Managers in Public Schools*

Carolyn Tsao†

November 4, 2024

Click here for latest version.

Abstract

To what extent do public school principals affect student outcomes, and how do managerial practices differ between more and less effectiveness principals? Using administrative data on teacher-principal-student links in two US states and a two-way manager-worker fixed effect framework, I estimate principal and teacher effectiveness. Variance decompositions show that principal effectiveness explains less of the variance in student outcomes than teacher effectiveness does. That said, switching to a more effective principal improves school outcomes: event studies around principal moves and retirements show receiving a more effective principal improves student outcomes (increased test scores, decreased absenteeism) and teacher outcomes (greater teacher retention, increased student test scores within teacher). Importantly, novel survey evidence suggests that principal effectiveness is attributable to particular managerial practices: linking the estimated principal effects to survey data on teachers' perceptions of their school leadership, I find that at schools that employ more effective principals, teachers are more likely to report the use of data-driven instructional practices but also a lack of trust and mutual respect between administration and staff.

*I thank Chris Campos, Sharada Dharmasankar, Adam Kapor, David Lee, Alex Mas, Pablo Muñoz, Chris Neilson, Karl Schulze, Jesse Silbert, and participants of the Industrial Relations Seminar at Princeton, the Industrial Organization Workshop at Princeton, and the Prize Fellowship Seminar at Princeton for their insights and suggestions. I also thank Angie Tombari and Barrett Ross at KYSTATS and Zhixiang Lu at the University of Texas Dallas Education Research Center (UTD ERC) for their support and assistance with the data. I am grateful for the financial support of the Industrial Relations Section at Princeton, the Prize Fellowship in Social Sciences at Princeton, and the Social Sciences and Humanities Council of Canada (SSHRC). Disclaimer: All errors are my own. The views and opinions expressed in this paper are my own and do not necessarily reflect the views of KYSTATS or the UTD ERC.

†Department of Economics, Princeton University. ctsao@princeton.edu

1 Introduction

Managers are responsible for making the personnel decisions critical to all production processes that rely on labor, including hiring/firing, mentoring, and allocating workers. Prior work shows that these personnel decisions can have large impacts on productivity in the private sector (e.g., Bloom and Van Reenen [2007], Ichniowski et al. [1997], Ichniowski and Shaw [2003], Bloom et al. [2013], Syverson [2011], Hoffman and Tadelis [2021]). Less attention however has been paid to managers in the public sector, where workplace institutions like union protections may restrict the range of personnel-related actions a manager can take and thus potentially limit the scope of their effectiveness.

This paper examines the effectiveness of and the actions taken by managers in one such setting: public schools. *Principals* are the managers of the school, and are specifically responsible for hiring and "firing" teachers, monitoring and mentoring teachers' work, and allocating teachers to classes. Given that teacher quality is a key determinant of student outcomes (e.g., Kane and Staiger [2008], Chetty et al. [2014], Jackson [2018], Gilraine and Pope [2021]), principals have the potential to play a key role in determining schools' output.

However, there are a number of challenges involved in quantifying principal productivity and identifying what makes for a productive principal. The first challenge is data availability. Public schools are a relatively advantageous setting to study managers' effects on output and/or workers, as state administrative data on schools typically contains links between principals, teachers, and students' outcomes. However, these data typically do not contain information on managerial styles that can be linked to manager performance. The second challenge is the limited number, and limited mobility, of principals. The primary strategy for identifying manager effectiveness in the literature leverages manager moves across workers (e.g., Bertrand and Schoar [2003], Lazear et al. [2015], Fenizia [2022]). But principals are few in number and change schools so infrequently that principal moves across schools may not provide sufficient variation to separately identify the effects of individual principals versus the effects of schools [Bartanen et al., 2024].

I circumvent these challenges by combining administrative schools data with novel survey data on school leadership practices and a method that identifies principal productivity

using principal moves across *teachers* instead of across schools. I implement the method on data from two statewide US K-12 school systems (Kentucky and Texas). Then, using the resulting estimates of principal productivity, I assess *whether* and *how* principals are able to influence student outcomes in three ways. First, using a variance decomposition approach, I evaluate how much of the variance in student outcomes can be explained by principals as opposed to teachers or schools. Second, leveraging the large number of observed principal moves across schools in Texas, I examine whether and by how much switching to a "more productive" (in terms of student achievement) principal affects other dimensions of a school's student and teacher outcomes. Third, linking the administrative data to novel survey data on teachers' perceptions of school leadership in Kentucky, I examine whether principal productivity is attributable to specific managerial practices as reported in the survey.

I first develop and validate a new method for estimating principal productivity. I propose a simple model in which student outcomes are additively determined by constant teacher effects and principal effects. The model is analogous to the model of worker and firm productivity in Abowd et al. [1999] (henceforth AKM) and Card et al. [2013] (henceforth CHK), but instead of examining worker and firm effects on wages, I examine manager and worker effects on output. Intuitively, the principal effects are identified off of moves of teachers across principals and vice versa, rather than only moves of principals across schools, which circumvents the issue of limited mobility among principals. Using a series of heuristic tests, I demonstrate that the assumptions that need to be satisfied for the principal and teacher effects to be identified hold in the context of Kentucky and Texas schools.

