ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 20-190

Does Principal Professional Development Improve Principal Persistence, Teacher Effectiveness and Student Achievement? 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 within five years of employment. Using panel data on all Pennsylvania students, teachers and principals, we leverage within-principal variation in the timing of PIL completion within the same school setting to estimate the impact of PIL induction on principal persistence, teacher effectiveness and student achievement. PIL induction increased principal tenure by 18 percent, corresponding to approximately half an additional year as principal of record; however, PIL induction had no impact on teacher effectiveness or student achievement. We discuss the implications of our findings for principal induction efforts.

VERSION: January 2020

Suggested citation: Steinberg, Matthew P., and Haisheng Yang. (2020). Does Principal Professional Development Improve Principal Persistence, Teacher Effectiveness and Student Achievement? Evidence from Pennsylvania's Inspired Leadership Induction Program. (EdWorkingPaper: 20-190). Retrieved from Annenberg Institute at Brown University: https://www.edworkingpapers.com/ai20-190

Does Principal Professional Development Improve Principal Persistence, Teacher Effectiveness and Student Achievement?

Evidence from Pennsylvania's Inspired Leadership Induction Program

Matthew P. Steinberg
George Mason University

Haisheng Yang
University of Pennsylvania

December 27, 2019

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; his recent work uses econometric and quasi-experimental methods to study the effects of school closures, the consequences of student discipline, and educator human capital (msteinb6@gmu.edu). Haisheng Yang (yanghais@upenn.edu) is a doctoral candidate at the University of Pennsylvania's Graduate School of Education; his research interests include teacher human capital and the effect of school accountability on student, teacher and principal outcomes.

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 within five years of employment. Using panel data on all Pennsylvania students, teachers and principals, we leverage within-principal variation in the timing of PIL completion within the same school setting to estimate the impact of PIL induction on principal persistence, teacher effectiveness and student achievement. PIL induction increased principal tenure by 18 percent, corresponding to approximately half an additional year as principal of record; however, PIL induction had no impact on teacher effectiveness or student achievement. We discuss the implications of our findings for principal induction efforts.

Keywords: Principal induction; professional development; principal tenure; student achievement; teacher effectiveness

Introduction

Stable and 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). While student achievement improves when principals have greater decision-making autonomy in their schools (Clark, 2009; Steinberg, 2014; Steinberg & Cox, 2017), student achievement has been shown to decline in the years following principal mobility events (Bartanen et al., 2019; Miller, 2013). 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 (Creemers & Reezigt, 1996; Hill, 1998; Leithwood et al., 2004). Yet, despite the widely acknowledged importance of principal human capital (Creemers & Reezigt, 1996; Hill, 1998; Ladd, 2011; Loeb et al., 2012; Leithwood et al., 2004), little empirical evidence exists linking principal professional development to principal, 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 was designed to endow new principals with the skills necessary to be effective school leaders and to extend principal longevity in their first school of record.[1] In this paper, we rely on rich administrative data on all students, teachers and principals in Pennsylvania public schools to address the following questions: (1) What are the patterns of PIL completion among PIL eligible principals? (2) Does PIL induction increase principal persistence? (3) Does PIL induction improve teacher effectiveness and student achievement?

To estimate the effect of principal professional development in Pennsylvania on principal, teacher and student outcomes, we leverage within-principal variation in the timing of completion of PIL induction within the same school setting. By restricting identification to principal*school cells, we aim to mitigate concern that unobserved determinants of the timing of PIL completion may be correlated with school-specific heterogeneity. We find that PIL induction increased principal persistence in the same school by 18 percent, on average, corresponding to approximately half of an additional year as principal of record in the same school. However, PIL induction had no discernible effect on student achievement or teacher effectiveness.

Given the extent of school- and geographic-specific heterogeneity across schools in Pennsylvania, the absence of any discernible impact of PIL induction on teacher and student outcomes may be attributable to PIL's one-size-fits-all approach to principal professional development. Indeed, prior research demonstrates that principals adjust their leadership behaviors based on the characteristics of their school settings (Goldring et al., 2008; May et al., 2012); as such, we might expect greater returns to principal professional development if

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

policymakers tailored their professional development efforts to specific school contexts. Indeed, evidence on the heterogeneous impact of PIL induction suggest that the returns to PIL induction may differ across school contexts. We find suggestive evidence that PIL induction improves student math achievement more at schools serving more economically disadvantaged and minority students.

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 and 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). Indeed, 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 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] Toward this end, this paper aims to 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 principal persistence, student achievement and teacher effectiveness. 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 (and limitations) of targeted principal induction for improving student and teacher performance.

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

[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.

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).

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 engage parents and local communities by situating the school in a wider community environment and 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]

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 is correlated with increased principal persistence (Goldring & Taie, 2014; Jacob et al., 2015). Evidence from the 2012-13 School and Staffing Survey (SASS) finds that 77.6 percent of principals who participated in professional developed remained as principals at their schools. In contrast, 55.8 percent of

[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).

principals who did not participate in professional development remained at their schools (Goldring & Taie, 2014). 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.

To the best of our knowledge, ours is the first study to focus exclusively on the effects of a 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

[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.

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.

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

[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).

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 (via a student and/or principal fixed effects, which we detail below in the empirical approach section), but which might also introduce 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 & Cardichon, 2017).[8] The policy expansion of principal professional development and, specifically, induction programs, motivate the current study's efforts to understand the effects of PIL induction on principal persistence, teacher effectiveness, and student achievement.

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

[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).
[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)

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 an 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* (see Table A2 for a summary of Pennsylvania's core and corollary leadership standards). The first course, World-Class Schooling, lasts for six days, and is divided into three units: (i) The Educational Challenge; (ii) Principal as Strategic Thinker; and (iii) Elements of Standards-Aligned Instructional Systems. This course counts for 60 professional development hours. The second course, Driving for Results, lasts for six days and is divided into the following three units: (i) Driving for Change; (ii) Leading for Results; and (iii) Culminating Simulation. This course counts for 60 hours of the professional development credits (see Table A3 for a description of the NISL courses).[10]

Pennsylvania adopted 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

complete 180 hours of continuing professional education programs, activities or learning experiences through a PDE approved provider; or (iv) any combination of the above.

[10] The curriculum for World-Class Schooling is associated with the first two core leadership standards while the curriculum for Driving for Results is associated with the third core leadership standard (see Table A3 for leadership standards). Notably, the number of professional development hours are not the same as actual contact hours. For example, World Class Schooling counts for 60 professional development hours, but the course is scheduled for 36 hours with 6 hours of pre-work assignments.

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. In the World-Class Schooling course, principals: (i) gain the knowledge and skills to plan strategically; and (ii) develop 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 Driving for Results course imbues school leaders with the ability to access and use appropriate data to inform decision-making at all levels of the system. These two courses emphasize the skills that Pennsylvania and PDE policymakers believed early career principals needed to succeed. Under *Act* 45, the newly implemented leadership standards and the NISL courses required of principals to meet those standards represented the state's effort to inject greater rigor and accountability into in-service principal professional development.

Although PIL ineligible principals (i.e., those hired prior to January 1, 2008) can take any combination of PDE-approved courses to fulfill professional development requirements, the PIL induction program required that all eligible principals – those hired any time between January 1, 2008 through the 2015-16 school year – complete World-Class Schooling and Driving for Results. 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.

Consistent with *Act 45*, we define principals hired after January 1, 2008 as PIL-eligible principals. Of these principals, those that complete both Driving for Results and World Class Schooling 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;[11] 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).[12]

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 effectiveness and student achievement outcomes. Use of student, teacher, and principal microdata allows us to leverage variation at the individual, rather than school level, which improves not just the precision of our estimates, but, more importantly, allows us to control for unobservable characteristics of principals (that do not very across time) that may be correlated

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

[12] 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. 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.

with both the completion of PIL induction as well as student and teacher outcomes. Second, the current study spans a longer time period and includes the population of Pennsylvania schools, as opposed to a select sample of 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.

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); 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 tested grades. 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). PIL induction required PIL eligible principals to complete two NISL courses – World-Class Schooling and Driving for Results – within the first five years of employment as a school principal.

