EdWorkingPaper No. 19-136

Identifying Principal Improvement

Brendan Bartanen
Texas A&M University

Using statewide data from Tennessee over more than a decade, this paper estimates the job performance returns to principal experience as measured by student, teacher, and principal outcomes. I find that principals improve substantially over time, evidenced by higher student achievement, higher ratings from supervisors, and lower rates of teacher turnover. However, improvement in student achievement as principals gain experience does not carry over when principals change schools. The returns to school-specific experience are largest for principals in high-poverty schools, highlighting the potential benefits of policies to improve the recruitment and retention of high-quality leaders in hard-to-staff environments.

VERSION: September 2019

Suggested citation: Bartanen, Brendan. (2019). Identifying Principal Improvement. (EdWorkingPaper: 19-136). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/kcrz-tm19

Identifying Principal Improvement

Brendan Bartanen

Texas A&M University

September 2019

Abstract

Using statewide data from Tennessee over more than a decade, this paper estimates the job performance returns to principal experience as measured by student, teacher, and principal outcomes. I find that principals improve substantially over time, evidenced by higher student achievement, higher ratings from supervisors, and lower rates of teacher turnover. However, improvement in student achievement as principals gain experience does not carry over when principals change schools. The returns to school-specific experience are largest for principals in high-poverty schools, highlighting the potential benefits of policies to improve the recruitment and retention of high-quality leaders in hard-to-staff environments.

Identifying Principal Improvement

Effective leadership is an important ingredient in school performance. High-quality principals are linked to a variety of school and student outcomes, including higher student achievement (Branch, Hanushek, & Rivkin, 2012; Chiang, Lipscomb, & Gill, 2016; Coelli & Green, 2012; Dhuey & Smith, 2018; Grissom, Kalogrides, & Loeb, 2015), lower teacher turnover (Boyd et al., 2011; Grissom, 2011; Grissom & Bartanen, 2019b; Ladd, 2011), and better school climate (Burkhauser, 2017; Sebastian & Allensworth, 2012). At the same time, there exists substantial variation in principal quality, and disadvantaged schools tend to be led by principals with less experience (Branch et al., 2012; Grissom, Bartanen, & Mitani, 2019; Loeb, Kalogrides, & Horng, 2010) and lower effectiveness ratings (Grissom & Bartanen, 2019a).

While it is clear that principals matter and that some principals are more effective than others, we have less knowledge of what drives the variation in principal quality. In particular, little research has considered on-the-job improvement. Despite the conventional wisdom that more experienced principals tend to be more effective, we have little evidence that explicitly identifies the job performance returns to principal experience. This lack of evidence stands in contrast to the robust literature on teacher improvement and worker productivity, more broadly. Numerous studies, for example, document substantial returns to experience for teachers, particularly in the first few years of teaching (e.g., Harris & Sass, 2011; Kane, Rockoff, & Staiger, 2008; Ladd & Sorensen, 2017; Papay & Kraft, 2015). While most of these studies focus on student achievement, some have also shown positive effects of teacher experience on non-test outcomes, such as student attendance (Ladd & Sorensen, 2017).

Understanding the returns to principal experience is an important issue for both policymakers and researchers, particularly given increased investments in school leadership at the district, state, and federal level. Prior studies have found a correlation between principal experience and performance (e.g., Bastian & Henry, 2015; Clark, Martorell, & Rockoff, 2009; Grissom, Blissett, & Mitani, 2018), but it is not clear whether this is driven by within-principal improvement or the selective attrition of less effective principals. Distinguishing between improvement and selection matters, as they lead to different policy prescriptions. If the returns to principal experience are small, for instance, resources may be better spent on identifying high-quality candidates for school leadership and removing ineffective leaders, rather than focusing on the development of the existing stock of leaders. However, if the returns to principal experience are large relative to the overall distribution of principal quality, policies focused on retaining leaders and providing them with opportunities for development would be more effective at raising the average quality of school leadership.

Using longitudinal administrative data from Tennessee, this study fills an important gap in the literature by estimating the job performance returns to principal experience. Specifically, my primary research question examines the extent to which principals improve at raising student achievement as they gain experience. I supplement this analysis with other outcomes, including rubric-based ratings of principals' practice from their supervisors and teacher turnover rates in the principal's school. Examining these outcomes is important because they potentially capture different dimensions of principal performance, but they may also help to explain the mechanisms that drive the relationship between principal experience and student achievement.

My second research question examines whether the returns to principal experience are portable–i.e., whether improvement in a principal's ability to raise student test scores carries over when the principal changes schools. Specifically, I estimate the extent to which the returns to principal experience are driven by *total* experience (i.e., number of years as principal) versus *school-specific* experience (i.e., number of years as principal in the same school). Understanding the difference between the returns to total versus school-specific principal experience is important for research and policy. Conceptually, this distinction can provide insight into the nature of principal improvement. What are the skills that principals build that lead to greater student performance? For example, principals may become more effective over time at using data to drive school-level policies (e.g., identifying students who need additional support), which could lead to higher test scores, even when the principal changes schools. Alternatively, principals may become more effective over time through fostering relationships with teachers, families, or the broader community, which ultimately improve student learning. Upon moving to a new school, these relationships have to be rebuilt, such that the accumulated "improvement" in the prior school does not help to increase student achievement in the new school.

From a policy perspective, understanding whether principal improvement reflects returns to total versus school-specific experience informs debates regarding the allocation of principals and the importance of promoting stability in school leadership. If improvement is largely school-specific, for instance, a policy that moves experienced principals into struggling schools may have unanticipated costs to both sending and receiving schools.

My final research question examines heterogeneity in the returns to principal experience by school context. Prior work establishes that principals in schools serving larger percentages of low-income and low-achieving students have less experience and lower ratings from supervisors, on average (Grissom et al., 2019). One potential driver of this pattern is that principals in these schools improve at lower rates, which could lead them to receive lower ratings and turn over more frequently. Further, understanding the average improvement trajectory of principals across different types of schools helps to quantify the extent to which differences in principal turnover rates may contribute to disparities in access to high-quality principals.

Estimating the causal effect of principal experience presents an empirical challenge. To overcome this, I estimate models that include both principal and school fixed effects. Because principal quality varies and likely influences whether principals remain in the principalship (Grissom & Bartanen, 2019a), it is critical to isolate within-principal variation in experience to ameliorate selection bias. Similarly, principals are not randomly assigned to schools and prior work shows they tend to sort to more advantaged schools over time (Béteille, Kalogrides, & Loeb, 2012; Loeb et al., 2010). The inclusion of school fixed effects helps ensure that the estimated returns to experience are not conflated with differences in school quality across a principal's career.

One existing paper has used longitudinal administrative data to examine the relationship between principal experience and student test scores. Clark et al. (2009) use a school fixed effects model, but their models do not contain principal fixed effects, meaning that they do not explicitly estimate the extent to which principals improve over time. Rather, their experience estimates capture both the returns to experience and "ability" bias generated by non-random attrition from the principalship. They find that, within a given school, performance is higher when led by a principal with more experience.

To my knowledge, this is the first study to estimate the returns to principal experience using both principal and school fixed effects. I find substantial returns to experience for student achievement. The returns are largest for math achievement—relative to a principal's first year in the principalship, the average student in the principal's school scores 0.065 SD higher on statewide exams when the principal has 5 years of experience. Student test scores continue to improve up to 14 years—the highest value of experience I can observe in my data. Additionally, principals' ratings from supervisors increase by more than 0.45 standard deviations in their first five years on the job, on average, which moves the typical principal from the 35th to the 53rd percentile in the statewide distribution of scores.

Leveraging principals who work in multiple schools across the study period, I find that the returns to principal experience for student achievement are driven by school-specific rather than total experience. Put differently, improvement as measured by higher student test scores does not transfer across schools—principals effectively "start over" at their new school. The non-portability of improvement highlights the cost of frequent principal turnover, as it may take several years for leaders—even those with prior principal experience—to drive improvements in student learning.

I proceed first by reviewing the existing literature that relates principal experience to school performance, which helps to frame the contribution of this study. I also provide a framework for examining the portability of principals' accumulated skills. Next, I describe the data, measures, and methods used to estimate the returns to principal experience. I then describe the results for the returns to experience, the portability of improvement, and heterogeneity in improvement by school context. The final section concludes with the implications of the study for policy and suggestions for future research.

Connecting Principal Experience and School Performance

An increasing body of work demonstrates that principals have substantial effects on student learning. For instance, recent estimates of principal effects on math achievement range from 0.05 to 0.20 student-level standard deviations (Branch et al., 2012; Dhuey & Smith, 2018; Grissom et al., 2015). In other words, a principal who is 1 standard deviation above the mean in terms of quality increases student growth by 0.05 to 0.20 standard deviations. While these estimates tend to be smaller in magnitude than teacher effects, principals affect the learning of *every* student in the school, which further underscores the importance of high-quality leadership for student success.

One limitation of existing work is that principal quality typically is treated as fixed, and principals are either "effective" or "ineffective." However, given the complex nature of school leadership, it is likely that effectiveness comes from a broad set of skills that principals develop over time. For instance, principals have responsibilities across many domains, including administrative tasks (e.g., managing student discipline, fulfilling compliance requirements), instruction management (e.g., conducting classroom observations), and internal relations (e.g., developing relationships with staff members).[1] Further, many of the skills important for effective leadership are not necessarily those developed from experience as a classroom teacher. While most principals complete certification programs and serve as assistant principals, these training experiences likely do not cover the full range of a principal's responsibilities. There is strong reason to expect that novice principals have yet to develop the full range of skills required for effective leadership, such that they become more effective as they gain experience on the job. Further, while the issue of principal improvement has received little attention, ample evidence demonstrates substantial job performance returns to experience for teachers, particularly in their first few years in the classroom (e.g., Ladd & Sorensen, 2017; Papay & Kraft, 2015; Rockoff, 2004).

[1] See, for example: Gates, Ringel, Santibañez, Chung, and Ross (2003); Horng, Klasik, and Loeb (2010); Urick and Bowers (2014)

Despite a strong conceptual basis for expecting that principals improve over time, there is little empirical evidence documenting returns to principal experience, and the findings from the handful of prior studies are mixed. One major reason for this inconsistency is that differences in empirical approaches among studies produce parameters that have different interpretations. For instance, using cross-sectional data, Eberts and Stone (1988) find a strong positive correlation between principal experience and student achievement, while Brewer (1993) find no evidence of a relationship. However, it is difficult to draw good inferences about principal improvement from these studies since they exploit both across- and between-principal variation in experience. Instead, any correlation between principal experience and student achievement likely conflates three processes: (1) the returns to principal experience (i.e., within-principal improvement), (2) systematic sorting of principals to certain types of schools over their careers, and (3) nonrandom attrition of less (or more) effective principals.

More recent studies have addressed nonrandom sorting of principals to schools using a school fixed effects approach (Bastian & Henry, 2015; Clark et al., 2009; Grissom et al., 2018). Inclusion of school fixed effects effectively compares the performance of principals who lead the same school in different years, which alleviates bias from unobserved school characteristics (to the extent they are fixed over time). Models using school fixed effects have consistently found a positive association between experience and principal effectiveness (Bastian & Henry, 2015; Clark et al., 2009; Grissom et al., 2018). As noted by Clark et al. (2009), however, these estimates reflect the combined effect of any returns to principal experience and "ability" bias induced by less effective leaders leaving the principalship. To my knowledge, no studies have employed models that account for both nonrandom attrition and principal-school sorting to isolate the returns to principal experience.[2]

[2] Grissom et al. (2018) demonstrate a within-principal correlation between supervisor ratings and experience, which is suggestive of positive returns to experience. However, the results are not definitive with respect to improvement because the models do not contain any other covariates.

The Portability of Principal Improvement

Human capital theory distinguishes between general and firm-specific or industry-specific skills (Becker, 1962). General human capital increases worker productivity not only in the current firm but in any other firm, while firm-specific human capital only increases productivity only in the current firm. In the case of principals, distinguishing between the returns to general versus firm-specific (school-specific) experience is important for two reasons. First, it tells us something about the nature of principal improvement. For instance, what are the actual skills that principals are building that lead them to better job performance? Second, distinguishing between general and school-specific returns to principal experience informs policy debates around the distribution of experienced principals. If the returns to school-specific experience are large relative to general experience, policies aimed at reallocating highly experienced principals to struggling schools may be less effective than those that promote the retention and development of existing leaders in such schools.

Assuming some amount of on-the-job learning for principals, it is unclear, *a priori*, the extent to which this improvement constitutes general or firm-specific (school-specific) human capital. On one hand, principals' responsibilities may not vary widely across schools, particularly schools of the same level (e.g., elementary schools), which would suggest that productivity increases from on-the-job learning carry over when the principal changes schools. On the other hand, schools are complex ecosystems that have specific strengths and challenges that principals must adapt to and learn from over time.

As an illustrative example, consider one of the primary channels through which principals affect student learning: hiring and retention of teachers (Jacob, 2011; Loeb, Kalogrides, & Beteille, 2012). Prior work finds that effective principals may improve the quality of their school's teachers by engaging in strategic hiring and retention of effective teachers (Grissom & Bartanen, 2019b; Loeb et al., 2012). Improving the composition of the teaching staff likely requires both general and school-specific human capital. For example, principals needs to be able to identify high and low performers (Grissom & Bartanen, 2019b). Even with the widespread adoption of multiple-measure teacher evaluation systems that explicitly aim to facilitate differentiation of teacher quality, principals must rely on their own judgment in weighing different signals of teacher performance (e.g., formal classroom observations, informal walk-throughs, value-added measures). As principals gain experience with hiring teachers and making retention decisions, they may learn which signals are more reliable in terms of predicting future effectiveness, which ultimately leads to a more effective teaching staff and increased student learning. In this situation, on-the-job learning constitutes an increase in principals' general human capital, as the ability to accurately predict teacher effectiveness should increase principal job performance at any school.

Alternatively, an important facet of retaining effective teachers, particularly in disadvantaged schools, is building a positive school climate—an atmosphere where teachers feel a sense a collegiality, trust, and support (e.g., Brown & Wynn, 2009; Johnson & Birkeland, 2003). Creating and maintaining a positive school climate may require principals to form individual relationships with teachers over time. Additionally, the dynamics of schools—even those in the same neighborhood—may vary widely. Even an experienced principal who changes schools must build relationships with the new teaching staff and adapt to the specific context over time. Here, on-the-job learning is school-specific, as the relationships and trust built among teachers in one school do not carry over to the next school.

To summarize, there are several gaps in our understanding of the returns to principal experience. First, prior work documents a positive correlation between principal experience and school performance. However, it is unclear the extent to which this relationship is explained by principal improvement as opposed to systematic attrition of less effective leaders from the principalship. Additionally, prior studies have almost exclusively focused on student achievement,[3] despite agreement that principal effects on test scores are indirect. Finally, the issue of portability has been almost completely ignored. The remainder of the study focuses on helping to fill these gaps.

[3] Clark et al. (2009) and Grissom et al. (2018) being the most notable exceptions.

Data, Sample, and Measures

This study analyzes longitudinal administrative data from Tennessee covering the 2001–02 through 2016–17 school years, provided by the Tennessee Department of Education (TDOE) via the Tennessee Education Research Alliance at Vanderbilt University. The Tennessee data contain detailed information about all employees in the K–12 public school system, including job title, school placement, and demographic information. I connect these staff data to student files beginning in 2006–07. The student data include demographic and enrollment information, as well as achievement scores on statewide end-of-year exams for grades 3–8 and end-of-course exams for high school students.

Measuring Principal Experience

Like many statewide administrative datasets, the Tennessee data do not contain measures of job-specific experience. This means that while I can observe how long a given principal has worked in K–12 public education in Tennessee, I cannot observe how long they have been a principal if they entered the principalship in 2002 or earlier, which is the first year of the staff data files. Similarly, I cannot observe years of school-specific experience for individuals who were a principal in 2002 until they move to a different school. There are two ways to address this data limitation. My preferred approach is to treat principal experience as missing in cases where I cannot definitively determine the true experience value. This effectively drops principals who entered the principalship in 2002 or earlier. The primary disadvantage of this approach is that it potentially limits external validity, because the returns to principal experience are identified only from those individuals who entered the principalship in 2003 or later. Relatedly, I am only able to estimate the returns to experience up to 14 years. An alternative approach is to retain the full sample of principals and top-code experience (e.g., 10+ years). While this avoids the external validity issue, the trade-off is increased risk of bias, particularly if principal effectiveness continues to increase (or decreases) beyond the first 10 or 15 years.[4] Additionally, because I

[4] I explain this reasoning more thoroughly in the methods section when discussing the censored growth model.

cannot observe the true experience value for these principals, there is no clear way to estimate a model that would recover the returns to experience after 14 years even when keeping the entire sample.

Figure 1 shows the statewide distribution of total principal experience and school-specific experience. In 2017, only 6.7% of principals in Tennessee had 15 or more years of total principal experience, and only 3.9% had been in the same school (as a principal) for 15 or more years. Roughly two-thirds of principals had six or fewer years of prior principal experience, and the typical principal was in their fifth year in their current school. Figure 1 demonstrates that the restriction of estimating the returns to experience up to 14 years is not a major limitation, as most principals do not remain in the principalship for very long.

