Who Becomes a School Leader? An Investigation of Teachers' Careers and Value-Added

Molly Schwarz*

Current draft: November 7, 2022†

Abstract

This paper investigates whether effective teachers are more likely to progress to school leadership positions. Using data on 4th through 8th grade teachers at elementary and middle schools in North Carolina, I estimate the relationship between a teacher's effectiveness in teaching (as measured by value-added to math and reading standardized test scores) and their likelihood of becoming an assistant principal in the next academic year. I find that one standard deviation higher value-added math (reading) teachers are on average 33.7% (7.9%) more likely to become an assistant principal in the next academic year after controlling for differences across educators, such as demographic characteristics, teaching experience, and training. There is important heterogeneity in the size of this effect across groups, with value-added being more predictive of promotion to assistant principal for both male teachers and non-white teachers. Finally, to explore the consequences of promoting teachers based on current performance, I estimate principal value-added and compare it to an educator's teacher value-added. Using a back of the envelope calculation, I find that these promotion practices on average lead to small positive impacts on student achievement.

Keywords: Value-added, principals, assistant principals, teachers, management

JEL codes: I20, J24, J45, M51

---

*Department of Economics, University of California Santa Barbara, 2127 North Hall, Santa Barbara, CA. Email: mcschwarz@ucsb.edu.

†I would like to thank Kara Bonneau and the North Carolina Education Research Data Center for access to data. I would also like to thank Kelly Bedard, Heather Royer, Dick Startz, Peter Kuhn, Cathy Weinberger, Sarah Robinson, Richard Uhrig, Luisa Wallossek, Antoine Dib, the UCSB Econ 290 Applied Working Group, and the UCSB Applied Micro Economics Lunch for their careful advice and feedback on versions of this paper. All errors are my own.

1 Introduction

A growing body of empirical research documents the important role of public school principals in improving student achievement, including raising student test scores and improving school attendance (Branch, Hanushek and Rivkin, 2008, 2012; Coelli and Green, 2012; Dhuey and Smith, 2014, 2018; Li, 2015; Bartanen, 2020). While we know principals can impact students, we know little about who becomes a principal and what makes someone an effective principal (Liebowitz and Porter, 2019, 2020). One thing we do know is that most principals were once teachers themselves. According to the 2011-2012 Schools and Staffing Survey from the National Center for Education Statistics, 98.3% of public school principals in the U.S. had teaching experience before becoming principals.

Since the principalship is filled almost entirely by former teachers, a natural question to ask is whether the teachers who are most effective at raising student achievement are the ones most likely to become principal. Teaching effectiveness may be at least weakly positively correlated with principal effectiveness since both require an understanding of what makes students successful. Teaching effectiveness may also be a relatively useful and objective signal of principal effectiveness relative to other factors considered in the hiring process, such as teaching evaluations. Thus, hiring committees may be interested in highly effective teachers as candidates for principal positions. On the other hand, superintendents and school boards may be aware of the cost of removing an effective teacher from the classroom to engage in school leadership. The "Peter principle," a management concept from Peter, Hull et al. (1969), states that because current performance is used for deciding promotions and different skills are often needed at the next level, employees are promoted to the point of incompetence. Thus, the Peter principle would suggest that effective teachers are more likely to get promoted, though it may not be optimal because teacher and principal effectiveness may be uncorrelated. Though superintendents and school boards have institutional knowledge about which teachers are most likely to succeed as school leaders, systematically examining this topic from an empirical lens can help to expand our knowledge on the current state of the school leadership pipeline and to understand areas for improvement and targeted policy.

The focus of this paper is the role of assistant principal (AP), which is an increasingly important stepping stone from teaching to becoming a principal. In the context studied in this paper, North Carolina (NC) public elementary and middle schools, 90% of principals have experience as an AP.[1] Most APs are no longer providing classroom instruction, so they are responsible for aiding the principal in school management responsibilities, such as observing teachers, scheduling testing, and student discipline (Goldring, Rubin and Herrmann, 2021). Understanding who becomes an AP is critical to understanding who eventually becomes principal and consequently the impact principals have on their schools. This paper explores the career transition from teacher to AP to understand the first step up the school leadership ladder.[2]

I investigate whether more effective teachers, as measured by having higher value-added (VA) to student achievement, are more likely to be promoted to AP compared to their less effective colleagues. Teacher VA modeling is used to quantify how much an individual teacher contributes to their students' achievement while controlling for other important factors impacting students' human capital accumulation up to that point. The most commonly used outcome in teacher VA modeling is student standardized test scores, which I employ in this paper, though non-cognitive outcomes such as attendance and suspensions are also used in the literature. Having a teacher with higher VA has been linked to improved long run outcomes for students, such as increased high school graduation, college attendance, and lifetime earnings (Chetty, Friedman and Rockoff, 2014b; Jackson, 2018). Thus, teacher VA has been shown to be a useful and meaningful proxy for teacher effectiveness.

Using administrative data on 4th through 8th grade teachers in elementary and middle schools in NC, I quantify the relationship between a teacher's VA and their likelihood of becoming an AP in the next academic year. In the main analysis, I employ a linear probability model and find that teachers who have one standard deviation (SD) higher math (reading) VA are 0.114 (0.025) percentage points likelier to become AP in the next academic

[1]The share of principals who were once APs ranges from 50-90% across several states (Austin et al., 2019). The recent national average has increased from 50% to over 75% in the last 30 years (Goldring, Rubin and Herrmann, 2021).

[2]In this paper, I define school leaders as APs and principals.

year, conditional on teacher experience and demographic factors, such as age, gender, and race/ethnicity. Relative to the outcome mean of 0.336% (0.323%) for teachers with average math (reading) VA, these results correspond to math (reading) teachers being 33.7% (7.9%) likelier to become AP in the next academic year. I also explore non-linearity in these results. Teachers in the top quintile of math VA are 153.2% likelier to become AP in the next academic year compared to those in the bottom quintile, whose likelihood is approximately 0.2%. Given that there are roughly 40,000 elementary and middle school teachers in NC, extrapolating these values across ten years suggests that 160 people in the bottom quintile would become AP and 400 people in the top quintile would become AP.

Additionally, I investigate heterogeneity in the main results by teacher gender and race/ethnicity. Most teachers in the sample are white and female, so there may be important differences across demographic groups that are not seen in the full sample. I find that math VA is a stronger predictor of becoming an AP for male teachers compared to female teachers and for non-white teachers compared to white teachers. A 1 SD higher math VA male teacher is 106.0% likelier to become an AP next year relative to the outcome mean, while the same value for women is only 23.2%. Similarly, a 1 SD higher VA non-white teacher is 74.9% likelier to become an AP next year relative to the outcome mean, while a 1 SD higher VA white teacher is only 24.3% likelier to become AP next year. I also test the robustness of the main results and find consistent results regardless of the estimation strategy or specification.

Finally, I explore the relationship between teacher VA and principal VA to understand whether the practice of promoting effective teachers leads to average test score gains or losses. I find that the correlation between teacher and principal VA is modest and positive at a value of 0.15. Using a back of the envelope calculation, I conclude that the current promotion practices likely lead to small positive gains in average student achievement.

This paper makes several important contributions. First, this paper extends the teacher retention literature by looking at transitions to other positions within the education system. This is the first paper to investigate the relationship between VA and transitions to school leadership positions. Prior work has devoted attention to the relationship between teacher

VA and teacher transitions to other schools, districts, or out of the state's public school system (Hanushek et al., 2005; Krieg, 2006; West and Chingos, 2009; Goldhaber, Gross and Player, 2011; Feng and Sass, 2017). A goal of these retention papers is to understand whether the best teachers are staying in teaching. This paper broadens our understanding of where effective teachers go by highlighting that the top teachers are leaving the classroom to work in school leadership. It is important to investigate the loss of teachers to school leadership because, depending on the relationship between teacher effectiveness and principal effectiveness, there may be a net benefit to students due to this specific type of attrition.

This paper also contributes to the management and personnel economics literature, especially the topic of the Peter principle. Addressing the first component of the principle, this paper finds that current job performance is indeed predictive of promotion to the role of AP. This is the first paper to document significant positive selection into school leadership. Two recent papers briefly address the issue of selection into school leadership but find statistically insignificant positive results in different states (Goldhaber, Holden and Chen, 2019; Liebowitz and Porter, 2020). Using a greater number of years of data and in a study designed to specifically investigate selection, I am able to uncover the first statistically significant evidence of positive selection into school leadership.

Finally, this paper addresses the second component of the Peter principle, which is related to the consequences of promoting based on current performance. In Section 6, I estimate principal VA and compare it to teacher VA for a small subgroup of educators. I then use a back of the envelope calculation to estimate whether the promotion of effective teachers impacts student achievement positively or negatively. I find that the current promotion practices on average lead to slight test score gains. This is due to both the positive correlation between teacher and principal VA and the fact that an educator reaches more students as a principal than as a teacher. This result suggests that the current school leader promotion practices, which favor more effective teachers, do not fall prey to the pitfalls outlined in the Peter principle in this context. These results improve upon prior work, which also correlates teacher and principal VA to understand whether the best teachers make the best principals but does not find as large or as significant of a relationship in other states (Goldhaber, Holden and Chen, 2019; Liebowitz and Porter, 2020; Grissom, Woo and Bartanen, 2020).

The rest of the paper proceeds as follows. Section 2 provides a brief background on the role of AP in NC and describes the administrative education data from NC. Section 3 discusses the VA model and the regression model used to quantify the relationship between teacher VA and becoming an AP. Section 4 presents the main results and an analysis of heterogeneity by teacher gender and race/ethnicity. Section 5 tests the results' robustness to alternative estimators and variations in VA calculation, regression specification, and sample. Section 6 discusses the implications of these results for student achievement, and Section 7 concludes.

2 Background and data

2.1 Becoming an assistant principal in NC

This project focuses on APs in NC. The overarching rules and guidelines for APs and principals in NC are governed by North Carolina General Statutes, Chapter 115C Article 19. The typical process for becoming a school leader in NC starts with a vacancy being posted through a state-wide system. Applicants are initially screened and interviewed by district administrators, teachers, staff, and parents at the school, who then recommend final candidates to the superintendent. Then, the superintendent recommends an individual to the local board of education (Miller, 2013). Both APs and principals are required to have the same certification, which involves passing an exam from the State Board of Education and meeting proper education and training requirements. The education requirement is satisfied by having a master's degree in public school administration or comparable education and training. The State Board of Education can also give out one-year provisional assistant principal certifications under special circumstances, such as a shortage of qualified principals. An assistant principal's duties are delegated by the principal from the principal's duties.

