0.15	0.14	0.15	0.14	0.01
Suburb	0.49	0.53	-0.09	0.55	0.53	0.04
Town	0.08	0.11	-0.11	0.10	0.10	0.00
Rural	0.23	0.21	0.05	0.20	0.22	-0.06
Teacher Characteristics						
Age	41.98	41.95	0.00	41.79	41.89	-0.01
Female	0.82	0.83	-0.01	0.83	0.83	0.02
White	0.94	0.95	-0.08	0.95	0.96	-0.02
Black	0.05	0.03	0.06	0.04	0.03	0.03
Hispanic	0.01	0.01	0.05	0.00	0.00	0.00
Asian	0.00	0.00	-0.01	0.00	0.00	-0.01
Other	0.00	0.00	0.02	0.00	0.00	0.00
Advanced Degree	0.56	0.61	-0.11	0.62	0.61	0.02
Years Experience	14.18	14.02	0.02	14.02	14.10	-0.01
Years						
2011	0.02	0.10	-0.36	0.06	0.03	0.14
2012	0.06	0.09	-0.10	0.10	0.08	0.09
2013	0.07	0.09	-0.08	0.10	0.09	0.02
2014	0.09	0.10	-0.02	0.11	0.11	-0.01

2015	0.36	0.26	0.20	0.31	0.35	-0.09
2016	0.40	0.26	0.30	0.30	0.34	-0.08

Notes. Treated units are teachers in school-year observations with a principal who completed PIL induction in that school year (and all subsequent years) in the same school. Matching procedure matches teacher-year observations across treated and untreated school-year cells using the variables reported above. School level achievement is standardized within subject-year at the school level. *Std. Diff* is the standardized difference between treated and untreated units.

Table A7. Propensity Score Matching (Teacher Turnover): Covariate Balance

	Unmatched			Matched		
	Treated	**Untreated**	**Std. Diff**	**Treated**	**Untreated**	**Std. Diff**
Outcome Variable and Lags						
Teacher Mobility	0.11	0.11	0.01	0.12	0.11	0.03
Teacher Mobility (t-1)	0.05	0.05	0.03	0.06	0.05	0.05
School Characteristics						
Enrollment	717.52	822.63	-0.21	748.72	752.51	-0.01
% Female	0.49	0.49	-0.02	0.49	0.49	0.00
Age	11.15	11.85	-0.21	11.40	11.46	-0.02
% Minority	0.25	0.22	0.09	0.23	0.22	0.01
% FRPL	0.48	0.41	0.24	0.43	0.43	-0.01
% IEP	0.16	0.16	0.07	0.16	0.16	-0.04
% ELL Participation	0.03	0.03	0.09	0.03	0.03	-0.02
% Gifted	0.04	0.04	-0.18	0.04	0.04	0.00
Math Achievement	-0.06	-0.06	-0.01	-0.03	-0.02	-0.01
Reading/ELA Achievement	-0.08	0.05	-0.13	0.01	0.03	-0.01
Math Achievement (t-1)	-0.07	-0.05	-0.01	-0.02	-0.01	-0.01
Reading/ELA (t-1)	-0.06	-0.06	-0.01	-0.03	-0.02	-0.01
Charter	0.02	0.03	-0.02	0.02	0.03	-0.02
City	0.22	0.18	0.10	0.18	0.18	0.00
Suburb	0.46	0.49	-0.06	0.50	0.50	0.00
Town	0.08	0.11	-0.12	0.10	0.09	0.04
Rural	0.24	0.22	0.05	0.22	0.23	-0.03
Teacher Characteristics						
Age	42.08	41.90	0.02	42.49	41.94	0.05
Female	0.75	0.72	0.07	0.71	0.74	-0.06
White	0.93	0.95	-0.06	0.93	0.95	-0.06
Black	0.05	0.04	0.05	0.05	0.04	0.06
Hispanic	0.01	0.01	0.03	0.01	0.01	0.01
Asian	0.01	0.01	0.00	0.01	0.01	0.01
Other	0.00	0.00	0.04	0.00	0.00	0.00
Advanced Degree	0.52	0.53	-0.02	0.54	0.54	-0.01
Years Experience	14.01	13.90	0.01	14.12	13.97	0.02
Years						
2011	0.04	0.16	-0.39	0.15	0.07	0.26
2012	0.11	0.15	-0.11	0.17	0.15	0.06
2013	0.15	0.14	0.04	0.16	0.18	-0.08
2014	0.21	0.14	0.18	0.17	0.20	-0.08
2015	0.22	0.13	0.22	0.16	0.20	-0.11