Table A1 shows, for each analysis year, the proportion of principals that have non-missing total and school-specific experience. In the earliest years, a majority of principals entered the principalship prior to 2002, and thus have missing values for total principal experience. For instance, only 29% of principals working in Tennessee schools in 2007 have observable prior experience. However, this proportion increases steadily over time, as the rate of principal attrition is high. By 2012, I can observe prior experience for two-thirds of principals, and 84% of principals in 2017. For school-specific experience, the missing data problem is less severe, as principals change schools relatively frequently. In 2007, I can observe school-specific experience for half of principals, up to 93% in 2017.

Table 1 shows descriptive statistics for Tennessee principals between 2006–07 and 2016–17. Among those for whom I can observe prior experience, the average principal has 3.1 years of prior principal experience and has been in her school as a principal for 2.7 years. A large portion of principals are both new to the principalship and new to their schools. Thirty-five percent of principals have fewer than two years of prior principal experience, and 41% have been in their current school (as a principal) for fewer than two years. The average principal in Tennessee is 50 years old and has worked in the public school system for roughly 23 years. Almost 60% of principals work in elementary schools, and more than half of Tennessee's schools are located in areas classified as town or rural.

To separate the returns to total experience as a principal and school-specific experience, there must be a sufficient number of principals who work in multiple schools across the study period. Figure A1 shows the distribution of total principal experience versus school-specific experience. Each dot is a principal-year observation, with random jitter added to show relative density. The diagonal (i.e., equal amounts of total experience and school-specific experience) have the most observations, demonstrating that many principals in the sample are observed in just a single school. Nevertheless, there are a fair number of principals who move between schools, and thus have different amounts of total and school-specific experience.

Measuring Principal Effectiveness

Student Achievement. The main outcome of interest in this study is student achievement. Specifically, I draw on achievement scores in math, reading, and science for students in grades 3–8 and end-of-course (EOC) exams for high school students. The grade 3–8 exams are required for every student across each year of the study period, while the EOC exams vary by year. In 2016–17, students took exams for Algebra I, Geometry, Algebra II, English I, English II, English III, Chemistry, and Biology. Earlier years had fewer tested subjects in high school. I construct a common measure of student achievement by standardizing exam scores within subject, grade, and year for grades 3–8. For EOC exams, which can have students from multiple grades (e.g., the Algebra I exam includes large numbers of ninth and tenth grade students), I standardize scores within each course and year.

Ratings from Supervisors. As an alternative to using changes in student outcomes as a proxy for principal performance, I also draw on (plausibly) more direct measure of principals' practice: ratings from their supervisors. These ratings are rubric-based scores that principals receive as part of Tennessee's statewide educator evaluation system (TEAM) implemented in 2011–12. Fifty percent of the TEAM evaluation for principals comes from ratings of principal performance on a rubric derived from the Tennessee Instructional Leadership Standards.[5] These ratings are based on formal observations conducted by the principal's supervisor. Prior work shows that principals' ratings across indicators are highly inter-related and can be reduced to a single underlying performance score using factor analysis (Grissom et al., 2018). In this analysis, I use principals' average yearly observation scores—the exact measure used by the state to calculate summative evaluation ratings. I refer to this measure as "supervisor ratings."[6]

Teacher Outcomes. The final outcome I examine is teacher turnover. Prior work demonstrates that high-quality principals retain teachers at higher rates (Boyd et al., 2011; Grissom & Bartanen, 2019b; Ladd, 2011). Therefore, I also examine the extent to which principals improve at retaining teachers as they gain experience. More specifically, I construct a binary and multinomial measure of teacher turnover. The binary measure takes a value of one if teacher s in school s in year t is no longer a teacher in school s in year $t+1$, and zero otherwise. The multinomial measure categorizes three types of teacher turnover: exits from the state education system, moves to a teaching position in a different school, and changes to a non-teaching position (e.g., assistant principal, instructional coach).

Methods

Research Question 1: To what extent do principals become more effective as they gain experience?

My first research question seeks to estimate the job performance returns to principal experience. I estimate via ordinary least squares models of the general form:

$$Y_{ist} = \delta Experience_{it} + \gamma \mathbb{X}_{st} + \mu_i + \psi_s + \tau_t + \epsilon_{ist} \tag{1}$$

[5] For more information about TEAM, see http://team-tn.org/evaluation/administrator-evaluation/

[6] Using the average observation score instead of the factor score described in Grissom et al. (2018) allows me to include principals in districts that used alternative observation rubrics (approximately one-quarter of principals in the state), as these districts do not report domain-specific scores for principals. However, for principals for whom I can calculate factor scores, the average observation score and the factor score are correlated at 0.95 or higher each year.

where Y is the performance of principal i in year t and δ is the marginal effect of principal experience. As discussed above, in addition to direct measures of principal performance, I also examine whether there are returns to principal experience for student- and teacher-level outcomes. These models follow the same form as principal-level models but include additional covariates, which I explain below. Equation 1 also includes a vector of school characteristics ($\mathbb{X}$): enrollment size and average student demographics (race/ethnicity, free/reduced-price lunch eligibility, gifted status, special education status). Finally, I include fixed effects for principal (μ_i), school (ψ_s), and year (τ_t).

The inclusion of principal, school, and year fixed effects are critical to the identification of δ. Principal fixed effects isolate within-principal variation in experience, such that the effects of additional experience are identified by comparing student outcomes under the same principal across years. An unbiased estimate of δ in a model without principal fixed effects requires that the accumulation of experience is uncorrelated with any fixed differences in principal quality. Prior work demonstrates that less effective principals are more likely to exit the principalship (Grissom & Bartanen, 2019a), which highlights the importance of including principal fixed effects.

School fixed effects control for time-invariant differences between schools, such as the quality of facilities or neighborhood effects. If principals were randomly assigned to schools, accounting for school heterogeneity would not be necessary. However, prior work demonstrates that principals may seek to sort to more advantaged schools over time (e.g., Béteille et al., 2012). Including school fixed effects helps to ensure that the returns to experience are not conflated with sorting to higher-quality schools.[7]

Year fixed effects account for any state-level factors that are correlated with both the given outcome and the accumulation of experience. In particular, I must account for any systematic *trends* in the outcome, which I would otherwise attribute to the returns to experience. The nature

[7] Here, I refer to differences in school quality in terms of factors that the principal cannot control. Clearly, principals themselves are an input to school quality. However, there are many school-level factors that are more less fixed over time, cannot be controlled by the principal, and contribute to student learning. Some examples are the neighborhood in which the school is located, the amount of resources the principal can access, and the quality of school facilities.

of the year fixed effects—and what they actually account for—depends on the particular outcome variable and whether it has been standardized within year. In the case of an unstandardized variable, such as teacher turnover, the year fixed effects will capture any time-varying factors that change the turnover propensity all teachers in the state, such as the implementation of a high-stakes educator evaluation system, law changes affecting teachers' tenure and due-process protections, or labor market conditions. For outcome variables that have been standardized within year, the year fixed effects account for average changes in the distribution of principal quality over time. If, for instance, the quality of new principals is increasing over time and the outcome variable is standardized within year, estimates of the returns to principal experience in a model without year fixed effects will be biased downwards.[8]

Separately Identifying Year Fixed Effects and the Returns to Experience

An important consideration in estimating equation 1 is how to parameterize principal experience. Modeling within-person returns to experience (i.e., including principal fixed effects) means that experience is perfectly collinear with year for most principals.[9] This collinearity means that identifying both the returns to experience and year fixed effects requires additional identification assumptions or sample restrictions. Papay and Kraft (2015) discuss three approaches for identifying identifying year fixed effects in the context of teacher fixed effects models, which I outline below.

Approach 1: Place Restrictions on the Experience Profile. The first approach—and the one that is most common in the literature—is to exploit regions of the experience profile where

[8] To see why, consider a simplified example where principal performance improves by x with each additional year of experience and the average quality of entering principals also improves by x each year. In this scenario, as long as principals leaving the profession are not systematically above average in terms of effectiveness, the distribution of principal quality increases across years. However, the outcome variable does not measure true principal performance, but rather reflects a principal's performance relative to the average principal in that year. A given principal, then, who improves by x each year, appears to improve less than x because of the global mean shift in the standardized outcome.

[9] More specifically, experience and year are perfectly collinear for principals who do not have discontinuous careers. For individuals who leave the principalship (i.e., move to central office or take a year off) and then return, experience and year will not be perfectly correlated. I discuss this case further below.

the marginal returns to experience are zero (or small in magnitude).[10] For instance, Rockoff (2004) implements a "censored growth model" whereby the returns to experience beyond a teacher's tenth year are restricted to be zero. Using this restriction, Rockoff (2004) identifies the year fixed effects from the subset of teachers who have more than ten years of experience. A related approach is the "indicator variable model," which places restrictions throughout the experience profile by constructing experience "bins" (e.g., 0, 1–2, 3–5, 6–10 years, etc.). The censored growth model is effectively a special case of the indicator variable model that uses a single bin (10+ years). The identifying assumptions of these models are similar; unbiased estimates of the returns to experience require that the marginal effect of experience is zero within the specified bins. Any productivity growth (decline) within these bins will lead to upward (downward) bias in the estimated year fixed effects, which will downwardly (upwardly) bias the estimated returns to experience.

The choice of how to construct the experience bins is arbitrary, though researchers typically draw on prior empirical findings. For example, numerous studies demonstrate that teachers improve most rapidly in their first few years on the job (e.g., Harris & Sass, 2011; Ladd & Sorensen, 2017; Papay & Kraft, 2015), which suggests that placing restrictions on productivity growth towards the beginning of the experience profile will lead to conservative estimates of the returns to experience. Of course, the true shape of the experience profile is unknown and could vary substantially across contexts (e.g., state-level versus district-level datasets, urban versus rural schools), which means that relying on prior findings is not a perfect solution. Papay and Kraft (2015) propose two checks for the plausibility of the restrictions on the experience profile. The first is to simply examine the estimates near the cutoff points; evidence of productivity growth near the censoring point(s) would suggest that the returns to experience within these bins are not zero.[11] The second check is to split the bins (e.g., 5–10 years becomes 5–7 and 8–10 years) and

[10] See Harris and Sass (2011); Kraft and Papay (2014); Ladd and Sorensen (2017); Papay and Kraft (2015); Rockoff (2004) for examples.

[11] Rockoff (2004) also implements this check and finds that his censoring point at 10 years is reasonable for most, though not all, of his outcomes.

compare the estimates to the initial model. If the zero growth assumption holds, narrow bins should produce an estimated experience profile similar to wider bins.

Approach 2: Leverage Discontinuous Careers. Whereas approach 1 relies on assumptions about the shape of the experience profile, a different approach is to circumvent the perfect collinearity between experience and year by leveraging individuals who have "discontinuous careers." Some teachers temporarily leave the profession such that they do not always accumulate additional experience each year. Without placing restrictions on the experience profile, one can identify both the returns to experience and year fixed effects. As noted by Papay and Kraft (2015), the discontinuous career approach faces both internal and external validity concerns. Teachers with discontinuous careers tend to be a small subset of the sample, which raises concerns that the returns to experience (or, equivalently, the year fixed effects) for these teachers are not generalizeable to teachers with continuous careers. In terms of internal validity, this approach assumes that temporarily leaving the profession has no effect on the returns to experience. This assumption is violated if, for instance, taking medical leave has a negative shock on teacher effectiveness upon returning.[12] An additional limitation of this approach is that even if the assumptions holds, the estimates can be very imprecise given the small number of individuals with discontinuous careers.

Approach 3: Leverage Between-Person Variation. A final approach proposed by Papay and Kraft (2015) is a two-stage model that produces estimates of the year fixed effects in the first stage then applies these coefficients to the second-stage model when estimating the returns to experience. Specifically, the first-stage model omits teacher fixed effects, such that the year fixed effects are identified from between-teacher variation. The key assumption of this approach is that there is no change in the quality of the teachers entering the profession (among those in the particular sample) over time. As Papay and Kraft (2015) discuss, there are many plausible reasons why this assumption would not hold. In particular, they suggest that policy reforms that

[12] Papay and Kraft (2015) find evidence of negative productivity shocks for teachers with temporary absences from the profession. Specifically, they estimate modified versions of the discontinuous career model with indicators for the year immediately before and after a discontinuity, finding that teacher effectiveness is lower in both of these years.

have lowered barriers to entry through alternative certification and improvements in teacher preparation programs could have led to increases in the quality of new teachers. Papay and Kraft (2015) show that under this scenario, the estimated returns to experience in the two-stage approach would be biased downwards.[13]

Given the dearth of evidence on principals, it is unclear which of these approaches is best-suited for estimating the returns to principal experience, which leads me to estimate models using each of them. Specifically, I estimate an indicator variable model using the following experience ranges: 0, 1, 2, 3, 4–6, 7–9, 10–14 years. While this choice of bins is based on the assumption that the returns to principal experience are largest in the first few years, I test the sensitivity of the estimates to different bins. Additionally, I estimate a model where growth is censored after five years. Using a cutoff at a higher experience level is not feasible given data limitations and how few principals remain in the principalship over time. I also implement the two-stage model proposed by Papay and Kraft (2015) by omitting principal fixed effects in the first-stage (but still including school fixed effects) and applying the estimated coefficients for the year fixed effects to the second-stage model to estimate the returns to principal experience. For the discontinuous career model, I include a fully non-parametric specification of principal experience up to 14 years. However, this approach yields imprecise estimates and, for the sake of brevity, I simply provide the results in Appendix Table A2.

To examine the returns to experience for student outcomes, I estimate the following specification of equation 1:

$$Y_{igjst} = \delta Experience_{it} + \gamma \mathbb{X}_{st} + \eta \mathbb{Z}_{jt} + \mu_i + \psi_s + \sigma_g + \tau_t + \epsilon_{igjst} \tag{2}$$

where Y_{igjst} is the achievement score of student j in grade g, with principal i, in school s, in year t. In addition to school characteristics, these models also adjust for student characteristics ($\mathbb{Z}_{jt}$):

[13] Both Papay and Kraft (2015) and Ladd and Sorensen (2017) both find evidence that the key identifying assumption of the two-stage model is not met for at least some subjects. In both cases, they find that the quality of new teachers is increasing over time, which biases downwards the estimated returns to experience in the two-stage model.

race/ethnicity, gender, free/reduced-price lunch eligibility, gifted and special education status, an indicator for grade repetition, and an indicator for whether the student was previously enrolled at a different school in the current year. Note that I do not control for students' prior achievement scores. Because most students remain in the same school between year $t-1$ and year t, they also tend to have the same principal in both years. The inclusion of prior-year achievement, then, is a violation of strict exogeneity in a model with principal fixed effects, as the principal in year t often affects the prior-year score. As a check, however, I also estimate (a) models that adjust for prior-year test scores and (b) models that adjust for a student's most recent test score in a prior school and find qualitatively similar results for the returns to experience.[14] I cluster standard errors at the principal-by-school level.

The model for principals' ratings from supervisors is:

$$Y_{ist} = \delta Experience_{it} + \gamma \mathbb{X}_{st} + \mu_i + \psi_s + \tau_t + \epsilon_{ist} \tag{3}$$

where Y is a principal's average score in year t, with scores standardized across the full sample of principals within each year to have a mean of zero and standard deviation of one. Besides the fixed effects, I also include controls, $\mathbb{X}_{st}$, for time-varying school characteristics (enrollment size and school-level averages of student demographics). I cluster standard errors by school district.

Finally, I estimate the following teacher-level model:

$$Y_{ijst} = \delta Experience_{it} + \gamma \mathbb{X}_{st} + \eta \mathbb{Z}_{jt} + \mu_i + \psi_s + \tau_t + \epsilon_{ijst} \tag{4}$$

where Y is a binary indicator for teacher turnover (i.e., takes a value of one in year t if teacher j does not remain a teacher in the same school in year $t+1$). As in the student models, I control for personal (teacher) characteristics, $\mathbb{Z}_{jt}$, which include race, gender, age, experience, and highest education level. I cluster standard errors at the principal-by-school level.

I make one very important modification when estimating the returns to principal experience

[14] These results are shown in Table A3.

for teacher turnover: I include an indicator for principal turnover. Prior work demonstrates that teacher turnover is greater in years where schools change principals (Bartanen, Grissom, & Rogers, 2019; Miller, 2013), and principals' efforts to retain teachers may be less effective if the principal is not returning to the school in the following year. Additionally, the likelihood of principal turnover increases the longer that principals remain in the school (Grissom & Bartanen, 2019a), such that including these principal turnover years may lead to the conclusion that more experienced principals are less effective at retaining teachers. Given that this analysis focuses on identifying improvement, adjusting for the final year of a school spell ensures that identification comes only from years when the principal should be actively working to retain teachers.

Research Question 2: Are the returns to principal experience driven by total or school-specific experience?