2.2 Data

The data for this project comes from the North Carolina Education Research Data Center, which partners with the North Carolina Department of Public Instruction to manage data on the state's public schools, teachers, and students. This administrative data contains information at the individual-level, such as individual students' test scores and individual teachers' employment records. It is recorded at the academic-year-level, which ranges from fall of one calendar year to spring of the following calendar year. I will subsequently identify academic years by the spring semester (e.g., 2018-2019 is henceforth 2019).

2.2.1 Teacher data

To construct a record of teachers' careers over time, I create an individual-level teacher panel using yearly teacher pay data. This data contains employment information for certified employees in NC public schools, which includes positions such as teachers, counselors, and principals.[3] It is available for the 1995 to 2019 academic years. The dataset contains budget codes for an employee's position. There are unique budget codes for teachers, assistant principals, principals, superintendents, psychologists, teaching support, and more. When individuals are employed in more than one position in a given year, I denote their position for that year based on the position code with the highest percent of employment. Using the unique employee ID, I track individuals across years to construct their employment histories. Then, I merge this data with additional information on teacher's demographic characteristics, education, and licensure. Importantly, years of experience in the NC public school system is reported starting in 2012. In Figure 1, I use the 2012-2019 years of data to plot the years of experience an individual has the first time they become AP. I include only APs at elementary, middle, or combined elementary and middle schools. It is clear from this figure that one's likelihood of becoming an AP varies with experience. The most frequent time to become AP is in one's 9th year. It is quite rare for very inexperienced or very experienced individuals to become AP. This indicates that controlling for experience

[3] Positions that do not require certification, such as teaching assistants, are not included in this dataset.

will be extremely important for understanding the relationship between VA and becoming AP.

Most individuals who become AP were once teachers—94% of the 2,614 individuals who are hired to be an AP in the 2012-2019 data are observed as teachers at least once before becoming AP. Those who are not observed as teachers in NC likely had teaching experience in another state or in private schools. Additionally, the most frequent position in the year before becoming AP is teacher. 63% of these individuals are teachers in the year before becoming AP.[4] Since most APs were once teachers, effectiveness as a teacher is likely a factor involved in one's career advancement and a relevant predictor to examine.

2.2.2 Student data

To calculate teacher VA to student achievement, I use the student-level end-of-grade test score data spanning from 2007 to 2019. In North Carolina, 3rd through 8th grade students are tested in reading and mathematics at the end of each year. To account for differences in the test across grade and year, I normalize students' test scores by subject, grade, and year to have a mean of zero and standard deviation of one. The VA model I will use relies on controlling for a student's one-year lagged test score in a given subject. Therefore, 3rd grade test scores and data for the year 2007 are only used for obtaining students' prior year test scores. A key variable in the dataset is the identifier for a student's reading and/or math teacher. For the data years 2008 to 2019, I use course membership data to link students to the exact instructor of their reading or math course.[5] My sample goes from 7.7 million observations to 5.3 million due to dropping students who can't be linked to their teacher, lack the prior year test score, repeat a grade, or lack demographic information. To ensure accuracy in my value-added model, I also drop classes with fewer than 10 students, and reach a final sample of approximately 5 million observations.

[4]The next most common position is instructional facilitators at 19%.

[5]For the data years 1998-2007, the data only allows for identification of the teacher listed as administering the math or reading exam to identify the student's math and/or reading teacher, respectively. Prior literature using the NC data suggests that this match is most reliable for only grades 4 through 6. Thus, I do not include data years that rely on an imprecise match so as to minimize incorrect attribution of student achievement to teacher.

In Table 1, I present summary statistics on my final sample of 4th through 8th grade students in North Carolina at the student-by-academic-year level for years 2008-2019. Of these students, 4.9% have limited English proficiency and 49.8% are economically disadvantaged. Just over half of students are White, about one quarter are Black, and about 15% are Hispanic. Though math and reading test scores are normalized by grade and year to have a mean of 0 and standard deviation of 1, this table only includes students who are successfully matched to their teacher and to demographic data.[6] The average test scores for reading students are slightly above average compared to the overall sample of reading students.

2.2.3 Sample restrictions for main analysis

I begin with 449,085 teacher-by-year observations for teachers at elementary, middle, or combined elementary and middle schools in the years 2012-2018.[7] After dropping the 6% of individuals with missing demographic data, I am left with 421,636 observations. The final and key restriction I make is that individuals must be linked to teaching 4th-8th grade students math or reading in the years 2008-2019 so I can calculate VA for them. This restriction reduces my final sample to 161,580 teacher-by-year observations.

Table 2 reports the summary statistics for the 421,636 teacher-by-year observations with demographic information and compares the differences between those for whom I can and cannot calculate VA. Of the 161,578 teacher-by-year observations with VA, 84,695 have both math and reading VA, while 76,883 have only math or only reading VA. Correspondingly, there are 260,058 teacher-by-year observations for whom I cannot link to VA. Teachers in the "No VA" column are those who across the years 2008-2019 never taught math or reading in grades 4-8. For example, a kindergarten teacher, a PE teacher, and a middle school social studies teacher will not have student standardized test scores for the grade and/or subject they teach.[8]

[6] 72% (66%) of students with math (reading) scores are successfully matched to their teacher and to demographic data.

[7] I do not include data from the year 2019 because my outcome of interest is becoming AP in the next year, which is not observed in the last data year.

[8] In Table A1, I present summary statistics comparing those with and without VA who taught 4th-8th

In general, these two samples are quite similar, but there are several differences between the teachers with VA and those without. Those with VA are more likely to be at middle schools and less likely to be at elementary schools than those without VA. This is likely because there are more untested grades at elementary schools, while all grades are tested in middle schools. Additionally, those with VA are more likely to become AP and have a principal license compared to those without VA.[9] 0.34% of teacher-year-observations with VA will become AP in the next year, while only 0.25% of those without VA will do so. A two-sided t-test of this difference is statistically significant (p-value of 0.000). 3.71% of teacher-year-observations with VA have a principal license while only 2.53% of those without VA have one (p-value of 0.000). So, while the investigation of VA and becoming AP unfortunately does not facilitate the inclusion of all teachers, it does target those relatively more likely and more qualified to become AP.[10]

3 Methods

To understand the relationship between a teacher's effectiveness and the likelihood of becoming an AP, I first need to obtain a measure of teacher quality. I estimate teacher VA separately for math and reading standardized test scores. The impact teachers have on their students' achievement in a given subject is conceptualized through the following student-level estimating equation:

$$A_{ijct} = X_{ijct}\beta + \nu_{ijct}, \text{ where } \nu_{ijct} = \mu_j + \theta_{jct} + \varepsilon_{ijct} \tag{1}$$

The dependent variable A_{ijct} is the end-of-grade standardized test score for student i taught by teacher j in classroom c in year t. The vector X_{ijct} contains controls for student and

graders over the sample period 2012-2018. This drops teachers from the lower grades who we may not even care about comparing to the 4th-8th grade math and reading teachers. The group of 4th-8th grade teachers with no VA only has 92,374 observations compared to 260,056 when including other grades.

[9] A teaching license is required as a pre-requisite for a principal license, so individuals with principal licenses also have teaching licenses.

[10] In Appendix Table A2, I compare teachers in the math VA sample to those in the reading VA sample and find few substantial differences, though the reading sample has more women than the math sample.

classroom characteristics, such as prior year standardized test scores, demographics, and classroom- and school-year-level means of those variables, to account for individual and peer factors influencing student achievement. The error term ν_{ijct} is the sum of three components: the teacher's time-invariant VA (μ_j), classroom-level shocks (θ_{jct}), and idiosyncratic student shocks (ε_{ijct}). For teacher VA to be interpreted as the causal impact of a teacher on student achievement, teacher assignment must be unrelated to the error term, ε_{ijct}. Thus, the vector X_{ijct} should include any factors that may impact student sorting into classrooms, such as student socio-economic status.[11]

To estimate VA for teachers using the 2008-2019 student standardized test score data, I follow the two-step method often attributed to Kane and Staiger (2008). The two steps are as follows:

1. First, estimate equation 1 using ordinary least squares and generate the residual, $\hat{\nu}_{ijct}$, for each student.

2. Average the student-level residuals within a classroom.

The intuition behind this method is that the classroom shock θ_{jct} and student shock ε_{ijct} will on average be close to zero, so the component that remains when averaging the residuals within a classroom is the teacher's impact, μ_j. Next, I use these classroom average residuals to recover a single VA estimate per teacher. Since some teachers lead more than one classroom per year (e.g., an 8th grade math teacher who teaches multiple math classes in an academic year), I first take a precision-weighted average of the average classroom residuals for a given teacher-year, as in Chetty, Friedman and Rockoff (2014a).[12] Since I then have one value for each teacher-year, I take a precision-weighted average of the teacher-year values for each teacher. At this point, I have one VA estimate for each teacher, but it is estimated with error. The final step is to construct an empirical Bayes estimate by multiplying this value

[11] A large literature has focused on the choice of controls, estimation method, and richness of data needed to reduce bias in the estimation of teacher VA in non-experimental settings (see Koedel, Mihaly and Rockoff (2015) for a review.)

[12] A precision-weighted average is a weighted average that uses weights based on the inverse of variance. Based on the formula for the precision weights, larger classrooms receive more weight and smaller classrooms receive less weight.

by a Bayesian shrinkage factor that shrinks the estimate toward the common Bayesian prior of the average teacher VA (which is 0 by construction). The use of the Bayesian shrinkage factor provides a conservative estimate of the standard deviation of VA and is especially important when using VA as an explanatory variable since it reduces the estimation-error variance.

To understand whether more effective teachers are more likely to become school leaders, I match my estimate of teacher j's VA, $\hat{\mu}_j$, to their employment data across the years 2012-2018. I then estimate the following linear probability model (LPM) using ordinary least squares:

$$\mathbb{1}\{\text{Become AP next year}\}_{jt} = \gamma_1 \hat{\mu}_j + \gamma_2 X_{jt} + \epsilon_{jt} \tag{2}$$

where $\mathbb{1}\{\text{Become AP next year}\}_{jt}$ is a binary indicator if teacher j becomes an AP in the NC public school system in year $t+1$. The variable $\hat{\mu}_j$ is teacher VA as estimated above, and X_{jt} is a vector of other factors that influence one's likelihood of becoming AP, such as age, gender, ethnicity, experience, training, and education. The outcome is a binary variable, so the estimate of $\hat{\gamma}_1$ indicates the average percentage point change in the likelihood of becoming AP next year when a teacher has 1 unit higher VA. The independent variable, $\hat{\mu}_j$, is constructed and thus may contain measurement error. If the error is independent of a teacher's true ability, the error likely biases the estimate $\hat{\mu}_j$ toward zero and may be an underestimate of the true coefficient.