2016	0.25	0.12	0.33	0.15	0.17	-0.05

Notes. Treated units are teachers in school-year observations with a principal who completed PIL induction in that school year (and all subsequent years) in the same school. Matching procedure matches teacher-year observations across treated and untreated school-year cells using the variables reported above. School level achievement is standardized within subject-year at the school level. *Std. Diff* is the standardized difference between treated and untreated units.

Table A8. Propensity Score Matching (Student Math Achievement): Covariate Balance

	Unmatched			Matched		
	Treated	**Untreated**	**Std. Diff**	**Treated**	**Untreated**	**Std. Diff**
Outcome Variable						
Std. Achievement	-0.05	0.01	-0.06	0.01	0.01	0.00
School Characteristics						
Enrollment	627.49	673.80	-0.17	654.64	652.09	0.01
% Female	0.49	0.49	-0.01	0.49	0.49	0.01
Age	10.14	10.49	-0.16	10.31	10.37	-0.03
% Minority	0.25	0.23	0.07	0.22	0.23	-0.03
% FRPL	0.48	0.42	0.21	0.42	0.44	-0.04
% IEP	0.16	0.16	0.02	0.16	0.16	-0.01
% ELL Participation	0.04	0.03	0.08	0.03	0.03	-0.03
% Gifted	0.04	0.04	-0.17	0.04	0.04	0.01
Math Achievement	0.12	0.22	-0.11	0.24	0.21	0.03
Reading/ELA Achievement	0.04	0.16	-0.12	0.16	0.15	0.00
Math Achievement (t-1)	0.15	0.23	-0.10	0.26	0.25	0.01
Reading/ELA (t-1)	0.02	0.15	-0.13	0.15	0.15	0.00
Charter	0.03	0.03	-0.04	0.03	0.03	-0.02
City	0.23	0.19	0.11	0.16	0.18	-0.04
Suburb	0.08	0.11	-0.10	0.09	0.09	-0.02
Town	0.22	0.20	0.04	0.20	0.21	-0.02
Rural	0.03	0.03	-0.04	0.03	0.03	-0.02
Student Characteristics						
Female	0.49	0.49	0.00	0.49	0.49	0.00
Age	10.92	11.12	-0.11	11.01	11.02	-0.01
4th	0.19	0.17	0.05	0.17	0.17	-0.01
5th	0.18	0.16	0.03	0.17	0.17	-0.01
6th	0.17	0.16	0.01	0.18	0.16	0.03
7th	0.15	0.17	-0.06	0.15	0.16	-0.01
8th	0.14	0.17	-0.08	0.15	0.16	-0.01
White	0.68	0.71	-0.06	0.72	0.71	0.00
Black	0.15	0.14	0.04	0.14	0.13	0.02
Hispanic	0.10	0.09	0.03	0.09	0.09	-0.03
Asian	0.04	0.04	0.00	0.04	0.04	0.00
Other	0.03	0.02	0.05	0.03	0.03	0.00
Poverty	0.46	0.42	0.10	0.42	0.43	-0.02
FRPL Elig	0.48	0.42	0.12	0.42	0.43	-0.02