My second research question seeks to examine whether principal improvement is driven by improvements in general or school-specific skills. Here, I exploit the fact that some principals work in multiple schools over their careers to separately identify the returns to *total* principal experience and *school-specific* principal experience:

$$Y_{ist} = \delta Experience_{it} + \theta ExperienceSchool_{ist} + \gamma \mathbb{X}_{st} + \mu_i + \psi_s + \tau_t + \epsilon_{ist} \tag{5}$$

If principal improvement over time reflects an increase in skills that are fully portable across schools, controlling for school-specific experience should not appreciably change estimates of δ relative to equation 1. Conversely, if improvement is not portable, estimates of δ will be attenuated while estimates of θ will be positive. As mentioned above, successfully separating the returns to total and school-specific experience requires principals who I observe in multiple schools. Across the study period (2007–2017) and sample (principals for whom I can determine total experience), 19% of the 2,500 unique principals worked in more than one school.[15]

[15] Specifically, 16% worked in two schools, 2.5% worked in three schools, and 0.5% worked in more than three schools. Figure A1 plots total experience versus school-specific experience for each principal-by-year observation,

Research Question 3: To what extent is there heterogeneity in the returns to principal experience across school contexts?

My third research question examines heterogeneity in the returns to principal experience across school contexts. To be specific, I examine whether principals in certain types of schools improve more or less rapidly over time. Because principals may sort to (or away from) certain types of schools over their careers (e.g., moving from elementary to high schools), I focus on heterogeneity in the returns to school-specific experience. I examine three contextual variables: student poverty (low-, medium-, and high-poverty),[16] school locale type (urban, suburban, and town/rural), and school level (elementary, middle, high).[17] To test for heterogeneity, I include an interaction between the given school contextual variable and school-specific experience:

$$\begin{aligned} Y_{igjst} = \; & \theta ExperienceSchool_{ist} + \eta(ExperienceSchool_{ist} \times Context_s) \\ & + \gamma \mathbb{X}_{st} + \mu_i + \psi_s + \omega(Year_t \times Context_s) + \epsilon_{ist} \end{aligned} \tag{6}$$

η represents the difference in the returns to principal experience relative to the arbitrary holdout group. Positive (negative) estimates of η would indicate that principals in the given school category improve more (less) rapidly than principals in the omitted school category. Note that the school context variables are time-invariant and thus the main effects are absorbed by the school fixed effect. I also replace year fixed effects with year-by-context effects, which allows any trends in unobserved determinants of principal effectiveness to be different by school contextual category. This modification is critical to avoid misattributing heterogeneity in productivity trends to heterogeneity in the returns to principal experience. If, for example, changes in school

with random jitter added to show density at discrete values of experience.

[16] I construct these categories to be time-invariant across the study period by taking the median of the proportion of students that qualify for free/reduced price lunch (FRPL) at the school in each year. The low-poverty group includes schools with fewer than 30% FRPL students, medium-poverty includes 30–80% FRPL, and high-poverty includes 80% or higher FRPL.

[17] I drop from the heterogeneity analysis the small number of schools that are classified as "other" school level by NCES.

accountability systems affect principal quality in high-poverty schools more than in low-poverty schools, including year fixed effects (which assume that time-varying shocks affect all principals/schools equally) will lead to bias in the estimates of heterogeneity in the returns to experience.

Results

The analysis proceeds in four parts. First, I present estimates of the returns to principal experience for student outcomes and supervisor ratings. Based on the findings for student achievement, I then propose an alternative approach to estimating the returns to experience that places restrictions on the year fixed effects rather than the experience profile. Second, I show results from models that separate the returns to total principal experience and school-specific experience. Third, I examine the relationship between principal experience and teacher turnover. Finally, I examine heterogeneity in the returns to experience for principals working in different school contexts.

The Returns to Principal Experience for Student Achievement

Table 2 shows estimates of the returns to experience for student achievement in math, English/language arts (ELA), and science. For each outcome, I show results from the indicator variable, censored growth, and two stage models. While I also estimated the discontinuous career model (see Appendix Table A2), there was an insufficient number of principals to produce precise estimates.[18]

[18] Further, the documented relationship between school/principal performance and principal turnover (Bartanen et al., 2019; Grissom & Bartanen, 2019a) strongly suggests that the discontinuous career approach is problematic both in terms of internal and external validity. Specifically, principal turnover (which is the primary reason why a principal would fall into the discontinuous career group) is preceded by a drop in school/principal performance and principals who are less effective are more likely to turn over. Thus, almost by definition, discontinuous career principals are systematically less effective than the typical principal in Tennessee. While this is not problematic, per se, it suggests that the improvement trajectory among these principals may also be unrepresentative of the population. In terms of internal validity, the identifying assumption of the discontinuous career model is that the performance of these principals when they returned to the principalship is the same (in expectation) as it would have been had they not left. Whereas teachers are more likely to have discontinuous careers for reasons plausibly orthogonal to performance (e.g., child-bearing), the reasons for principals temporarily exiting the principalship are more likely performance-related. Finally, the gap in experience typically corresponds to a change in school, making it difficult to assume that principal

Beginning with the estimates from the indicator variable model (IVM), I find positive returns to principal experience in math and science. For ELA, the coefficients are positive and increasing over time, but they are not statistically significant at conventional levels. Similar to findings for teachers, the IVM results show that principals improve most rapidly in the first few years. However, I find that the marginal returns to experience are positive throughout the experience profile. This implies that the identifying assumption of both the IVM and censored growth model——that there is a "flat" region of the experience profile which can be used to identify year fixed effects—is not met. By consequence, both the IVM and censored growth models should produce conservative estimates of the returns to principal experience. The censored growth model results in columns 4–6 are consistent with this expectation. Placing a restriction of zero returns to experience after a principal's fifth year leads to upward bias in the year fixed effects and downward bias in the returns to experience.[19] I return to this issue below.

Columns 7–9 show the estimated returns to principal experience from the two-stage model. Compared to the IVM model, the coefficients are substantially smaller in magnitude. The identifying assumption of the two-stage is that the quality of new principals between 2007 and 2017 (the study period) is unchanging. That the returns to experience are smaller compared to the IVM results suggests that this assumption does not hold. Specifically, that the estimates appear to be biased downwards suggests the quality of new principals *increased* over time.

An Alternative Approach to Estimating the Returns to Principal Experience

Table 2 suggests that while the returns to principal experience for student achievement are positive, the estimates from each of the models are conservative. However, the magnitude of the bias is unclear. To further explore this problem, I propose an alternative approach that places restrictions on the year fixed effects instead of the experience profile. The logic of this approach is to exploit flat regions of the *time* profile to identify the returns to principal experience instead

performance would have been the same in the absence of the move.

[19] Appendix Table A4 shows the results of the IVM using narrower experience bins. As expected given the apparent presence of within-bin growth in Table 2, narrowing the bins leads to larger estimates of the returns to experience.

of using flat regions of the experience profile to identify year fixed effects. In essence, this is a modified version of the indicator variable model, where the “buckets” are groups of adjacent years. Here, the identifying assumption is that, conditional on principal experience, there is no positive or negative productivity growth within the specified bins. Of course, identifying *which* years should be grouped together is critical to this approach. Whereas the choice of experience bins in the IVM approach is motivated by the hypothesis of decreasing marginal returns to experience, there is no theory to guide the choice of year bins. Instead, I rely on empirical estimates of the year fixed effects from the models estimated in Table 2. The validity of this approach relies on the assumption that the estimated year fixed effects from these models accurately identify the true *shape* of the productivity trend across years.[20]

Table A5 shows the estimated year fixed effects from the student achievement models. In contrast to the experience profile, there appear to be regions where the estimated time trend is flat. Guided by these estimates, I specify bins that group adjacent sets of years, which I use to re-estimate the achievement models. For example, in the math models I replace year fixed effects with the following year bins: 2007, 2008–2009, 2010, 2011, 2012–2013, 2014–2015, 2016–2017.[21] Table A6 shows the achievement results replacing year fixed effects with the constructed year bins (but not changing the experience variables). The estimated returns to experience are very similar to the estimates from the models with year fixed effects, supporting the validity of the modified year indicators.

Next, I replace the principal experience bins with a fully flexible set of year indicators, with the results shown in Table 3. As expected, placing restrictions on the year fixed effects effects

[20] Here I distinguish *shape* from *level* in saying that it is not necessary that the year fixed effects from the IVM are unbiased as long as the magnitude and direction of the bias between adjacent years is small enough to make an accurate judgment of whether the year-to-year productivity growth is small or large.

[21] I use slightly different bins for ELA and science: (ELA) 2007–2008, 2009, 2010–2011, 2012–2013, 2014, 2015, 2016, 2017; (Science) 2007, 2008, 2009, 2010–2011, 2012–2013, 2014, 2015–2016, 2017. There is no reason to assume that the time trends should follow the exact same pattern across subjects. For instance, there may be changes in standards or the end-of-year exams that affect math achievement differently than reading achievement. Nevertheless, the patterns in the estimated year effects in Table A5 are substantially similar across subjects, suggesting that the year effects are picking up general productivity trends that affect test score performance similarly in all subjects.

rather than the experience profile leads to larger estimates of the returns to experience, suggesting that the estimates in Table 2 are indeed conservative. Further, the magnitude of difference between the estimates in Table 3 and Table 2 is fairly substantial, particularly at higher levels of experience. The returns to experience at 1 year, for instance, are 25–40% larger (depending on the subject) using my modified approach as compared to the indicator variable model. At 4–6 years, the difference in magnitude is more than 30% in math and more than 60% in ELA and science.

At a minimum, the results from the IVM approach and the approach using modified year bins demonstrate that a principal's ability to raise student achievement increases substantially with experience. Even conservative estimates using the IVM approach imply that the magnitude of principal improvement is large. In other work I estimate that the standard deviation of principal effects on math (ELA) achievement (i.e., the amount that student math achievement increases for a 1 SD increase in principal quality) is roughly 0.20 (0.10) SD in Tennessee (Bartanen & Grissom, 2019). Thus, the amount of within-principal improvement relative to the overall distribution of principal quality is quite large. Put another way, these results imply that the average principal substantially increases their rank in the effectiveness distribution over time. Next, I examine the extent to which principals' apparent improvement over time is corroborated using alternative outcomes.

The Returns to Principal Experience for Supervisor Ratings

Table 4 shows estimates of the returns to principal experience for supervisor ratings. Different from the estimates for student achievement, the IVM, censored growth, and two-stage approaches produce very similar results for the returns to experience. Across each model, principals improve substantially in their first five years on the job (the returns to experience up to five years are roughly 0.50 SD), with little to no evidence of returns to experience beyond five years. Given the lack of growth beyond five years, it is no surprise that the censored growth and indicator variable models produce similar estimates, since they both assume that the returns to experience are largest in a principal's first few years.

Why are the results from the two-stage model similar to those from the censored growth and indicator variable models? Whereas the year fixed effects are statistically significant (and large in magnitude) in the student achievement models, they have no explanatory power in the supervisor ratings models, as evidenced by the large *p*-values for *F*-tests of their joint significance shown in the bottom of Table 4.[22] Further, the year fixed effects in the first stage of the two-stage approach are small in magnitude and not jointly significant, meaning that the second stage model is effectively equivalent to estimating a model that simply omits year fixed effects.

As with student test scores, the magnitude of the estimated returns to experience for supervisor ratings are substantial. On average, principals in Tennessee move from the 35th percentile to the 53rd percentile in supervisor ratings between their first and sixth year in the principalship. That said, it is important to note that the analysis of supervisor ratings faces an important limitation: part of the observed effect is likely driven by raters' knowledge of a principal's experience level and their expectation that principals improve with experience and/or that experienced principals are more effective than inexperienced principals. Still, the improvement in supervisor ratings supports the findings from the student achievement models that there are substantial returns to principal experience.

Is Principal Improvement General or School-Specific?

The previous section establishes that principals improve substantially over time as measured by increases in student test scores and ratings from supervisors. Next, I investigate the extent to which this improvement is driven by general or school-specific skills. Specifically, I leverage principals who have worked in multiple schools across the study period to separate the returns to *total* experience and *school-specific* experience for student achievement.

The odd columns in Table 5 show the relationship between school-specific experience and student test scores in math, ELA, and science. The estimates are similar to those for the returns to

[22] This is not driven by large standard errors. The coefficient estimates are uniformly small in magnitude. For instance, the coefficients for the year fixed effects in the IVM are (ascending by year with 2017 omitted): 0.015, 0.018, -0.033, -0.018, -0.01 SD for 2012–2016.

total principal experience in Table 2, which is expected given that the majority of principals are only observed in a single school.[23] The even columns show the parameters of interest. Across all three subjects, the returns to principal experience are driven by school-specific rather than total experience. Put another way, the results in Table 5 suggest that principal improvement (as measured by changes in student achievement) is not portable—that the skills principals acquire with experience in one school do not help them raise test scores in a different school.

Given my reliance on principals who I observe in multiple schools to separate the returns to total and school-specific experience, it is important to understand the extent to which these principals are representative of the broader sample. In Table A8, I examine whether the returns to school-specific experience are different between principals observed in a single school versus those observed in multiple schools. I find no evidence that the experience trajectories differ between these groups of principals.

A related concern, which applies to the analysis more broadly, is the high rate of attrition among principals. As mentioned above, roughly 80% of principals in the sample are observed in only a single school and the median tenure length is three years. By consequence, estimates of the returns to experience rely on an increasingly small subset of principals as experience increases. This may undermine the generalizability of the estimates to the broader sample if those who remain in the principalship have systematically steeper (or flatter) experience profiles. This is relevant for the policy implications of the study. If, for instance, the returns to experience merely reflect the fact that principals who improve most rapidly are those who stay in the principalship, then policies that curb principal attrition may not yield the expected benefits. To explore this possibility, I estimate a series of models that test for heterogeneity in the experience trajectory by how long the principal stays in their school.[24] The results, shown in Table A7, show no evidence of heterogeneity. In other words, I do not find that those who only remain in the principalship a

[23] The correlation between total experience and school-specific experience is 0.78.

[24] Specifically, I estimate models that interact the returns to school-specific experience with an indicator for whether the principal stays in the school for at least x years, with separate models for x=2,3,5,10. Further, I exclude from the model principal-by-school spells shorter than x if they are left-censored, since I cannot determine how long these principals will stay in their school.

few years improve at different rates than those who stay longer.

Principal Improvement and Teacher Outcomes

A key channel through which principals affect student outcomes is human capital management—the hiring and retention of teachers. Given that teachers are the most important school-level input to student learning (e.g., McCaffrey, Lockwood, Koretz, & Hamilton, 2003; Rivkin, Hanushek, & Kain, 2005), the positive returns to principal experience for student achievement are likely to be mediated (at least in part) by principals' effects on teacher-level outcomes. In this section, I explore this potential pathway by estimating the returns to principal experience (total and school-specific) for teacher outcomes.

Tables 6 shows the results for teacher turnover. Prior work demonstrates that teacher turnover has negative effects on student achievement (Ronfeldt, Loeb, & Wyckoff, 2013), which suggests that principals who are able to retain teachers at higher rates may see improvements in student learning. Further, an important factor in teachers' decisions to remain in a school is the quality of school leadership (Boyd et al., 2011; Grissom, 2011; Kraft, Marinell, & Yee, 2016). In addition to a binary measure of teacher turnover (i.e., whether the teacher remains in the school between year t and $t + 1$), I also three types of turnover: exits from the education system, transfers to a different school, and position changes (e.g., becoming a school counselor or an assistant principal).[25] Motivated by the results in Table 5, I estimate models examining both total and school-specific principal experience.

Columns 1–3 show the results for a binary measure of teacher turnover. All models are estimated via OLS such that the coefficients reflect the marginal change in the probability that a teacher turns over. While the coefficients in column 1 (total experience) and column 2 (school-specific experience) are negative and increasing in magnitude at higher levels of principal experience (which would indicate lower teacher turnover rates), most are not statistically

[25] Specifically, I analyze this categorical outcome as a series of linear probability models where the base category is stayers. This is preferable to alternative modeling approaches (e.g., multinomial logistic regression) because of the inclusion of high-dimensional fixed effects.

significant with large standard errors. Column 3, which includes both experience types, again suggests that the patterns are driven by school-specific rather than total principal experience.

The remaining columns examine specific types of teacher turnover. Column 5 shows that there are fairly substantial returns to school-specific experience in terms of lowering the probability that teachers transfer to another school. Compared to a principal's first year in the school, teachers are 0.6 percentage points less likely to transfer following the principal's second year, up to 2.2 and 5.3 percentage points at 4–6 and 10–14 years of school-specific experience, respectively. These effects correspond to 7%, 24%, and 59% of the average teacher transfer rate in Tennessee across the study period. For exits, the point estimates are suggestive of a pattern for total experience in column 7, but they are not statistically significant. For position changes, the point estimates are consistently close to zero and not statistically significant.

An additional consideration in estimating the returns to principal experience for teacher turnover is that the composition of the teaching staff changes over time. New-to-school principals inherit the teaching staff of their predecessor(s), which they subsequently shape through hiring new teachers and retaining (or failing to retain) existing teachers. The returns to principal experience for teacher turnover, then, may be misleading (in terms of examining improvement) if the the average latent turnover propensity (i.e., individual factors that drive turnover decisions that the principal cannot control) of the teaching staff is correlated with a principal's length of tenure in the school. Further, any improvement in their ability to retain teachers may vary by whether the teacher was inherited or hired. Inherited teachers—who by definition have been in the school for a longer time and who tend to be more experienced—may be considerably less responsive to school leadership with respect to their mobility decisions.[26]

In Table 7, I examine whether the relationship between principal experience and teacher turnover varies between inherited and hired teachers.[27] For ease of interpretation, I show the

[26] Selection may also be a factor here. Inherited teachers presumably had the option to leave the school when the prior principal turned over. Those that remain may be systematically more committed to the particular school.