This regression model estimates the relationship between VA and one's likelihood of promotion, holding important determinants of promotion constant. A causal interpretation of γ_1 relies on assuming conditional independence. In other words, a teacher with a VA of 0 needs to be a good counterfactual for a teacher with a 1 unit higher VA after conditioning on a rich set of observable characteristics. An important threat to identification is omitted variable bias. To tend to this concern, in Section 5 I explore how the coefficient on VA changes with the inclusion of various observable factors and find no qualitative difference in the results across specification. This exercise suggests stability in the relationship regardless of which observable factors are included, which may mitigate concern regarding omitted variable bias.

4 Results

4.1 Teacher value-added

To estimate teacher VA for 4th-8th grade teachers, I first estimate equation 1 using ordinary least squares. The outcome variable is standardized test scores from end-of-grade state-wide testing, normalized by subject, grade, and year. The vector of control variables, X_{it}, contains a cubic polynomial of the one year lagged standardized test score in the subject, indicators for sex, age, race and ethnicity, grade, economic disadvantage, and limited English proficiency, academic year fixed effects, school fixed effects, and class- and school-by-year means of the demographic controls. Importantly, the regression includes teacher fixed effects to reduce bias in the estimation of β, but these are not subtracted out in the creation of the residual. The estimation sample includes only 4th through 8th graders so that all observations have a lagged test score. I include all students in the model, even if their teacher can't be linked to the 2012-2018 employment data, to reduce bias in estimating β.

Following the method detailed in section 3, I estimate one value per teacher and merge this information into the 2012-2018 employment panel to investigate the relationship between VA and becoming AP. Summary statistics for VA are found in Table 3, separated by subject. This table first reports VA for all 2008-2019 teachers and then VA for the subset of teachers that I can link to the 2012-2018 employment data. Consistent with prior literature, math VA has a larger standard deviation than reading VA. These values indicate that a 1 SD higher VA teacher raises student test scores by about 16% of a standard deviation for math teachers and 10.4% of a standard deviation for reading teachers. These values fall well within typical ranges, so I proceed with the subsequent analyses. For brevity, I henceforth use the term "math teacher" to describe all teachers in the math VA analysis and "reading teacher" to describe all teachers in the reading VA analysis, even though some of these teachers may also teach other subjects, such as a 5th grade teacher who teaches all subjects.

4.2 Value-added and becoming assistant principal

Next, I estimate equation 2 with a binary indicator for becoming AP in the next academic year as the outcome and $\hat{\mu}_j$ (estimated VA) as the independent variable of interest. For ease of interpretation, VA is normalized to have a mean of zero and standard deviation of one. I also provide the outcome mean for teachers in the sample and report the results as percents relative to this value. The control variables include indicator variables for race/ethnicity, gender, education, national board certification, licensing level, school type, and academic year, as well as a second order polynomial of age and indicators for teaching experience in 5 year intervals.[13] In all regressions, robust standard errors are clustered at the school district because most first time APs are promoted within the same district.[14]

Table 4 reports the results for equation 2 for math in column (1) and for reading in column (3). Turning to the results in column (1), I find that a 1 SD higher VA math teacher is 0.113 percentage points (pp) likelier to become AP next year. In percent terms, this corresponds to a 33.7% difference relative to the average likelihood of becoming AP next year, 0.336%. This result is highly significant. Column (3) reports the corresponding results using reading VA. A 1 SD higher reading teacher is 0.025 pp or 7.9% likelier to become AP next year. However, this result is not statistically significant. These results indicate that both math and reading teachers with higher VA are on average more likely to become AP, holding other important factors constant. Being a higher math VA teacher, however, is a much stronger and more significant predictor than being a higher reading VA teacher.

To understand how large the impact of VA is on one's likelihood of becoming AP, I compare the magnitude of the coefficient on VA to other important coefficients in the regression results. In Table A3, I also report the coefficients for the indicators for being a female teacher

[13] School type is elementary, middle, or combined elementary and middle school.

[14] 83% of new APs in the sample were employed in the same district in the prior year. Only 25% were employed in the same school in the prior year.

or a Black teacher . For math VA, the impact of VA is about 26% of the gender gap and 41% of the Black-white gap. For reading VA, the impact of VA is about 7.5% of the gender gap and 10.5% of the Black-white gap.

Equation 2 imposes linearity in the relationship between VA and one's likelihood of becoming AP next year. To relax this assumption, I use VA quintiles as my independent variables of interest. In Table 4, I use math VA quintiles in column (2) and reading VA quintiles in column (4). These results are also reported in Figure 2. The first quintile is used as the reference group, so to get results in percent terms relative to the bottom quintile, I provide the outcome mean for the first quintile. For math teachers, there is a monotonic and somewhat linear relationship between VA quintile and becoming AP. From quintile two to quintile five, the percent increase in one's likelihood of becoming AP, relative to the bottom quintile, is 55.7, 78.5, 87.4, and 153.2%, respectively. The left panel of Figure 2 highlights this relationship. For reading teachers in column (4), none of the quintiles can be distinguished from the lowest quintile at the 5% level, as seen in the right panel of Figure 2. Thus, the results for math and reading VA do not suggest a highly non-linear relationship between VA and becoming AP, so I proceed with using the linearly specified VA in subsequent analyses.

4.3 Heterogeneity

Next, I investigate the heterogeneity of the main results based on teacher characteristics. This analysis is important for several reasons. First, elementary and middle school teachers in NC are predominately female (88%) and white (83%), though both men and non-white teachers are more represented in school leadership than in the general teacher population.[15] Second, if we care about improving student outcomes, it is important to understand the promotion propensity for different groups because minority teachers and school leaders can positively impact minority student achievement (Dee, 2005; Gershenson et al., 2018; Bartanen and Grissom, 2021).

I return to the LPM with VA included as a linear independent variable and interact VA

[15] 67% of APs are female and 68% are white.

with indicators for gender or race/ethnicity.[16] Due to the predominance of female and white teachers in the sample, I am not able to precisely estimate the coefficients in a regression with the triple interaction of VA, gender, and race/ethnicity; I instead investigate heterogeneity separately by gender and race/ethnicity.

I first interact VA with an indicator for being a female teacher. These results are reported in Table 5. Using math teachers in column (1), I find that there is a difference in the relationship between VA and becoming principal based on teachers' gender. For male teachers, being a 1 SD higher VA math teacher equates to being 0.357 pp likelier to become an AP next year; for female teachers, the same value is 0.078 pp. The difference between these two values is captured in the coefficient on the interaction Female×VA, which is statistically significant at the 1% level. Relative to the outcome mean of 0.336%, the effect of being 1 SD higher math VA on becoming AP equals 106.0% for men and 23.2% for women. For reading teachers in column (3), the same pattern appears. While men with a 1 SD higher reading VA are 0.119 pp (36.9%) likelier to become AP next year, female teachers with a 1 SD higher reading VA are only 0.015 pp (4.8%) likelier to become AP next year. However, the percentage point difference is not statistically significant at conventional levels. These results show that VA is a stronger predictor of becoming AP for men compared to women, though the relationship is only statistically significant for math teachers.

In columns (2) and (4), I interact VA with an indicator for being white, using non-white teachers as the reference category. I pool together all non-white teachers into the same category for precision since NC teachers are predominately white.[17] For white math teachers, being a 1 SD higher VA teacher corresponds to being 0.082 pp (24.3%) likelier to become AP next year. For non-white math teachers, this value is 0.252 pp or 74.9% relative to the outcome mean. The percentage point difference between white and non-white teachers is significant at the 1% level. For reading teachers, the results go in the same direction, though they are very small and insignificant. A 1 SD higher reading VA white

[16]The heterogeneity results are robust to fully interacting the characteristic of interest with all controls.

[17]In Appendix Table A4, I instead use separate indicators for teachers who are Black, Hispanic, and other race/ethnicity. For math VA, the race/ethnicity differences are driven by Black teachers, who make up the highest share of non-white teachers.

teacher is 0.023 pp (7.0%) likelier to become AP next year, but the same value is 0.033 pp (10.0%) for non-white reading teachers. These results suggest that for math teachers, VA is a stronger predictor of becoming AP for non-white teachers compared to white teachers, but for reading teachers, there is not enough precision to detect strong racial differences in the relationship between VA and becoming AP.

In general, the results for gender and race/ethnicity go in the same direction when considering which groups are in the minority of teachers. Though the results are only statistically significant for math teachers, they suggest that VA is a stronger predictor of promotion for people in the gender minority group (male teachers) and the racial/ethnic minority group (non-white teachers). This aligns with discrimination theory, which would suggest that members of minority groups have to work harder to prove themselves "worthy" of promotion. As indicated by the coefficient on Female in columns (1) and (3) and Non-White in columns (2) and (4), male teachers have a higher likelihood of becoming AP compared to female teachers and non-white teachers have a higher likelihood of becoming AP compared to white teachers. Thus, the gap in the likelihood of becoming AP between male and female teachers (the gender gap) and white and non-white teachers (the race gap) is larger for more effective teachers than average teachers.

4.4 Alternative outcomes

Thus far, I have investigated the transition from teaching to becoming AP since it is the first (and most common) position change in the process of becoming principal. Given the necessity of holding a principal license to become an AP in most circumstances, obtaining this license could instead be seen as the first step.[18] Thus, I also investigate whether teacher VA is related to a teacher's likelihood of obtaining a principal license. These results are reported in columns (1) and (4) of Table 6. One SD higher math VA teachers are 20.2% likelier to obtain a principal license in the next year, while the same value for reading teachers

[18]Provisional principal licenses can be given to individuals to serve as APs under special circumstances, such as a shortage of qualified individuals with proper licensing.

is 11.0%. These results highlight that even in earlier steps of the pipeline, higher VA teachers are more likely to take the steps to become licensed to become a school leader. This also suggests the relevance of the supply side impact of VA on becoming AP since obtaining this license takes time, money, and effort to obtain.

Another important outcome is becoming principal. So far I have focused on year-to-year employment transitions of teachers, and the role of AP is a stepping stone from teaching to becoming principal. However, it is important to investigate whether the relationship between VA and becoming a school leader carries through to the next step. If teacher VA is only weakly predictive of management skills and those can be observed better once someone becomes an AP, the relationship could weaken across the next step. To investigate whether teacher VA is also predictive of becoming principal, I include all employees with teacher VA in the regression and use an outcome that is an indicator for becoming principal in the next academic year. These results are reported in columns (2) and (5) of Table 6. Employees with 1 SD higher math VA are 32.3% likelier to become principal in the next academic year, and the same value is 7.7% for reading VA. These values are nearly identical to those from the main results looking at transitions from teacher to AP. However, this might not be the proper comparison given the importance of the role of AP.[19] Thus, in columns (3) and (6), I restrict the sample to look at only APs. The sample size falls a great deal and the results lose significance due to loss of power, but the point estimates are still positive. These results suggests that for the transition from AP to principal, teacher VA is still positively related to promotion, though the relationship is weaker. This makes intuitive sense if teacher VA is a positive but weak predictor of management success.