I estimate the model on the largest connected sets of managers (principals) and employees (teachers) in each state to obtain fixed effect estimates of principal and teacher effectiveness.[1] I use a variance decomposition approach to illustrate how much of the variance in student outcomes is explained by principals versus teachers. To address the concern in the prior literature that there may not be enough principal moves to separately identify principal effects from school effects, I also estimate the model on the largest connected set of firms (schools) and employees (teachers) in each state to evaluate whether the result-

[1]The method also delivers a set of principal fixed effects but also a set of teacher fixed effects, which have a similar interpretation to the classic teacher value-add measures (e.g. Kane and Staiger [2008], Chetty et al. [2014]) but net out the mean effect the principal may directly have on the outcome in question.

ing school fixed effect estimates are meaningfully different from the principal fixed effect estimates.

My main finding from the variance decompositions is that principal fixed effects explain some of the variance in student outcomes, but less than teacher or school fixed effects do. In both states, principal fixed effects explain 20-25% of the variance in student outcomes—including student test scores, absences, and suspensions—that teacher fixed effects do. School fixed effects (from the schools-teachers model) explain 1-1.5 times the variance in student outcomes that the principal fixed effects do (from the principals-teachers model), suggesting that schools have persistent effects on their own. Importantly, the variance decompositions also suggest that the method credibly separates principal effects from school effects, in that the decompositions from the schools-teachers and principals-teachers models are meaningfully different.[2] The results are robust to shrinkage, limited mobility bias corrections [Card et al., 2013, Andrews et al., 2008], and more recent approaches to estimating the AKM model, including the k-means clustering approach proposed by Bonhomme et al. [2019] and the multiple connected sets approach used by Fenizia [2022] and Best et al. [2023].

Next, using an event study design that leverages the long panel of data in Texas and the timing of principal retirements and moves, I examine how school outcomes change following switches to more and less effective principals. The principal retirements design bears a causal interpretation in particular, as retirements are determined by age and/or experience and are thus plausibly exogenous to other determinants of school outcomes.

My main finding from the event studies is that switching to a higher-FE principal significantly affects both student outcomes and teacher outcomes on average. Student outcomes improve: schools that switch to higher-FE principals experience increased math and reading test scores within as little as one year, reduced student absenteeism within two to three years, and no change in suspension rates. The stability and quality of teaching staff also improve: schools that switch to higher-FE principals have higher teacher retention rates, more positive selection on teacher value add, and increased within-teacher productivity within three years. Taken together, the event study findings are consistent with principals affecting student

[2]Furthermore, the principal fixed effect estimates from the two models are positively, but far from perfectly, correlated, with a Pearson correlation coefficient of 0.3, further supporting the idea that the estimated school and principal effects are distinct.

outcomes through personnel selection and/or by directly increasing teachers' productivity.

Finally, motivated by the event study finding that switching to a higher-FE principal can increase within-teacher productivity, I investigate *how* principals are able to impact teachers' productivity. Do effective principals raise teachers' productivity directly by implementing specific instructional practices, mentorship, or monitoring? Or do effective principals raise teachers' productivity more indirectly by fostering desirable workplace conditions that teachers work more effectively in?

To explore these questions, I leverage the availability of novel survey data from Kentucky on teachers perceptions of their working conditions. Specifically, in the style of Bender et al. [2018], I take teachers' (anonymous) survey responses, average them to the school-year level, and link the averaged responses to the estimated principal fixed effects from the administrative data. These data provide a unique opportunity to study the relationship between principal productivity and teachers' ratings of their principal's managerial style and of the workplaces their principals influence.

My main finding from the linked survey analysis is that higher-FE principals are strongly associated with instructional practices that emphasize standardized testing performance, but are not necessarily associated with favorable leadership qualities. Teachers who work under higher-FE principals are more likely to report that the school leadership "facilitates using data to improve student learning," but are less likely to agree that "there is an atmosphere of trust and mutual respect" or that "leadership makes a sustained effort to address teachers' concerns" at their school. These results are consistent with higher-FE principals being managers who take active steps to increase productivity (in this case, test scores), but potentially at the expense of establishing a trusting relationship with their employees. That said, I also find that teachers who work under higher-FE principals are more likely to report that they intend to continue working at the same school one year later, which suggests that issues that teachers have with school leadership are not sufficient to deter them from working with the principals in question.

Overall, my findings suggest that public school principals explain a small fraction of variation in student outcomes above and beyond teacher effects. However, despite the restricted set of actions principals can take with regards to personnel management in public

schools, I find evidence that personnel management—through hiring and retention, as well as through direct effects on teacher productivity—is a key channel through which principal effects can arise.

This paper primarily relates to two strands of literature. The first is an extensive body of work studying the role of managers in the firm, from assessing the extent to which managers affect output (e.g., Bloom and Van Reenen [2007], Bertrand and Schoar [2003], Bushnell and Wolfram [2009], Syverson [2011], Ichniowski and Shaw [2003], Lazear et al. [2015], Metcalfe et al. [2023]) to identifying what it is that productive managers do differently from less productive managers (e.g., Ichniowski et al. [1997], Bloom et al. [2013], Fenizia [2022], Hoffman and Tadelis [2021], Bandiera et al. [2020], Bender et al. [2018]). A primary bottleneck in the literature is finding data that allows one to link managers to their workers' output *and* their managerial practices. Many studies overcome this challenge using data from a single private sector firm in which output is measured in a standardized fashion, the manager-worker assignment practices are well understood, and survey data on the firm's managerial practices exists or can be collected. As a result however, little is known about how managerial practices relate to management efficacy in the public sector, where effects may be differ due given public sector worker protections and the nature of public sector bureaucracy. My unique combination of administrative and survey data from Kentucky schools in particular allows me to contribute to a growing literature studying the role of public sector managers' management practices in firm performance (e.g., Fenizia [2022], Best et al. [2023], Bloom et al. [2015]).