[27] I construct a binary indicator of "hired" teacher using the combination of teacher and principal's length of tenure in school. If a teacher entered the school in the same year or later than the principal, I code them as "hired," otherwise they are "inherited."

marginal effects of school-specific experience (of the principal) for these two groups of teachers.[28] Column 1 shows that while principals improve over time at lowering turnover rates of both inherited and hired teachers, the effect is much larger for hired teachers. As with the baseline models, these patterns are driven by lower rates of transfer among both hired and inherited teachers. However, column 3 shows that, as they gain experience in the school, principals see lower rates of exits among teachers that they hired.

While examining average teacher turnover rates is informative given the documented relationship between turnover and lower student performance, effective leaders may focus their retention efforts on effective teachers (Grissom & Bartanen, 2019b), for whom the costs of turnover are greater (Adnot, Dee, Katz, & Wyckoff, 2017). Next, I examine whether principals may become better at *strategic retention*—retaining effective teachers and "failing to retain" ineffective teachers—as they gain experience. To measure teacher effectiveness, I first estimate value-added (VA) models following the methodology of Chetty, Friedman, and Rockoff (2014), then construct a categorical measure of teachers' rank in the statewide VA distribution (bottom 20%, middle 60%, top 20%).[29] In total, I can produce VA estimates for roughly 45% of teachers in the sample.

Table 8 shows the estimated marginal effects of principal experience for turnover among low, middle, and high value-added teachers.[30] While the marginal effects are largest for high VA teachers, in general the estimates are noisy and I cannot reject the null hypothesis that they are

[28] The main effect plus interaction models are shown in Table A9. Except for position changes, the differences between hired and inherited teachers are statistically significant across the individual coefficients and in joint tests of the interaction terms.

[29] The estimation steps are as follows. First, I residualize student test scores (separately by subject) on a vector of prior-year test scores, student characteristics (race/ethnicity, gender, FRPL eligibility, gifted status, special education status, lagged absences, grade repetition, and whether the student changed schools at least once during the year), school- and grade-level averages of these student characteristics, grade-by-year fixed effects, and teacher fixed effects. After computing the student residuals, I add back the teacher fixed effects and estimate the best linear predictor of a teacher's average student residuals in the current year based on their residuals from prior and future years. The coefficients from this best linear predictor are then used to predict a teacher's value-added in the current year. For teachers with value-added estimates in multiple subjects, I average these scores to construct a single measure of teacher effectiveness.

[30] Table A10 shows the main effects with interactions instead of marginal effects.

equal across high, middle, and low VA teachers (see Table A10). Finally, Table A11 directly estimates the relationship between teacher value-added and principal experience. Perhaps unsurprisingly given the results in Table 8, there is no apparent improvement in teacher quality as principals gain experience.

Heterogeneity in the Returns to Principal Experience by School Context

My final research question examines the extent to which principals improve more or less rapidly in certain types of schools. Specifically, I estimate whether the returns to school-specific experience for student achievement vary across three measures of context: school poverty, school level, and school locale.

Table 9 shows the estimated returns to experience for principals in low-poverty (0-30% FRPL), medium-poverty (30-70% FRPL), and high-poverty (70-100% FRPL) schools.[31] Across each subject, the returns to experience are largest in high-poverty schools. For instance, average math achievement improves by 0.042 SD between a principal's first and second year in a high-poverty school, compared to 0.002 SD (n.s.) and 0.009 SD (n.s.) in low- and medium-poverty schools, respectively. By 4–6 years of school-specific experience in high-poverty schools, student achievement has increased by 0.10 to 0.12 SD depending on the subject.

In Table A12, I show the results replacing year-by-poverty fixed effects with year fixed effects. The difference in the results between these two specifications is substantial. Whereas the returns to experience are driven by principals in high-poverty schools in my preferred models, the year fixed effects models lead to the opposite conclusion. This difference highlights the importance of allowing for flexibility in terms of identifying productivity trends that are correlated with the acquisition of experience. As an additional check, I estimate separate models by poverty group, which are shown in Table A13. Whereas the main models account for heterogeneity in the year fixed effects, separate models by poverty group account for

[31] Table A12 shows the main effects and interactions that correspond to Table 9.

heterogeneity in all of the covariates that might otherwise be conflated with heterogeneity in the returns to experience. Here, I find that the split sample results are very similar to the pooled models that include year-by-poverty fixed effects.

Tables A14 and A15 show estimates of heterogeneity by school level (elementary, middle, high) and locale (urban, suburban, town/rural), respectively. As with the poverty models, my preferred specification includes year-by-level or year-by-locale fixed effects, though I also show the results with year fixed effects. Again, accounting for heterogeneity in the year fixed effects matters, particularly for the results by school level. In my preferred models, I find no evidence that the returns to experience vary by school level or locale. In each case, joint tests of the interaction terms are not statistically significant at conventional levels, though in some cases the standard errors on the interactions terms are quite large, particularly at higher experience levels.

Discussion

Increased recognition of the pivotal role of principals in driving school performance has spurred policy attention at the local, state, and federal levels to developing of high-quality school leadership. Yet we know little about what makes some principals more effective than others, and, by extension, the extent to which differences in principal quality are driven by on-the-job improvement.

My results demonstrate that principals become substantially more effective at raising student achievement as they gain experience. These returns to experience are largest for math achievement, where the difference between a principal's first and fifth year in the principalship is roughly 25% of a standard deviation in terms of the distribution of principal quality in Tennessee. The returns to principal experience for student achievement are corroborated by ratings from supervisors. Within-principal improvement in supervisor ratings is substantial—the average principal improves their average score by 0.20 SD between their first and second years in the principalship and an additional 0.25 SD between their second and sixth years.

An important addendum to these results is that the returns to experience for student achievement are driven by school-specific experience rather than total experience. Put another way, improvement in a principal's ability to raise test scores in one school does not carry over when they change schools. There are multiple potential explanations for this finding. The first is that a principal's skill acquisition over time (with respect to skills that lead to student test score growth) constitutes school-specific rather than general learning. Principals may be adapting to the unique needs of their school or forming relationships with teachers, students, and parents which ultimately leads to improvements in student learning. Such improvements in principals' practice may not help them in a new school context.

An alternative explanation—and one that I cannot definitively rule out given my data and method—is that these results reflect the benefits of stable leadership over time. More specifically, principals may not substantially improve (in terms of concrete changes in their practice) with additional experience in the school, but student achievement may increase simply because the principal has time to have an impact. Prior work, for instance, has pointed out that new principals inherit the school conditions shaped by their predecessor, such that any effects they might have will take time to manifest in student test score gains (e.g., Coelli & Green, 2012; Grissom et al., 2015). Without direct measures of principal skills and behaviors, it is very difficult to test which of these explanations holds. However, the improvement in supervisor ratings and the decrease in teacher turnover with experience are consistent with some amount of on-the-job learning. From a policy perspective, both explanations lead to the prescription that stable leadership is important for school performance.

These results have implications for policy and research. First, this study suggests that the increased investments to principal development and coaching are warranted, as the average principal improves quite substantially over time. To the extent that these supports also increase the likelihood that principals remain in their schools, the findings here suggest that increased stability in the principal's office is likely to benefit student learning. Unfortunately, a large number of principals in Tennessee and nationally have been in their schools for only a few years (Fuller & Young, 2009; Grissom et al., 2019), meaning that only a small percentage of schools are reaping the benefits of having a highly experienced principal. Second, in light of the finding that the returns to principal experience are largely school-specific, district administrators should consider policies that aim to reduce the shuffling of principals across schools. This is particularly relevant for larger urban districts, which have the highest rates of within-district principal transfers and more schools led by inexperienced leaders (Grissom et al., 2019). Finally, this study suggests that we need to think differently about the nature of principal effects on student outcomes. Previous work most often treats principal quality as a fixed quantity, whereas the results here demonstrate that a given principal's effectiveness varies substantially as they gain experience and change schools.

This study has some important limitations. Perhaps most importantly, estimating the returns to principal experience involves addressing several obstacles to identification using imperfect methods. As with any analysis using observational data, there is no guarantee that all confounding factors have been addressed. Of particular note in this study is that isolating the returns to experience from unobserved time-varying productivity trends requires imposing strong identification assumptions. While others have proposed approaches for navigating these obstacles in the context of teacher productivity (e.g., Papay & Kraft, 2015), my results suggest that these approaches may not be suitable for principals.

An additional issue is that the majority of principals I observe only worked in one school during the study period, and many are observed for a small number of years. While this is a reflection of the nature of the principal labor market, it has two implications for my findings. First, identification of the returns to experience are only based on principals who remain in the principalship. As principal experience increases, this subset of "stayers" becomes increasingly small, which raises the possibility that the estimates are not representative of the full sample. This is particularly concerning for school-specific experience, where my estimates could be driven by principals who are particularly well matched to their schools. That said, I do not find evidence of differences in the early-career returns to experience for principals who exit early, which suggests they would have improved at similar rates to those who stayed in the principalship longer.

The second implication is that the returns to total experience and school-specific experience are identified only from principals who changed schools. To the extent that this mobility is influenced by general or school-specific improvement, these findings may not generalize to the broader population of principals. Finally, because the data do not contain job-specific experience measures, I cannot estimate the returns to principal experience beyond the length of the panel—14 years in this case. Future work should seek to leverage increasingly available longitudinal datasets that span across many years and include job-specific measures of experience.

In sum, the findings of this study suggest that principals have substantial capacity for improvement, such that the quality of the school leadership could be increased through promoting leadership stability and professional growth. Future work should continue to focus on identifying ways to better support principals, particularly in disadvantaged or under-resourced schools. Additionally, this study raises important questions about the nature of principal improvement. What are the actual skills that principals build as they accumulate more experience as leaders? Why do some principals improve at greater rates than others? What training or preparation experiences lead to faster on-the-job learning? Answers to these questions will inform policies that can increase the quality of school leadership.

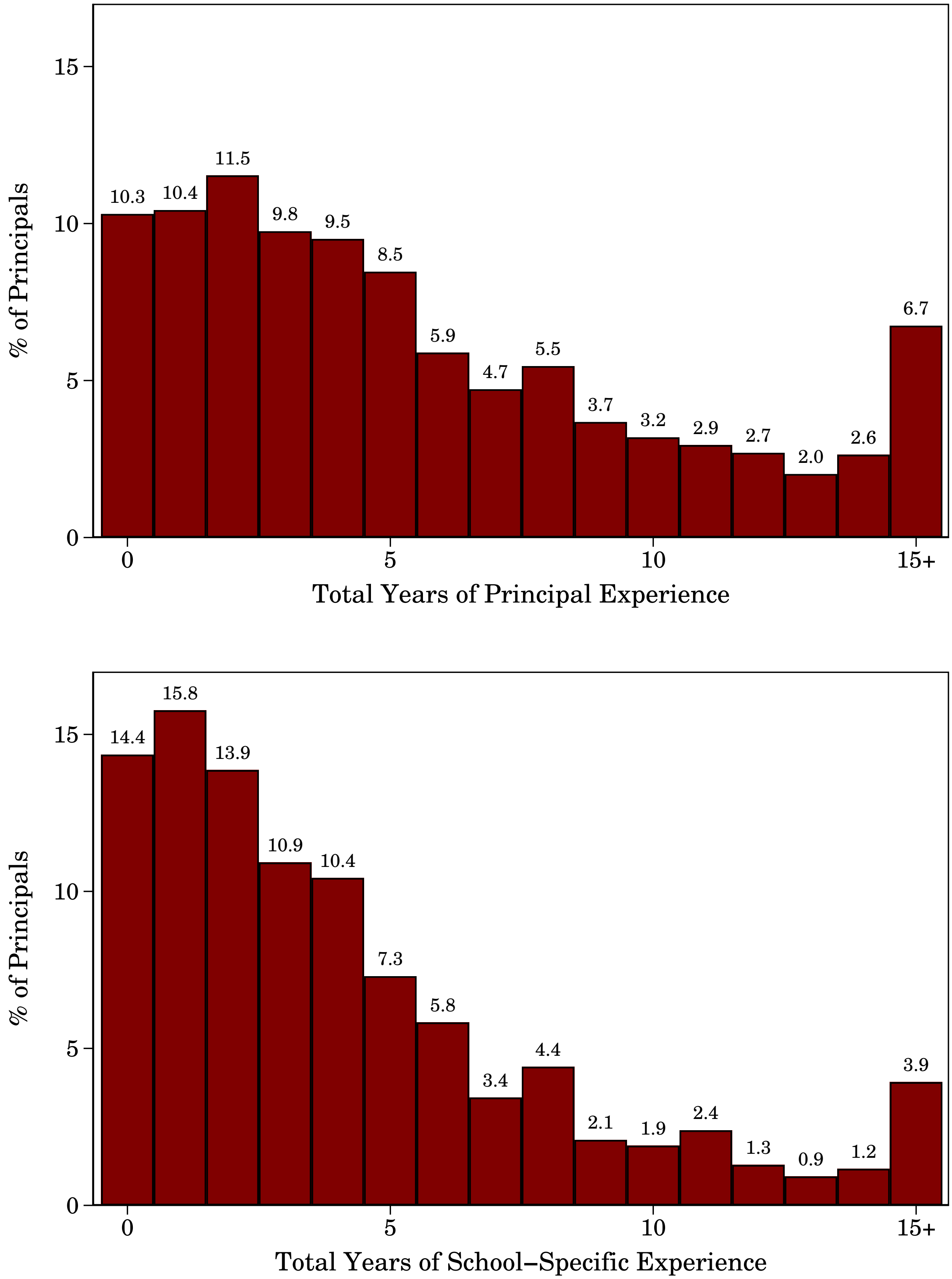

Figure 1. Distribution of Principal Experience in Tennessee (2017)

Notes: Zero experience refers to a principal in their first year as a principal. School-specific experience only counts years served as the principal.

Table 1
Descriptive Statistics

	Mean	SD	Min	Max	N
Principal Experience					
Total years	3.1	2.7	0	14	11577
0 years	0.17				11577
1 year	0.18				11577
2–3 years	0.28				11577
4–6 years	0.24				11577
7–9 years	0.10				11577
10–14 years	0.03				11577
Tenure in School					
Total years	2.7	2.6	0	14	14475
0 years	0.21				14475
1 year	0.20				14475
2–3 years	0.29				14475
4–6 years	0.21				14475
7–9 years	0.07				14475
10–14 years	0.02				14475
Principal Demographics					
Black	0.19				18079
Male	0.45				18079
Age	50.0	9.1	19	93	17788
Experience in TN system	22.6	9.4	0	66	17992
School Demographics					
Enrollment size (100s)	6.45	3.82	0.14	40.65	18044
Proportion FRPL	0.57	0.26	0.00	1.00	18017
Proportion Black	0.25	0.31	0.00	1.00	18017
Proportion Hispanic	0.06	0.09	0.00	0.74	18017
Proportion Gifted	0.02	0.03	0.00	0.56	18017
Proportion SPED	0.15	0.08	0.00	1.00	18017
School Level					
Elementary	0.59				17992
Middle	0.19				17992
High	0.18				17992
Other	0.05				17992
School Locale					
Urban	0.31				17995
Suburban	0.15				17995
Town	0.16				17995
Rural	0.39				17995

Notes: Includes principals in Tennessee from 2006–07 to 2016–17. Unit of observation is principal-by-year.