[19] 86% of first time principals are employed as APs in the year prior.

5 Robustness

5.1 Alternative estimators

Importantly, I have thus far estimated LPMs for understanding the relationship between VA and becoming AP. While the LPM has several benefits, such as ease of interpretability and estimation, it also has limitations. I discuss these limitations and the alternative estimators I use to address whether my main results are robust to the choice of estimator. The alternative estimators I consider include the logistic regression, survival analysis, and competing risk analysis.

First, I consider the use of a logistic regression. A standard logistic regression estimates the relationship between linear independent variables and the log odds ratio of a binary outcome occurring. Since the outcome in the main analysis is a binary variable with a very small mean, a LPM may not provide a good fit and the logistic regression may be more suitable. I estimate a logistic regression using maximum likelihood. As in the LPM, I use $\hat{\mu}_j$ and X_{jt} as explanatory variables. The estimated coefficient on VA represents the change in the log odds ratio of becoming AP next year when a teacher has a VA of 1 compared to a VA of 0. Since the odds ratio is not directly comparable to the outcome in the main analysis, I also calculate the average marginal effect of a 1 unit increase in VA, which can be directly compared to the percent impact of VA in the main results using the coefficient on VA divided by the outcome mean.

Second, I turn to survival analysis to investigate whether the main results are biased due to attrition. In this setting with state-level administrative data, attrition is due to no longer teaching at an elementary or middle school in NC. Even if both the LPM and logistic regression provide similar results, they might both provide biased estimates of the relationship between VA and becoming AP if attrition is significant. Survival analysis is a natural method choice when dealing with attrition due to its accommodation of censoring.[20]

[20] For example, a teacher may leave the data after a few years because they leave the profession or move to another state, so we never see them become AP. Additionally, the data is censored for all individuals after 2019, since that is the last year of data.

Survival analysis is a type of time to event analysis, where the event is called the "failure". In this context, the failure is becoming an AP for the first time in the next academic year. I investigate whether one's hazard of becoming AP next year is related to their VA.

Using the Cox Proportional Hazard Model (CPHM) (Cox, 1972), the hazard function $h_j(t)$ for teacher j takes the form

$$h_j(t) = h_0(t)e^{\gamma_1 \hat{\mu}_j + \gamma_2 X_j} \tag{3}$$

where $h_0(t)$ is an arbitrary baseline hazard representing the probability of failure conditional on surviving to time period t. The estimated coefficient $\hat{\gamma}_1$ indicates the change in the hazard rate when a teacher has a VA of 1 compared to a VA of 0. If the coefficient is greater than 1, then higher VA increases one's likelihood of becoming AP. If for example the coefficient is 1.5, then teachers with a VA of 1 are 50% likelier to become AP next year compared to teachers with a VA of 0. This value of 50% can be compared to the percent impact of VA from the main results. I estimate the CPHM using maximum likelihood.

Third, I estimate a variant of the hazard function that accommodates competing risk. The CPHM assumes censoring is uninformative, or orthogonal to the independent variable of interest. The more prevalent a competing risk is, the greater the divergence between the estimates from a CPHM and a competing risk model. Evidence from Goldhaber, Gross and Player (2011) suggests early career attrition does indeed vary by VA, so the results from the CPHM may be biased. I use a competing risk model, which is an extension of the hazard function with multiple risks to failure. The risks I consider are 1. becoming AP next year and 2. leaving the data in a given year. The competing risk model takes the form

$$h_j c(t) = h_{0c}(t)e^{\gamma_1 \hat{\mu}_{jc} + \gamma_2 X_{jc}} \text{where } c = 1, 2 \tag{4}$$

with competing risks c. The subhazard model for $c = 1$ is estimated using maximum likelihood per the methods in Fine and Gray (1999). The estimated coefficient $\hat{\gamma}_1$ will indicate the change in the subhazard rate of becoming AP next year for a teacher with a VA

1 of compared to a teacher with a VA of 0. Similar to the CPHM, a coefficient greater than 1 indicates an increased risk of failure and can be compared to the main results in the same way.

In Table 7, I report results using these methods. The logistic regressions contain the same controls as mentioned previously. Since the survival analysis and competing risk analysis are estimated as functions of experience, I remove experience as a control variable.

In column (1), I find that for math teachers, being a 1 SD higher VA teacher increases the odds ratio for becoming AP next year by 36.0%. I also compute the average marginal effect in percent terms for comparability to the LPM results. Using a logistic regression, 1 SD higher VA math teachers are on average 28.9% likelier to become AP next year relative to the average teacher. In column (2), I employ the survival analysis and estimate how VA impacts one's risk of becoming AP next year. Based on the survival analysis, a 1 SD higher VA teachers' risk of becoming AP next year is 28.3% higher. In column (3), I further account for the competing risk of leaving the NC data. These results imply a 1 SD higher VA math teacher's risk of becoming AP next year is 36.1% higher. These values are similar to the value of 33.7% that I find using the LPM.

Similarly for reading teachers, I report the results for logit, survival analysis, and competing risk analysis in columns (3) through (5). The average marginal effect from the logit model implies that 1 SD higher reading teachers are 6.8% likelier to become AP next year relative to the average teacher. For the survival analysis and competing risk analysis, a 1 SD higher reading VA teacher's risk of becoming AP next year is 3.6% and 9.5% higher. Similar to the result of 7.9% that I find using the LPM, these results are small statistically insignificant. In summary, I find that the results using alternative estimators for both math and reading teachers are all in line with LPM results I report in columns (1) and (3) of Table 4.

5.2 Additional robustness checks

Finally, I consider a variety of robustness checks on the main specification. These checks include investigating the robustness of the LPM to the method of estimating VA, choice of controls, and choice of sample.

Table 8 explores different ways of estimating VA. First, I calculate VA in the same way as the main specification without the school fixed effect. The choice to use a school fixed effect in the VA estimating equation has benefits and drawbacks. While it can reduce bias in the estimate of VA, it also restricts the identification to be based on within-school differences and thus relies on interconnected networks of teachers working at multiple schools ("switchers") for comparison. With few years of data, this can cause issues, but given that I have 12 years of data, my estimation likely doesn't suffer from this problem. The results in column (1) for math teachers and column (4) for reading teachers highlight that the main results still hold without using school fixed effects in the VA calculation, though the magnitude of the coefficient is larger and more significant for reading teachers.

Second, I revert back to the main specification using the school fixed effect but remove controls for classroom and school-by-academic-year means of student characteristics. Use of these controls can over-attribute the teacher's impact on students to the coefficients on those variables. The results in column (2) for math teachers and in column (5) for reading teachers show that the main results still hold when removing these controls in the VA calculation.

Third, I employ the drift-adjusted VA model from Chetty, Friedman and Rockoff (2014a) that accounts for changes in a teacher's VA over time. While the estimating equation for this method matches the one used in the main specification, the Chetty, Friedman and Rockoff (2014a) method uses a jackknife (leave one year out) estimator to find a teacher-by-year measure of VA rather than a precision weighted average and Bayesian shrinkage factor. Using this jackknife method forces the sample size to drop because it requires that teachers be observed in more than one year across 2008-2019. However, VA with this method can closer approximate one's effectiveness in each year, rather than the constant component of their effectiveness over time. Regardless, the results in columns (3) and (6) are very similar to the main results.

Next, I investigate the choice of controls used in the estimation of equation 2. Columns (1) and (4) of Table 9 report the results for math and reading, respectively, with VA as the only independent variable. Columns (2) and (5) report the main results, which include controls for demographic factors, education, licensing, etc. Columns (3) and (6) also include school fixed effects to control for time-invariant school factors that might impact a teacher's employment in NC, such as a school being close to a border with another state. For both math and reading teachers, the results are robust to the inclusion of additional covariates.

Finally, I investigate robustness to choices in the sample restrictions and employment panel creation in Table 10. Columns (1) and (3) no longer restrict to just using teachers when estimating equation 2. This means the sample now includes employees in other positions, such as instructional facilitators and instructional support personnel.[21] This is an important restriction to check since about one third of APs were in a position other than teacher in the year before becoming AP. I find that the results are similar for both math and reading teachers when including additional employees. In columns (2) and (4), I define teachers and APs by whether they were employed at all in one of those positions in a school year rather than by the highest percent of employment. The main results hold when this definition is altered.

I summarize these robustness checks in Figure 3. None of these choices qualitatively impact the main results.

6 Discussion

This paper has shown that in North Carolina public elementary and middle schools, more effective teachers are more likely to become school leaders. What are the consequences of this phenomenon for student achievement? If teacher and principal effectiveness are positively correlated, then this likely leads to net improvements in student achievement since the best

[21]Most of people in these positions were once teachers, which allows for calculation of VA for those individuals.

teachers (who are also the best principals) are promoted. However, if there is not a strong positive relationship between the two, then this could be detrimental since classrooms are losing their best teachers. In this section, I explore the average consequences of the current promotion practices. I abstract away from the time component of promotion from teacher to principal, which takes several years and involves the intermediate role of AP, both for ease of calculation and because there is currently no literature measuring AP effectiveness.

First, I need to obtain a measure of principal effectiveness. As previously mentioned, there is a growing body of research focused on estimating principal VA to student achievement (Branch, Hanushek and Rivkin, 2008, 2012; Coelli and Green, 2012; Dhuey and Smith, 2014, 2018; Li, 2015; Bartanen, 2020). Though there are some methodological issues with the estimation of principal VA, it is currently the best available method for estimating the impact an individual principal has on their students.[22] I estimate principal VA to student standardized test scores across the years 2008-2019 for 4,352 educators, which is summarized in Table A5. I only focus on math VA for this exercise since it had the strongest relationship with promotion for teacher VA. I follow the same steps as I did when estimating teacher VA, including regressing student standardized test scores on observable determinants of student achievement, averaging residuals at the principal-year, taking a precision-weighted average across years, and multiplying by a Bayesian shrinkage factor. The standard deviation is similar to what is found in the existing principal VA literature. The interpretation is that a principal who is 1 SD above the average improves math standardized test scores by 15.7% of a SD at their school.

Given the amount of time it takes to progress from the role of teacher to principal, I am only able to calculate both teacher and principal math VA for 301 educators. I first find the correlation between the two measures to be 0.1527. I plot the teacher and principal math VA in Figure A1 along with the line of best fit, which highlights the modest positive correlation between the two measures.