Sample

Among principals employed in Pennsylvania's traditional and charter public schools during the 2008-09 through 2015-16 period, we restrict the principal sample to those principals who were employed at a single school with available demographic and academic achievement data in each year.[13] We further limit the sample to schools with just one principal employed in a single year. Schools with multiple assistant principals were retained in the sample. Thus, our analytic sample includes 4,893 unique principals and 20,689 principal-year observations.[14]

[13] For example, if Principal A was employed at School A and B in 2008-09, we drop Principal A from the analytic sample. However, if Principal A was employed only at School A in 2009-10, we retain Principal A in the analytic sample.

[14] In 2008-09, we observe 2,622 principals; 2,724 in 2009-10; 2,694 in 2010-11; 2,636 in 2011-12; 2,530 in 2012-13; 2,549 in 2013-14; 2,474 in 2015; and 2,460 in 2015-16. Not all principals are at schools with tested students; 4,714 unique principals (19,838 principal-year observations) are at schools with tested students in both math and ELA. In our student-level regressions, we further restrict this sample to include only principals in schools that enroll students in grades 3-8 (3,994 unique principals; 16,487 principal*year observations).

Among our sample of 4,893 unique principals, 1,879 (38 percent) were PIL eligible, and 3,014 (62 percent) principals were ineligible to participate in the PIL induction program.[15] PIL ineligible principals included those who were first employed as a principal prior to the 2008-09 school year or who completed either PIL courses – World Class Schooling and Driving for Results – as an assistant principal (prior to becoming a school principal).[16]

Table 2 (Panel A) summarizes the demographic characteristics of principals by PIL eligibility status. Consistent with the fact that PIL induction targeted novice and early-career principals, PIL Eligible principals are less experienced and more racially/ethnically diverse than PIL ineligible principals. Compared to PIL ineligible principals, PIL principals are approximately five years younger, more likely to be a racial minority (i.e., black or Hispanic), and have approximately 5 fewer years of educational experience in Pennsylvania. PIL eligible principals are also less likely to hold an advanced degree (i.e., Master's or PhD) than their PIL ineligible counterparts.

<Table 2 about here>

Table 2 (Panel B) summarizes the characteristics of schools that principals lead, by PIL eligibility status. PIL eligible principals lead schools that are more racially/ethnically diverse; specifically, 27 percent of students in schools led by PIL eligible principals are black or

[15] In 2008-09, we observe 392 PIL Eligible and 2,230 PIL Ineligible principals; 625 PIL Eligible and 2,099 PIL Ineligible in 2009-10; 801 PIL Eligible and 1,893 PIL Ineligible in 2010-11; 919 PIL Eligible and 1,717 PIL Ineligible in 2011-12; 986 PIL Eligible and 1,544 PIL Ineligible in 2012-13; 1,078 PIL Eligible and 1,471 PIL Ineligible in 2013-14; 1,109 PIL Eligible and 1,365 PIL Ineligible in 2014-15; and 1,173 PIL Eligible and 1,287 PIL Ineligible in 2015-16. Our analytic sample includes 7,083 PIL Eligible principal-year observations and 13,606 PIL Ineligible principal-year observations.

[16] 657 principals took a PIL course as an assistant principal. Notably, PIL ineligible principals could take PIL coursework to fulfill their professional development requirements. Of the 3,014 PIL ineligible principals, 752 completed both World Class Schooling and Driving for Results (25 percent); 425 completed just World Class Schooling (14 percent); and 65 completed just Driving for Results (2 percent). We include any PIL ineligible principal who completed any PIL coursework in the comparison group; this is akin to allowing for treatment crossover.

Hispanic, on average, compared to 22 percent in schools led by PIL ineligible principals. Further, PIL eligible principals lead schools where 48 percent of students, on average, are free/reduced-price lunch eligible, compared to 42 percent of students, on average, in schools led by PIL ineligible principals. Academically, schools with PIL eligible principals serve lower-achieving students. On average, 60 percent of students attending a school led by a PIL eligible principal are academically proficient in math and 64 percent in ELA; this compares to math (ELA) proficiency of 68 percent (69 percent), on average, among students attending a school led by a PIL ineligible principal.

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 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). Given the imperfect compliance, a natural question is whether differences exist between PIL compliers and their non-complier counterparts. Indeed, PIL compliers are, on average, younger and less experienced than their non-compliant counterparts (see Table 2). Further, PIL compliant principals led schools with fewer minority students and higher achieving students (in mathematics) than schools with non-compliant principals (Table 2). Evidence on the non-random selection of PIL eligible principals into completing PIL induction suggests that cross-principal (and cross-school) estimates of PIL induction are likely biased (Nunnery et al., 2010; Nunnery et al., 2011B). Below, we detail our empirical approach which relies not on cross-principal (school) variation, but rather variation in PIL completion within principal*school cells.

Empirical Approach

To examine whether PIL participation is associated with an increase in principal persistence as school leader, we estimate variants of the following model:

$$(1)\ log(Persistence_{pst}) = \beta_1 PIL_{pt} + \boldsymbol{X}_{pt}\boldsymbol{\Gamma} + \boldsymbol{Z}_{st}\boldsymbol{\gamma} + \phi_{sp} + \delta_t + \mu_{pst}$$

where *Persistence* is the number of consecutive years principal *p* has been a principal at school *s* by the end of school year *t.*[17] The natural log of the dependent variable (*Persistence*) allows us to interpret the coefficient of interest (*PIL*) as a percent change. *PIL* is a binary indicator for whether principal *p* completed PIL induction by the end of school year *t* (i.e., *PIL*=1 in the year of completion and all subsequent years). **X** is vector of time-varying principal characteristics, including 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). The variable ϕ_{sp} is a principal*school fixed effect, allowing us to control for all time-invariant (and unobserved) principal-by-school-level characteristics and thereby restricting identification to principal*school cells; δ_t is a year fixed effect, which controls for year-specific differences common to all schools; and μ_{pst} represents a random error term. In alternative estimates of equation (1), we control for the average level of teacher effectiveness of a school's teachers (*School VAM*) to account for the possibility that principal persistence may be greater in schools

[17] There are principals in the sample who leave and return to the same school (approximately 1% of principal*year observations). For the purpose of this study, we restart the tenure counter the year a principal returns to a school (e.g., Principal A leaves School A in 2008, is not observed in 2009, and returns to School A in 2010; in the data, Principal A's tenure in 2010 is 1 year).

with more effective teachers, independent of the effect of principal professional development. We cluster the standard errors at the principal*school level (i.e., the level of treatment).

Next, we estimate the effect of PIL on teacher effectiveness (see Appendix B for details on constructing the measure of teacher effectiveness). To do so, we estimate variants of the following model:

$$(2)\ VAM_{jspt} = \beta_1 PIL_{pt} + \boldsymbol{X}_{pt}\boldsymbol{\Gamma} + \boldsymbol{Z}_{st}\boldsymbol{\gamma} + \overline{VAM}_{j-1,spt} + \phi_{sp} + \delta_t + \mu_{jpst}$$

where *VAM* is the measure of teacher effectiveness ($\widehat{\Omega}_{jt}$ from equation 1 in Appendix B), in either math or ELA, of teacher *j* in school *s* led by principal *p* in school year *t*. Given prior evidence on teacher peer effects (Jackson & Bruegmann, 2009; Sun et al., 2017), we include $\overline{VAM}_{j-1,spt}$, which we measure in two ways: (i) the average effectiveness (measured by VAM) of teacher *j's* peers (leaving out teacher *j*) at the school*subject*year level (which we refer to as *Subject Peer VAM*); and (ii) the average effectiveness (measured by VAM) of teacher *j's* peers (leaving out teacher *j*) at the school*subject*grade*year level (which we refer to as *Subject*Grade Peer VAM*). All other variables are defined as in equation (1), and we cluster the standard errors at the principal*school level.