The second strand studies principals specifically, with a focus on understanding how important principals are in determining student outcomes (e.g., Branch et al. [2012], Miller [2013], Dhuey and Smith [2014], Muñoz and Prem [2024], Bloom et al. [2015], Bartanen et al. [2024], Hanushek et al. [2024], Coelli and Green [2012], Grissom et al. [2015]) and how principals' personnel management affects school outcomes (e.g., Bryson et al. [2023], Engel et al. [2018], Fryer et al. [2017], Grissom et al. [2013], Grissom and Bartanen [2019], Ingersoll et al. [2017], Jacob [2011], Liebowitz and Porter [2019]). Many of these studies borrow the two-way firm and manager fixed effects approach from the management literature to estimate principal effects. However, credibly separating the effect of the principal from that of the

school requires either randomness in the assignment of principals to schools or observing a high number of principal moves across schools, neither of which tend to occur in the U.S. setting [Bartanen et al., 2024]. My approach circumvents this challenge by leveraging moves of principals across *teachers* as well as across schools, allowing me to use similar public school administrative data as previous papers but with significantly more variation. I contribute a new set of estimates to the literature which suggest that principal quality varies less than previously thought. The only other paper that takes a similar approach to measuring manager quality happens to also study the school setting: Muñoz and Prem [2024], who study the effects of a recruitment reform in Chile on the quality of recruited principals, measure quality by leveraging moves of principals across teachers.

The rest of the paper proceeds as follows. Section 2 describes the administrative and survey data sources used, as well as the general process by which principal-teacher matches occur. Section 3 presents the main econometric framework and Section 4 presents the resulting estimated variance decompositions. Section 5 and 6 present the event study analysis while Section 7 presents the survey-based analysis. Section 8 concludes.

2 Data and Institutional Background

I use administrative and survey data on teachers and principals in the state of Kentucky. The administrative dataset, the Kentucky Data Longitudinal System (KLDS), consists of individual-level panel data on all public school students and staff in Kentucky from 2008 to 2022, obtained through the Kentucky Center for Statistics. The main variables used in the analysis are provided by the Kentucky Department of Education, including: student demographics (age, race/ethnicity, gender, grade, free or reduced price lunch status), student outcomes (e.g. standardized test scores in math and reading; disciplinary events like in-school removal, out-of-school suspension, and expulsion), staff roles and responsibilities, teacher and principal assignments to schools, and teacher and principal demographics (including gender, age, race, salary, and years of experience). I link teachers to their students by constructing unique classroom identifiers using student transcript data, which lists the main classroom teacher associated with each course section by semester. I merge in county-level controls

on median house value and crime from the American Community Survey and the National Historical Geographic Information System (NHGIS). I also merge in district-level controls for school funding from the public use Kentucky "School Report Cards" (SRCs) published online by the KDE from 2011 onwards.

The survey, called the TELL Survey, was administered by the KDE every two years between 2011 and 2019 to all Kentucky teachers to gather information about their perceptions of their working conditions.[3] The questions cover a wide range of topics, including: community engagement and support, teacher leadership, school leadership, managing student conduct, use of time, professional development, facilities and resources, instructional practices and support, and new teacher support. The survey is anonymous, meaning that response cannot be traced back to the educator in question. However, the anonymous nature of the survey has also resulted in consistently high response rates over time, with over 80, 89, and 91% of educators responding in 2011, 2015, and 2017. In addition, while the responses cannot be linked to individuals across years, the responses can be aggregated to the school level and linked to schools across year, allowing me to create a school-level panel on teachers' perceptions of each schools' working conditions. I link the panel survey data with the administrative data to obtain survey-based measures of teachers' perceptions of each school and its corresponding principal.

I also use administrative data on teachers and principals in the state of Texas, obtained through the Education Research Center at the University of Texas at Dallas. The Texas administrative data is similar to the Kentucky administrative data in that it consists of individual-level panel data on all public school students, principals, and teachers in Texas. However, the key strength of the Texas data compared to the Kentucky data is its size: not only does Texas have a greater number of principals, teachers, and students, but the coverage of the data also extends further back in time, spanning from 1995 to 2022. The main data, provided by the Texas Education Agency's (TEA) Public Education Information Management System (PEIMS), contains information on student demographics (gender,

[3]Starting in 2019, the TELL Survey was renamed the Impact Survey. The questionnaire changed as well. The initial release of the survey data however was delayed due to COVID-19. We therefore only focus on the consistent TELL Survey data up to 2017 and exclude the Impact Survey data from 2019 in the main analysis.

race/ethnicity, age, grade), student outcomes (e.g. standardized test scores in math and reading; disciplinary events including suspensions), staff roles and responsibilities, teacher and principal assignments, and teacher and principal demographics (including gender, race, age[4], salary, and years of experience).

Most U.S. states have similar statutes governing the hiring and firing of teachers. In Kentucky and Texas, teachers apply to schools directly on designated centralized websites run by the state departments of education that post information on vacancies. While schools may have hiring committees that interview the candidates, principals typically have the final say in hiring decisions. Teachers who are new to a district work on one-year probationary contracts which principals can choose not to renew, thus effectively letting go of the teacher. However, teachers can earn tenure protections after working for four (three) consecutive years in the same district in Kentucky (Texas), which grants them significant job stability.[5]

Principals do not have the same tenure protections as teachers do, but there are still state mandated guidelines on hiring and firing principals. In most states, principals must have previously been certified teachers for a minimum number of two years and must have completed additional principal training, which can take a minimum of one year. When principal vacancies open, whether due to a principal retirement or a move to a different administrative position, designated school committees consisting of teachers, parents, and district administrators are tasked with reviewing applications and interviewing candidates. The superintendent has the final say in who the new principal is.