Next, I breakdown educators into four groups based on teacher and principal VA, which

[22]See Grissom, Kalogrides and Loeb (2015); Bartanen and Husain (2021) for discussions of methodological issues.

I denote as VA^t and VA^p, respectively. These groups are:

1. $VA^t > 0$, $VA^p > 0$: above average as both teacher and principal

2. $VA^t > 0$, $VA^p \leq 0$: above average as teacher, below average as principal

3. $VA^t \leq 0$, $VA^p > 0$: below average as teacher, above average as principal

4. $VA^t \leq 0$, $VA^p \leq 0$: below average as both teacher and principal

The share of educators I observe in each of the four categories is summarized in Table A6. This group of educators is likely skewed toward having a higher share of more effective principals because these individuals would need to quickly progress from teaching to becoming principal to be observed in both samples. This is evidenced by 60% of the educators having above average principal VA.

How can we quantify the impact on student achievement when an educator leaves teaching to become a school leader? When this happens, two groups of students are impacted. First, there is a group of students that is no longer taught by that teacher. Second, there is a group of students who now has that educator as their principal. Whether this transition leads to test score gains or losses depends on 1. the sign of the educator's principal and teacher VA, 2. the relative magnitude of the educator's principal and teacher VA, and 3. the share of students impacted by that teacher at the school they're leaving. I summarize these two effects with two equations. First, I define the average change in student test scores at school s_1 as:

$$(0 - VA^t) \times N_c/N_{s_1} = -VA^t \times N_c/N_{s_1} \tag{5}$$

where the size of the classroom the teacher leaves is N_c, the size of the school is N_{s_1}, and the teacher is replaced by someone with average math VA (which equals 0). I define the average change in student test scores at school s_2 as

$$(VA^p - 0) \times N_{s_2}/N_{s_2} = VA^p \tag{6}$$

where the size of the school the teacher goes to is N_{s_2} and they are replacing a principal with average effectiveness (which equals 0). Assuming all schools are the same size for simplicity, the conclusion of whether a teacher becoming principal leads to test score gains or losses reduces to comparing the magnitude of equations 5 and 6.

For the four groups of educators, we can assign whether there is on average a net positive or negative impact on overall student achievement when the teacher becomes principal. For group 2, the effect is always negative because both equations 5 and 6 will be negative. For group 4, the effect is always positive because equations 5 and 6 are both positive. For groups 1 and 3, the sign of the effect depends on a few factors. For group 1, the gain to the school with the new principal exceeds the loss to the classroom if $VA^p/VA^t < N_c/N_{s_1}$. For group 4, the gain to the classroom exceeds the loss to the school with the new principal if the same condition holds.

I quantify whether the current promotion practices lead to average gains or losses to student achievement. For simplicity, I assume an average class size of 30 students and an average school size of 500 students. I estimate the following equation:

$$\frac{1}{301}\sum_{j=1}^{301}(VA_j^p - VA_j^t \times 30/500) \tag{7}$$

which calculates the average change in student achievement for each educator j when they leave teaching to become a principal and averages that value across all educators. I find that this value equals 0.029. This means that the current promotion practices on average lead to test score gains of 2.9% of a standard deviation. Whether this is large depends on comparing it to alternative promotion practices, which is beyond the scope of this paper.

In summary, teacher effectiveness is an imperfect determinant of principal effectiveness. Though 63% of the above average teachers promoted to principal also end up being above average principals, 37% end up being below average principals. Additionally, 56% of the below average teachers promoted to principal end up being above average principals. Thus, the practice of promoting more effective teachers does not deterministically lead to improvements in student achievement. However, teacher effectiveness may be a better signal than other observable factors. Future work investigating this question would be of great use to both schools and policy makers.

7 Conclusion

More effective teachers are more likely to become school leaders. A 1 SD higher math (reading) VA teacher is 33.7% (7.9%) likelier to become an AP in the next academic year relative to the outcome mean of about 0.3%, holding factors such as demographic characteristics, teaching experience, and training constant. These results are estimated using a linear probability model, but the main findings are robust to the use of alternative estimation strategies, including logistic regression, survival analysis, and competing risk analysis, as well as variations in the calculation of VA and controls used in the LPM. Additionally, heterogeneity analyses point to variation in the strength of the relationship between VA and becoming AP by both gender and race/ethnicity of teachers, especially for math teachers. A 1 SD higher male math teacher is 106% likelier to become AP next year compared to the outcome mean, while for female math teachers, the same value is only 23.2%. Additionally, relative to the outcome mean, 1 SD higher VA non-white math teachers are more likely to become AP next year (74.9%) than 1 SD higher VA white math teachers (24.3%).

An important limitation of this study is the non-random assignment of teachers into school leadership. From a policy perspective, we would likely not want school leaders to be randomly assigned, but this limits the ability of the researcher to disentangle the mechanism behind the observed phenomenon. In future work, it would be helpful to obtain more detailed data on the AP and principal hiring process, such as applications and interviews, to disentangle whether higher VA teachers being promoted more often is driven by supply versus demand side factors. Additionally, creative future work that is able to find cases of quasi-random promotion of teachers to school leadership could also increase our understanding of this area.

In terms of policy implications, these results suggest that the current promotion practices in NC elementary and middle schools, which tends to favor the promotion of teachers who skilled at raising student math achievement, can result in gains to student math achievement as long as the relationship between teacher and principal effectiveness is strong enough to offset the loss of the teacher from the classroom and there are enough qualified teachers to replace them once they've been promoted. More work can be done to understand the relationship between teacher and principal effectiveness, since the observed relationship is limited by only observing a small, selected sample of teachers as school leaders.

References

Austin, Wes, Bingjie Chen, Dan Goldhaber, Eric Hanushek, Kris Holden, Cory Koedel, Helen Ladd, Jin Luo, Eric Parsons, Gregory Phelan, et al. 2019. "Path to the Principalship and Value Added: A Cross-State Comparison of Elementary and Middle School Principals. Working Paper No. 213-0119-1." National Center for Analysis of Longitudinal Data in Education Research (CALDER).

Bartanen, Brendan. 2020. "Principal quality and student attendance." Educational Researcher, 49(2): 101–113.

Bartanen, Brendan, and Aliza N Husain. 2021. "Connected Networks in Principal Value-Added Models."

Bartanen, Brendan, and Jason A Grissom. 2021. "School Principal Race, Teacher Racial Diversity, and Student Achievement." Journal of Human Resources, 0218–9328R2.

Branch, Gregory F, Eric A Hanushek, and Steven G Rivkin. 2008. "Principal turnover and effectiveness." Unpublished manuscript.

Branch, Gregory F, Eric A Hanushek, and Steven G Rivkin. 2012. "Estimating the effect of leaders on public sector productivity: The case of school principals." National Bureau of Economic Research.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014a. "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates." American Economic Review, 104(9): 2593–2632.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014b. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." American Economic Review, 104(9): 2633–79.

Coelli, Michael, and David A Green. 2012. "Leadership effects: School principals and student outcomes." Economics of Education Review, 31(1): 92–109.

Cox, David R. 1972. "Regression models and life-tables." Journal of the Royal Statistical Society: Series B (Methodological), 34(2): 187–202.

Dee, Thomas S. 2005. "A teacher like me: Does race, ethnicity, or gender matter?" American Economic Review, 95(2): 158–165.

Dhuey, Elizabeth, and Justin Smith. 2014. "How important are school principals in the production of student achievement?" Canadian Journal of Economics/Revue canadienne d'économique, 47(2): 634–663.

Dhuey, Elizabeth, and Justin Smith. 2018. "How school principals influence student learning." Empirical Economics, 54(2): 851–882.

Feng, Li, and Tim R Sass. 2017. "Teacher quality and teacher mobility." Education Finance and Policy, 12(3): 396–418.

Fine, Jason P, and Robert J Gray. 1999. "A proportional hazards model for the subdistribution of a competing risk." Journal of the American statistical association, 94(446): 496–509.

Gershenson, Seth, Cassandra MD Hart, Joshua Hyman, Constance Lindsay, and Nicholas W Papageorge. 2018. "The long-run impacts of same-race teachers." National Bureau of Economic Research.

Goldhaber, Dan, Betheny Gross, and Daniel Player. 2011. "Teacher career paths, teacher quality, and persistence in the classroom: Are public schools keeping their best?" Journal of Policy Analysis and Management, 30(1): 57–87.

Goldhaber, Dan, Kristian Holden, and Bingjie Chen. 2019. "Do more effective teachers become more effective principals." CALDER Working Paper.

Goldring, Ellen, Mollie Rubin, and Mariesa Herrmann. 2021. "The Role of Assistant Principals: Evidence and Insights for Advancing School Leadership." Wallace Foundation.

Grissom, Jason A, David S Woo, and Brendan Bartanen. 2020. "Ready to lead on day one: Predicting novice principal effectiveness with information available at time of hire."

Grissom, Jason A, Demetra Kalogrides, and Susanna Loeb. 2015. "Using student test scores to measure principal performance." Educational evaluation and policy analysis, 37(1): 3–28.

Hanushek, Eric A, John Kain, Daniel O'Brien, and Steven G Rivkin. 2005. "The market for teacher quality."

Jackson, C. Kirabo. 2018. "What do test scores miss? The importance of teacher effects on non–test score outcomes." Journal of Political Economy, 126(5): 2072–2107.

Kane, Thomas J, and Douglas O Staiger. 2008. "Estimating teacher impacts on student achievement: An experimental evaluation." National Bureau of Economic Research.

Koedel, Cory, Kata Mihaly, and Jonah E Rockoff. 2015. "Value-added modeling: A review." Economics of Education Review, 47: 180–195.

Krieg, John M. 2006. "Teacher quality and attrition." Economics of Education review, 25(1): 13–27.

Li, Danielle. 2015. School accountability and principal mobility: How No Child Left Behind affects the allocation of school leaders. Harvard Business School Boston, MA.

Liebowitz, David D, and Lorna Porter. 2019. "The effect of principal behaviors on student, teacher, and school outcomes: A systematic review and meta-analysis of the empirical literature." Review of Educational Research, 89(5): 785–827.

Liebowitz, David D, and Lorna Porter. 2020. "Descriptive evidence on school leaders' prior professional experiences and instructional effectiveness."