Table 2

The Returns to Principal Experience (Student Achievement)

	IVM			Cen. Growth			2-Stage		
	Math	ELA	Sci	Math	ELA	Sci	Math	ELA	Sci
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Total Principal Experience									
0 years (base)									
1 year	0.017***	0.005	0.010*	0.013**	0.003	0.003	0.009*	0.005*	0.004
	(0.006)	(0.004)	(0.005)	(0.006)	(0.004)	(0.005)	(0.005)	(0.003)	(0.004)
2 years	0.023**	0.005	0.024***	0.014	0.000	0.011	0.007	0.004	0.012**
	(0.010)	(0.007)	(0.008)	(0.009)	(0.006)	(0.008)	(0.006)	(0.004)	(0.006)
3 years	0.041***	0.012	0.039***	0.028**	0.005	0.022*	0.018**	0.011**	0.023***
	(0.013)	(0.010)	(0.012)	(0.012)	(0.009)	(0.012)	(0.007)	(0.005)	(0.007)
4–6 years	0.065***	0.015	0.048***				0.028***	0.013**	0.027***
	(0.019)	(0.013)	(0.017)				(0.009)	(0.006)	(0.009)
7–9 years	0.066**	0.015	0.059**				0.010	0.013	0.030**
	(0.027)	(0.020)	(0.024)				(0.013)	(0.008)	(0.013)
10–14 years	0.099***	0.024	0.071**				0.023	0.021	0.031
	(0.037)	(0.025)	(0.033)				(0.018)	(0.015)	(0.019)
4 years				0.048***	0.009	0.027*			
				(0.015)	(0.011)	(0.015)			
5 years				0.038*	-0.001	0.022			
				(0.020)	(0.015)	(0.019)			
N	3034743	3328312	3029986	3034743	3328312	2819174	3034743	3328312	2819174
R^2	0.301	0.319	0.341	0.301	0.319	0.339	0.302	0.319	0.339
Joint Test of Year FE	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Notes: Standard errors clustered by principal-school shown in parentheses. The dependent variables are standardized student achievement scores. The omitted category is principals who have zero years of experience. Models include fixed effects for principal, school, year, and grade. Covariates include student characteristics and time-varying school characteristics. "Joint Test of Year FE" shows the p-value from an F-test that the estimated year fixed effects are jointly zero. For the 2-stage model, this test refers to the year FE from the first-stage model. IVM = Indicator Variable Model. N refers to the total number of student-by-year observations.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3

The Returns to Principal Experience Using Modified Year Bins

	Math (1)	ELA (2)	Sci (3)
Total Principal Experience			
0 years (base)			
1 year	0.021***	0.007**	0.014**
	(0.006)	(0.004)	(0.005)
2 years	0.031***	0.009	0.033***
	(0.008)	(0.006)	(0.009)
3 years	0.053***	0.018**	0.054***
	(0.012)	(0.008)	(0.013)
4 years	0.083***	0.026***	0.071***
	(0.015)	(0.010)	(0.017)
5 years	0.082***	0.020	0.074***
	(0.018)	(0.012)	(0.021)
6 years	0.092***	0.029**	0.088***
	(0.021)	(0.014)	(0.024)
7 years	0.096***	0.027	0.110***
	(0.025)	(0.017)	(0.029)
8 years	0.098***	0.038**	0.105***
	(0.028)	(0.019)	(0.033)
9 years	0.105***	0.030	0.104***
	(0.032)	(0.022)	(0.037)
10 years	0.141***	0.044*	0.116**
	(0.037)	(0.026)	(0.045)
11 years	0.132***	0.031	0.148***
	(0.044)	(0.030)	(0.046)
12 years	0.179***	0.085**	0.183***
	(0.045)	(0.036)	(0.056)
13 years	0.230***	0.120	0.194***
	(0.080)	(0.109)	(0.060)
14 years	0.075	0.034	0.185**
	(0.086)	(0.055)	(0.083)
N	3034743	3328312	2819174
R^2	0.301	0.319	0.339

Notes: Standard errors clustered by principal-school shown in parentheses. The dependent variables are standardized student achievement scores. Models include fixed effects for principal, school, year, and grade. Covariates include student characteristics and time-varying school characteristics. The year fixed effects are replaced with the following year bins: (math) 2007, 2008–2009, 2010, 2011, 2012–2013, 2014–2015, 2016–2017; (ELA) 2007–2008, 2009, 2010–2011, 2012–2013, 2014, 2015, 2016, 2017; (Science) 2007, 2008, 2009, 2010–2011, 2012–2013, 2014, 2015–2016, 2017. N refers to the total number of student-by-year observations.

* p < 0.1, ** p < 0.05, *** p < 0.01.

Table 4
The Returns to Principal Experience (Supervisor Ratings)

	IVM (1)	Cen. Growth (2)	2-Stage (3)
Total Principal Experience			
0 years (base)			
1 year	0.208***	0.218***	0.215***
	(0.039)	(0.039)	(0.034)
2 years	0.319***	0.340***	0.333***
	(0.064)	(0.053)	(0.054)
3 years	0.367***	0.398***	0.387***
	(0.071)	(0.065)	(0.065)
4–6 years	0.462***		
	(0.092)		
7–9 years	0.443***		
	(0.123)		
10–14 years	0.448**		
	(0.187)		
4 years		0.487***	0.471***
		(0.081)	(0.081)
5 years		0.534***	0.510***
		(0.108)	(0.099)
6 years			0.515***
			(0.105)
7 years			0.501***
			(0.132)
8 years			0.501***
			(0.147)
9 years			0.475***
			(0.170)
10 years			0.476**
			(0.188)
11 years			0.600***
			(0.218)
12 years			0.552**
			(0.249)
13 years			0.447*
			(0.270)
14 years			0.157
			(0.386)
N	6924	6986	6986
R^2	0.710	0.712	0.713
Joint Test of Year FE	0.9839	0.9797	0.9886

Notes: Standard errors clustered by school district shown in parentheses. The dependent variables is a principal's average rating from their supervisor, standardized within year. Models include fixed effects for principal, school, and year. Covariates include time-varying school characteristics. "Joint Test of Year FE" shows the p-value from an F-test that the estimated year fixed effects are jointly zero. For the 2-stage model, this test refers to the year FE from the first-stage model. IVM = Indicator Variable Model. N refers to the total number of principal-by-year observations.
* p < 0.1, ** p < 0.05, *** p < 0.01.

Table 5

The Returns to Total and School-Specific Principal Experience

	Math		ELA		Sci	
	(1)	(2)	(3)	(4)	(5)	(6)
Total Principal Experience						
0 years (base)						
1 year		0.012		-0.002		0.004
		(0.011)		(0.007)		(0.009)
2 years		-0.002		-0.019*		0.001
		(0.017)		(0.011)		(0.013)
3 years		0.009		-0.016		0.012
		(0.022)		(0.015)		(0.018)
4–6 years		0.030		-0.014		0.011
		(0.027)		(0.019)		(0.024)
7–9 years		0.008		-0.036		0.008
		(0.037)		(0.026)		(0.031)
10–14 years		0.031		-0.039		0.010
		(0.048)		(0.033)		(0.039)
School-Specific Experience						
0 years (base)						
1 year	0.018***	0.008	0.009**	0.011*	0.012**	0.009
	(0.006)	(0.010)	(0.004)	(0.006)	(0.005)	(0.009)
2 years	0.032***	0.034**	0.016**	0.032***	0.031***	0.030**
	(0.010)	(0.016)	(0.007)	(0.010)	(0.008)	(0.012)
3 years	0.053***	0.045**	0.027***	0.040***	0.046***	0.037**
	(0.014)	(0.022)	(0.009)	(0.015)	(0.013)	(0.018)
4–6 years	0.075***	0.051*	0.032**	0.042**	0.061***	0.052**
	(0.019)	(0.027)	(0.013)	(0.019)	(0.017)	(0.025)
7–9 years	0.095***	0.090**	0.051**	0.080***	0.079***	0.074**
	(0.028)	(0.038)	(0.020)	(0.026)	(0.026)	(0.033)
10–14 years	0.129***	0.104*	0.069**	0.099***	0.095**	0.089*
	(0.042)	(0.054)	(0.029)	(0.037)	(0.038)	(0.046)
N	3034743	3034743	3328312	3328312	3029986	3029986
R^2	0.301	0.301	0.319	0.319	0.341	0.341

Notes: Standard errors clustered by principal-school shown in parentheses. The dependent variables are standardized student achievement scores. The omitted category is principals who have zero years of experience. Models include fixed effects for principal, school, year, and grade. Covariates include student characteristics and time-varying school characteristics.

* p < 0.1, ** p < 0.05, *** p < 0.01.

Table 6

The Returns to Principal Experience (Teacher Turnover)

	All Turnover			Transfer			Exit			Position Change		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Total Principal Experience												
0 years (base)												
1 year	-0.003		0.000	-0.002		0.005	-0.002		-0.004	0.001		-0.002
	(0.004)		(0.008)	(0.003)		(0.007)	(0.003)		(0.006)	(0.001)		(0.002)
2 years	-0.005		-0.005	-0.005		0.007	-0.001		-0.014*	0.000		-0.002
	(0.006)		(0.011)	(0.005)		(0.009)	(0.004)		(0.007)	(0.001)		(0.002)
3 years	-0.009		0.000	-0.006		0.010	-0.007		-0.011	0.001		-0.003
	(0.008)		(0.013)	(0.006)		(0.011)	(0.005)		(0.009)	(0.002)		(0.003)
4–6 years	-0.012		0.010	-0.004		0.024	-0.011		-0.013	-0.000		-0.003
	(0.011)		(0.019)	(0.009)		(0.016)	(0.008)		(0.013)	(0.003)		(0.004)
7–9 years	-0.012		0.008	-0.005		0.030	-0.012		-0.023	0.000		-0.003
	(0.017)		(0.026)	(0.013)		(0.023)	(0.012)		(0.019)	(0.004)		(0.006)
10–14 years	-0.022		0.012	-0.011		0.042	-0.020		-0.033	0.002		-0.000
	(0.026)		(0.039)	(0.020)		(0.034)	(0.018)		(0.027)	(0.005)		(0.008)
School-Specific Experience												
0 years (base)												
1 year		-0.004	-0.004		-0.006*	-0.010		-0.000	0.003		0.001	0.003**
		(0.004)	(0.008)		(0.003)	(0.007)		(0.003)	(0.005)		(0.001)	(0.002)
2 years		-0.007	-0.002		-0.012**	-0.017*		0.004	0.015**		0.001	0.002
		(0.006)	(0.011)		(0.005)	(0.009)		(0.004)	(0.007)		(0.001)	(0.002)
3 years		-0.014*	-0.014		-0.015**	-0.023**		-0.003	0.005		0.002	0.004
		(0.009)	(0.014)		(0.007)	(0.012)		(0.006)	(0.009)		(0.002)	(0.003)
4–6 years		-0.023*	-0.031		-0.022**	-0.040**		-0.006	0.003		0.001	0.004
		(0.012)	(0.021)		(0.010)	(0.017)		(0.008)	(0.013)		(0.003)	(0.004)
7–9 years		-0.022	-0.027		-0.028*	-0.050*		0.000	0.019		0.001	0.003
		(0.020)	(0.030)		(0.016)	(0.026)		(0.013)	(0.020)		(0.004)	(0.006)
10–14 years		-0.044	-0.053		-0.053**	-0.085**		-0.004	0.022		0.003	0.003
		(0.029)	(0.044)		(0.022)	(0.038)		(0.020)	(0.029)		(0.006)	(0.008)
N	350843	350843	350843	320877	320877	320877	318333	318333	318333	296924	296924	296924
R^2	0.073	0.073	0.073	0.085	0.085	0.085	0.068	0.069	0.069	0.025	0.025	0.025

Notes: Standard errors clustered by principal-school shown in parentheses. The dependent variable is a binary indicator for whether the teacher left their position in the current year, with the type of turnover listed in the column header. Models include fixed effects for principal, school, and year. Models include teacher and school characteristics, and an indicator for whether the principal left the school at the end of the year.

* p < 0.1, ** p < 0.05, *** p < 0.01.

Table 7

The Returns to Principal Experience for Turnover of Hired and Inherited Teachers (Marginal Effects)

	All Turnover (1)	Transfer (2)	Exit (3)	Position Change (4)
School-Specific Experience				
Principal Inherited Teacher				
0 years (base)				
1 year	-0.008*	-0.007**	-0.002	0.000
	(0.004)	(0.003)	(0.003)	(0.001)
2 years	-0.014**	-0.016***	0.000	0.000
	(0.006)	(0.005)	(0.004)	(0.001)
3 years	-0.025***	-0.022***	-0.007	-0.000
	(0.009)	(0.007)	(0.006)	(0.002)
4–6 years	-0.030**	-0.025**	-0.010	0.000
	(0.013)	(0.010)	(0.008)	(0.003)
7–9 years	-0.032	-0.034**	-0.003	0.001
	(0.021)	(0.016)	(0.014)	(0.004)
10–14 years	-0.024	-0.035	0.006	0.001
	(0.031)	(0.023)	(0.022)	(0.007)
Principal Hired Teacher				
0 years (base)				
1 year	-0.023***	-0.021***	-0.010**	0.004**
	(0.006)	(0.005)	(0.005)	(0.002)
2 years	-0.030***	-0.029***	-0.010*	0.000
	(0.008)	(0.007)	(0.006)	(0.002)
3 years	-0.038***	-0.033***	-0.018***	0.003
	(0.010)	(0.008)	(0.007)	(0.002)
4–6 years	-0.055***	-0.046***	-0.025***	0.001
	(0.013)	(0.011)	(0.009)	(0.003)
7–9 years	-0.053**	-0.050***	-0.018	0.000
	(0.021)	(0.017)	(0.014)	(0.005)
10–14 years	-0.085***	-0.083***	-0.026	0.002
	(0.030)	(0.023)	(0.023)	(0.007)
N	350843	320877	318333	296924
R^2	0.078	0.089	0.070	0.026

Notes: Standard errors clustered by principal-school shown in parentheses. The dependent variable is a binary indicator for whether the teacher left their position in the current year, with the type of turnover listed in the column header. Models include fixed effects for principal, school, and year. Models include teacher and school characteristics, an indicator for whether the principal left the school at the end of the year, and an indicator for being a "hired" teacher. "Inherited" teachers are those who have more school-specific experience than the principal. "hired" teachers are those with the same or less school-specific experience than the principal. The model results showing the main effect and interactions are in Table A9.

* p < 0.1, ** p < 0.05, *** p < 0.01.

Table 8

The Returns to Principal Experience for Turnover of Effective and Ineffective Teachers (Marginal Effects)

	All Turnover (1)	Transfer (2)	Exit (3)	Position Change (4)
School-Specific Experience				
Low Value-Added Teacher				
0 years (base)				
1 year	-0.014*	-0.014**	-0.003	0.000
	(0.008)	(0.007)	(0.005)	(0.002)
2 years	-0.019*	-0.018**	-0.003	-0.001
	(0.010)	(0.009)	(0.006)	(0.002)
3 years	-0.023*	-0.027***	0.000	-0.001
	(0.012)	(0.010)	(0.008)	(0.003)
4–6 years	-0.028*	-0.027**	-0.007	-0.001
	(0.016)	(0.014)	(0.010)	(0.004)
7–9 years	-0.030	-0.036	0.004	-0.004
	(0.027)	(0.022)	(0.017)	(0.006)
10–14 years	-0.040	-0.062*	-0.001	0.013
	(0.046)	(0.034)	(0.031)	(0.014)
Middle Value-Added Teacher				
0 years (base)				
1 year	-0.002	-0.004	0.003	-0.001
	(0.005)	(0.004)	(0.003)	(0.001)
2 years	-0.007	-0.011*	0.005	-0.002
	(0.008)	(0.006)	(0.005)	(0.002)
3 years	-0.013	-0.013	-0.000	-0.002
	(0.010)	(0.009)	(0.007)	(0.003)
4–6 years	-0.024*	-0.023*	-0.002	-0.005
	(0.015)	(0.012)	(0.009)	(0.003)
7–9 years	-0.023	-0.026	0.004	-0.007
	(0.025)	(0.021)	(0.015)	(0.006)
10–14 years	-0.057*	-0.056**	-0.000	-0.014
	(0.033)	(0.028)	(0.024)	(0.009)
High Value-Added Teacher				
0 years (base)				
1 year	-0.022***	-0.018***	-0.006	0.001
	(0.007)	(0.006)	(0.004)	(0.002)
2 years	-0.021**	-0.020**	0.001	-0.005**
	(0.010)	(0.008)	(0.006)	(0.002)
3 years	-0.039***	-0.032***	-0.009	-0.005
	(0.012)	(0.010)	(0.008)	(0.003)
4–6 years	-0.043***	-0.031**	-0.011	-0.008**
	(0.016)	(0.013)	(0.010)	(0.004)
7–9 years	-0.042	-0.036	0.001	-0.014**
	(0.026)	(0.022)	(0.016)	(0.007)
10–14 years	-0.063	-0.074**	0.006	-0.006
	(0.039)	(0.031)	(0.025)	(0.011)
N	164585	154140	147697	140613
R^2	0.089	0.105	0.075	0.039

Notes: Standard errors clustered by principal-school shown in parentheses. The dependent variable is a binary indicator for whether the teacher left their position in the current year, with the type of turnover listed in the column header. Models include fixed effects for principal, school, and year. Models include teacher and school characteristics, an indicator for whether the principal left the school at the end of the year, and indicators for being high and low value-added. Value-added categories correspond to the top 20%, middle 60%, and bottom 20% of the statewide distribution. The model results showing the main effect and interactions are in Table A10.
* p < 0.1, ** p < 0.05, *** p < 0.01.