We then estimate the effect of PIL on student academic achievement. To do so, we estimate variants of the following student-level model:

$$(3)\ Achievement_{ipst} = \beta_1 PIL_{pt} + \boldsymbol{X}_{pt}\boldsymbol{\lambda} + \boldsymbol{V}_{it}\boldsymbol{\Gamma} + \phi_{sp} + \delta_t + \eta_i + \mu_{ipst}$$

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*.[18] **V** is a vector of time-

[18] Students' scaled scores are standardized within year-subject-grade in our sample to account for test differences across years.

varying student characteristics, including: age, race, gender, grade level, free/reduced price-lunch eligibility status, special education status, English language learner (ELL) status, and gifted status. The variable η is a student fixed effect, which allows us to control for all time-invariant characteristics of students (e.g., motivation, ability) that may be correlated with both sorting to a school led by a PIL principal and student achievement. All other variables are defined as in equation (1), and we cluster the standard errors at the principal*school level. Estimates of the effect of PIL on student academic achievement are for students in grades 3-8 with available PSSA test data (in either math or ELA).

Across models of principal persistence, teacher effectiveness and student achievement, a key identifying assumption is that the timing of PIL completion is exogeneous, conditional on principal, school, student, teacher, and year effects. While we include a rich set of student, school, and principal controls, we are concerned that unobserved school or principal characteristics could be associated with both the timing of PIL completion and the outcomes of interest. For example, principals may complete PIL coursework in schools (and years) where both fewer demands may be placed on a principal's time and where student achievement may be better. To the extent that these demands are unobserved and time-varying, concern rests on the possibility that such unobserved, time-varying shocks to principals (and schools) might be correlated with the timing of PIL completion.

By including principal*school fixed effects, we restrict identification to changes in a principal's PIL status (i.e., completion) within the same school setting. Doing so allows us to control for unobserved, time-invariant characteristics—such as principal motivation or school culture—unique to school*principal combinations. By restricting identification to changes in a principal's PIL completion status within the same school setting, we control for unobserved

variation in principal characteristics across school contexts (e.g., principal effectiveness or principal motivation, for the same principal, might vary across school contexts). Nonetheless, while we cannot account for the presence of unobserved, time-varying shocks to principals (and schools) that might be correlated with the timing of PIL completion, the robustness of our estimates across a range of specifications (discussed below) suggest that the consequences of any such unobserved shocks are likely minimal.

Results

To what extent did the completion of PIL induction improve principal persistence? Table 3 summarizes these results. In Panel A, we present estimates based on the sample of principals that includes all schools (*Full Sample*), including those with non-tested students; Panel B restricts the analytic sample to those principals at schools with tested students in grades 3-8 (*Achievement Sample*). We find consistent returns to PIL induction for principal persistence. In our preferred model, which restricts comparison to principal*school cells, we find that PIL induction increased principal persistence by 18 percent among the full sample (Table 3, Panel A, Column 5) and by 19 percent increase among the achievement sample (Table 3, Panel B, Column 5). These effects correspond to approximately half of an additional year as the principal of record and are robust to the inclusion of controls for the average level of a school's teacher effectiveness (i.e., *School VAM*) (Table 3, Column 6).

<Table 3 about here>

While completion of PIL induction improves principal persistence, we find no evidence that PIL completion improved teacher effectiveness or student achievement. Table 4 summarizes estimates of the effect of PIL induction on teacher effectiveness. The effect of PIL induction on teacher effectiveness – both for math and ELA – are precisely estimated and indistinguishable

from zero across multiple models where we condition on, alternatively, school, principal, or principal*school effects. Estimates of the effect of PIL induction on student achievement mirror our findings for teacher effectiveness, indicating that PIL induction had no discernible effect on student math or ELA achievement (Table 5).

<Tables 4 and 5 about here>

Dynamic and Heterogeneous Effects of PIL Induction

The potential improvements to principal human capital, and the consequences for teacher effectiveness and student achievement, may not emerge until years after the completion of principal professional development. Alternatively, if the human capital benefits of principal professional development (via PIL induction) are short term and fade over time, estimates of the average effect of PIL induction may mask potentially important dynamic effects of PIL induction in the years after completion of principal professional development requirements. To examine the potentially dynamic effects of PIL induction, we re-estimates models (1)-(3) by disaggregating the *PIL* treatment indicator into year-specific effects.[19]

Table 6 summarizes estimates of the dynamic effects of PIL induction on principal persistence. In the years following the completion of PIL induction, the effect on principal persistence fades over time. Specifically, in the first two years after principals complete PIL induction, principal persistence increases by approximately 20 percent; yet by the fourth year after PIL induction, the effect on principal persistence is statistically indistinguishable from zero.

[19] Specifically, we re-estimate equations (1)-(3) by replacing PIL_{pt} with $\sum_{r=0}^{l_{max}} 1\left(t = t_{sp}^{*} + r\right)\beta_r$, where the variable t_{sp}^{*} is the school year in which principal *p* completed PIL induction. The coefficient β_r indexes the year of PIL completion in year t_{sp}^{*} on outcomes *r* years after the event, relative to the year of completion of PIL induction. We estimate $\beta_{r=0}$, $\beta_{r=1}$, $\beta_{r=2}$, $\beta_{r=3}$, and $\beta_{r=4+}$, where $\beta_{r=4+}$ includes all years that are four (or more) after completion of PIL induction (which we collapse into a single post-completion indicator). The reference category includes all years prior to the completion of PIL induction.

These results suggest that, over time, the additional professional development principals received through PIL induction had little effect on their likelihood of remaining in their original school of record, a result that reflects the fact that principals in Pennsylvania during this study period remain in their schools for, on average, 3 years (see Table 3). In contrast, we find no short-term or longer-term effects of principal professional development on either teacher effectiveness (Table 6) or student achievement (Table 7). These results indicate that PIL induction did little (to nothing) to improve principal human capital in ways that had any discernible effect on teacher and student productivity.

<Tables 6, 7 and 8 about here>

Schools in Pennsylvania vary in the characteristics of the students that they serve and the geographic settings in which the schools are located (see Table 2). As such, and in light of the one-size-fits-all design of PIL induction, the effect of PIL induction might vary across principals in schools that serve different student populations. This is in light of prior evidence that principals adjust their leadership behaviors based on the characteristics of their school settings (Goldring et al., 2008; May et al., 2012). We examine four dimensions of school settings for which the consequences of principal professional development might differentially impact principal, teacher and student outcomes: (i) poverty; (ii) racial/ethnic minority; (iii) achievement; and (iv) geographic location.

On average, PIL induction increases principal persistence more among principals in schools that serve the least economically disadvantaged students and the fewest students identified as racial/ethnic minorities (Table A4, Panels A and B). Among principals in the least economically disadvantaged schools – those serving, on average, 13 percent of students who are free/reduced-price lunch eligible – PIL induction increases principal persistence by 20 percent;

this compares to an increase of 15 percent among principals serving the most economically disadvantaged students (an average of 84 percent of students who are free/reduced-price lunch eligible). Further, the effect of PIL induction on principal persistence is smallest among principals in urban schools. However, given that principal mobility rates tend to be highest in the least advantaged, urban schools (Grissom & Bartanen, 2018), variation in the impact of principal professional development on principal persistence is not unexpected.

Notably, however, we find suggestive evidence that PIL induction improves student math achievement more at schools serving more economically disadvantaged and minority students (Table A6). Among schools serving the most economically disadvantaged schools, PIL induction increased student math achievement by (a marginally significant) 0.02 standard deviations; this compares to an effect statistically indistinguishable from zero among schools serving the least economically disadvantaged students. This evidence is consistent with the potentially greater returns to investments in the human capital of school leaders who serve the most disadvantaged students.

The effects of professional development might also depend on the timing of when novice and 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. Though we find that the effect of PIL induction on principal persistence is greater among principals who complete PIL induction earlier in their novice careers, we find no evidence that any human capital benefits of principal professional development differentially affect teacher or student outcomes (see Table A7).

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. Leveraging within-principal variation in the timing of PIL participation and completion, we estimate the effect of PIL induction on principal persistence, teacher effectiveness and student achievement. We find that while the completion of PIL induction improved principal longevity, PIL induction had no discernible effect on student achievement or teacher effectiveness.