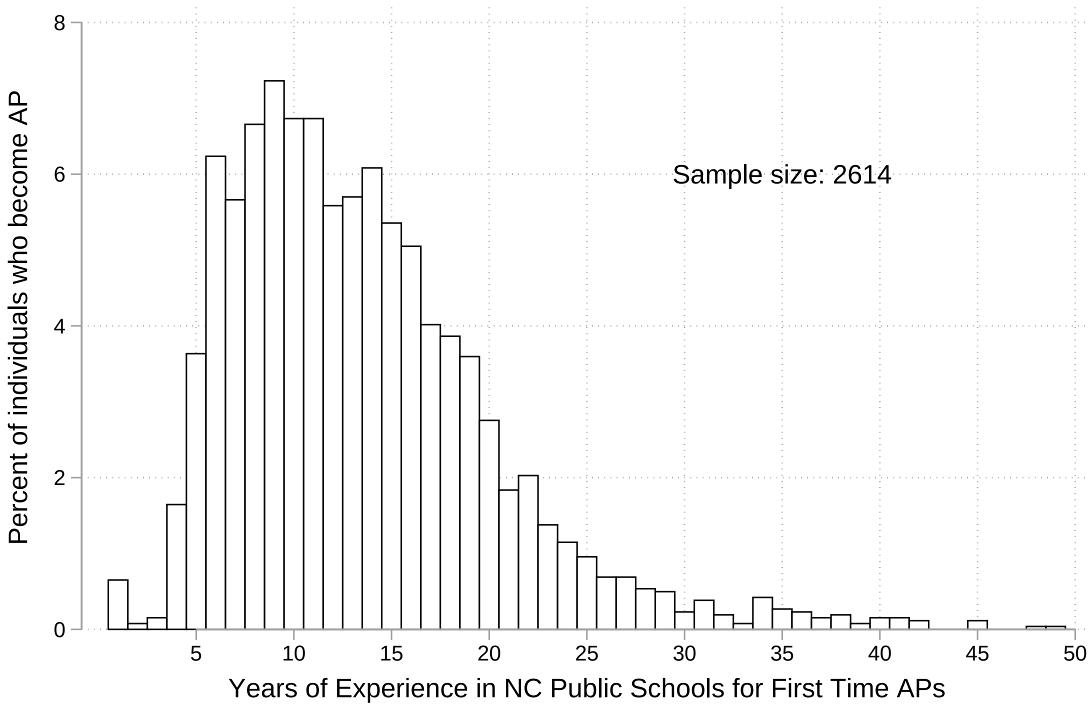
Miller, Ashley. 2013. "Principal turnover and student achievement." Economics of Education Review, 36: 60–72.

Peter, Laurence J, Raymond Hull, et al. 1969. The peter principle. Vol. 4, Souvenir Press London.

West, Martin R, and Matthew M Chingos. 2009. "Teacher effectiveness, mobility, and attrition in Florida." Performance incentives: Their growing impact on American K-12 education, 251–71.

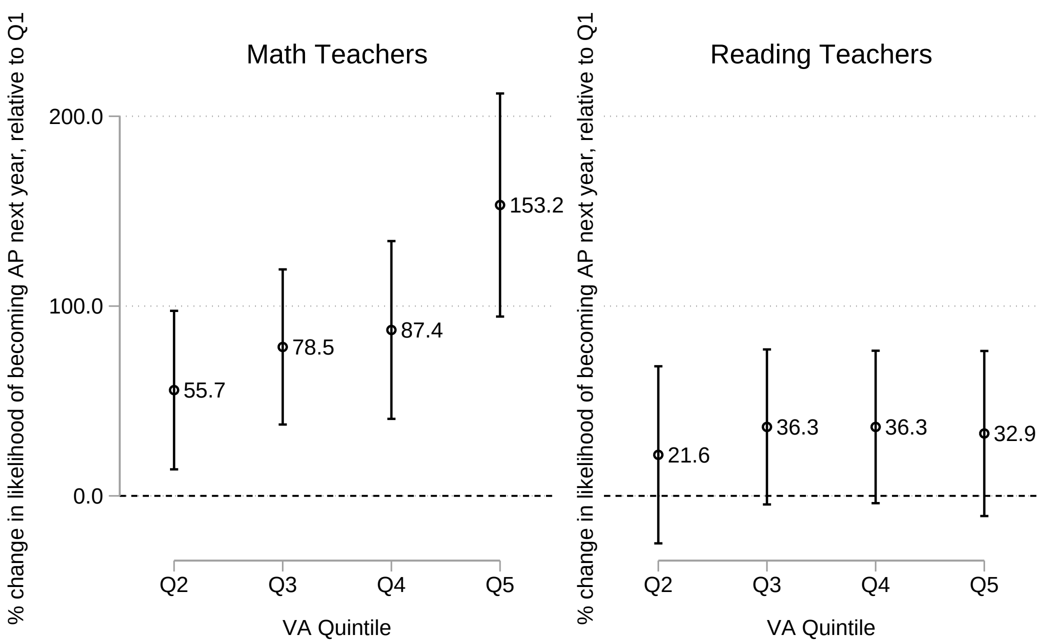
Figures and Tables

Figure 1: Experience when Becoming AP for the First Time



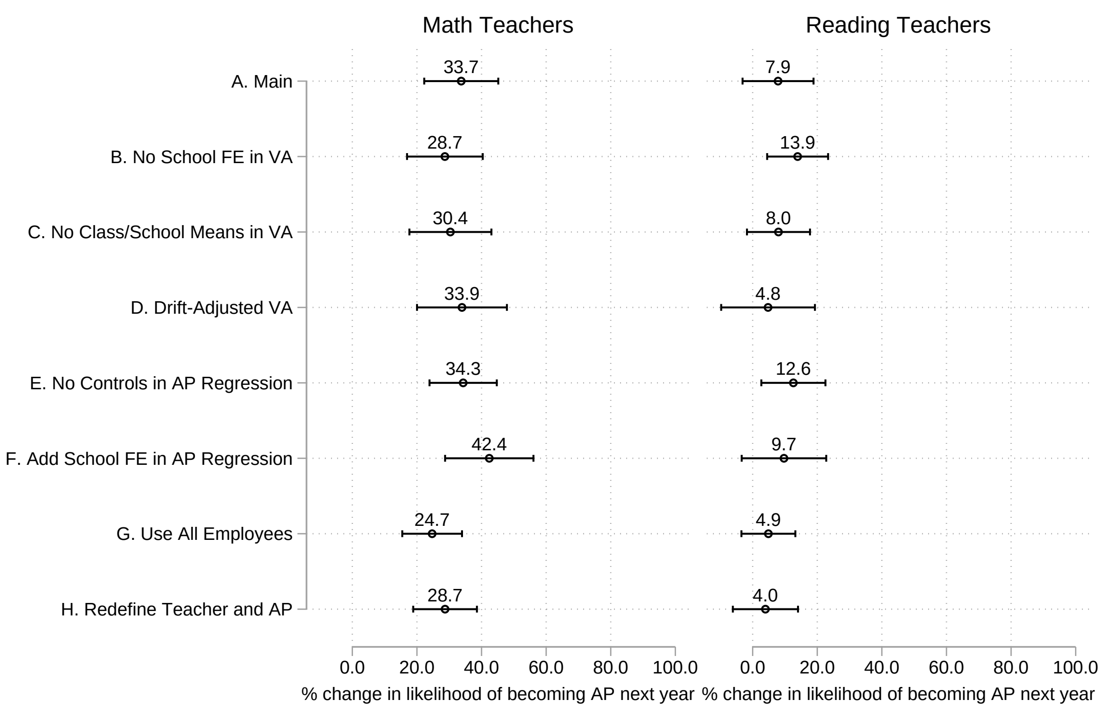
Note: This figure uses employment data from 2012-2019. Each of the 2,614 observations represents one employee who become AP for the first time in the data years 2012-2019.

Figure 2: 95% Confidence Intervals for VA Quintiles



Note: Point estimates and 95% confidence intervals are reported using columns (2) and (4) from Table 4.

Figure 3: 95% Confidence Intervals Summarizing Robustness Checks



Note: Point estimates and 95% confidence intervals are reported. Row A. reports the results from Table 4 column (1) for math and (3) for reading. Rows B. through D. report the results from Table 8 columns (1)-(3) for math and (4)-(6) for reading. Rows E. and F. report the results from Table 9 columns (1) and (3) for math and (4) and (6) for reading. Rows G. and H. report the results from Table 10 columns (1) and (2) for math and (3) and (4) for reading.

Table 1: Summary Statistics for 4th-8th Grade Students, 2008-2019

	Mean	SD
Normalized math score	0.004	0.946
Normalized reading score	0.097	0.955
Female	0.501	0.500
Age	12.420	1.483
Limited English proficiency	0.049	0.216
Economic disadvantage	0.498	0.500
White	0.530	0.499
Black	0.242	0.428
Hispanic	0.148	0.355
Other race/ethnicity	0.081	0.272

Note: This table contains the sample of 5,305,100 students in 4th-8th grade in the years 2008-2019 who have end of year math and/or reading test scores and are successfully matched to their math and/or reading teachers as well as demographic data.

Table 2: Summary Statistics for Teachers 2012-2018

	Math or Reading VA	No VA	P-value
Female	0.8885	0.8655	0.000
	(0.315)	(0.341)	
White	0.8420	0.8331	0.001
	(0.365)	(0.373)	
Black	0.1294	0.1261	0.168
	(0.336)	(0.332)	
Hispanic	0.0112	0.0214	0.000
	(0.105)	(0.145)	
Other race/ethnicity	0.0173	0.0194	0.033
	(0.130)	(0.138)	
Age	40.6778	42.1782	0.000
	(10.740)	(11.456)	
Experience	12.4216	13.6583	0.000
	(8.219)	(9.161)	
Bachelors	0.6164	0.6399	0.000
	(0.486)	(0.480)	
Masters	0.3770	0.3528	0.000
	(0.485)	(0.478)	
Advanced deg	0.0040	0.0043	0.505
	(0.063)	(0.065)	
Doctorate	0.0026	0.0030	0.239
	(0.051)	(0.055)	
National board cert	0.1233	0.1142	0.000
	(0.329)	(0.318)	
Teaching license	0.9615	0.9732	0.000
	(0.192)	(0.161)	
Principal license	0.0371	0.0253	0.000
	(0.189)	(0.157)	
Superintendent license	0.0014	0.0014	0.862
	(0.037)	(0.038)	
Elem school	0.5865	0.7145	0.000
	(0.492)	(0.452)	
Elem and middle	0.0372	0.0343	0.033
	(0.189)	(0.182)	
Middle school	0.3763	0.2512	0.000
	(0.484)	(0.434)	
AP next year	0.0034	0.0025	0.000
	(0.058)	(0.050)	
Observations	161,578	260,058	
Teachers	35,778	65,036	

Note: This table contains the sample of teachers in the years 2012-2018 who worked at elementary, middle, or combined elementary and middle schools. Means are presented with standard deviations below in parentheses. P-value comes from a regression of the variable of interest on an indicator for having math or reading VA. Standard errors are clustered at the teacher level since there are multiple observations per teacher, and p-value comes from the statistical test of whether the coefficient on the indicator for having math or reading VA is different from zero.