Table 9

Heterogeneity in the Returns to Experience by School Poverty (Marginal Effects)

	Math (SD) (1)	ELA (SD) (2)	Sci (SD) (3)
School-Specific Experience			
Low-Poverty School			
0 years (base)			
1 year	0.002	-0.021*	-0.008
	(0.017)	(0.012)	(0.015)
2 years	0.005	-0.014	0.014
	(0.027)	(0.022)	(0.027)
3 years	0.028	-0.028	0.019
	(0.039)	(0.032)	(0.041)
4–6 years	0.064	-0.038	0.032
	(0.053)	(0.043)	(0.055)
7–9 years	0.054	-0.112	0.023
	(0.082)	(0.075)	(0.101)
10–14 years	0.052	-0.114	0.040
	(0.110)	(0.097)	(0.130)
Medium-Poverty School			
0 years (base)			
1 year	0.009	0.003	0.002
	(0.008)	(0.005)	(0.007)
2 years	0.023*	0.006	0.011
	(0.012)	(0.008)	(0.011)
3 years	0.035**	0.012	0.016
	(0.018)	(0.011)	(0.017)
4–6 years	0.053**	0.011	0.020
	(0.024)	(0.016)	(0.024)
7–9 years	0.079**	0.032	0.020
	(0.036)	(0.024)	(0.034)
10–14 years	0.120**	0.058	0.007
	(0.055)	(0.036)	(0.051)
High-Poverty School			
0 years (base)			
1 year	0.042***	0.036***	0.032***
	(0.012)	(0.007)	(0.010)
2 years	0.058***	0.048***	0.060***
	(0.019)	(0.012)	(0.016)
3 years	0.101***	0.082***	0.102***
	(0.027)	(0.017)	(0.023)
4–6 years	0.123***	0.103***	0.122***
	(0.037)	(0.024)	(0.033)
7–9 years	0.124**	0.151***	0.184***
	(0.053)	(0.037)	(0.052)
10–14 years	0.122	0.123*	0.262***
	(0.079)	(0.064)	(0.067)
N	3034717	3328229	2819110
R^2	0.302	0.319	0.339

Notes: Standard errors clustered by principal-school shown in parentheses. In each column, I show the estimated marginal effects for school-specific principal experience (i.e., main effect plus interaction term) by school-poverty level. Models include fixed effects for principal, school, year-by-poverty group, and grade. Covariates include student characteristics and time-varying school characteristics. The model results showing the main effect and interactions are in Table A12. High, medium, and low-poverty refer to the percentage of students in the school who qualify for free/reduced price lunch: 0-30%, 30-70%, 70-100%. These categories are time-invariant and absorbed by the school fixed effect.

* p < 0.1, ** p < 0.05, *** p < 0.01.

References

Adnot, M., Dee, T., Katz, V., & Wyckoff, J. (2017). Teacher Turnover, Teacher Quality, and Student Achievement in DCPS. *Educational Evaluation and Policy Analysis*, *39*(1), 1–23.

Bartanen, B., & Grissom, J. A. (2019). *School Principal Race and the Hiring and Retention of Racially Diverse Teachers.*

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

Bastian, K. C., & Henry, G. T. (2015). The Apprentice: Pathways to the Principalship and Student Achievement. *Educational Administration Quarterly*, *51*(4), 600–639.

Becker, G. S. (1962). Investment in Human Capital : A Theoretical Analysis. *Journal of Political Economy*, *70*(5), 9–49.

Béteille, T., Kalogrides, D., & Loeb, S. (2012). Stepping stones: Principal career paths and school outcomes. *Social Science Research*, *41*(4), 904–919.

Boyd, D., Grossman, P., 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.*

Brewer, D. J. (1993). Principals and Student Outcomes: Evidence From U.S. High Schools. *Economics of Education Review*, *12*(4), 281–292.

Brown, K. M., & Wynn, S. R. (2009). Finding, Supporting, and Keeping: The Role of the Principal in Teacher Retention Issues. *Leadership and Policy in Schools*, *8*(1), 37–63.

Burkhauser, S. (2017). How Much Do School Principals Matter When It Comes to Teacher Working Conditions? *Educational Evaluation and Policy Analysis*, *39*(1), 126–145.

Chetty, R., Friedman, J., & Rockoff, J. (2014). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, *104*(9), 2633–2679.

Chiang, H., Lipscomb, S., & Gill, B. (2016). Is School Value Added Indicative of Principal

Quality? *Education Finance and Policy*, *11*(3), 283–309.

Clark, D., Martorell, P., & Rockoff, J. (2009). *School Principals and School Performance.* Washington, D.C.

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

Dhuey, E., & Smith, J. (2018). How school principals influence student learning. *Empirical Economics*, *54*, 851–882.

Eberts, R. W., & Stone, J. A. (1988). Student achievement in public schools: Do principals make a difference? *Economics of Education Review*, *7*(3), 291–299.

Fuller, E., & Young, M. D. (2009). *Tenure and Retention of Newly Hired Principals in Texas* (Tech. Rep.). University Council for Educational Administration.

Gates, S. M., Ringel, J. S., Santibañez, L., Chung, C. H., & Ross, K. E. (2003). *Who is Leading our Schools? An Overview of School Administrators and Their Careers* (Tech. Rep.). Santa Monica: RAND Corporartion.

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. (2019a). Principal Effectiveness and Principal Turnover. *Education Finance and Policy*, *14*(3), 355–382.

Grissom, J. A., & Bartanen, B. (2019b). Strategic Retention: Principal Effectiveness and Teacher Turnover in Multiple-Measure Teacher Evaluation Systems. *American Educational Research Journal*, *56*(2), 514–555.

Grissom, J. A., Bartanen, B., & Mitani, H. (2019). Principal Sorting and the Distribution of Principal Quality. *AERA Open*, *5*(2), 1–21.

Grissom, J. A., Blissett, R. S. L., & Mitani, H. (2018). Evaluating School Principals: Supervisor Ratings of Principal Practice and Principal Job Performance. *Educational Evaluation and Policy Analysis*, *40*(3), 446–472.

Grissom, J. A., Kalogrides, D., & Loeb, S. (2015). Using Student Test Scores to Measure Principal Performance. *Educational Evaluation and Policy Analysis*, *37*(1), 3–28.

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics*, *95*(7-8), 798–812.

Horng, E., Klasik, D., & Loeb, S. (2010). Principal's Time Use and School Effectiveness. *American Journal of Education*, *116*.

Jacob, B. (2011). Do Principals Fire the Worst Teachers? *Educational Evaluation and Policy Analysis*, *33*(January), 403–434.

Johnson, S. M., & Birkeland, S. E. (2003). Pursuing a ”Sense of Success”: New Teachers Explain Their Career Decisions. *American Educational Research Journal*, *40*(3), 581–617.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, *27*(6), 615–631.

Kraft, M. A., Marinell, W. H., & Yee, D. (2016). School Organizational Contexts, Teacher Turnover, and Student Achievement: Evidence from Panel Data. *American Educational Research Journal*, *53*(5), 1411–1449.

Kraft, M. A., & Papay, J. P. (2014). Can professional environments in schools promote teacher development? Explaining heterogeneity in returns to teaching experience. *Educational Evaluation and Policy Analysis*, *30*(4), 476–500.

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.

Ladd, H. F., & Sorensen, L. C. (2017). Returns to Teacher Experience: Student Achievement and Motivation in Middle School. *Education Finance and Policy*, *12*(2).

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

Loeb, S., Kalogrides, D., & Horng, E. L. (2010). Principal Preferences and the Uneven

Distribution of Principals Across Schools. *Educational Evaluation and Policy Analysis*, *32*(2), 205–229.

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating Models for Teacher Accountability* (Tech. Rep.). Santa Monica, CA: RAND Corporation.

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

Papay, J. P., & Kraft, M. A. (2015). Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement. *Journal of Public Economics*, *130*, 105–119.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, *73*(2), 417–458.

Rockoff, J. E. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *Papers and Proceedings of the One Hundered Sixteenth Annual Meetings of the American Economic Assocation San Diego, CA*, *94*(2), 247–252.

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How Teacher Turnover Harms Student Achievement. *American Educational Research Journal*, *50*(1), 4–36.

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.

Urick, A., & Bowers, A. J. (2014). What Are the Different Types of Principals Across the United States? A Latent Class Analysis of Principal Perception of Leadership. *Educational Administration Quarterly*, *50*(1), 96–134.

Appendix

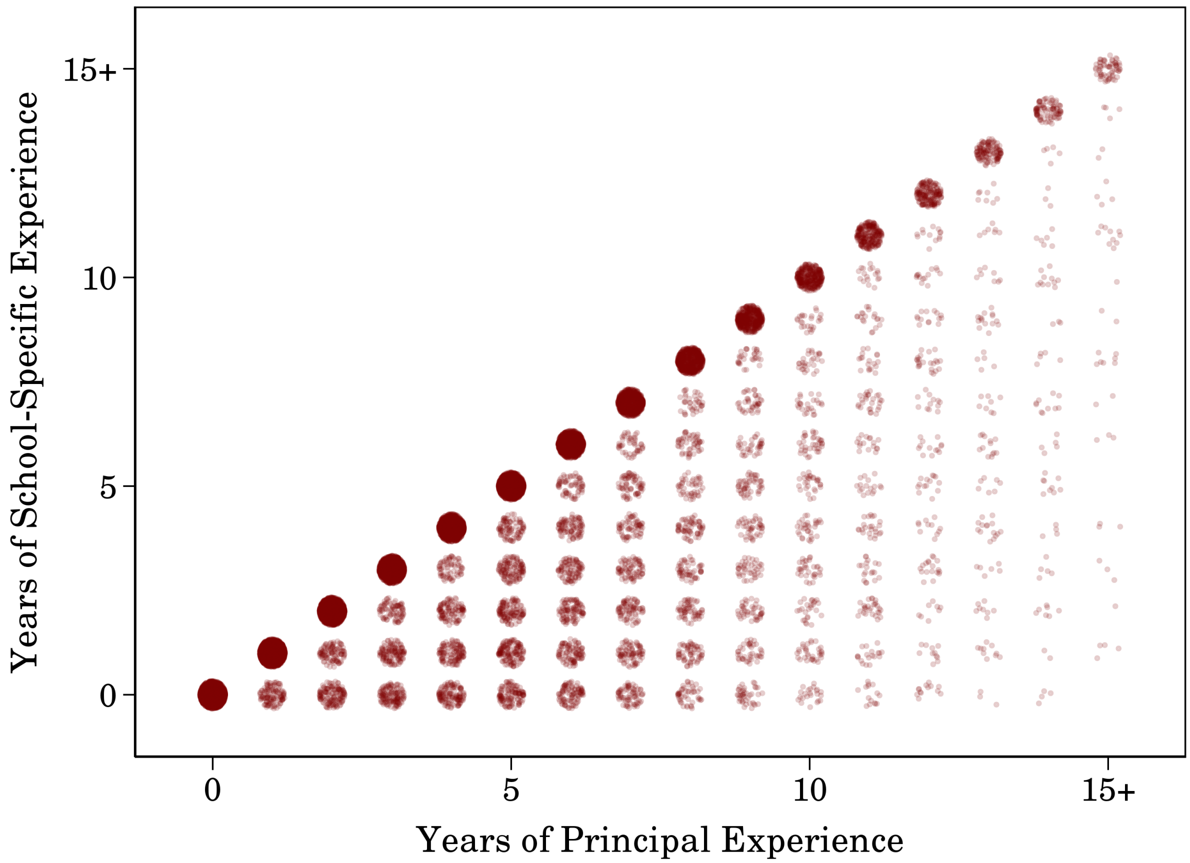

Figure A1. The Distribution of Total vs. School-Specific Principal Experience

Notes: Experience and tenure are discrete values from 0 to 15+. Random jitter added to show density.

Table A1

Proportion of Principals with Observed Experience and Tenure by Year

Year	# of Principals	Experience is Non-missing	Tenure is Non-missing
2007	1620	0.29	0.50
2008	1643	0.38	0.59
2009	1651	0.47	0.67
2010	1677	0.55	0.74
2011	1696	0.61	0.79
2012	1703	0.67	0.83
2013	1698	0.72	0.87
2014	1707	0.75	0.88
2015	1679	0.79	0.91
2016	1681	0.82	0.92
2017	1670	0.84	0.93

Table A2

Returns to Principal Experience from Discontinuous Career Model

	Math (1)	ELA (2)	Sci (3)	Sup. Ratings (4)
Total Principal Experience				
0 years (base)				
1 year	0.102***	0.024**	0.023	0.341
	(0.032)	(0.012)	(0.016)	(0.247)
2 years	0.192***	0.042*	0.051	0.587
	(0.064)	(0.024)	(0.032)	(0.506)
3 years	0.294***	0.067*	0.080*	0.768
	(0.096)	(0.035)	(0.048)	(0.758)
4 years	0.400***	0.088*	0.104	0.980
	(0.128)	(0.047)	(0.064)	(1.013)
5 years	0.477***	0.098*	0.116	1.147
	(0.160)	(0.059)	(0.080)	(1.257)
6 years	0.572***	0.122*	0.135	1.282
	(0.191)	(0.070)	(0.096)	(1.499)
7 years	0.658***	0.137*	0.163	1.396
	(0.224)	(0.082)	(0.112)	(1.764)
8 years	0.740***	0.160*	0.166	1.526
	(0.256)	(0.093)	(0.128)	(2.007)
9 years	0.820***	0.168	0.174	1.629
	(0.287)	(0.105)	(0.143)	(2.266)
10 years	0.942***	0.197*	0.188	1.753
	(0.320)	(0.117)	(0.161)	(2.512)
11 years	1.013***	0.203	0.234	2.004
	(0.352)	(0.129)	(0.176)	(2.769)
12 years	1.138***	0.270*	0.276	2.086
	(0.384)	(0.141)	(0.193)	(3.051)
13 years	1.265***	0.298	0.266	2.100
	(0.420)	(0.182)	(0.211)	(3.337)
14 years	1.195***	0.251	0.292	1.931
	(0.452)	(0.169)	(0.231)	(3.580)
N	3262309	3583834	3029986	6986
R^2	0.303	0.321	0.341	0.713

Notes: Standard errors clustered by principal-school (district in column 4) shown in parentheses. The dependent variables is listed in the column header. Models include fixed effects for principal, school, and year. Columns 1–3 also include fixed effects for grade. Covariates include student characteristics (columns 1–3) and time-varying school characteristics.

* p < 0.1, ** p < 0.05, *** p < 0.01.

Table A3

The Returns to Principal Experience Including Prior Test Scores

	Prior-year Score			Prior-school Score		
	Math (1)	ELA (2)	Sci (3)	Math (4)	ELA (5)	Sci (6)
Total Principal Experience						
0 years (base)						
1 year	0.017**	0.005	0.009	0.023***	0.003	-0.001
	(0.007)	(0.004)	(0.006)	(0.009)	(0.005)	(0.008)
2 years	0.021**	0.009	0.025***	0.030**	0.008	0.007
	(0.010)	(0.006)	(0.009)	(0.013)	(0.008)	(0.012)
3 years	0.035**	0.013	0.030**	0.051***	0.010	0.004
	(0.014)	(0.009)	(0.013)	(0.018)	(0.011)	(0.018)
4–6 years	0.047**	0.012	0.041**	0.064**	0.012	0.009
	(0.021)	(0.013)	(0.019)	(0.025)	(0.016)	(0.025)
7–9 years	0.060**	0.012	0.057**	0.075**	0.011	0.007
	(0.030)	(0.020)	(0.028)	(0.036)	(0.024)	(0.034)
10–14 years	0.098**	0.031	0.064	0.135***	0.026	-0.025
	(0.040)	(0.028)	(0.039)	(0.051)	(0.031)	(0.050)
N	2320609	2612118	2106567	1477933	1728090	1295608
R^2	0.579	0.625	0.600	0.549	0.607	0.573
Joint Test of Year FE	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Notes: Standard errors clustered by principal-school shown in parentheses. The dependent variables are standardized student achievement scores. The omitted category is principals who have zero years of experience. Models include fixed effects for principal, school, year, and grade. Covariates include student characteristics and time-varying school characteristics. "Joint Test of Year FE" shows the p-value from an F-test that the estimated year fixed effects are jointly zero.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A4

The Returns to Principal Experience Using Narrow Experience Bins

	Math			ELA			Sci		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Total Principal Experience									
0 years (base)									
1 year	0.021***		0.014	0.013***		0.003	0.004		-0.004
	(0.007)		(0.012)	(0.005)		(0.007)	(0.007)		(0.011)
2 years	0.031**		0.001	0.020**		-0.008	0.014		-0.015
	(0.012)		(0.019)	(0.009)		(0.012)	(0.011)		(0.015)
3 years	0.053***		0.012	0.035***		-0.000	0.026*		-0.008
	(0.018)		(0.026)	(0.012)		(0.017)	(0.016)		(0.021)
4–5 years	0.082***		0.040	0.048***		0.013	0.031		-0.010
	(0.024)		(0.032)	(0.017)		(0.021)	(0.022)		(0.028)
6–7 years	0.093***		0.019	0.064***		0.005	0.037		-0.019
	(0.032)		(0.044)	(0.023)		(0.028)	(0.030)		(0.035)
8–9 years	0.099**		0.019	0.081***		0.011	0.026		-0.035
	(0.042)		(0.055)	(0.029)		(0.035)	(0.038)		(0.044)
10–11 years	0.136***		0.037	0.096***		0.012	0.032		-0.033
	(0.052)		(0.065)	(0.037)		(0.047)	(0.051)		(0.061)
12–14 years	0.172***		0.075	0.153***		0.059	0.068		-0.041
	(0.067)		(0.088)	(0.053)		(0.073)	(0.065)		(0.074)
School-Specific Experience									
0 years (base)									
1 year		0.024***	0.013		0.017***	0.015**		0.010	0.012
		(0.007)	(0.011)		(0.005)	(0.007)		(0.007)	(0.010)
2 years		0.043***	0.043**		0.033***	0.039***		0.028**	0.039***
		(0.012)	(0.017)		(0.008)	(0.011)		(0.011)	(0.014)
3 years		0.069***	0.060**		0.052***	0.052***		0.043***	0.048**
		(0.017)	(0.025)		(0.012)	(0.016)		(0.016)	(0.020)
4–5 years		0.096***	0.066**		0.066***	0.055***		0.053**	0.059**
		(0.024)	(0.032)		(0.017)	(0.021)		(0.023)	(0.028)
6–7 years		0.129***	0.116***		0.097***	0.092***		0.070**	0.083**
		(0.032)	(0.044)		(0.023)	(0.028)		(0.031)	(0.036)
8–9 years		0.136***	0.123**		0.121***	0.111***		0.063	0.088*
		(0.043)	(0.057)		(0.029)	(0.035)		(0.041)	(0.046)
10–11 years		0.191***	0.160**		0.151***	0.139***		0.070	0.091
		(0.057)	(0.071)		(0.037)	(0.048)		(0.054)	(0.066)
12–14 years		0.175**	0.110		0.192***	0.135*		0.185**	0.212**
		(0.072)	(0.099)		(0.045)	(0.076)		(0.076)	(0.086)
N	3034743	3034743	3034743	3328312	3328312	3328312	2819174	2819174	2819174
R^2	0.301	0.301	0.301	0.319	0.319	0.319	0.339	0.339	0.339