Table 1 summarizes the main samples of principals and teachers in the Kentucky and Texas datasets, with side-by-side comparisons of similarly defined variables where possible. The school populations in Kentucky are much smaller than those in Texas: in the 2011/12 academic year, there were nearly 39,000 teachers and just over 2,000 principals in Kentucky,

[4]The ERC began making teacher age available in 2023. However, the primary analysis in this paper was conducted prior to the availability of this variable. In the primary analysis, staff age is determined using data on post-secondary schooling provided by the Texas Higher Education Coordinating Board (THECB). The THECB data contains the age reported for students enrolled anywhere in the Texas post-secondary system from 1995 onwards. Using these data, I compute the birthyear for all individuals who were enrolled in the system since 1995, and merge the resulting birthyears on the teacher staff data. Doing so delivers a non-missing birthyear value for around 75% of all school principals between 1995 and 2022.

[5]Tenure status can be inferred from the the data based on the number of consecutive years a teacher is observed working in the same district.

while there were over 338,000 teachers and 7,800 principals in Texas. Educators are also less racially diverse in Kentucky than Texas, in line with the Kentucky population being predominantly white (87.5%).

However, state-level differences aside, the average principal and teacher in Kentucky are largely demographically comparable to the average principal and teacher in Texas. In both states, the average principal in 2011/12 was around 45 years old with 15-20 years of experience and earned an annual salary of $92,000 (in 2018 USD). The average teacher was around 40 years old with 11 years of experience and earned an annual salary of $55,000 (in 2018 USD). By 2018/19, around 30% of the principals and teachers in 2011/12 had made at least one move between schools. The only notable differences are that principals are more likely to be women in Texas (60% compared to 46% in Kentucky), and are also more likely to exit the public school system entirely by 2018/19 (with 54% exiting, compared to 30% in Kentucky).

Table 1: Summary statistics on principal and teachers in Kentucky and Texas

	Principals		Teachers	
	KY	TX	KY	TX
	(1)	(2)	(3)	(4)
Panel A: Demographics in 2011/12				
Age	45.63	46.21	41.19	38.33
	(8.29)	(8.27)	(10.43)	(10.50)
Female	0.46	0.60	0.79	0.77
	(0.50)	(0.49)	(0.41)	(0.42)
White	0.93	0.86	0.95	0.86
	(0.26)	(0.35)	(0.21)	(0.35)
Experience (years)	16.47	20.50	11.86	11.31
	(7.68)	(8.86)	(8.23)	(9.60)
Salary (thousands, 2018 USD)	92.04	90.52	53.45	55.46
	(18.31)	(16.20)	(11.45)	(9.85)
Panel B: Moves made by 2018/19				
Made a between-school transfer	0.28	0.27	0.26	0.32
	(0.45)	(0.44)	(0.44)	(0.47)
Exited public schools for 3+ consecutive years	0.30	0.54	0.41	0.40
	(0.46)	(0.50)	(0.49)	(0.49)
N	2,061	7,856	38,995	338,529

Note: Sample includes all public school principals and teachers from Kentucky and Texas in the academic year 2011/12.

3 Econometric Framework for Assessing Principal Productivity

This section describes the proposed strategy for estimating principal (and teacher) productivity, which leverages the fact documented above that over 25% of teachers and principals move between schools between 2011 and 2019.

3.1 Identification Challenge

Consider first the ideal setting in which one could credibly identify principal productivity. Suppose that schools were pre-populated with independently and identically distributed student characteristics. If principals and teachers were randomly assigned to said schools, then an accurate model of student outcomes would be

$$A_i = \alpha_{j(i)} + \theta_{p(i)} + \varepsilon_i, \tag{1}$$

where A_i is student i's outcome of interest, such as a test score, $\alpha_{j(i)}$ is the fixed component associated with teacher j teaching student i, and $\theta_{p(i)}$ is the fixed component associated with principal p of student i's school. Because student characteristics are balanced across schools, the error terms would be orthogonal to the teacher and principal components. Estimating (1) via ordinary least squares (OLS) would thus give teacher and principal fixed effect estimates that could be interpreted as the average "treatment effect" on student outcomes associated with each teacher and principal. More specifically, the estimated teacher fixed effect α_j would be the average value that teacher j adds to their students' performance *net of* the average teacher effect, or θ_p. Thus, each estimated α_j could be interpreted as the teacher's "value-added" independent of the principal, and each estimated θ_p could be interpreted as the principals' effect above and beyond their teachers' effects.

In our setting however, teachers and principals are not randomly assigned to schools. In fact, since teacher moves and principal moves are officially voluntary in public school systems, and principals play a key role in deciding it comes to which teachers to employ, there is reason to believe that teachers and principals may systematically move between

schools. As a result, naively estimating the fixed effects model (1) on the observed data may produce biased estimates of principal and teacher effects.