ELL	0.03	0.03	0.02	0.03	0.03	-0.01
IEP	0.15	0.15	0.00	0.15	0.15	0.01
Gifted	0.05	0.05	-0.03	0.05	0.05	0.00
Years						
2010	0.01	0.14	-0.52	0.08	0.05	0.13
2011	0.05	0.14	-0.33	0.14	0.07	0.22
2012	0.12	0.13	-0.05	0.16	0.14	0.07
2013	0.16	0.13	0.07	0.16	0.18	-0.05
2014	0.21	0.13	0.22	0.16	0.20	-0.11
2015	0.21	0.12	0.25	0.15	0.19	-0.10
2016	0.25	0.12	0.34	0.15	0.17	-0.06

Notes. Treated units are students in school-year observations with a principal who completed PIL induction in that school year (and all subsequent years) in the same school. Matching procedure matches student-year observations across treated and untreated school-year cells using the variables reported above. *Std. Achievement* is standardized achievement with grade*subject*year level at the student level. School level achievement is standardized within subject-year at the school level. *Std. Diff* is the standardized difference between treated and untreated units.

Table A9. Propensity Score Matching (Student ELA Achievement): Covariate Balance

	Unmatched			Matched		
	Treated	**Untreated**	**Std. Diff**	**Treated**	**Untreated**	**Std. Diff**
Outcome Variable						
Std. Achievement	-0.05	0.01	-0.06	0.01	0.01	0.00
School Characteristics						
Enrollment	627.36	674.56	-0.17	654.80	652.05	0.01
% Female	0.49	0.49	-0.01	0.49	0.49	0.01
Age	10.14	10.50	-0.16	10.31	10.37	-0.02
% Minority	0.25	0.23	0.07	0.22	0.22	-0.02
% FRPL	0.48	0.42	0.21	0.42	0.44	-0.04
% IEP	0.16	0.16	0.02	0.16	0.16	-0.01
% ELL Participation	0.04	0.03	0.08	0.03	0.03	-0.03
% Gifted	0.04	0.04	-0.17	0.04	0.04	0.01
Math Achievement	0.13	0.22	-0.11	0.24	0.21	0.03
Reading/ELA Achievement	0.05	0.16	-0.12	0.16	0.16	0.00
Math Achievement (t-1)	0.15	0.23	-0.09	0.26	0.25	0.01
Reading/ELA (t-1)	0.03	0.15	-0.13	0.15	0.15	0.00
Charter	0.03	0.03	-0.04	0.03	0.03	-0.02
City	0.23	0.19	0.11	0.16	0.18	-0.04
Suburb	0.47	0.50	-0.07	0.54	0.52	0.05
Town	0.08	0.11	-0.10	0.09	0.09	-0.01
Rural	0.22	0.20	0.04	0.21	0.21	-0.01
Student Characteristics						
Female	0.49	0.49	0.00	0.49	0.49	0.00
Age	10.92	11.13	-0.12	11.01	11.02	-0.01
4th	0.18	0.16	0.05	0.17	0.17	-0.01
5th	0.18	0.16	0.03	0.17	0.17	-0.01
6th	0.17	0.17	0.00	0.18	0.16	0.03
7th	0.15	0.17	-0.07	0.15	0.16	-0.01
8th	0.14	0.17	-0.07	0.15	0.16	-0.01
White	0.69	0.71	-0.06	0.72	0.72	0.00
Black	0.15	0.14	0.04	0.14	0.13	0.02
Hispanic	0.10	0.09	0.03	0.08	0.09	-0.03
Asian	0.04	0.04	0.00	0.04	0.04	0.00