Notes: Standard errors clustered by principal-school shown in parentheses. The dependent variables are standardized student achievement scores. The omitted category is principals who have zero years of experience. Models include fixed effects for principal, school, year, and grade. Covariates include student characteristics and time-varying school characteristics.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A5

Year Fixed Effects Estimates from Achievement Models

	IVM			Cen. Growth		
	Math (1)	ELA (2)	Sci (3)	Math (4)	ELA (5)	Sci (6)
Year Fixed Effects						
2007 (base)						
2008	-0.012	-0.009	-0.020**	-0.011	-0.011	-0.017*
	(0.009)	(0.007)	(0.009)	(0.009)	(0.007)	(0.009)
2009	-0.018	-0.017*	-0.035***	-0.017	-0.020**	-0.028**
	(0.013)	(0.010)	(0.012)	(0.012)	(0.010)	(0.012)
2010	-0.088***	-0.053***	-0.072***	-0.083***	-0.055***	-0.061***
	(0.017)	(0.013)	(0.015)	(0.015)	(0.012)	(0.015)
2011	-0.122***	-0.082***	-0.109***	-0.113***	-0.083***	-0.095***
	(0.020)	(0.015)	(0.018)	(0.018)	(0.014)	(0.018)
2012	-0.108***	-0.080***	-0.104***	-0.097***	-0.078***	-0.087***
	(0.023)	(0.017)	(0.020)	(0.020)	(0.015)	(0.020)
2013	-0.104***	-0.064***	-0.095***	-0.090***	-0.062***	-0.074***
	(0.026)	(0.020)	(0.023)	(0.023)	(0.017)	(0.022)
2014	-0.107***	-0.057***	-0.091***	-0.089***	-0.053***	-0.066***
	(0.029)	(0.022)	(0.026)	(0.025)	(0.019)	(0.025)
2015	-0.097***	-0.036	-0.068**	-0.075***	-0.030	-0.041
	(0.033)	(0.024)	(0.028)	(0.028)	(0.021)	(0.027)
2016	-0.102***	-0.058**	-0.090***	-0.077**	-0.052**	-0.059*
	(0.038)	(0.029)	(0.033)	(0.030)	(0.024)	(0.030)
2017	-0.101**	-0.034	-0.078**	-0.071**	-0.027	-0.043
	(0.040)	(0.029)	(0.034)	(0.032)	(0.024)	(0.030)
Joint Test of Year FE	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Notes: Standard errors clustered by principal-school shown in parentheses. These estimates correspond to the models shown in Table 2.
* p < 0.1, ** p < 0.05, *** p < 0.01.

Table A6
IVM Results Using Modified Year Bins

	Math (1)	ELA (2)	Sci (3)
Total Principal Experience			
0 years (base)			
1 year	0.019***	0.006*	0.009*
	(0.005)	(0.003)	(0.005)
2 years	0.026***	0.006	0.024***
	(0.008)	(0.005)	(0.008)
3 years	0.046***	0.014**	0.041***
	(0.010)	(0.007)	(0.011)
4–6 years	0.073***	0.018*	0.054***
	(0.014)	(0.010)	(0.015)
7–9 years	0.080***	0.021	0.071***
	(0.019)	(0.014)	(0.021)
10–14 years	0.118***	0.032*	0.091***
	(0.027)	(0.019)	(0.030)
Modified Year Bins			
2008–2009	-0.017*		
	(0.010)		
2010	-0.092***		
	(0.014)		
2011	-0.128***		
	(0.016)		
2012–2013	-0.115***		
	(0.018)		
2014–2015	-0.115***		
	(0.022)		
2016–2017	-0.119***		
	(0.027)		
2008–2009		-0.014*	
		(0.007)	
2010		-0.054***	
		(0.010)	
2011–2012		-0.083***	
		(0.012)	
2013–2014		-0.065***	
		(0.014)	
2015		-0.041**	
		(0.016)	
2016		-0.065***	
		(0.021)	
2017		-0.041**	
		(0.019)	
2008			-0.020**
			(0.009)
2009			-0.035***
			(0.012)
2010			-0.070***
			(0.015)
2011–2012			-0.107***
			(0.018)
2013–2014			-0.097***
			(0.021)
2015			-0.075***
			(0.024)
2016–2017			-0.087***
			(0.028)
N	3034743	3328312	2819174
R^2	0.301	0.319	0.339

Notes: Standard errors clustered by principal-school shown in parentheses. The dependent variables are standardized student achievement scores. The omitted category is principals who have zero years of experience. Models include fixed effects for principal, school, year, and grade. Covariates include student characteristics and time-varying school characteristics.
* p < 0.1, ** p < 0.05, *** p < 0.01.

Table A7

The Returns to School-Specific Experience by Length of Stay in School

	Math				ELA				Sci			
Length of Spell (x) =	2	3	5	10	2	3	5	10	2	3	5	10
School-Specific Experience												
0 years (base)												
1 year	0.034***	0.026***	0.033***	0.024***	0.013*	0.008	0.015***	0.011**	0.008	0.012	0.015**	0.006
	(0.011)	(0.008)	(0.007)	(0.008)	(0.007)	(0.005)	(0.005)	(0.005)	(0.010)	(0.008)	(0.006)	(0.007)
2 years	0.042***	0.037***	0.040***	0.032**	0.017***	0.015*	0.016**	0.013	0.030***	0.030**	0.028***	0.015
	(0.010)	(0.014)	(0.012)	(0.013)	(0.006)	(0.008)	(0.007)	(0.008)	(0.009)	(0.012)	(0.010)	(0.011)
3 years	0.065***	0.060***	0.062***	0.050**	0.028***	0.029***	0.023**	0.018	0.044***	0.041***	0.048***	0.021
	(0.015)	(0.014)	(0.017)	(0.020)	(0.009)	(0.009)	(0.011)	(0.012)	(0.013)	(0.013)	(0.016)	(0.017)
4–6 years	0.097***	0.087***	0.087***	0.070***	0.032***	0.032**	0.026*	0.012	0.057***	0.052***	0.057***	0.016
	(0.021)	(0.020)	(0.022)	(0.027)	(0.012)	(0.013)	(0.014)	(0.016)	(0.019)	(0.019)	(0.020)	(0.024)
7–9 years	0.129***	0.112***	0.112***	0.081**	0.050***	0.048**	0.040**	0.017	0.073***	0.064**	0.070**	0.008
	(0.032)	(0.029)	(0.031)	(0.038)	(0.019)	(0.019)	(0.020)	(0.026)	(0.028)	(0.027)	(0.029)	(0.038)
10–14 years	0.182***	0.156***	0.158***	0.178***	0.068**	0.065**	0.058**	0.090**	0.086**	0.073*	0.078*	0.044
	(0.048)	(0.044)	(0.044)	(0.064)	(0.027)	(0.027)	(0.027)	(0.040)	(0.040)	(0.040)	(0.041)	(0.072)
Interactions												
1 year x Spell >= x	-0.008	-0.001	-0.021**	-0.011	-0.002	0.006	-0.005	0.012	0.004	0.001	-0.008	0.006
	(0.011)	(0.009)	(0.009)	(0.021)	(0.007)	(0.006)	(0.006)	(0.022)	(0.010)	(0.008)	(0.008)	(0.029)
2 years x Spell >= x		0.004	0.001	-0.007		0.004	-0.001	-0.007		-0.002	0.002	0.012
		(0.013)	(0.011)	(0.022)		(0.008)	(0.007)	(0.025)		(0.011)	(0.009)	(0.043)
3 years x Spell >= x			-0.005	-0.018			0.000	0.014			-0.001	0.030
			(0.013)	(0.023)			(0.008)	(0.028)			(0.011)	(0.045)
4–6 years x Spell >= x				0.008				0.048				0.069
				(0.030)				(0.031)				(0.051)
7–9 years x Spell >= x				0.033				0.048				0.033
				(0.036)				(0.033)				(0.061)
N	3072943	2907300	2539881	1830425	3394732	3219404	2823009	2031196	2862682	2712853	2378006	1737399
R^2	0.299	0.299	0.300	0.304	0.319	0.318	0.317	0.320	0.339	0.340	0.343	0.350

Notes: Standard errors clustered by principal-school shown in parentheses. The dependent variables are standardized student achievement scores. Models include fixed effects for principal, school, year, and grade. Covariates include student characteristics and time-varying school characteristics. The interaction group is defined by an indicator for whether a principal stays in their school for at least x years, where x is defined at the top of the column. Right-censored principal-school spells are dropped from the model if the highest observed year is less than the length of the spell. For example, if x = 10, I only keep school-principal spells that ended prior to 2017 (the last year of the data stream) or at least 10 years long by 2017.

* p < 0.1, ** p < 0.05, *** p < 0.01.

Table A8

The Returns to School-Specific Experience by Whether Principal is Observed in Multiple Schools

	Math (1)	ELA (2)	Sci (3)
School-Specific Experience			
0 years (base)			
1 year	0.021***	0.009**	0.006
	(0.007)	(0.004)	(0.006)
2 years	0.034***	0.013*	0.025***
	(0.011)	(0.007)	(0.010)
3 years	0.060***	0.028***	0.041***
	(0.015)	(0.010)	(0.014)
4–6 years	0.090***	0.035**	0.057***
	(0.021)	(0.014)	(0.020)
7–9 years	0.113***	0.055***	0.074**
	(0.031)	(0.020)	(0.029)
10–14 years	0.153***	0.069**	0.086**
	(0.047)	(0.029)	(0.041)
Interactions			
1 year x Multiple Schools	-0.000	0.000	0.011
	(0.008)	(0.005)	(0.008)
2 years x Multiple Schools	0.004	0.007	0.007
	(0.011)	(0.007)	(0.009)
3 years x Multiple Schools	-0.011	-0.006	0.001
	(0.014)	(0.009)	(0.013)
4–6 years x Multiple Schools	-0.028*	-0.017	-0.015
	(0.016)	(0.011)	(0.016)
7–9 years x Multiple Schools	-0.027	-0.041**	-0.034
	(0.028)	(0.018)	(0.029)
10–14 years x Multiple Schools	0.010	-0.016	-0.053
	(0.037)	(0.030)	(0.051)
N	3262311	3583835	3029986
R^2	0.303	0.321	0.341

Notes: Standard errors clustered by principal-school shown in parentheses. The dependent variables are standardized student achievement scores. Models include fixed effects for principal, school, year, and grade. Covariates include student characteristics and time-varying school characteristics. The interaction group is defined by an indicator for whether the principal is observed in two or more schools across the study period.

* p < 0.1, ** p < 0.05, *** p < 0.01.

Table A9

The Returns to Principal Experience for Hired and Inherited Teachers (Main Effect + Interactions)

	All Turnover (1)	Transfer (2)	Exit (3)	Position Change (4)
School-Specific Experience				
0 years (base)				
1 year	-0.008*	-0.007**	-0.002	0.000
	(0.004)	(0.003)	(0.003)	(0.001)
2 years	-0.014**	-0.016***	0.000	0.000
	(0.006)	(0.005)	(0.004)	(0.001)
3 years	-0.025***	-0.022***	-0.007	-0.000
	(0.009)	(0.007)	(0.006)	(0.002)
4–6 years	-0.030**	-0.025**	-0.010	0.000
	(0.013)	(0.010)	(0.008)	(0.003)
7–9 years	-0.032	-0.034**	-0.003	0.001
	(0.021)	(0.016)	(0.014)	(0.004)
10–14 years	-0.024	-0.035	0.006	0.001
	(0.031)	(0.023)	(0.022)	(0.007)
Interactions				
1 year x Prin Hired Teacher	-0.015***	-0.015***	-0.008*	0.004**
	(0.006)	(0.005)	(0.004)	(0.002)
2 years x Prin Hired Teacher	-0.016***	-0.012**	-0.010**	0.000
	(0.006)	(0.005)	(0.005)	(0.002)
3 years x Prin Hired Teacher	-0.013**	-0.011**	-0.011**	0.003
	(0.006)	(0.005)	(0.005)	(0.002)
4–6 years x Prin Hired Teacher	-0.025***	-0.021***	-0.015***	0.000
	(0.006)	(0.005)	(0.004)	(0.002)
7–9 years x Prin Hired Teacher	-0.021**	-0.017***	-0.015**	-0.001
	(0.008)	(0.006)	(0.006)	(0.002)
10–14 years x Prin Hired Teacher	-0.061***	-0.048***	-0.032	0.002
	(0.019)	(0.011)	(0.020)	(0.005)
N	350843	320877	318333	296924
R^2	0.078	0.089	0.070	0.026
Joint Test of Hired Int	0.001	0.000	0.036	0.161

Notes: Standard errors clustered by principal-school shown in parentheses. The dependent variable is a binary indicator for whether the teacher left their position in the current year, with the type of turnover listed in the column header. Models include fixed effects for principal, school, and year. Models include teacher and school characteristics, an indicator for whether the principal left the school at the end of the year, and an indicator for being a "hired" teacher. "Inherited" teachers are those who have more school-specific experience than the principal. "hired" teachers are those with the same or less school-specific experience than the principal.

* p < 0.1, ** p < 0.05, *** p < 0.01.

Table A10

The Returns to Principal Experience for Turnover of Effective and Ineffective Teachers (Main Effect + Interactions)

	All Turnover (1)	Transfer (2)	Exit (3)	Position Change (4)
School-Specific Experience				
0 years (base)				
1 year	-0.002	-0.004	0.003	-0.001
	(0.005)	(0.004)	(0.003)	(0.001)
2 years	-0.007	-0.011*	0.005	-0.002
	(0.008)	(0.006)	(0.005)	(0.002)
3 years	-0.013	-0.013	-0.000	-0.002
	(0.010)	(0.009)	(0.007)	(0.003)
4–6 years	-0.024*	-0.023*	-0.002	-0.005
	(0.015)	(0.012)	(0.009)	(0.003)
7–9 years	-0.023	-0.026	0.004	-0.007
	(0.025)	(0.021)	(0.015)	(0.006)
10–14 years	-0.057*	-0.056**	-0.000	-0.014
	(0.033)	(0.028)	(0.024)	(0.009)
Interactions				
1 year x Low VA Teacher	-0.011	-0.010	-0.006	0.001
	(0.007)	(0.006)	(0.005)	(0.002)
2 years x Low VA Teacher	-0.011	-0.007	-0.008	0.001
	(0.008)	(0.007)	(0.005)	(0.002)
3 years x Low VA Teacher	-0.010	-0.014*	0.000	0.002
	(0.009)	(0.007)	(0.006)	(0.002)
4–6 years x Low VA Teacher	-0.003	-0.004	-0.005	0.004*
	(0.008)	(0.007)	(0.005)	(0.002)
7–9 years x Low VA Teacher	-0.007	-0.010	-0.001	0.003
	(0.013)	(0.010)	(0.009)	(0.003)
10–14 years x Low VA Teacher	0.017	-0.005	-0.001	0.026**
	(0.030)	(0.025)	(0.022)	(0.012)
1 year x High VA Teacher	-0.019***	-0.014**	-0.009**	0.002
	(0.007)	(0.006)	(0.005)	(0.002)
2 years x High VA Teacher	-0.014*	-0.008	-0.005	-0.003
	(0.007)	(0.006)	(0.005)	(0.002)
3 years x High VA Teacher	-0.026***	-0.019***	-0.009	-0.002
	(0.008)	(0.007)	(0.005)	(0.003)
4–6 years x High VA Teacher	-0.019**	-0.008	-0.009*	-0.003
	(0.007)	(0.006)	(0.005)	(0.003)
7–9 years x High VA Teacher	-0.019*	-0.011	-0.004	-0.007*
	(0.012)	(0.010)	(0.008)	(0.004)
10–14 years x High VA Teacher	-0.006	-0.018	0.006	0.008
	(0.023)	(0.016)	(0.017)	(0.010)
N	164585	154140	147697	140613
R^2	0.089	0.105	0.075	0.039
Joint Test of Low VA Int	0.628	0.572	0.678	0.203
Joint Test of High VA Int	0.043	0.166	0.404	0.207

Notes: Standard errors clustered by principal-school shown in parentheses. The dependent variable is a binary indicator for whether the teacher left their position in the current year, with the type of turnover listed in the column header. Models include fixed effects for principal, school, and year. Models include teacher and school characteristics, an indicator for whether the principal left the school at the end of the year, and indicators for being high and low value-added. Value-added categories correspond to the top 20%, middle 60%, and bottom 20% of the statewide distribution.
* p < 0.1, ** p < 0.05, *** p < 0.01.