Do workers (teachers) move across managers (principals) in our setting in a such way that (1) would be misspecified? To answer this question, I draw on the insights of the extensive literature documenting the extent to which workers and firms contribute to wage inequality (e.g. Abowd et al. [1999], Card et al. [2013]). The workhorse model in this literature is a two-way fixed effects model used to approximate the relationship between workers, firms, and outcomes (wages), rather than between workers, managers, and outcomes.

I adapt the Card et al. [2013] framework to the teacher-principal (or more generally, the worker-manager) environment. I define exogenous mobility as being satisfied when the probability a teacher-principal pair is observed is independent of the error term r, where

$$a = D\alpha + P\theta + X\beta + r, \tag{2}$$

where a is the student or school outcome of interest that serves as a measure of teacher or principal output, such as student test scores, $[D, P]$ are dummies for teachers and principals, X contains student and extensive school controls. The fixed effects of interest, α and θ, are identified off of moves of principals over teachers and vice versa.

An important feature of my framework is that the model excludes firm (school) fixed effects. This exclusion is necessary for the model to be tractable: identifying a school's fixed effect would require observing the school over time with multiple principals and teachers, and for the very same principals and teachers to also overlap at other schools, and so on, a pattern that rarely occurs even within large data like Texas'. That said, one may be concerned that the model (2) without school fixed effects is misspecified, i.e. if "the school" itself makes an important fixed contribution to student outcomes beyond the contributions of principals, teachers, and students themselves. I address this concern by observing that once one takes out the staff and students from the school, the primary channel through which the school itself can affect student outcomes is through (1) its physical location, meaning the catchment area of students in the neighborhood and the economic conditions of the area, and (2) the funding the school receives, which is determined by a formula

based on the demographics of the student population. I therefore include in the model an extensive set of time-varying school-level controls, including: student body composition (population size, age, shares by gender, race, qualifying for free/reduced price lunch, lagged average test scores and behavioral outcomes), funding (federal, state, and local), and county characteristics (labor force participation, share below poverty line, median house value). I also demonstrate that the estimated principal fixed effects that would result from estimating a two-way teacher-principal fixed effect model using the rich set of school controls are distinct from the school fixed effects that would result from estimating a two-way teacher-school fixed effect model (Table 2).

Finally, I assume that r can be broken down into two components,

$$r = \text{match effect} + \text{unit root component}, \tag{3}$$

such that there are two distinct forms of endogenous mobility that could arise in the school setting.[6] The first form would occur if teachers and principals sort across each other based on match effects; that is, if there were multiplicative productivity gains to certain teacher-principal-school combinations and if teachers and principals purposely selected into these advantageous combinations. Such match effects could arise if teachers whose comparative advantage is working with low-income students choose to work with low-income students, or if principals whose comparative advantage is training low-VA teachers tend to hire low-VA teachers.[7] The second form of endogenous mobility could occur if changes in principal-teacher-school assignments are predicted by drifts in performance prior to the change: for example, if teachers who have improved dramatically in their productivity are more likely to move to higher-performing schools, or if principals with falling productivity are more likely

[6]The literature on worker-firm effects typically includes a third "transitory" component in the error term, which would lead to a third form of endogenous mobility that would occur if transitory fluctuations predicted moves across firms. However, this form of endogeneity is not relevant in the public school setting. Demand for teachers and principals in public schools is stable and only changes gradually over time with the population of students. Furthermore, moves by full-time teachers and principals are initiated by the staff themselves, and typically only once per year; districts do not reallocate teachers and principals each year to higher- or lower-performing schools, even when there could be efficiency gains to doing so [Biasi et al., 2021].

[7]Biasi et al. [2021] find that while teachers have comparative advantages in working with low- or high-income students, there is not stark sorting on said comparative advantages, in part because moving costs between districts are high.

to move to lower-performing schools, or if schools with falling productivity are more likely to receive a higher-performing principal.

3.2 Heuristic Checks of Identifying Assumptions

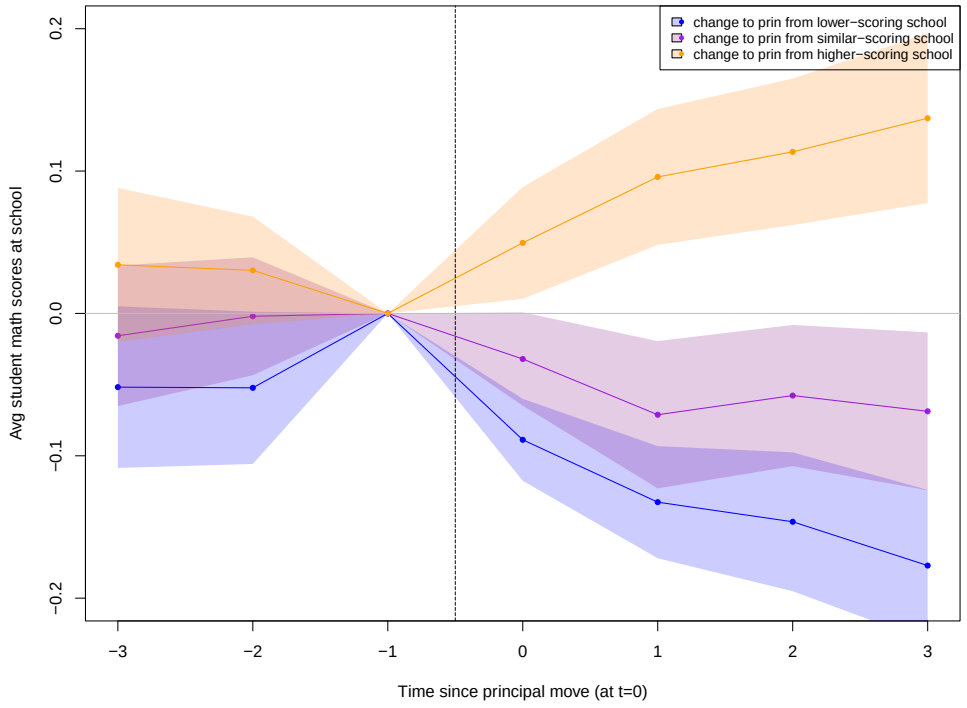
This section assesses the extent to which both forms of endogenous mobility—match effects and drift—are prevalent in the data.