Other	0.03	0.02	0.05	0.03	0.03	0.00
Poverty	0.46	0.42	0.10	0.42	0.43	-0.02
FRPL	0.48	0.42	0.12	0.42	0.43	-0.02
ELL	0.03	0.03	0.02	0.02	0.03	-0.01
IEP	0.15	0.15	0.00	0.15	0.15	0.01
Gifted	0.05	0.05	-0.03	0.05	0.05	0.00
Years						
2010	0.01	0.14	-0.52	0.08	0.05	0.12
2011	0.05	0.14	-0.33	0.14	0.07	0.22
2012	0.12	0.13	-0.04	0.16	0.14	0.06
2013	0.16	0.13	0.08	0.16	0.18	-0.05
2014	0.21	0.13	0.22	0.16	0.20	-0.11
2015	0.21	0.12	0.25	0.15	0.19	-0.10
2016	0.25	0.12	0.34	0.15	0.17	-0.06

Notes. Treated units are students in school-year observations with a principal who completed PIL induction in that school year (and all subsequent years) in the same school. Matching procedure matches student-year observations across treated and untreated school-year cells using the variables reported above. *Std. Achievement* is standardized achievement with grade*subject*year level at the student level. School level achievement is standardized within subject-year at the school level. *Std. Diff* is the standardized difference between treated and untreated units.

Table A10. Propensity Score Matching (Student Math Achievement, Teacher VAM Sample): Covariate Balance

	Unmatched			Matched		
	Treated	**Untreated**	**Std. Diff**	**Treated**	**Untreated**	**Std. Diff**
Outcome Variable						
Std. Achievement	-0.06	-0.01	-0.05	-0.03	-0.01	-0.01
School Characteristics						
Enrollment	652.36	724.05	-0.25	697.17	679.88	0.06
% Female	0.49	0.49	-0.04	0.49	0.49	-0.02
Age	10.52	11.35	-0.39	11.04	10.87	0.08
% Minority	0.26	0.25	0.05	0.25	0.24	0.04
% FRPL	0.50	0.44	0.20	0.45	0.45	0.00
% IEP	0.16	0.16	-0.05	0.16	0.16	-0.03
% ELL Participation	0.04	0.03	0.09	0.03	0.03	0.00
% Gifted	0.04	0.05	-0.24	0.04	0.04	0.02
Math Achievement	0.07	0.12	-0.06	0.13	0.14	-0.02
Reading/ELA Achievement	0.04	0.15	-0.11	0.13	0.15	-0.02
Math Achievement (t-1)	0.10	0.13	-0.03	0.16	0.19	-0.04
Reading/ELA (t-1)	0.01	0.13	-0.13	0.12	0.13	-0.01
Charter	0.03	0.04	-0.06	0.03	0.03	0.00
City	0.26	0.21	0.10	0.19	0.19	0.00
Suburb	0.46	0.50	-0.07	0.53	0.51	0.03
Town	0.08	0.10	-0.08	0.09	0.09	0.00
Rural	0.21	0.19	0.05	0.19	0.20	-0.03
Student Characteristics						
Female	0.49	0.49	0.00	0.49	0.49	0.00
Age	11.18	11.72	-0.31	11.48	11.37	0.06
4^{th}	0.16	0.10	0.17	0.13	0.14	-0.04
5^{th}	0.16	0.12	0.12	0.13	0.15	-0.04
6^{th}	0.18	0.19	-0.04	0.19	0.18	0.03
7^{th}	0.18	0.24	-0.16	0.21	0.20	0.03
8^{th}	0.17	0.25	-0.18	0.21	0.20	0.03
White	0.67	0.69	-0.05	0.68	0.70	-0.04
Black	0.16	0.15	0.03	0.15	0.14	0.04
Hispanic	0.10	0.10	0.02	0.10	0.10	0.00
Asian	0.04	0.04	0.00	0.04	0.04	0.00
Other	0.03	0.02	0.05	0.03	0.03	0.01
Poverty	0.47	0.43	0.09	0.44	0.44	0.00
FRPL Elig	0.49	0.44	0.11	0.45	0.45	0.01