Table A11
Principal Experience and Teacher Quality

	(1)	(2)	(3)
Total Principal Experience			
0 years (base)			
1 year	-0.008		-0.017
	(0.007)		(0.015)
2 years	0.005		-0.012
	(0.011)		(0.018)
3 years	0.013		-0.002
	(0.016)		(0.024)
4–6 years	0.003		-0.034
	(0.022)		(0.030)
7–9 years	-0.013		-0.069*
	(0.031)		(0.040)
10–14 years	-0.019		-0.066
	(0.044)		(0.055)
School-Specific Experience			
0 years (base)			
1 year		-0.001	0.012
		(0.007)	(0.014)
2 years		0.016	0.025
		(0.011)	(0.018)
3 years		0.025	0.025
		(0.016)	(0.024)
4–6 years		0.030	0.055*
		(0.023)	(0.031)
7–9 years		0.034	0.087**
		(0.033)	(0.043)
10–14 years		0.009	0.056
		(0.062)	(0.075)
N	164585	164585	164585
R^2	0.186	0.186	0.186

Notes: Standard errors clustered by principal-school shown in parentheses. The dependent variable is a teacher's value-added score in the given year. Models include fixed effects for principal, school, and year. Models include teacher and school characteristics, and an indicator for whether the principal left the school at the end of the year.
* p < 0.1, ** p < 0.05, *** p < 0.01.

Table A12

Heterogeneity in the Returns to Experience by School Poverty (Main Effect + Interactions)

	Math (SD)		ELA (SD)		Sci (SD)	
	(1)	(2)	(3)	(4)	(5)	(6)
School-Specific Experience						
0 years (base)						
1 year	0.011	0.009	0.005	0.003	0.006	0.002
	(0.007)	(0.008)	(0.004)	(0.005)	(0.006)	(0.007)
2 years	0.026**	0.023*	0.011	0.006	0.022**	0.011
	(0.010)	(0.012)	(0.007)	(0.008)	(0.009)	(0.011)
3 years	0.041***	0.035**	0.019*	0.012	0.033**	0.016
	(0.015)	(0.018)	(0.010)	(0.011)	(0.014)	(0.017)
4–6 years	0.062***	0.053**	0.022	0.011	0.045**	0.020
	(0.020)	(0.024)	(0.014)	(0.016)	(0.020)	(0.024)
7–9 years	0.093***	0.079**	0.047**	0.032	0.057**	0.020
	(0.029)	(0.036)	(0.021)	(0.024)	(0.028)	(0.034)
10–14 years	0.140***	0.120**	0.074**	0.058	0.062	0.007
	(0.047)	(0.055)	(0.032)	(0.036)	(0.043)	(0.051)
Interactions						
1 year x Low-Poverty	0.011	-0.007	-0.002	-0.024*	-0.002	-0.010
	(0.013)	(0.018)	(0.009)	(0.013)	(0.013)	(0.017)
2 years x Low-Poverty	0.020	-0.018	0.028**	-0.021	0.018	0.003
	(0.017)	(0.030)	(0.012)	(0.023)	(0.015)	(0.029)
3 years x Low-Poverty	0.046**	-0.007	0.031*	-0.039	0.023	0.002
	(0.021)	(0.043)	(0.016)	(0.034)	(0.023)	(0.044)
4–6 years x Low-Poverty	0.095***	0.012	0.065***	-0.049	0.049**	0.012
	(0.020)	(0.058)	(0.016)	(0.045)	(0.022)	(0.060)
7–9 years x Low-Poverty	0.119**	-0.025	0.048	-0.144*	0.074	0.003
	(0.053)	(0.090)	(0.036)	(0.079)	(0.050)	(0.106)
10–14 years x Low-Poverty	0.119**	-0.067	0.073**	-0.172*	0.083	0.033
	(0.048)	(0.124)	(0.035)	(0.104)	(0.054)	(0.140)
1 year x High-Poverty	0.020*	0.033**	0.016**	0.033***	0.012	0.030**
	(0.010)	(0.014)	(0.007)	(0.009)	(0.009)	(0.012)
2 years x High-Poverty	0.006	0.035	0.003	0.041***	0.009	0.050**
	(0.014)	(0.023)	(0.009)	(0.015)	(0.012)	(0.020)
3 years x High-Poverty	0.020	0.066**	0.013	0.071***	0.024	0.086***
	(0.016)	(0.033)	(0.011)	(0.021)	(0.015)	(0.029)
4–6 years x High-Poverty	-0.007	0.070	0.003	0.092***	0.008	0.102**
	(0.019)	(0.045)	(0.014)	(0.030)	(0.019)	(0.041)
7–9 years x High-Poverty	-0.063**	0.045	-0.017	0.118***	0.031	0.164***
	(0.028)	(0.065)	(0.019)	(0.045)	(0.036)	(0.063)
10–14 years x High-Poverty	-0.171***	0.002	-0.112*	0.065	0.058	0.254***
	(0.053)	(0.098)	(0.058)	(0.074)	(0.056)	(0.086)
Year Fixed Effects	✓		✓		✓	
Year x FRPL Fixed Effects		✓		✓		✓
N	3034717	3034717	3328229	3328229	2819110	2819110
R^2	0.301	0.302	0.319	0.319	0.339	0.339
Joint Test of Low-Pov Int	0.000	0.480	0.001	0.099	0.298	0.932
Joint Test of High-Pov Int	0.001	0.007	0.012	0.002	0.426	0.075

Notes: Standard errors clustered by principal-school shown in parentheses. In each column, I show the estimated main effect for school-specific principal experience and interactions by school poverty level. Models include fixed effects for principal, school, and grade. Covariates include student characteristics and time-varying school characteristics. High, medium, and low-poverty refer to the percentage of students in the school who qualify for free/reduced price lunch: 0-30%, 30-70%, 70-100%. These categories are time-invariant and absorbed by the school fixed effect.
* p < 0.1, ** p < 0.05, *** p < 0.01.

Table A13

Heterogeneity in the Returns to Experience by School Poverty (Separate Models)

	Math			ELA			Sci		
	Low-Pov (1)	Med-Pov (2)	High-Pov (3)	Low-Pov (4)	Med-Pov (5)	High-Pov (6)	Low-Pov (7)	Med-Pov (8)	High-Pov (9)
School-Specific Experience									
0 years (base)									
1 year	-0.005	0.008	0.037***	-0.025**	0.002	0.031***	-0.015	0.001	0.026***
	(0.016)	(0.007)	(0.012)	(0.013)	(0.005)	(0.007)	(0.015)	(0.007)	(0.010)
2 years	-0.002	0.020*	0.048**	-0.019	0.005	0.040***	0.004	0.010	0.050***
	(0.026)	(0.012)	(0.020)	(0.022)	(0.008)	(0.012)	(0.026)	(0.011)	(0.016)
3 years	0.016	0.032*	0.089***	-0.034	0.010	0.071***	0.007	0.016	0.086***
	(0.037)	(0.017)	(0.028)	(0.031)	(0.011)	(0.017)	(0.040)	(0.017)	(0.024)
4–6 years	0.048	0.049**	0.101***	-0.047	0.009	0.086***	0.013	0.020	0.097***
	(0.051)	(0.024)	(0.037)	(0.043)	(0.016)	(0.024)	(0.055)	(0.024)	(0.034)
7–9 years	0.043	0.073**	0.095*	-0.114	0.029	0.123***	0.019	0.021	0.147***
	(0.085)	(0.035)	(0.055)	(0.076)	(0.024)	(0.037)	(0.098)	(0.034)	(0.053)
10–14 years	0.043	0.108**	0.102	-0.119	0.052	0.098	0.042	0.005	0.222***
	(0.110)	(0.054)	(0.084)	(0.101)	(0.036)	(0.063)	(0.121)	(0.052)	(0.070)
N	392955	1907131	734631	435035	2135345	757849	364507	1749871	704732
R^2	0.291	0.248	0.215	0.299	0.257	0.224	0.279	0.249	0.245

Notes: Standard errors clustered by principal-school shown in parentheses. In each column, I estimate a model for the returns to school-specific experience on the sample defined by the column header (high, medium, or low-poverty). Models include fixed effects for principal, school, year, and grade. Covariates include student characteristics and time-varying school characteristics. High, medium, and low-poverty refer to the percentage of students in the school who qualify for free/reduced price lunch: 0-30%, 30-70%, 70-100%.

* p < 0.1, ** p < 0.05, *** p < 0.01.

Table A14

Heterogeneity in the Returns to Experience by School Level (Main Effect + Interactions)

	Math (SD)		ELA (SD)		Sci (SD)	
	(1)	(2)	(3)	(4)	(5)	(6)
School-Specific Experience						
0 years (base)						
1 year	0.022***	0.022***	0.011**	0.010*	0.013**	0.015**
	(0.007)	(0.008)	(0.005)	(0.005)	(0.006)	(0.007)
2 years	0.049***	0.049***	0.025***	0.021**	0.041***	0.043***
	(0.011)	(0.013)	(0.008)	(0.009)	(0.010)	(0.011)
3 years	0.078***	0.076***	0.049***	0.044***	0.064***	0.066***
	(0.015)	(0.018)	(0.011)	(0.012)	(0.014)	(0.016)
4–6 years	0.104***	0.104***	0.061***	0.054***	0.074***	0.077***
	(0.021)	(0.025)	(0.015)	(0.018)	(0.020)	(0.022)
7–9 years	0.121***	0.121***	0.085***	0.072***	0.095***	0.102***
	(0.032)	(0.036)	(0.022)	(0.026)	(0.031)	(0.033)
10–14 years	0.159***	0.159***	0.108***	0.087**	0.130***	0.140***
	(0.049)	(0.054)	(0.034)	(0.039)	(0.045)	(0.047)
Interactions						
1 year x Middle School	0.003	-0.009	0.004	0.006	-0.008	-0.012
	(0.010)	(0.013)	(0.007)	(0.008)	(0.009)	(0.012)
2 years x Middle School	-0.014	-0.038*	-0.005	-0.003	-0.020*	-0.027
	(0.012)	(0.020)	(0.008)	(0.014)	(0.010)	(0.019)
3 years x Middle School	-0.018	-0.056**	-0.017	-0.015	-0.031**	-0.039
	(0.015)	(0.028)	(0.010)	(0.020)	(0.014)	(0.028)
4–6 years x Middle School	-0.010	-0.069*	-0.022*	-0.017	-0.017	-0.030
	(0.017)	(0.039)	(0.012)	(0.029)	(0.017)	(0.040)
7–9 years x Middle School	0.012	-0.083	-0.041**	-0.031	-0.036	-0.060
	(0.032)	(0.056)	(0.019)	(0.044)	(0.035)	(0.063)
10–14 years x Middle School	0.012	-0.104	-0.023	-0.005	-0.099*	-0.124
	(0.063)	(0.091)	(0.039)	(0.065)	(0.052)	(0.086)
1 year x High School	-0.021*	-0.001	-0.010	-0.008	-0.011	-0.013
	(0.012)	(0.016)	(0.007)	(0.010)	(0.012)	(0.017)
2 years x High School	-0.053***	-0.010	-0.013	-0.009	-0.033**	-0.036
	(0.016)	(0.026)	(0.009)	(0.017)	(0.014)	(0.026)
3 years x High School	-0.079***	-0.020	-0.042***	-0.034	-0.049***	-0.053
	(0.018)	(0.036)	(0.011)	(0.024)	(0.018)	(0.038)
4–6 years x High School	-0.105***	-0.013	-0.056***	-0.041	-0.056***	-0.061
	(0.022)	(0.051)	(0.012)	(0.034)	(0.020)	(0.052)
7–9 years x High School	-0.115***	0.027	-0.053***	-0.025	-0.056	-0.066
	(0.031)	(0.074)	(0.018)	(0.052)	(0.034)	(0.077)
10–14 years x High School	-0.159**	0.049	-0.080	-0.036	-0.067	-0.080
	(0.073)	(0.113)	(0.067)	(0.075)	(0.069)	(0.111)
Year Fixed Effects	✓		✓		✓	
Year x Level Fixed Effects		✓		✓		✓
N	2884580	2884580	3157297	3157297	2684642	2684642
R^2	0.299	0.300	0.316	0.316	0.338	0.338
Joint Test of Middle Sch Int	0.623	0.324	0.147	0.527	0.199	0.494
Joint Test of High Sch Int	0.000	0.584	0.000	0.177	0.112	0.800

Notes: Standard errors clustered by principal-school shown in parentheses. In each column, I show the estimated main effect for school-specific principal experience and interactions by school level. Models include fixed effects for principal, school, and grade. Covariates include student characteristics and time-varying school characteristics. School level is time-invariant and absorbed by the school fixed effect.
* p < 0.1, ** p < 0.05, *** p < 0.01.

Table A15

Heterogeneity in the Returns to Experience by School Locale (Main Effect + Interactions)

	Math (SD)		ELA (SD)		Sci (SD)	
	(1)	(2)	(3)	(4)	(5)	(6)
School-Specific Experience						
0 years (base)						
1 year	0.021***	0.023***	0.009**	0.009*	0.008	0.016**
	(0.007)	(0.008)	(0.004)	(0.005)	(0.007)	(0.008)
2 years	0.035***	0.038***	0.013**	0.013*	0.027***	0.040***
	(0.011)	(0.013)	(0.007)	(0.008)	(0.009)	(0.011)
3 years	0.057***	0.060***	0.029***	0.029***	0.037***	0.057***
	(0.015)	(0.018)	(0.010)	(0.011)	(0.014)	(0.017)
4–6 years	0.090***	0.090***	0.037***	0.038**	0.045**	0.077***
	(0.020)	(0.025)	(0.014)	(0.015)	(0.019)	(0.023)
7–9 years	0.119***	0.119***	0.057***	0.060**	0.044	0.099***
	(0.031)	(0.037)	(0.022)	(0.025)	(0.029)	(0.035)
10–14 years	0.161***	0.162***	0.096***	0.107***	0.089*	0.159***
	(0.050)	(0.056)	(0.033)	(0.037)	(0.048)	(0.054)
Interactions						
1 year x Urban School	-0.006	-0.007	0.001	0.003	0.006	-0.009
	(0.009)	(0.012)	(0.006)	(0.008)	(0.009)	(0.012)
2 years x Urban School	-0.012	-0.009	-0.001	-0.000	-0.001	-0.026
	(0.012)	(0.019)	(0.008)	(0.013)	(0.011)	(0.018)
3 years x Urban School	-0.014	-0.007	-0.011	-0.010	0.005	-0.034
	(0.015)	(0.026)	(0.010)	(0.018)	(0.014)	(0.026)
4–6 years x Urban School	-0.051***	-0.032	-0.021*	-0.022	0.008	-0.053
	(0.017)	(0.037)	(0.012)	(0.025)	(0.018)	(0.037)
7–9 years x Urban School	-0.056**	-0.026	-0.013	-0.021	0.067**	-0.039
	(0.027)	(0.054)	(0.017)	(0.037)	(0.034)	(0.059)
10–14 years x Urban School	-0.061	-0.014	-0.062	-0.085	0.003	-0.133*
	(0.062)	(0.083)	(0.048)	(0.058)	(0.051)	(0.080)
1 year x Suburban School	-0.004	-0.016	-0.005	-0.004	-0.008	-0.021
	(0.011)	(0.015)	(0.007)	(0.008)	(0.010)	(0.013)
2 years x Suburban School	0.007	-0.019	0.019**	0.021	-0.003	-0.028
	(0.014)	(0.022)	(0.009)	(0.014)	(0.011)	(0.019)
3 years x Suburban School	0.003	-0.031	0.009	0.008	0.012	-0.025
	(0.017)	(0.029)	(0.011)	(0.019)	(0.016)	(0.028)
4–6 years x Suburban School	0.006	-0.035	0.006	0.007	0.019	-0.034
	(0.018)	(0.037)	(0.011)	(0.023)	(0.017)	(0.037)
7–9 years x Suburban School	-0.023	-0.078	-0.007	-0.004	0.022	-0.062
	(0.030)	(0.055)	(0.018)	(0.031)	(0.031)	(0.055)
10–14 years x Suburban School	-0.041	-0.134	-0.022	-0.016	-0.034	-0.145*
	(0.052)	(0.082)	(0.030)	(0.048)	(0.046)	(0.079)
Year Fixed Effects	✓		✓		✓	
Year x Locale Fixed Effects		✓		✓		✓
N	3011760	3011760	3304432	3304432	2798667	2798667
R^2	0.300	0.301	0.318	0.318	0.338	0.338
Joint Test of Urban Sch Int	0.101	0.902	0.477	0.477	0.424	0.387
Joint Test of Suburb Sch Int	0.872	0.664	0.368	0.272	0.646	0.371

Notes: Standard errors clustered by principal-school shown in parentheses. In each column, I show the estimated main effect for school-specific principal experience and interactions by school locale (urban, suburban, town/rural). Models include fixed effects for principal, school, and grade. Covariates include student characteristics and time-varying school characteristics. School locale is time-invariant and absorbed by the school fixed effect.

* p < 0.1, ** p < 0.05, *** p < 0.01.