Though we do not find any effect of PIL induction on student achievement or teacher effectiveness, improving principal longevity via targeted induction is important for several reasons. First, high rates of principal mobility (i.e., exit from schools) disrupt student learning (Miller, 2013). Second, the retention of effective school leaders reduces teacher mobility and increases teacher satisfaction (Grissom, 2011). Third, schools led by experienced principals have higher rates of student proficiency on standardized tests (Branch et al., 2012). Our findings indicate that completing PIL induction increased principal tenure by approximately 18 percent—increasing a principal's tenure in their initial school of record by approximately half an additional year, on average. These effects are consistent with evidence on the association between principal professional development and principal retention among a national sample of principals. In the 2012-13 school year, Goldring and Taie (2014) found that principals who participated in professional development were 39 percent more likely to remain in their school as a principal, compared to principals who did not participate in professional development.

The relatively short tenure of Pennsylvania principals, however, may explain (in part) the absence of an effect of PIL induction on student achievement and teacher effectiveness. The average tenure of a Pennsylvania principal in his/her school is about 3 years, and the completion of PIL induction is found to increase average tenure by approximately half a year. Recent evidence demonstrates that not only does principal turnover negatively affect student achievement, but that student achievement returns to the same level following a principal transition only after the fifth year of the new principal's tenure in the school (Miller, 2013). Given that, on average, a Pennsylvania principal remains in his/her school for fewer than five years, even after completing PIL induction, principal turnover may limit any improvements to student achievement that could otherwise result from PIL induction.

To what extent might the programmatic design of PIL induction in Pennsylvania limit its efficacy for improving teacher and student performance? PIL induction required that novice Pennsylvania principals complete two NISL-designed courses – World Class Schooling and Driving for Results. However, the extent to which the skills taught in these courses were (mis)aligned with the principal leadership practices that have been shown to improve student and teacher outcomes might further explain the limited efficacy of PIL induction. For example, the Driving for Results NISL course emphasizes the use of data-driven decision making within a school. Yet, this course does not support the development of instructional (Grissom et al., 2013; Robinson et al. 2008) or managerial (Grissom & Loeb, 2011) leadership practices that have a demonstrated relationship with student achievement. Indeed, a recent review of the research literature identifies three types of leadership practices that significantly influence student outcomes: (i) goal setting and planning; (ii) promoting and participating in teacher learning; and (iii) planning, coordinating, and evaluating teaching and the curriculum (Robinson et al., 2008, p.

656). Further, research shows that developing a principal's ability to improve a school's learning climate, develop rigorous curriculum, and enact performance accountability has the potential to improve student outcomes (Sebastian & Allensworth, 2012; Reardon, 2011).[20]

The findings from our study of Pennsylvania's PIL induction differ in important ways from prior evidence on principal induction. For example, Gates et al. (2014) find that student achievement improves in schools with principals who participated in in-service induction. Yet, the induction program that Gates et al. (2014) study included a mentorship component for novice principals that was absent from the PIL induction program. And, while most states with a formal principal induction requirement incorporate a principal mentorship component (Goldrick, 2016), PIL induction does not. Instead, PIL required the completion of just two NISL courses—World Class Schooling and Driving for Results. This distinguishes PIL induction both from other states' induction programs and other studies which have considered the effectiveness of NISL EDP, from which PIL borrows course-specific elements.

Notably, our study offers a methodological advancement over prior studies of NISL-specific coursework and principal induction. Prior studies 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.

[20] Beginning in the 2016-17 school year (the year after the end of this study period), PDE amended PIL induction course requirements by replacing the Driving for Results course with a different NISL course, Sustaining Transformation through Capacity and Commitment, which emphasizes distributed and instructional leadership and focuses on developing a principal's capacity as the organizational leader of schools. Given that instructional and distributive leadership practices are associated with improvements in student achievement (Grissom et al, 2013; Heck & Hallinger, 2009), this recent change in PIL induction coursework requirements may lead to future improvements in student outcomes as a result of the closer alignment of PIL induction coursework with the type of instructional and distributive leadership practices that have demonstrated effects on student outcomes.

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 controls for unit-level heterogeneity – via a rich set of fixed effects – that may be correlated with both the timing of a principal's completion of PIL induction and the principal, student and teacher outcomes we study. The inclusion of principal-specific fixed effects in a more rigorous econometric framework allows us to avoid the type of bias due to unobserved, individual-level heterogeneity likely present in existing studies of principal induction. Indeed, this fixed effects approach allows us to account for selection bias due to factors such as principal motivation that, in the absence of such controls, would likely overstate the efficacy of PIL induction. 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.

Do our null findings call into question the efficacy of principal induction programs to improve teacher effectiveness and student achievement? Though the lack of a mentorship component in the PIL induction program limits the generalizability of our findings, our results reveal important limitations in the design and implementation of PIL induction. Indeed, most principal induction programs nationally include a mentorship component in which novice principals are paired with more experienced principals. While all PIL-eligible principals in Pennsylvania were required to complete the same two NISL courses, individualized mentorship for novice principals may allow principals to develop leadership skills specific to their school contexts. The ability of principals to adapt leadership skills specific to their school contexts has

been shown to be a necessary skill of effective principals (Leithwood, Jantzi, & Steinbach, 1999). In contrast, PIL induction relied solely upon NISL coursework, which not only lacks a mentorship component but also applies the same leadership standards across school contexts. Furthermore, most induction programs focus professional development resources on the first two years of a principal's tenure, while PIL induction can be completed any time within a principal's first five years. As such, we might expect intensive, early career coaching to have a more pronounced effect on both the development of principal human capital and a range of schooling outcomes than the completion of two courses spread over the first five years of a principal's tenure. Nonetheless, we believe the results from this study of Pennsylvania's approach to principal induction should be instructive for state and local policymakers throughout the nation when designing and implementing professional development for novice and early career principals.

Our null findings also 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 reveals heterogeneous returns to principal persistence 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 achievement in Pennsylvania's most disadvantaged schools. Together, these findings suggest the importance of considering school context and the local human capital needs of principals in very different school settings. Indeed, policymakers should focus on supporting principals who lead the lowest performing, most disadvantaged schools, where suggestive evidence from this study indicates the greatest potential to improve student achievement outcomes.

Limitations & Further Research

While we focus on principal persistence, teacher effectiveness and student achievement as outcomes in this study, we recognize the complex role of the principalship and the inherent difficulty in measuring principal improvement across the various domains of principals' leadership responsibilities. Future research could examine dimensions of principal practice not explicitly measured by student achievement outcomes – such as principal instructional leadership – via observations of principal leadership practices, interviews with district administrators and surveys of teachers (e.g., the UChicago Consortium on School Research annual teacher survey). These additional measures of principal practices would complement the student achievement outcomes captured in this study and would enable an even richer understanding of how in-service induction might shape additional dimensions of principal performance.

Though this study offers a rigorous empirical investigation of a principal induction program in Pennsylvania, a deeper understanding of the implementation and effects of principal induction would likely benefit from additional quantitative and qualitative (i.e., mixed methods) 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 no discernible impact on teacher effectiveness or student achievement. Further, the qualitative evidence generated from principal interviews, coupled with principal surveys about principals' experiences in PIL induction, could reveal important insights from school leaders

about the value they place on the NISL coursework, information that is unavailable from individual-level microdata of the type used in this study.

Notably, we do not observe important aspects of PIL implementation and the extent of principals' experiences in PIL coursework. For example, we do not observe the quality of PIL instructors who deliver the course content, the extent of principal participation and engagement in the PIL coursework, nor the quality of course instruction or specific course materials (e.g, the content of the Action Learning Projects and their relevance to principals' experiences as school leaders). 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 – such as distributive leadership or instructional leadership – 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 student, teacher, and principal outcomes.

References

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

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.

Creemers, B. P., & Reezigt, G. J. (1996). School level conditions affecting the effectiveness of instruction. *School Effectiveness and School Improvement*, *7*(3), 197-228.

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.

Goldring, R., & Taie, S. (2014). Principal Attrition and Mobility: Results from the 2012-13 Principal Follow-Up Survey. First Look. NCES 2014-064. *National Center for Education Statistics*.

Grissom, J. A. (2011). Can good principals keep teachers in disadvantaged schools? Linking principal effectiveness to teacher satisfaction and turnover in hard-to-staff environments. *Teachers College Record*, *113*(11), 2552-2585.