I conduct two heuristic checks for the presence of match effects between teachers, principals, and schools, the results of which suggest that match effects are not a central determinant of teacher-principal allocations. First, I check for patterns in how individuals move across firms: namely, to check whether principals/teachers who move to more productive (i.e. higher performing) schools exhibit symmetric gains in their within-person productivity compared to those who move to less productive (i.e. lower performing) schools.[8] I identify all principal moves that are preceded and followed by stable employment, i.e. where the individual is employed as a principal in origin school A for at least three years, moves to destination school B, and is then employed as a principal in school B for at least four years. I stack the data such that $t = 0$ is normalized as the year of a principal's move to a new school. I calculate the distribution of average school-level test scores in $t = -1$, i.e. one year *prior to* a principal change. For each school that experiences a principal change, I compute the difference between the school's $t = -1$ performance under the old principal (the "destination" performance), and the $t = -1$ performance of the school that the new principal originates from (the "origin" performance). Finally, I classify the school principal changes into 3 cells based on terciles of the difference in destination and origin school performance. Figure 1 plots trends in the school-level average math test scores separately for each cell.

[8] Given our setting has the added dimension of two types of workers—employees and managers—one could also check for patterns in how individuals move across individuals, to check whether teachers (principals) who move to more productive principals (teachers) compared to less productive ones exhibit symmetric gains in their within-person productivity. Ideally, such a test would be conducted using principal-specific and teacher-specific measures of productivity. However, unlike a worker's wage which is primarily associated with the worker, a teacher's or principal's student's test score is primarily associated with the student and only partly associated with the teacher or principal, making such a test difficult to implement. One approach could be to add an intermediate step to obtain estimates of principal productivity that net out the contribution of the students or teachers—i.e. by estimating (2) to obtain principal and teacher fixed effects—and to check whether test scores symmetrically change when teachers start working with higher-fixed effect principals versus lower-fixed effect principals.

Asymmetric gains would suggest that principals are moving based on their own comparative advantages, which could arise if e.g. principals who move to lower-performing schools are those with comparative advantages raising test scores. However, the gains are largely symmetric. If anything, schools that receive a principal from a similar-performing school experience a slight decline in test score performance in the following years, rather than a gain that one might expect if there were sorting on comparative advantages.

Figure 1: Trends in school math test scores around a principal change

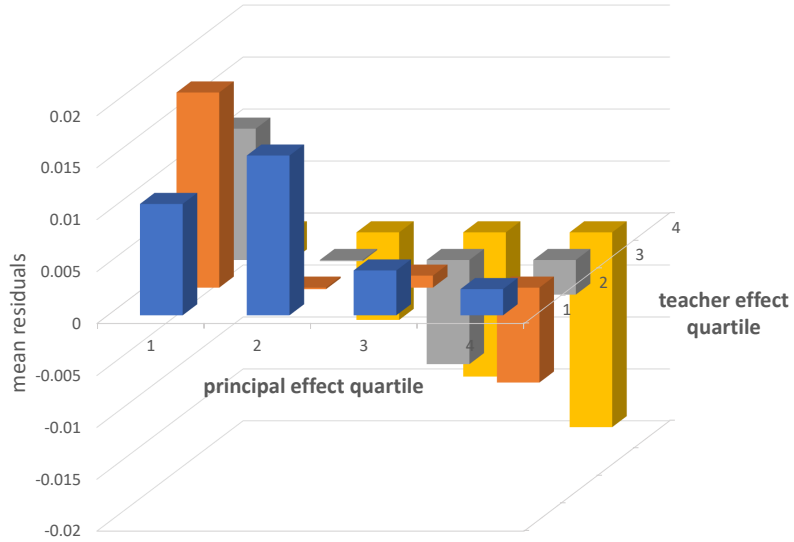


Note: This figure plots normalized trends in student math scores around the event of a principal change in $t = 0$. Schools are classified based on how the $t = -1$ performance of the school (i.e. under the old principal) compares to the $t = -1$ performance of the school that the new principal originated from. Data are from Texas public schools.

Second, I estimate (2) and examine the distribution of mean residuals within 16 cells of teacher by principal fixed effect quartiles. If match effects are large for certain combinations of teachers and principals, then the residuals, which contain the match-specific component of the error term, should be large for certain combinations of teacher by principal fixed effects. Figure 2 shows, however, that the resulting mean residuals are small and similar in

magnitude to those found by Card et al. [2013], which further support the assumption that match effects are not empirically important in our setting.

Figure 2: Distribution of mean residuals by principal and teacher fixed effects



Note: This figure plots the mean residuals resulting from estimating Equation (2), over 16 cells of principal fixed effect and teacher fixed effect quartiles. Data are from Kentucky schools.