ELL	0.03	0.03	0.03	0.03	0.03	0.00
IEP	0.15	0.15	0.00	0.15	0.15	0.00
Gifted	0.05	0.05	-0.04	0.05	0.05	0.00
Years						
2010	0.00	0.11	-0.47	0.02	0.01	0.05
2011	0.02	0.11	-0.34	0.08	0.03	0.20
2012	0.08	0.10	-0.08	0.12	0.09	0.09
2013	0.09	0.10	-0.04	0.12	0.11	0.04
2014	0.25	0.18	0.16	0.23	0.26	-0.09
2015	0.26	0.17	0.21	0.21	0.25	-0.09
2016	0.30	0.17	0.31	0.22	0.24	-0.05

Notes. Treated units are students in school-year observations with a principal who completed PIL induction in that school year (and all subsequent years) in the same school. Matching procedure matches student-year observations across treated and untreated school-year cells using the variables reported above. *Std. Achievement* is standardized achievement with grade*subject*year level at the student level. School level achievement is standardized within subject-year at the school level. *Std. Diff* is the standardized difference between treated and untreated units.

Table A11. Propensity Score Matching (Student ELA Achievement, Teacher VAM Sample): Covariate Balance

	Unmatched			Matched		
	Treated	**Untreated**	**Std. Diff**	**Treated**	**Untreated**	**Std. Diff**
Outcome Variable						
Std. Achievement	-0.05	0.01	-0.06	0.00	0.01	-0.01
School Characteristics						
Enrollment	655.01	725.05	-0.24	700.13	680.99	0.07
% Female	0.49	0.49	-0.03	0.49	0.49	-0.02
Age	10.55	11.45	-0.42	11.09	10.93	0.07
% Minority	0.25	0.23	0.08	0.23	0.22	0.03
% FRPL	0.49	0.42	0.23	0.44	0.44	-0.01
% IEP	0.16	0.16	-0.07	0.16	0.16	-0.03
% ELL Participation	0.04	0.03	0.12	0.03	0.03	0.02
% Gifted	0.04	0.05	-0.26	0.04	0.04	0.03
Math Achievement	0.09	0.16	-0.07	0.16	0.17	-0.01
Reading/ELA Achievement	0.07	0.21	-0.14	0.18	0.19	-0.01
Math Achievement (t-1)	0.13	0.17	-0.05	0.19	0.21	-0.03
Reading/ELA (t-1)	0.04	0.19	-0.16	0.16	0.17	-0.01
Charter	0.03	0.03	-0.03	0.03	0.04	-0.02
City	0.25	0.19	0.14	0.17	0.18	-0.01
Suburb	0.47	0.52	-0.10	0.54	0.52	0.04
Town	0.08	0.10	-0.10	0.10	0.09	0.02
Rural	0.21	0.19	0.04	0.18	0.21	-0.05
Student Characteristics						
Female	0.49	0.49	0.00	0.49	0.49	0.00
Age	11.22	11.84	-0.37	11.60	11.45	0.09
4^{th}	0.16	0.09	0.20	0.12	0.13	-0.05
5^{th}	0.15	0.09	0.18	0.12	0.13	-0.05
6^{th}	0.18	0.20	-0.06	0.21	0.19	0.04
7^{th}	0.18	0.26	-0.20	0.23	0.21	0.04
8^{th}	0.18	0.26	-0.20	0.22	0.21	0.04
White	0.68	0.71	-0.08	0.70	0.72	-0.04
Black	0.15	0.13	0.05	0.14	0.13	0.03
Hispanic	0.10	0.09	0.03	0.10	0.09	0.01
Asian	0.04	0.04	0.01	0.04	0.04	0.01