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

Grissom, J. A., & Loeb, S. (2011). Triangulating principal effectiveness how perspectives of parents, teachers, and assistant principals identify the central importance of managerial skills. *American Educational Research Journal*, *48*(5), 1091-1123.

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.*

Hill, P. W. (1998). Shaking the foundations: Research driven school reform. *School Effectiveness and School Improvement, 9*(4), 419-436.

Jackson, C. K., & Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics*, *1*(4), 85-108.

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., Jantzi, D., & Steinbach, R. (1999). *Changing leadership for changing times*. McGraw-Hill Education (UK).

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

Reardon, R. M. (2011). Elementary school principals' learning-centered leadership and educational outcomes: Implications for principals' professional development. *Leadership and Policy in Schools*, *10*, 63–83.

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*.

Sebastian, J., & Allensworth, E. (2012). The influence of principal leadership on classroom instruction and student learning: A study of mediated pathways to learning. *Educational Administration Quarterly*, 48(4), 626–663.

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.

Sun, M., Loeb, S., & Grissom, J. A. (2017). Building teacher teams: Evidence of positive spillovers from more effective colleagues. *Educational Evaluation and Policy Analysis*, *39*(1), 104-125.

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 Eligibility Status and Compliance

			PIL Eligible		
	All Principals	**PIL Ineligible**	**All**	**Non-Compliers**	**Compliers**
Panel A: Principal Characteristics					
Age	46.66 (8.62)	48.24 (8.42)	43.61*** (8.16)	44.53 (8.65)	42.35*** (7.25)
Female	0.45	0.44	0.47***	0.43	0.52***
White	0.88	0.90	0.85***	0.85	0.85
Black	0.10	0.09	0.12***	0.12	0.11
Hispanic	0.01	0.01	0.02***	0.02	0.02
Other Race	0.01	0.00	0.01***	0.01	0.02***
Experience	19.06 (9.17)	20.99 (9.04)	15.33*** (8.23)	15.96 (8.69)	14.48*** (7.48)
Bachelor's Degree	0.13	0.11	0.18***	0.17	0.18
Advanced Degree	0.86	0.89	0.82***	0.82	0.82
Panel B: School Characteristics					
Enrollment	610.25 (368.3)	620.74 (368.09)	590.11*** (367.89)	616.73 (379.73)	553.57*** (347.76)
Female	0.48 (0.04)	0.48 (0.04)	0.49** (0.05)	0.49 (0.05)	0.49 (0.04)
Age	10.83 (3.32)	10.81 (3.31)	10.87 (3.34)	11.19 (3.39)	10.43*** (3.21)
White	0.71 (0.31)	0.73 (0.30)	0.67*** (0.34)	0.67 (0.34)	0.68* (0.33)
Minority	0.23 (0.30)	0.22 (0.29)	0.27*** (0.33)	0.28 (0.33)	0.26** (0.32)
FRPL	0.44 (0.27)	0.42 (0.27)	0.48*** (0.28)	0.47 (0.28)	0.49** (0.28)
IEP	0.16 (0.06)	0.16 (0.06)	0.16*** (0.07)	0.17 (0.07)	0.16*** (0.05)
ELL	0.03 (0.05)	0.03 (0.05)	0.03*** (0.06)	0.03 (0.06)	0.03*** (0.06)
Gifted	0.04 (0.05)	0.04 (0.05)	0.03*** (0.04)	0.03 (0.05)	0.03*** (0.03)
Math Proficiency	0.65 (0.24)	0.68 (0.22)	0.60*** (0.25)	0.59 (0.26)	0.61*** (0.25)
ELA Proficiency	0.67 (0.19)	0.69 (0.18)	0.64*** (0.20)	0.64 (0.20)	0.64 (0.19)

Charter	0.04	0.02	0.08***	0.10	0.04***
Urban	0.19	0.18	0.23***	0.23	0.23
Suburban	0.45	0.45	0.43***	0.45	0.40***
Rural	0.25	0.24	0.25	0.23	0.28***
Town	0.11	0.12	0.09***	0.09	0.10
Principals	4,893	3,014	1,879	1,279	600
Schools	3,187	2,695	1,727	1,197	725
Principal*Years	20,689	13,606	7,083	4,098	2,985

Notes. In Panel A, proportions are reported, except for age and experience, which report mean (standard deviation). In Panel B, school-level mean (standard deviation) reported, except for charter status and urbanicity, which report proportions. The sample includes principals in K-12 traditional and charter public schools present in any school year during the 2008-09 through 2015-16 school years. *PIL Ineligible* includes principals hired prior to the 2008-09 school year or who completed either Driving for Results or World Class Schooling as assistant principals; *PIL Eligible* includes principals hired during or after the 2008-09 school year and who did not complete World Class Schooling or Driving for Results as assistant principals. *Non-Compliers* includes principals who were PIL Eligible but did not complete both Driving for Results and World Class Schooling prior to the 2015-16 school year. *Compliers* includes principals who were PIL Eligible and completed both Driving for Results and World Class Schooling prior to the 2015-16 school year. The sample of principals at schools with standardized tests is 4,720 for mathematics (2,926 PIL Ineligible; 1,794 PIL Eligible) and 4,714 (2,923 PIL Ineligible; 1,791 PIL Eligible) for ELA. Differences between *PIL Eligible* and *PIL Ineligible* and *Compliers* and *Non-Compliers* principals statistically significant at the *10%, **5% and ***1% levels.

Table 3. The Effect of PIL Induction on Principal Persistence

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: ***Full Sample***						
PIL Induction	0.16*** (0.02)	0.23*** (0.02)	0.22*** (0.02)	0.18*** (0.01)	0.18*** (0.01)	0.18*** (0.02)
Mean (SD) Tenure (years)			2.94 (1.89)			
Principals	4,714	4,714	4,714	4,714	4,714	3,133
Schools	3,044	3,044	3,044	3,044	3,044	2,281
Principal*Year Observations	19,838	19,838	19,838	19,838	19,838	9,595
Panel B: ***Achievement Sample***						
PIL Induction	0.16*** (0.02)	0.23*** (0.02)	0.23*** (0.02)	0.19*** (0.01)	0.19*** (0.01)	0.18*** (0.02)
Mean (SD) Tenure (years)			2.95 (1.90)			
Principals	3,994	3,994	3,994	3,994	3,994	3,133
Schools	2,559	2,559	2,559	2,559	2,559	2,281
Principal*Year Observations	16,487	16,487	16,487	16,487	16,487	9,595
Year FE	X	X	X	X	X	X
School FE		X		X		
Principal FE			X	X		
Principal*School FE					X	X
School VAM						X

Notes. Each column (within a panel) represents a separate regression. Coefficients are reported with robust standard errors (clustered at the principal*school level). *Principal Persistence* is the natural log of the number of consecutive years a principal led a given school. In Panel A, *Full Sample* includes all principals. In Panel B, *Achievement Sample* includes principals at schools who enroll students in grades 3 through 8 with available testing data. All models include controls for principal characteristics, including: age, gender, race, highest degree earned and years of experience in education; and school-level characteristics, including: school size (enrollment), percent of students free/reduced-price lunch eligible, percent of students receiving special education services, percent of students receiving English language learner (ELL) services, percent of gifted students, percent of minority students (Black or Hispanic), and student achievement (percent of students proficient on math and ELA exams). In column (6), we include *School VAM*, which is the average ELA and mathematic VAM among a school's teachers. Coefficients statistically significant at the *10%, **5% and ***1% levels.