Finally, I conduct a heuristic check for whether principal-teacher-school allocations are predicted by drift in performance. Namely, I examine whether there is suggestive evidence of districts re-assigning principals to schools, or of principals selecting schools to switch to, based on trends in student performance. Figure 1 shows that such patterns do not bear out in the data: school average test scores prior to a principal change are relatively flat three years prior to a change.

Taken together, our heuristic checks support the assumption that the two-way fixed effects model over principals and teachers, with the inclusion of rich controls for school location characteristics, is a reasonable approximation of the public school data.

3.3 Variance Decompositions

From (2), I decompose the variance in student outcomes—test scores, absences, and frequency of disciplinary events—as follows:

$$\begin{aligned} V(y_{it}) =& V(\alpha_{p(i,t)}) + V(\beta_{j(i,t)}) + V(x'_{it}\gamma) \\ &+ 2Cov(\alpha_{p(i,t)}, \beta_{j(i,t)}) + 2Cov(\alpha_{p(i,t)}, x'_{it}\gamma) \\ &+ 2Cov(\beta_{j(i,t)}, x'_{it}\gamma) + V(r_{it}), \end{aligned} \tag{4}$$

In the following section, I report the results of our estimation in terms of the sample analogues of each of the variance components in the decomposition and separately by state.

As noted in prior work estimating the worker-firm TWFE model, the estimated covariance between worker and firm fixed effects may be downward biased if there is "limited mobility," i.e. too few moves of workers across firms. Corrections for limited mobility bias are provided by Andrews et al. [2008]. Recent studies such as Bonhomme et al. [2019] also suggest alternative models that deliver the variance decomposition of interest but not individual fixed effects, thus circumventing the issue of limited mobility. Our results are largely robust to these various specifications.

4 How Much Principals Matter Relative to Teachers and Schools

Table 2, Columns (2) and (3) reports the variance decompositions that result from estimating (2) on largest connected sets of principals and teachers in each state. In both states, principal fixed effects explain smaller fractions of the variance in student achievement, absenteeism, and behavioral incidents than teacher fixed effects do. In Kentucky (Texas), principal fixed effects explain 18% (25%) of the variance in student test scores that teacher fixed effects do, 18% (27%) of the variance in absent days that teacher fixed effects do, and 20-26% (28%) of the variance in behavioral incidents that teacher fixed effects. Overall, teacher and principal fixed effects in Texas explain more of the variance in student test scores than those

in Kentucky do, but the teacher and principal fixed effects in Kentucky explain more of the variance in behavioral outcomes than those in Texas do. However, the relative magnitudes of the variance in teacher and principal fixed effects are similar in magnitude between the two states.

Table 2: Variance decompositions of student outcomes

		Model 1: Teacher-Principal		Model 2: Teacher-School	
		% of Var(Y_{ij}) attributable to...		% of Var(Y_{ij}) attributable to...	
Outcome Y_{ij}	Var(Y_{ij}) (1)	Var(teacher FE) (2)	Var(prin. FE) (3)	Var(teacher FE) (4)	Var(sch. FE) (5)
Kentucky schools					
Test scores	0.98	9.6%	1.8%	9.3%	4.3%
Absent days	63.53	6.4%	1.2%	6.5%	2.4%
Nr incidents	4.68	19.9%	5.4%	18.6%	7.2%
Nr resolutions	0.46	11.0%	2.9%	14.3%	6.1%
Total obs.		7,436,165		2,259,516	
Nr teachers		11,401		5,918	
Nr principals		1,494		198	
Texas schools					
Test scores	0.90	19.2%	4.8%	20.4%	6.0%
Absent days	32.42	5.9%	1.6%	6.4%	6.8%
Nr incidents	0.91	7.7%	2.2%	9.3%	15.5%
Nr suspensions	0.66	7.7%	2.2%	9.2%	12.1%
Total obs.		28,833,863		35,336,303	
Nr teachers		132,317		149,572	
Nr principals		6,206		5,184	

Note: Models are separately estimated on the largest connected set of principals and teachers over students in Grades 3-8 in administrative data from Kentucky and Texas. All specifications of Model 1 include school controls.

Given that a primary identification concern is that principals do not move between schools often enough, I also evaluate whether our design is meaningfully distinguishes betwe up variation driven by principal moves across teachers, as opposed to simply variation driven by teacher moves across schools. My primary approach is to estimate a similar two-way fixed effects model as (2), but on the largest connected sets of teachers and *schools* in each state as opposed to the largest sets of teachers and principals. If there is insufficient principal

movement in our data such that (2) only identifies school averages, mislabelled as principal fixed effects, then we would expect that the principal fixed effects from Model 1 and the school fixed effects in Model 2 would be almost perfectly correlated and result in similar variance decompositions. However, I find neither to be the case. Columns (4) and (5) show the resulting variance decompositions from estimating the teacher-school model. The fact that the variance of the teacher fixed effects are largely similar, but the variance of the school fixed effects exceeds that of the principal fixed effects, indeed suggests that there is sufficient variation in our data to identify fixed effects that are principal-specific in our main specification. Table 3 shows the correlations between school fixed effects and principal effects estimated using the two separate models for different student outcomes. The Pearson correlation coefficient in each case is around 0.3. Thus, while there appears to be some sorting based on principal and school fixed effects, our approach seems to identify principal fixed effects that are distinct from the school fixed effects.