Other	0.03	0.02	0.05	0.03	0.03	0.01
Poverty	0.47	0.41	0.10	0.43	0.43	0.00
FRPL	0.48	0.42	0.13	0.44	0.43	0.01
ELL	0.03	0.03	0.04	0.03	0.02	0.01
IEP	0.15	0.15	0.00	0.15	0.15	0.00
Gifted	0.05	0.06	-0.04	0.05	0.05	0.00
Years						
2010	0.00	0.10	-0.46	0.02	0.01	0.04
2011	0.02	0.10	-0.32	0.07	0.03	0.18
2012	0.07	0.09	-0.08	0.11	0.08	0.09
2013	0.08	0.09	-0.05	0.11	0.09	0.04
2014	0.26	0.19	0.15	0.23	0.27	-0.09
2015	0.26	0.18	0.20	0.22	0.26	-0.09
2016	0.30	0.18	0.30	0.23	0.25	-0.04

Notes. Treated units are students in school-year observations with a principal who completed PIL induction in that school year (and all subsequent years) in the same school. Matching procedure matches student-year observations across treated and untreated school-year cells using the variables reported above. *Std. Achievement* is standardized achievement with grade*subject*year level at the student level. School level achievement is standardized within subject-year at the school level. *Std. Diff* is the standardized difference between treated and untreated units.

Table A12. Effect of PIL Induction on Student Achievement (VAM Teacher Sample)

	Mathematics						ELA					
	All Schools		PIL Eligible		Matched		All Schools		PIL Eligible		Matched	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PIL	0.02* (0.01)		0.03** (0.02)		0.03*** (0.01)		0.01 (0.01)		0.01 (0.01)		0.02*** (0.01)	
4+ Years Before		-0.03 (0.06)		-0.07 (0.08)		0.02 (0.12)		-0.03 (0.06)		-0.00 (0.09)		0.06 (0.12)
3 Years Before		0.03 (0.03)		-0.01 (0.04)		0.04 (0.03)		-0.02 (0.02)		-0.02 (0.03)		-0.00 (0.02)
2 Years Before		-0.01 (0.02)		-0.01 (0.03)		-0.00 (0.02)		-0.00 (0.01)		-0.01 (0.02)		0.01 (0.02)
Year of PIL		0.02* (0.01)		0.03* (0.02)		0.03*** (0.01)		0.01 (0.01)		0.01 (0.01)		0.03*** (0.01)
1 Year After		0.01 (0.01)		0.02 (0.02)		0.02 (0.01)		-0.00 (0.01)		-0.00 (0.02)		0.01 (0.01)
2 Years After		0.02 (0.02)		0.04 (0.03)		0.03** (0.01)		-0.00 (0.01)		-0.00 (0.02)		0.01 (0.01)
3 Years after		0.03 (0.02)		0.05 (0.03)		0.04** (0.02)		0.01 (0.02)		0.01 (0.03)		0.02 (0.02)
4+ Years After		0.03 (0.02)		0.04 (0.03)		0.04* (0.02)		-0.01 (0.02)		-0.02 (0.03)		0.01 (0.02)
P-Values from F-Statistic:												
Before PIL		0.38		0.69		0.42		0.48		0.92		0.58
Year of/After PIL		0.58		0.48		0.53		0.11		0.14		0.25
Outcome Mean (SD)	-0.02 (1.01)		-0.08 (1.00)		-0.02 (1.00)		0.00 (1.00)		-0.05 (1.01)		0.00 (1.00)	

Student* Years	3,504,901	3,504,901	1,302,566	1,302,566	2,780,020	2,780,020	3,303,340	3,303,340	1,226,498	1,226,498	2,638,310	2,638,310
Students	1,587,635	1,587,635	711,998	711,998	1,316,406	1,316,406	1,563,598	1,563,598	695,914	695,914	1,300,560	1,300,560
School* Years	9,462	9,462	3,742	3,742	8,129	8,129	9,437	9,437	3,726	3,726	8,126	8,126
Schools	2,256	2,256	1,207	1,207	2,207	2,207	2,258	2,258	1,204	1,204	2,198	2,198
R^2	0.42	0.42	0.43	0.43	0.42	0.42	0.44	0.44	0.45	0.45	0.43	0.43