Table 4. The Effect of PIL Induction on Teacher Effectiveness

	Teacher Math VAM							Teacher ELA VAM						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**
PIL Induction	0.01* (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Year FE	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Teacher FE	X	X	X	X	X	X	X	X	X	X	X	X	X	X
School FE		X		X					X		X			
Principal FE			X	X						X	X			
Principal*School FE					X	X	X					X	X	X
Subject Peer VAM						X							X	
Subject*Grade Peer VAM							X							X
Principals	3,120	3,120	3,120	3,120	3,120	3,023	2,850	3,119	3,119	3,119	3,119	3,119	3,040	2,199
Schools	2,278	2,278	2,278	2,278	2,278	2,239	2,161	2,275	2,275	2,275	2,275	2,275	2,248	2,092
Principal*Year Observations	3,444	3,444	3,444	3,444	3,444	3,311	3,094	3,443	3,443	3,443	3,443	3,443	3,339	3,169
Teacher*Year Observations	60,171	60,171	60,171	60,171	60,171	59,677	54,056	71,474	71,474	71,474	71,474	71,474	71,070	66,619

Notes. Each column (within a panel) represents a separate regression. Coefficients are reported with robust standard errors (clustered at the principal*school level). All models include controls for principal characteristics, including: age, gender, race, highest degree earned and years of experience in education; and school-level characteristics, including: school size (enrollment), percent of students free/reduced-price lunch eligible, percent of students receiving special education services, percent of students receiving English language learner (ELL) services, percent of gifted students, percent of minority students (Black or Hispanic), and student achievement at the school level (percent of students proficient/advanced on math and ELA exams). In columns (6) and (13), *Subject Peer VAM* is a control for the average teacher peer VAM in a given subject at the school level (either mathematics or ELA); in columns (7) and (14), *Subject*Grade Peer VAM* is a control for the average teacher peer VAM at the subject*grade level (either mathematics or ELA). In academic years prior to the 2013-14 academic year, teacher VAM is only available for grades 6-8. Average math VAM is 0.002 with a standard deviation of 0.09; average ELA VAM is 0.001 with a standard deviation of 0.061. Coefficients statistically significant at the *10%, **5% and ***1% levels.

Table 5. The Effect of PIL Induction on Student Achievement

	Math Achievement						ELA Achievement					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
PIL Induction	0.01** (0.00)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	-0.00 (0.00)	0.00 (0.00)	0.00 (0.01)	-0.00 (0.01)	-0.00 (0.01)	0.00 (0.01)
Year FE	X	X	X	X	X	X	X	X	X	X	X	X
Student FE	X	X	X	X	X	X	X	X	X	X	X	X
School FE		X		X				X		X		
Principal FE			X	X					X	X		
Principal* School FE					X	X					X	X
School VAM						X						X
Principals	3,994	3,994	3,994	3,994	3,994	3,120	3,994	3,994	3,994	3,994	3,994	3,119
Schools	2,559	2,559	2,559	2,559	2,559	2,278	2,559	2,559	2,559	2,559	2,559	2,275
Principal*Year Observations	4,772	4,772	4,772	4,772	4,772	3,444	4,772	4,772	4,772	4,772	4,772	3,443
Students	1,673,099	1,673,099	1,673,099	1,673,099	1,673,099	1,591,294	1,661,463	1,661,463	1,661,463	1,661,463	1,661,463	1,579,392
Student*Year Observations	5,064,933	5,064,933	5,064,933	5,064,933	5,064,933	3,516,873	5,073,555	5,073,555	5,073,555	5,073,555	5,073,555	3,552,190

Notes. Each column represents a separate regression. Coefficients are reported with robust standard errors (clustered at the principal*school level). Student achievement is standardized within school year, subject, and grade. All models include controls for principal characteristics, including: age, gender, race, highest degree earned and years of experience in education; student-level characteristics, including: age, gender, race/ethnicity, grade level, free/reduced-price lunch status, special education status, English language learner (ELL) status and gifted status; and school-level characteristics, including: school size (enrollment), percent of students free/reduced-price lunch eligible, percent of students receiving special education services, percent of students receiving English language learner (ELL) services, percent of gifted students, percent of minority students (Black or Hispanic), and student achievement at the school level (percent of students proficient/advanced on math and ELA exams). In columns (6) and (12), we include *School VAM*, which is the average ELA or mathematics VAM among a school's teachers (e.g., if the outcome of interest is mathematics achievement, we control for mathematics VAM, but not ELA VAM). Coefficients statistically significant at the *10%, **5% and ***1% levels.

Table 6. Dynamic Effects of PIL Induction on Principal Persistence

	(1)	(2)	(3)	(4)	(5)	(6)
Year of PIL Induction	0.01 (0.02)	0.14*** (0.02)	0.23*** (0.02)	0.18*** (0.01)	0.18*** (0.01)	0.17*** (0.01)
1 Year After PIL Induction	0.15*** (0.03)	0.25*** (0.02)	0.23*** (0.03)	0.21*** (0.02)	0.20*** (0.01)	0.22*** (0.02)
2 Years After PIL Induction	0.24*** (0.03)	0.31*** (0.03)	0.23*** (0.03)	0.19*** (0.02)	0.19*** (0.02)	0.20*** (0.03)
3 Years After PIL Induction	0.33*** (0.04)	0.35*** (0.04)	0.18*** (0.04)	0.12*** (0.03)	0.13*** (0.02)	0.12*** (0.03)
4+ Years After PIL Induction	0.32*** (0.06)	0.34*** (0.05)	0.07 (0.06)	0.06 (0.04)	0.06* (0.03)	0.07 (0.04)
P-Value from F-Statistic:						
Year of PIL=1 Year After = … = 4 Years After	0.00	0.00	0.04	0.00	0.00	0.00
Year FE	X	X	X	X	X	X
School FE		X		X		
Principal FE			X	X		
Principal*School FE					X	X
School VAM						X
Principals	4,714	4,714	4,714	4,714	4,714	3,133
Schools	3,044	3,044	3,044	3,044	3,044	2,281
Principal*Year Observations	19,838	19,838	19,838	19,838	19,838	9,595

Notes. Each column represents a separate regression. Coefficients are reported with robust standard errors (clustered at the principal*school level). Reference category includes years prior to a principal's completion of PIL induction. In all columns, we include controls for principal characteristics, including: age, gender, race, highest degree earned and years of experience in education; and school-level characteristics, including: school size (enrollment), percent of students free/reduced-price lunch eligible, percent of students receiving special education services, percent of students receiving English language learner (ELL) services, percent of gifted students, percent of minority students (black and Hispanic), and school-level proficiency rates in mathematics and ELA. In columns (6), we include *School VAM*, which is the average ELA and mathematic VAM among a school's teachers. Coefficients statistically significant at the *10%, **5% and ***1% levels.

Table 7. Dynamic Effects of PIL Induction on Teacher Effectiveness

	Teacher Math VAM							Teacher ELA VAM						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**
Year of PIL Induction	0.01*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
1 Year After PIL Induction	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
2 Years After PIL Induction	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
3 Years After PIL Induction	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
4+ Years After PIL Induction	-0.00	0.00	-0.00	-0.00	-0.00	0.00	0.00	0.00	0.00	-0.00	0.00	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
P-Value from F-Statistic:														
Year of PIL=1 Year After = … = 4 Years After	0.85	0.98	0.94	0.98	0.99	0.98	0.94	1.00	0.98	0.98	0.99	0.99	0.87	0.93
Year FE	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Teacher FE	X	X	X	X	X	X	X	X	X	X	X	X	X	X
School FE		X		X					X		X			
Principal FE			X	X						X	X			
Principal*School FE					X	X	X					X	X	X
Subject Peer VAM						X							X	
Subject*Grade Peer VAM							X							X
Principals	3,120	3,120	3,120	3,120	3,120	3,023	2,850	3,119	3,119	3,119	3,119	3,119	3,040	2,199
Schools	2,278	2,278	2,278	2,278	2,278	2,239	2,161	2,275	2,275	2,275	2,275	2,275	2,248	2,092
Teacher*Year Observations	60,171	60,171	60,171	60,171	60,171	59,677	54,056	71,474	71,474	71,474	71,474	71,474	71,070	66,619

Notes. Each column represents a separate regression. Coefficients are reported with robust standard errors (clustered at the principal*school level). Reference category includes years prior to a principal's completion of PIL induction. In all columns, we include controls for principal characteristics, including: age, gender, race, highest degree earned and years of experience in education; and school-level characteristics, including: school size (enrollment), percent of students free/reduced-price lunch eligible, percent of students receiving special education services, percent of students receiving English language learner (ELL) services, percent of gifted students, percent of minority students (black and Hispanic), and school-level proficiency rates in mathematics and ELA. In columns (6) and (13), *Subject Peer VAM* is a control for the average teacher peer VAM in a given subject at the school level (either mathematics or ELA); in columns (7) and (14), *Subject*Grade Peer VAM* is a control for the average teacher peer VAM at the subject*grade level (either mathematics or ELA). In academic years prior to the 2013-14 academic year, teacher VAM is only available for grades 6-8. Average math VAM is 0.002 with a standard deviation of 0.09; average ELA VAM is 0.001 with a standard deviation of 0.061. Coefficients statistically significant at the *10%, **5% and ***1% levels.