Table 3: VA Summary Statistics

	Mean	SD	Min	Max	N
Math VA	-0.013	0.160	-0.838	0.826	33,815
Math VA - Analysis Sample	-0.011	0.163	-0.838	0.826	26,077
Reading VA	0.003	0.104	-0.564	0.620	35,377
Reading VA - Analysis Sample	0.002	0.106	-0.564	0.595	27,239

Note: Math (Reading) VA include all teachers from 2008-2019 who taught math (reading) to 4th-8th graders in at least one year. Math (Reading) VA - Analysis Sample includes only the subsample of math (reading) teachers who are in the employment data from 2012-2018, do not have missing demographic data, and teach in elementary, middle, or combined elementary and middle schools.

Table 4: VA and Likelihood of Becoming AP Next Year

	Math Teachers		Reading Teachers	
	(1)	(2)	(3)	(4)
VA	0.00113***		0.00025	
	(0.00020)		(0.00018)	
VA Q2		0.00110***		0.00050
		(0.00042)		(0.00055)
VA Q3		0.00155***		0.00084*
		(0.00041)		(0.00048)
VA Q4		0.00172***		0.00084*
		(0.00047)		(0.00047)
VA Q5		0.00302***		0.00076
		(0.00059)		(0.00051)
Pct. Effect	33.7		7.9	
Pct. Effect - Q2		55.7		21.6
Pct. Effect - Q3		78.5		36.3
Pct. Effect - Q4		87.4		36.3
Pct. Effect - Q5		153.2		32.9
Outcome Mean	0.00336		0.00323	
Outcome Mean, Q1		0.00197		0.00232
N	120,659	120,659	125,614	125,614

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are clustered at the school district. In all columns, the outcome is a binary indicator for becoming AP in the next academic year. In columns (1) and (2), only math teachers and math VA are used. In columns (3) and (4), only reading teachers and reading VA are used. In columns (1) and (3), the independent variable of interest is math and reading VA, respectively, which is normalized to have a mean of zero and standard deviation of one. In columns (2) and (4), the independent variables of interest are math and reading VA quintile, respectively. The reference category is the first quintile. In all columns, there are also indicators for race/ethnicity, gender, education, national board certification, licensing level, school type, and academic year, as well as a second order polynomial of age and indicators for teaching experience in 5 year intervals.

Table 5: VA and Becoming AP - Heterogeneity by Gender and Race/Ethnicity

	Math Teachers		Reading Teachers	
	(1)	(2)	(3)	(4)
VA	0.00357***	0.00252***	0.00119	0.00033
	(0.00065)	(0.00064)	(0.00076)	(0.00049)
Female	-0.00450***		-0.00344***	
	(0.00059)		(0.00086)	
Female X VA	-0.00279***		-0.00104	
	(0.00065)		(0.00076)	
White		-0.00192**		-0.00197**
		(0.00080)		(0.00085)
White X VA		-0.00170***		-0.00010
		(0.00063)		(0.00046)
Pct. Effect - Women	23.2		4.8	
Pct. Effect - Men	106.0		36.9	
Pct. Effect - White		24.3		7.0
Pct. Effect - Non-White		74.9		10.2
Outcome Mean	0.00336	0.00336	0.00323	0.00323
N	120,659	120,659	125,614	125,614

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are clustered at the school district. For brevity, the variables Black, Hispanic, and Other race/ethnicity are used in the regressions in columns (1) and (3) but the coefficient is not reported in this table. Similarly, the variable for Female is used in the regressions in columns (2) and (4) but the coefficient is not reported in this table. In all columns, the outcome is a binary indicator for becoming AP in the next academic year. In columns (1) and (2), only math teachers and math VA are used. In columns (3) and (4), only reading teachers and reading VA are used. Both math and reading VA are normalized to have a mean of zero and standard deviation of one. In all columns, there are also indicators for education, national board certification, licensing level, school type, and academic year, as well as a second order polynomial of age and indicators for teaching experience in 5 year intervals.

Table 6: VA and Alternative Outcomes

	Math Teachers			Reading Teachers		
	(1) Principal License	(2) Become Principal	(3) Become Principal, Conditional on AP	(4) Principal License	(5) Become Principal	(6) Become Principal, Conditional on AP
VA	0.00146***	0.00048***	0.01000	0.00077***	0.00011	0.00686
	(0.00026)	(0.00012)	(0.00745)	(0.00025)	(0.00009)	(0.00705)
Pct. Effect	20.2	32.3	11.8	11.0	7.7	8.6
Outcome Mean	0.00721	0.00148	0.08479	0.00700	0.00138	0.07985
N	116,044	127,617	1,663	120,843	133,280	1,603

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors are clustered at the school district. In all columns, the outcome is a binary indicator. In columns (1) and (4), the indicator equals 1 if the teacher obtains a principal license in the next academic year. In columns (2), (3), (5), and (6), the indicator equals one if the individual becomes principal in the next academic year. Columns (3) and (6) restrict the sample to only be APs. In columns (1)-(3), only math teachers and math VA are used. In columns (4)-(6), only reading teachers and reading VA are used. VA is normalized to have a mean of zero and standard deviation of one. In all columns, there are also indicators for race/ethnicity, gender, education, national board certification, school type, and academic year, as well as a second order polynomial of age and indicators for teaching experience in 5 year intervals. Licensing is only used as a control in columns (2), (3), (5), and (6).

Table 7: VA and Becoming AP - Alternative Estimators

	Math Teachers			Reading Teachers		
	(1) Logistic	(2) Hazard	(3) Competing Hazard	(4) Logistic	(5) Hazard	(6) Competing Hazard
VA	1.360***	1.283***	1.361***	1.074	1.036	1.095
	(0.069)	(0.060)	(0.065)	(0.062)	(0.054)	(0.061)
Marginal Effect (Pct.)	28.9			6.8		
N	120,659	120,659	120,659	125,614	125,614	125,614

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Standard errors are clustered at the school district. In columns (1) and (4), I use a logistic regression with the outcome as the odds ratio of becoming AP next year. In columns (2) and (5), I estimate a hazard function with the outcome as becoming AP next year, conditional on surviving to the current year. In columns (3) and (6), I estimate a competing hazard model, focusing on the subhazard of becoming AP and accounting for the competing risk of leaving the data, where the outcome is becoming AP next year, conditional on surviving to the current year. In columns (1)-(3), only math teachers and math VA are used. In columns (4)-(6), only reading teachers and reading VA are used. In columns (1)-(3) and (4)-(6), the independent variable of interest is math and reading VA, respectively, which is normalized to have a mean of zero and standard deviation of one. In columns (2) and (4), the independent variables of interest are math and reading VA quintile, respectively. All models are estimated using maximum likelihood. All models additionally contain indicators for race/ethnicity, gender, education, national board certification, licensing level, school type, and academic year, as well as a second order polynomial of age. The logistic regressions also include indicators for teaching experience in 5 year intervals.

Table 8: VA and Becoming AP - Robustness to Choice of VA Method

	Math Teachers			Reading Teachers		
	(1) No School FE	(2) No Class and School Means	(3) Drift Adjusted	(4) No School FE	(5) No Class and School Means	(6) Drift Adjusted
VA	0.00096***	0.00102***	0.00110***	0.00045***	0.00026	0.00015
	(0.00020)	(0.00022)	(0.00023)	(0.00016)	(0.00016)	(0.00023)
Pct. Effect	28.7	30.4	33.9	13.9	8.0	4.8
Outcome Mean	0.00336	0.00337	0.00325	0.00323	0.00323	0.00309
N	120,656	120,636	65,548	125,619	125,609	67,247

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Standard errors are clustered at the school district. In columns (1) and (4), VA is calculated without the use of school fixed effects. In columns (2) and (5), VA is calculated without the use of classroom and school-by-academic-year means of student characteristics. In columns (3) and (6), VA is calculated using the drift-adjusted method from Chetty, Friedman and Rockoff (2014a). In all columns, the outcome is a binary indicator for becoming AP in the next academic year. In columns (1)-(3), only math teachers and math VA are used. In columns (4)-(6), only reading teachers and reading VA are used. Both math and reading VA are normalized to have a mean of zero and standard deviation of one. In all columns, there are also indicators for race/ethnicity, gender, education, national board certification, licensing level, school type, and academic year, as well as a second order polynomial of age and indicators for teaching experience in 5 year intervals.

Table 9: VA and Becoming AP - Robustness to Choice of Controls in AP Regression

	Math Teachers			Reading Teachers		
	(1)	(2)	(3)	(4)	(5)	(6)
VA	0.00115***	0.00113***	0.00143***	0.00041**	0.00025	0.00031
	(0.00018)	(0.00020)	(0.00023)	(0.00016)	(0.00018)	(0.00022)
Covariates		X	X		X	X
School FE			X			X
Pct. Effect	34.3	33.7	42.4	12.6	7.9	9.7
Outcome	0.00336	0.00336	0.00336	0.00323	0.00323	0.00323
N	120,659	120,659	120,659	125,614	125,614	125,614

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Standard errors are clustered at the school district. In all columns, the outcome is a binary indicator for becoming AP in the next academic year. In columns (1)-(3), only math teachers and math VA are used. In columns (4)-(6), only reading teachers and reading VA are used. Both math and reading VA are normalized to have a mean of zero and standard deviation of one. In columns (1) and (4), there are no additional covariates. In columns (2) and (5), there are also indicators for race/ethnicity, gender, education, national board certification, licensing level, school type, and academic year, as well as a second order polynomial of age and indicators for teaching experience in 5 year intervals. In columns (3) and (6), there are additionally school fixed effects.