Notes. Each column within a panel represents a separate regression. Coefficients are reported with robust standard errors (clustered at the school level). All regressions include school and district-year fixed effects. Estimates based on samples that include school-grade-year observations with teacher VAM (i.e., *VAM Teacher Sample*). *PIL* is an indicator variable, defined at the school-year level, which equals one in the year (and all subsequent years) in which a principal has completed PIL induction in the same school (i.e., principal-school cells). *P-Value from F-Statistic* displays the p-values of F-tests of the joint significance of the pre-treatment effects (i.e., *Before PIL*) and of the post-treatment effects (i.e., *Year of/After PIL*). Coefficients statistically significant at *10% **5% and ***1% levels.

Table A13. Heterogeneous Effects of PIL on Student Achievement, by School Characteristics (VAM Teacher Sample)

	Math			ELA		
	Low	**Medium**	**High**	**Low**	**Medium**	**High**
Panel A: FRPL						
PIL	0.03 (0.02)	-0.02 (0.02)	0.03** (0.02)	-0.00 (0.01)	0.00 (0.01)	0.01 (0.01)
Tercile Mean (SD)	0.16 (0.07)	0.38 (0.06)	0.78 (0.18)	0.15 (0.07)	0.38 (0.06)	0.77 (0.18)
Student*Years	1,112,677	1,193,429	1,198,795	1,098,222	1,145,317	1,059,801
Students	568,159	628,643	587,877	559,881	618,960	570,501
School*Years	2,406	3,384	3,672	2,403	3,376	3,658
Schools	731	1,011	1,052	727	1,014	1,050
R^2	0.36	0.33	0.37	0.36	0.35	0.40
Panel B: Minority						
PIL	-0.01 (0.02)	0.04*** (0.01)	0.02 (0.02)	-0.04** (0.02)	0.02 (0.01)	-0.00 (0.01)
Tercile Mean (SD)	0.03 (0.01)	0.1 (0.04)	0.6 (0.28)	0.03 (0.01)	0.1 (0.04)	0.58 (0.28)
Student*Years	1,082,130	1,192,378	1,230,393	1,054,880	1,168,727	1,079,733
Students	545,914	605,435	571,266	541,032	596,247	554,111
School*Years	3,361	2,706	3,395	3,366	2,698	3,373
Schools	893	805	836	897	804	835
R^2	0.34	0.37	0.41	0.37	0.38	0.44
Panel C: Achievement						
PIL	0.01 (0.01)	0.02 (0.02)	0.04 (0.03)	-0.01 (0.01)	0.01 (0.03)	0.05*** (0.01)
Tercile Mean (SD)	0.40 (0.16)	0.74 (0.06)	0.88 (0.03)	0.40 (0.17)	0.74 (0.06)	0.88 (0.03)
Student*Years	1,525,528	1,119,442	859,931	1,433,212	1,045,582	824,546
Students	877,555	746,334	539,418	865,006	721,808	525,692
School*Years	4,607	3,090	1,765	4,599	3,075	1,763
Schools	1,829	1,469	823	1,827	1,471	822
R^2	0.43	0.36	0.33	0.46	0.38	0.34

Notes. Each column within a panel represents a separate regression. Coefficients are reported with robust standard errors (clustered at the school level). All regressions include school and district-year fixed effects. Estimates based on *All Schools* sample that includes school-grade-year observations with teacher VAM (i.e., *VAM Teacher Sample*). *FRPL* is the proportion of students in a school*year who are eligible to receive free-or-reduced price lunch. *Minority* is the proportion of students in a school*year who are either Black or Hispanic. *Achievement* is the proportion of

students in school*year who tested either proficient or advanced on the mathematics standardized exam. Terciles (*Low*, *Medium*, *High*) are based on school-by-year characteristics. *PIL* is an indicator variable, defined at the school-year level, which equals one in the year (and all subsequent years) in which a principal has completed PIL induction in the same school (i.e., principal-school cells). Coefficients statistically significant at *10% **5% and ***1% levels.