Table 8. Dynamic Effects of PIL Induction on Student Achievement

	Math Achievement						ELA Achievement					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Year of PIL Induction	0.01*	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	-0.00	-0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
1 Year After PIL Induction	0.01**	0.01*	0.01	0.01	0.01	0.02	0.00	0.01	0.00	0.00	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
2 Years After PIL Induction	0.00	0.00	-0.00	-0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
3 Years After PIL Induction	0.02**	0.02	0.02	0.02	0.02	0.03	0.00	0.01	-0.00	0.00	0.00	0.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
4+ Years After PIL Induction	0.01	0.01	0.02	0.01	0.01	0.03	-0.02	0.00	-0.00	0.00	0.00	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)
P-Value from F-Statistic:												
Year of PIL=1 Year After = … = 4 Years After	0.14	0.14	0.25	0.23	0.20	0.23	0.73	0.91	0.99	0.99	0.99	0.80
Year FE	X	X	X	X	X	X	X	X	X	X	X	X
Student FE	X	X	X	X	X	X	X	X	X	X	X	X
School FE		X		X				X		X		
Principal FE			X	X					X	X		
Principal*School FE					X	X					X	X
School VAM						X						X
Schools	2,559	2,559	2,559	2,559	2,559	2,278	2,559	2,559	2,559	2,559	2,559	2,275
Students	1,673,099	1,673,099	1,673,099	1,673,099	1,673,099	1,591,294	1,661,463	1,661,463	1,661,463	1,661,463	1,661,463	1,579,392
Student*Years	5,064,933	5,064,933	5,064,933	5,064,933	5,064,933	3,516,873	5,073,555	5,073,555	5,073,555	5,073,555	5,073,555	3,552,190

Notes. Each column represents a separate regression. Coefficients are reported with robust standard errors (clustered at the principal*school level). Reference category includes years prior to a principal's completion of PIL induction. In all columns, we include controls for principal characteristics, including: age, gender, race, highest degree earned and years of experience in education; and school-level characteristics, including: school size (enrollment), percent of students

free/reduced-price lunch eligible, percent of students receiving special education services, percent of students receiving English language learner (ELL) services, percent of gifted students, percent of minority students (black and Hispanic), and school level proficiency rates in mathematics and ELA; student characteristics, including: age, gender, race/ethnicity, grade level, free/reduced-price lunch status, special education status, English language learner (ELL) status and gifted status. In columns (6) and (12), we include *School VAM*, which is the average ELA or mathematics VAM among a school's teachers (e.g., if the outcome of interest is mathematics achievement, we control for average mathematics VAM, but not ELA VAM). Coefficients statistically significant at the *10%, **5% and ***1% levels.

Appendix A. Supplemental Tables

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 website (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 website (www.education.pa.gov). The Pennsylvania Inspired Leadership (PIL) induction program focuses on the 3 core leadership standards.

Table A3. PIL Induction Program Courses

Course Title (credit hours)	Course Description	Course Units
World Class Schooling: Vision and Goals (60)	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.
Focus on Teaching and Learning (90)	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.
Driving for Results (60)		• Driving for Change • Leading for Results • Culminating Simulation
Sustaining Transformation through Capacity and Commitment (60)	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 website (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. Heterogeneous Effects of PIL Induction on Principal Persistence

	Q1	Q2	Q3	Q4
Panel A: ***% FRPL***				
PIL Induction	0.20*** (0.03)	0.16*** (0.02)	0.18*** (0.02)	0.15*** (0.02)
Principals	1,445	1,714	1,740	1,597
Schools	1,076	1,286	1,261	1,051
Principal*School Obs.	4,987	4,933	4,903	5,015
Panel B: ***% Minority***				
PIL Induction	0.20*** (0.02)	0.21*** (0.02)	0.16*** (0.02)	0.14*** (0.02)
Principals	1,464	1,608	1,424	1,454
Schools	1,034	1,137	965	892
Principal*School Obs.	4,971	4,983	4,928	4,956
Panel C: ***% Proficient/Advanced in Mathematics***				
PIL Induction	0.13*** (0.02)	0.14*** (0.02)	0.12*** (0.02)	0.14*** (0.02)
Principals	2,341	2,406	2,236	1,665
Schools	1,900	2,061	1,839	1,311
Principal*School Obs.	4,954	4,961	4,962	4,961
Panel D: ***School Urbanicity***				
	City	Suburb	Town	Rural
PIL Induction	0.13*** (0.03)	0.18*** (0.02)	0.16*** (0.04)	0.20*** (0.02)
Principals	1,132	2,181	653	1,355
Schools	673	1,413	443	914

Principal*Year Observations	3,936	8,800	2,243	4,859

Notes. Each cell represents a separate regression. Coefficients are reported with robust standard errors (clustered at the principal*school level). *Principal Persistence* is the natural log of the number of consecutive years a principal led a given school. All models include controls for principal characteristics, including: age, gender, race, highest degree earned and years of experience in education; and school-level characteristics, including: school size (enrollment), percent of students free/reduced-price lunch eligible, percent of students receiving special education services, percent of students receiving English language learner (ELL) services, percent of gifted students, percent of minority students (Black or Hispanic), and student achievement (percent of students proficient on math and ELA exams). All models include principal*school and year fixed effects. Quartiles are constructed using school level characteristics, pooled across years. Mean (SD) for FRPL quartiles are: Q1) 0.13 (0.06); Q2) 0.31 (0.04); Q3) 0.48 (0.05); Q4) 0.84 (0.15). Mean (SD) for school minority quartiles are: Q1) 0.02 (0.01); Q2) 0.06 (0.01); Q3) 0.16 (0.06); Q4) 0.71 (0.23). Mean (SD) for math proficiency quartiles are: Q1) 0.31 (0.14); Q2) 0.62 (0.06); Q3) 0.78 (0.04); Q4) 0.90 (0.04). Coefficients statistically significant at the *10%, **5% and ***1% levels.

Table A5. Heterogenous Effects of PIL Induction on Teacher Effectiveness

	Teacher Math VAM				Teacher ELA VAM			
	Q1	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**
Panel A: ***% FRPL***								
PIL	0.00 (0.01)	-0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
Principals	700	863	972	932	708	883	996	958
Schools	571	701	781	720	578	710	799	748
Teacher*Year	14,407	13,881	12,916	12,852	18,234	17,388	16,257	14,737
Panel B: ***% Minority***								
PIL	0.00 (0.01)	-0.01 (0.01)	-0.00 (0.01)	0.01* (0.01)	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)
Principals	773	855	829	829	796	877	838	850
Schools	610	681	656	606	625	696	665	626
Teacher*Year	10,507	14,799	15,850	12,900	13,754	18,668	19,630	14,567
Panel C: ***% Proficient/Advanced in Mathematics***								
PIL	0.00 (0.01)	-0.01 (0.01)	-0.00 (0.01)	0.01 (0.01)	0.00 (0.00)	-0.01* (0.01)	0.01 (0.00)	0.01* (0.01)
Principals	1,457	1,389	1,198	733	1,483	1,451	1,220	752
Schools	1,285	1,270	1,072	659	1,308	1,317	1,084	678
Teacher*Year	16,165	14,413	13,272	10,206	19,182	17,248	16,837	13,352
Panel D: ***School Urbanicity***								
	City	Suburb	Town	Rural	City	Suburb	Town	Rural
PIL	0.01 (0.01)	-0.00 (0.00)	-0.01 (0.01)	0.01 (0.01)	0.00 (0.00)	-0.00 (0.00)	-0.00 (0.01)	0.01* (0.00)
Principals	646	1,303	309	746	659	1,323	320	766
Schools	463	1,011	235	551	480	1,022	240	561