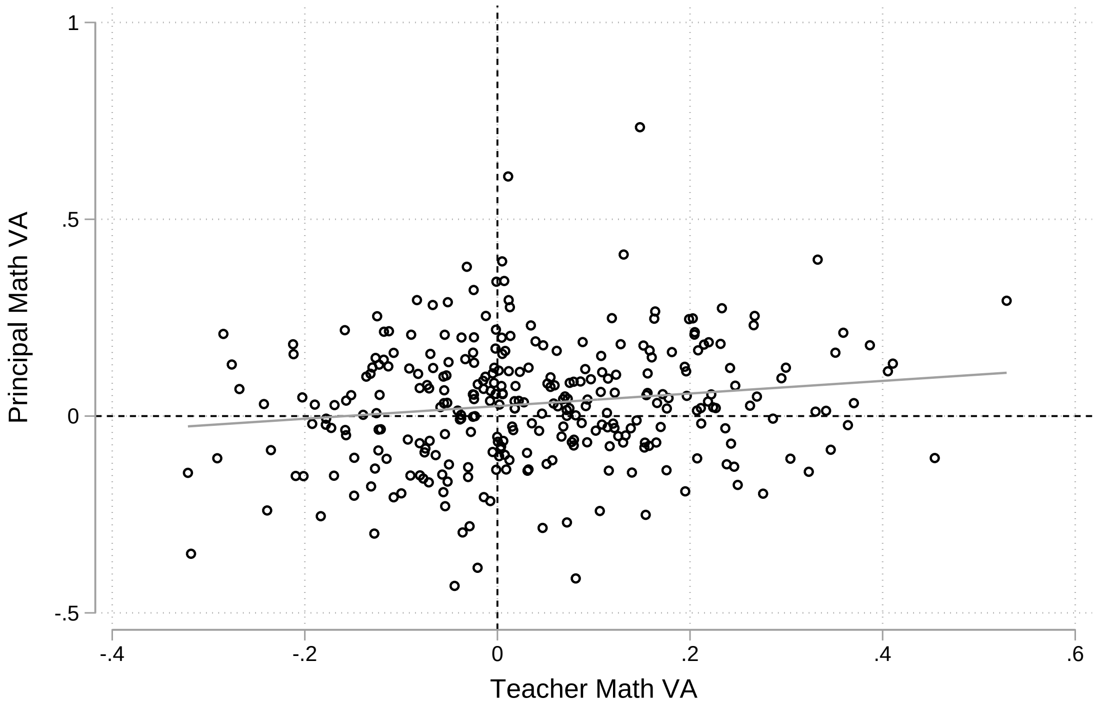
Table 10: VA and Becoming AP - Robustness to Data Choices

	Math Teachers		Reading Teachers	
	(1) Use All Employ-ees	(2) Redefine Teacher and AP	(3) Use All Employ-ees	(4) Redefine Teacher and AP
VA	0.00114***	0.00099***	0.00022	0.00013
	(0.00022)	(0.00017)	(0.00019)	(0.00017)
Pct. Effect	24.7	28.7	4.9	4.0
Outcome Mean	0.00463	0.00344	0.00450	0.00325
N	124,935	122,291	130,678	127,402

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Standard errors are clustered at the school district. In all columns, the outcome is a binary indicator for becoming AP in the next academic year. In columns (1)-(2), only math teachers and math VA are used. In columns (3)-(4), only reading teachers and reading VA are used. Both math and reading VA are normalized to have a mean of zero and standard deviation of one. In columns (1) and (3), there are indicators for race/ethnicity, gender, education, national board certification, licensing level, school type, and academic year, as well as a second order polynomial of age and indicators for teaching experience in 5 year intervals. In columns (2) and (4), there the same controls except for school type.

A Appendix

Figure A1: Scatterplot of Teacher and Principal VA



Note: Includes 301 educators with teacher and principal VA across 2008-2019. Each circle represents one educator. Line of best fit included.

Table A1: Summary Statistics for 4th-8th Grade Teachers 2012-2018

	Math or Reading VA	No VA	P-value
Female	0.8885	0.6954	0.000
	(0.315)	(0.460)	
White	0.8420	0.8109	0.000
	(0.365)	(0.392)	
Black	0.1294	0.1438	0.001
	(0.336)	(0.351)	
Hispanic	0.0112	0.0247	0.000
	(0.105)	(0.155)	
Other race/ethnicity	0.0173	0.0206	0.016
	(0.130)	(0.142)	
Age	40.6778	42.7052	0.000
	(10.740)	(11.406)	
Experience	12.4216	13.8031	0.000
	(8.219)	(9.226)	
Bachelors	0.6164	0.6392	0.003
	(0.486)	(0.480)	
Masters	0.3770	0.3517	0.001
	(0.485)	(0.478)	
Advanced deg	0.0040	0.0044	0.389
	(0.063)	(0.066)	
Doctorate	0.0026	0.0047	0.000
	(0.051)	(0.068)	
National board cert	0.1233	0.1036	0.000
	(0.329)	(0.305)	
Teaching license	0.9615	0.9656	0.036
	(0.192)	(0.182)	
Principal license	0.0371	0.0328	0.027
	(0.189)	(0.178)	
Superintendent license	0.0014	0.0015	0.582
	(0.037)	(0.039)	
Elem school	0.5865	0.3573	0.000
	(0.492)	(0.479)	
Elem and middle	0.0372	0.0334	0.096
	(0.189)	(0.180)	
Middle school	0.3763	0.6093	0.000
	(0.484)	(0.488)	
AP next year	0.0034	0.0033	0.717
	(0.058)	(0.057)	
Observations	161,578	92,374	
Teachers	35,778	24,355	

Note: This table contains the sample of teachers in the years 2012-2018 who worked at elementary, middle, or combined elementary and middle schools and taught 4th-8th graders. Means are presented with standard deviations below in parentheses. P-value comes from a regression of the variable of interest on an indicator for having math or reading VA. Standard errors are clustered at the teacher level since there are multiple observations per teacher, and p-value comes from the statistical test of whether the coefficient on the indicator for having math or reading VA is different from zero.

Table A2: Summary Statistics for Teachers 2012-2018 with Math and Reading VA

	Math VA	Reading VA	P-value
Female	0.8846	0.9113	0.000
	(0.320)	(0.284)	
White	0.8476	0.8468	0.000
	(0.359)	(0.360)	
Black	0.1230	0.1266	0.664
	(0.328)	(0.333)	
Hispanic	0.0118	0.0115	0.000
	(0.108)	(0.107)	
Other race/ethnicity	0.0177	0.0151	0.000
	(0.132)	(0.122)	
Age	40.3807	40.6118	0.000
	(10.611)	(10.706)	
Experience	12.2813	12.3575	0.000
	(8.084)	(8.115)	
Bachelors	0.6259	0.6074	0.000
	(0.484)	(0.488)	
Masters	0.3682	0.3866	0.000
	(0.482)	(0.487)	
Advanced deg	0.0036	0.0035	0.042
	(0.060)	(0.059)	
Doctorate	0.0024	0.0025	0.180
	(0.049)	(0.050)	
National board cert	0.1192	0.1287	0.000
	(0.324)	(0.335)	
Teaching license	0.9617	0.9621	0.000
	(0.192)	(0.191)	
Principal license	0.0368	0.0365	0.000
	(0.188)	(0.188)	
Superintendent license	0.0015	0.0013	0.560
	(0.039)	(0.036)	
Elem school	0.6892	0.6849	0.000
	(0.463)	(0.465)	
Elem and middle	0.0361	0.0371	0.104
	(0.186)	(0.189)	
Middle school	0.2748	0.2780	0.000
	(0.446)	(0.448)	
AP next year	0.0034	0.0032	0.005
	(0.058)	(0.057)	
Observations	120,659	125,614	
Teachers	26,077	27,239	

Note: This table contains the sample of teachers in the years 2012-2018 who worked at elementary, middle, or combined elementary and middle schools. Means are presented with standard deviations below in parentheses. P-value comes from a regression of the variable of interest on an indicator for having math or reading VA. Standard errors are clustered at the teacher level since there are multiple observations per teacher, and p-value comes from the statistical test of whether the coefficient on the indicator for having math or reading VA is different from zero.

Table A3: VA and Likelihood of Becoming AP Next Year

	Math Teachers	Reading Teachers
	(1)	(2)
VA	0.00113***	0.00025
	(0.00020)	(0.00018)
Female	-0.00432***	-0.00334***
	(0.00066)	(0.00089)
Black	0.00273***	0.00238**
	(0.00099)	(0.00096)
Outcome Mean	0.00336	0.00323
N	120,659	125,614

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Standard errors are clustered at the school district. In all columns, the outcome is a binary indicator for becoming AP in the next academic year. In columns (1) and (2), only math teachers and math VA are used. In columns (3) and (4), only reading teachers and reading VA are used. In columns (1) and (3), the independent variable of interest is math and reading VA, respectively, which is normalized to have a mean of zero and standard deviation of one. In columns (2) and (4), the independent variables of interest are math and reading VA quintile, respectively. The reference category is the first quintile. In all columns, there are also indicators for race/ethnicity, gender, education, national board certification, licensing level, school type, and academic year, as well as a second order polynomial of age and indicators for teaching experience in 5 year intervals.

Table A4: VA and Becoming AP - Heterogeneity by Gender and Race/Ethnicity

	Math Teachers		Reading Teachers	
	(1)	(2)	(3)	(4)
VA	0.00357***	0.00082***	0.00119	0.00023
	(0.00065)	(0.00017)	(0.00076)	(0.00015)
Female	-0.00450***		-0.00344***	
	(0.00059)		(0.00086)	
Female X VA	-0.00279***		-0.00104	
	(0.00065)		(0.00076)	
Black		0.00261***		0.00236**
		(0.00093)		(0.00098)
Hispanic		0.00038		-0.00017
		(0.00149)		(0.00135)
Other race/ethnicity		-0.00184*		0.00011
		(0.00103)		(0.00158)
Black X VA		0.00223***		0.00013
		(0.00080)		(0.00064)
Hispanic X VA		-0.00019		-0.00092
		(0.00145)		(0.00128)
Other X VA		0.00149**		0.00041
		(0.00062)		(0.00049)
Pct. Effect - Women	23.2		4.8	
Pct. Effect - Men	106.0		36.9	
Pct. Effect - Black		90.7		11.1
Pct. Effect - Hispanic		18.6		-21.4
Pct. Effect - Other		68.5		19.8
Pct. Effect - White		24.3		7.1
Outcome Mean	0.00336	0.00336	0.00323	0.00323
N	120,659	120,659	125,614	125,614

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Standard errors are clustered at the school district. For brevity, the variables Black, Hispanic, and Other race/ethnicity are used in the regressions in columns (1) and (3) but the coefficient is not reported in this table. Similarly, the variable for Female is used in the regressions in columns (2) and (4) but the coefficient is not reported in this table. In all columns, the outcome is a binary indicator for becoming AP in the next academic year. In columns (1) and (2), only math teachers and math VA are used. In columns (3) and (4), only reading teachers and reading VA are used. Both math and reading VA are normalized to have a mean of zero and standard deviation of one. In all columns, there are also indicators for education, national board certification, licensing level, school type, and academic year, as well as a second order polynomial of age and indicators for teaching experience in 5 year intervals.

Table A5: Principal VA Summary Statistics

	Mean	SD	Min	Max	N
Math VA	0.020	0.157	-0.795	0.734	4,352

Note: Includes elementary and middle school principals over the years 2008-2019.

Table A6: Share of Educators and Sign of VA

	$VA^t > 0$	$VA^t \leq 0$
$VA^p > 0$	0.36	0.24
$VA^p \leq 0$	0.21	0.19

Note: Includes 301 educators with teacher and principal VA across 2008-2019.