Appendix B. Measuring Teacher Effectiveness

We measure teacher effectiveness based on a teacher's value-added contribution to student achievement. Although teacher value-added estimates only capture certain aspects of teaching practice and behaviors (Steinberg & Kraft, 2017; Grossman et al., 2013), students assigned to higher value-added teachers have been shown to have higher college-going rates, earn higher salaries later in life, and be less likely to have children as teenagers (Chetty, Friedman, & Rockoff, 2014b). And, while much has been written on the consequences of different modeling choices with respect to estimating teacher value-added, we follow Kraft (2019) and estimate teacher effectiveness using a restricted maximum likelihood approach. We specify the model as:

$$(1)\ Achievement_{ijst} = \beta_1 Ach_{ijs(t-1)} + \beta_2 Ach^{other}_{ijs(t-1)} + \beta_3 \boldsymbol{X_{it}} + \beta_4 \boldsymbol{C_{jt}} + \beta_5 \boldsymbol{Z_{st}} + \Omega_{jt} + \mu_{ijst}$$

where *Achievement* for student *i* with teacher *j* at school *s* in year *t* is modeled as a function of a student's prior year test score in the same subject ($Ach_{ijs(t-1)}$) and prior year test score in the other subject ($Ach^{other}_{ijs(t-1)}$). For example, if we are estimating teacher value-added for teacher *j* in math in school year *t*, $Ach_{ijs(t-1)}$ will be student *i*'s math test score from the prior school year, and $Ach^{other}_{ijs(t-1)}$ will be student *i*'s ELA test score from the prior school year. **X** is a vector of time-varying student characteristics, including: age, race, gender, grade level, free or reduced price-lunch eligibility status, special education status, English language learner (ELL) status, and gifted status. **C** is a vector of time-varying classroom characteristics, which are the student characteristics aggregated to the classroom level; and **Z** is a vector of time-varying school characteristics, which are the student characteristics aggregated to the school-level.

The parameter estimate $\widehat{\Omega}_{jt}$ is the teacher*year random effect, capturing teacher *j's* estimated value-added contribution to student achievement (in either math or ELA) in school year *t*. Given that we model teacher effectiveness as a function of lagged student test scores, teacher effectiveness measures will just be for teachers who, in a given school year, teach in grades 4-8.

There is ongoing debate about the most appropriate (i.e., least biased and most efficient) approach for estimating teacher effectiveness using student test scores – i.e., teacher value-added measures, or VAMs. Critics cite the lack of random assignment, potential unobserved confounders, and the lack of clear modeling guidelines as reasons to avoid using VAMs to assign teacher effectiveness (Morganstein & Wasserstein, 2014). Proponents argue that, on average, VAM results accurately predict long-term student outcomes (Chetty et al., 2014; Chetty et al., 2014c) and have been cross-validated with experimental evidence (Kane & Staiger, 2008). Particularly relevant to this paper is the finding that different methods estimate similar teacher effects, conditional on controlling for students' prior achievement (Chetty et al., 2014). Prior research has relied on a teacher fixed effects approach to estimate teacher VAMs in which the estimated teacher fixed effects are adjusted using the Empirical Bayes post-estimation shrinkage estimator (see e.g., Atteberry, Loeb, and Wyckoff, 2015; Grissom and Loeb, 2017). In other research, the random effects estimates recovered from restricted maximum likelihood is preferred to a teacher fixed effect that has been shrunk via the Empirical Bayes method because it produces efficient and consistent estimators for the variance of true teacher effects (Kraft, 2019). While one concern related to the random effects estimator is that teacher assignment may be correlated with student characteristics, we control for prior student achievement and observable student and peer characteristics; these controls mitigate concern that the sorting of

students to teachers is based on prior academic performance and other observable student characteristics. Second, as sample size increases, the random effects and fixed effects estimators converge (Guarino et al., 2015; Kane & Staiger, 2008).