Principal*Year Observations	9,423	28,408	5,289	10,936	10,706	35,129	6,912	13,872

Notes. Each cell represents a separate regression. Coefficients are reported with robust standard errors (clustered at the principal*school level). All models include controls for principal characteristics, including: age, gender, race, highest degree earned and years of experience in education; school-level characteristics, including: school size (enrollment), percent of students free/reduced-price lunch eligible, percent of students receiving special education services, percent of students receiving English language learner (ELL) services, percent of gifted students, percent of minority students (Black or Hispanic), student achievement (percent of students proficient on math and ELA exams), and average peer VAM at the subject*school level (*Subject Peer VAM*) (either mathematics or ELA); and principal*year, teacher, and year fixed effects. Quartiles are constructed using school level characteristics, pooled across years. Mean (SD) for FRPL quartiles are: Q1) 0.13 (0.06); Q2) 0.31 (0.04); Q3) 0.48 (0.05); Q4) 0.84 (0.15). Mean (SD) for school minority quartiles are: Q1) 0.02 (0.01); Q2) 0.06 (0.01); Q3) 0.16 (0.06); Q4) 0.71 (0.23). Mean (SD) for math proficiency quartiles are: Q1) 0.31 (0.14); Q2) 0.62 (0.06); Q3) 0.78 (0.04); Q4) 0.90 (0.04). In academic years prior to the 2013-14 academic year, teacher VAM is only available for grades 6-8. Average math VAM is 0.002 with a standard deviation of 0.09; average ELA VAM is 0.001 with a standard deviation of 0.061. Coefficients statistically significant at the *10%, **5% and ***1% levels.

Table A6. Heterogenous Effects of PIL Induction on Student Achievement

	Math Achievement				ELA Achievement			
	Q1	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**
Panel A: ***% FRPL***								
PIL	-0.01 (0.02)	0.00 (0.02)	-0.01 (0.02)	0.02* (0.01)	-0.01 (0.02)	0.01 (0.01)	-0.01 (0.01)	0.01 (0.01)
Schools	860	1,026	1,073	910	860	1,026	1,073	910
Students	543,805	526,775	526,614	503,291	542,434	561,061	525,750	499,116
Student*Year	1,396,947	1,246,197	1,180,549	1,241,240	1,405,416	1,251,514	1,182,188	1,234,437
Panel B: ***% Minority***								
PIL	-0.02 (0.02)	-0.00 (0.02)	0.03** (0.01)	0.02 (0.01)	-0.01 (0.01)	0.00 (0.02)	-0.00 (0.01)	-0.00 (0.01)
Schools	874	950	821	745	874	950	821	745
Students	428,864	576,378	565,378	474,237	428,047	575,255	562,283	469,620
Student*Year	1,056,141	1,363,514	1,404,975	1,240,303	1,062,277	1,368,945	1,407,225	1,235,108
Panel C: ***% Proficient/Advanced in Mathematics***								
PIL	0.02 (0.02)	0.00 (0.02)	0.00 (0.02)	0.00 (0.02)	0.01 (0.01)	-0.01 (0.02)	-0.00 (0.01)	0.01 (0.01)
Schools	1,461	1,678	1,651	1,276	1,461	1,678	1,651	1,276
Students	637,527	693,840	774,770	655,602	634,546	692,686	775,457	653,284
Student*Year	1,068,191	1,089,665	1,402,520	1,504,557	1,063,086	1,090,924	1,411,909	1,507,636
Panel D: ***School Urbanicity***								
	City	Suburb	Town	Rural	City	Suburb	Town	Rural
PIL	0.01 (0.02)	0.02* (0.01)	-0.01 (0.03)	-0.02 (0.01)	0.01 (0.01)	-0.00 (0.01)	-0.03** (0.01)	0.00 (0.01)
Schools	560	1,190	354	777	560	1,190	354	777
Students	365,335	895,490	215,006	435,108	360,891	890,954	214,475	434,349

Student*Year	977,823	2,515,198	522,791	1,049,121	972,141	2,522,448	525,259	1,053,707

Notes. Each cell represents a separate regression. Coefficients are reported with robust standard errors (clustered at the principal*school level). All models include controls for principal characteristics, including: age, gender, race, highest degree earned and years of experience in education, student-level characteristics, including: age, gender, race/ethnicity, grade level, free/reduced-price lunch status, special education status, English language learner (ELL) status and gifted status; school-level characteristics, including: school size (enrollment), percent of students free/reduced-price lunch eligible, percent of students receiving special education services, percent of students receiving English language learner (ELL) services, percent of gifted students, percent of minority students (Black or Hispanic), and percent of students testing proficient or advanced in mathematics and ELA; and principal*school, student, and year fixed effects. Quartiles are constructed using school level characteristics, pooled across years. Mean (SD) for FRPL quartiles are: Q1) 0.13 (0.06); Q2) 0.31 (0.04); Q3) 0.48 (0.05); Q4) 0.84 (0.15). Mean (SD) for school minority quartiles are: Q1) 0.02 (0.01); Q2) 0.06 (0.01); Q3) 0.16 (0.06); Q4) 0.71 (0.23). Mean (SD) for math proficiency quartiles are: Q1) 0.31 (0.14); Q2) 0.62 (0.06); Q3) 0.78 (0.04); Q4) 0.90 (0.04). Coefficients statistically significant at the *10%, **5% and ***1% levels.

Table A7. Heterogeneous Effects of Timing of PIL Completion, by Year of PIL Induction Completion

	Year of PIL Induction Completion				
	2	**3**	**4**	**5**	**6**
Panel A. ***Principal Persistence***					
Persistence	0.46*** (0.02)	0.24*** (0.02)	0.04 (0.03)	-0.10*** (0.03)	-0.01 (0.06)
Principals	136	178	107	66	28
Schools	166	217	137	86	41
Principal*Years	573	858	610	424	195
Panel B. ***Teacher Effectiveness***					
Math VAM	0.00 (0.01)	0.01 (0.01)	0.00 (0.01)	0.01 (0.01)	0.00 (0.02)
ELA VAM	0.00 (0.00)	-0.00 (0.00)	0.00 (0.01)	-0.01* (0.01)	-0.01 (0.01)
Principals	115	145	86	50	23
Schools	126	159	98	59	23
Teacher*Years	3,197	4,822	3,340	1,713	637
Panel C. ***Student Achievement***					
Math	0.02 (0.02)	0.02 (0.02)	-0.03 (0.02)	0.02 (0.02)	-0.02 (0.04)
ELA	-0.01 (0.02)	0.01 (0.01)	-0.01 (0.02)	0.01 (0.02)	-0.06* (0.03)
Students	76,697	121,934	81,025	46,785	19,245
Schools	145	190	121	73	31
Student*Years	137,432	225,412	149,198	87,169	35,679

Notes. Each cell represents a separate regression. Coefficients are reported with robust standard errors (clustered at the principal*school level). Only *PIL Eligible* principals who completed PIL induction included in models. *PIL Eligible* includes schools and principals led by principals hired during or after the 2008-09 school year and did not complete World Class Schooling or Driving for Results as assistant principals. *Principal Persistence* is the natural log of the number of consecutive years a principal led a given school. All models include controls for principal characteristics, including: age, gender, race, highest degree earned and years of experience in education; school-level characteristics, including: school size (enrollment), percent of students free/reduced-price lunch eligible, percent of students

receiving special education services, percent of students receiving English language learner (ELL) services, percent of gifted students, percent of minority students (Black or Hispanic), and percent of students testing proficient or advanced in mathematics and ELA; and principal*school, unit-level (teachers for Panel B, students for Panel C), and year fixed effects. Student regressions include student-level characteristics, including: age, gender, race/ethnicity, grade level, free/reduced-price lunch status, special education status, English language learner (ELL) status and gifted status. Teacher-level regressions include average peer VAM at the school*subject level (*Subject Peer VAM*). School characteristics are based on a principal's first school of record. Coefficients 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).