Table 3: Correlations between principal and school fixed effects across models

	Est. principal FE from model of...		
	Test scores	Absent days	Incidents/year
Est. school FE from model of...			
Test scores	0.29***		
	(0.02)		
Absent days		0.40***	
		(0.02)	
Incidents/year			0.34***
			(0.02)
Adjusted R^2	0.33	0.29	0.26
N	840	1,065	1,065

Note: Coefficients are from a regression of estimated principal fixed effects from teacher-principal model on school fixed effects from teacher-schools model, for three models with different student outcomes. Data are from Kentucky schools. Standard errors are in parentheses.

5 Principal Productivity and Student Outcomes

Using the estimated principal fixed effects as measures of principal productivity, I exploit the timing of when schools change principals to explore how more productive principals affect student outcomes: namely test scores, absenteeism, and suspensions. I focus only on

principal changes in Texas, as the smaller sample size and shorter panel length in Kentucky precludes implementing the same designs.

I also implement the same event study design focusing only on the sample of principal turnover events to only those where the leaving principal *retires*, where a retirement is defined as an event in which the leaving principal does not appear at another school in the following three consecutive years and/or, following Texas' state mandates for public school educators, the leaving principal is of retirement age (60) or has enough garnered enough experience such that the sum of their age and their number of years of experience exceeds 80. Leveraging the subset of principal retirements addresses concerns about assignment bias of principals across schools that one may have with the main design: aside from retiring principals potentially changing their own behavior in the year or two leading up to their retirement, the reason for hiring a new principal ostensibly would be unrelated to the performance of the retiring principal. The estimates from the retirement design are shown in the Appendix, and are largely qualitatively similar to the figures shown in the main text. However, because only a quarter of all principal changes can be classified as retirements, the sample size is significantly smaller under the restricted retirement design. My preferred design, described below, therefore includes all principal changes.

The goal is to examine how student outcomes change at schools when one principal leaves and a principal of lower, similar, or higher productivity comes in. I identify all of the school-years in the Texas data in which the principal changed *and* the managerial team of principals was otherwise stable for three years before and after the event.[9] I classify the principal changes into three categories: those where the school changed to a more productive principal, to a less productive principal, or to a similarly productive principal, where the latter is defined as a change where the difference in estimated fixed effects between the old and new principal falls in the middle tercile of the empirical distribution of differences. I stack the school-year data such that the dates of each school's principal change align at $t = 0$.

For each principal change category, I estimate the following event study specification

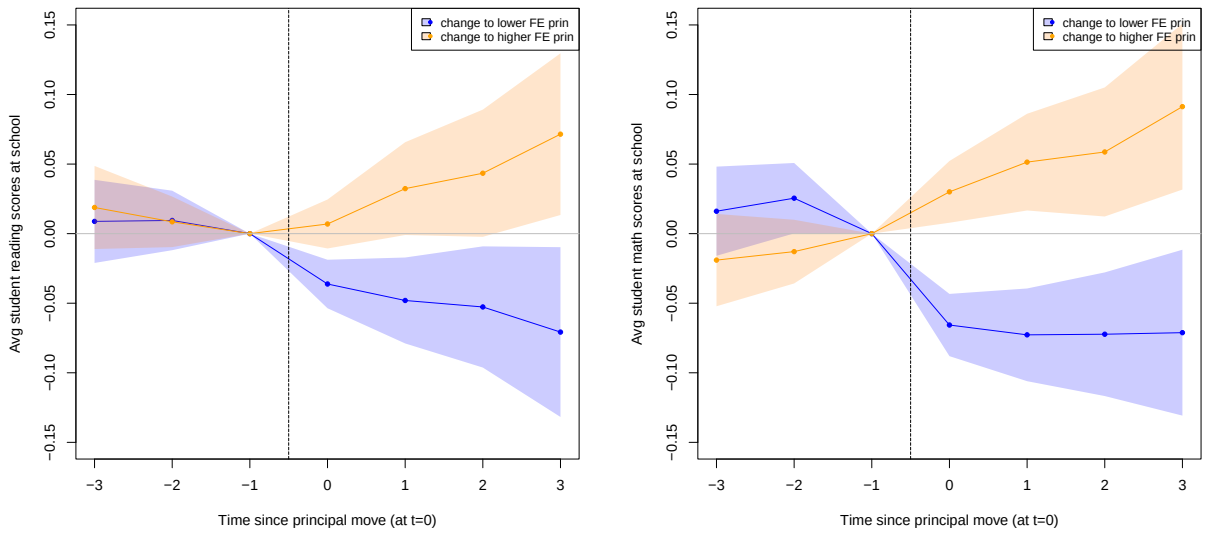
[9]The main reason I focus only on the Texas data in this analysis is because the number of school-years in which a principal changes in Kentucky *and* the school experiences managerial stability within a short window around the principal change is very small.

via OLS:

$$Y_{it} = \beta_0 + \sum_{\tau \in \{-3,-2,0,1,2,3\}} \alpha_\tau \cdot \mathbb{I}(t = \tau) + \gamma_t + \varepsilon_{it}, \tag{5}$$

where Y_{it} denotes the average student outcome of school i in year t, γ_t denotes year fixed effects, and the coefficients α_τ capture the adjusted difference in student outcomes between year $t = -1$ and year τ. I consider four student outcomes at the school level: average standardized math test scores, average standardized reading test scores, average number of days absent in a school year, and average number of suspensions in a school year. As a result, there are 36 different event study specifications: one for each combination of the four student outcomes, three principal change categories, and three different fixed effects that can be used to construct the principal change categories.

Figure 3: Event study estimates of changes in test scores around principal change



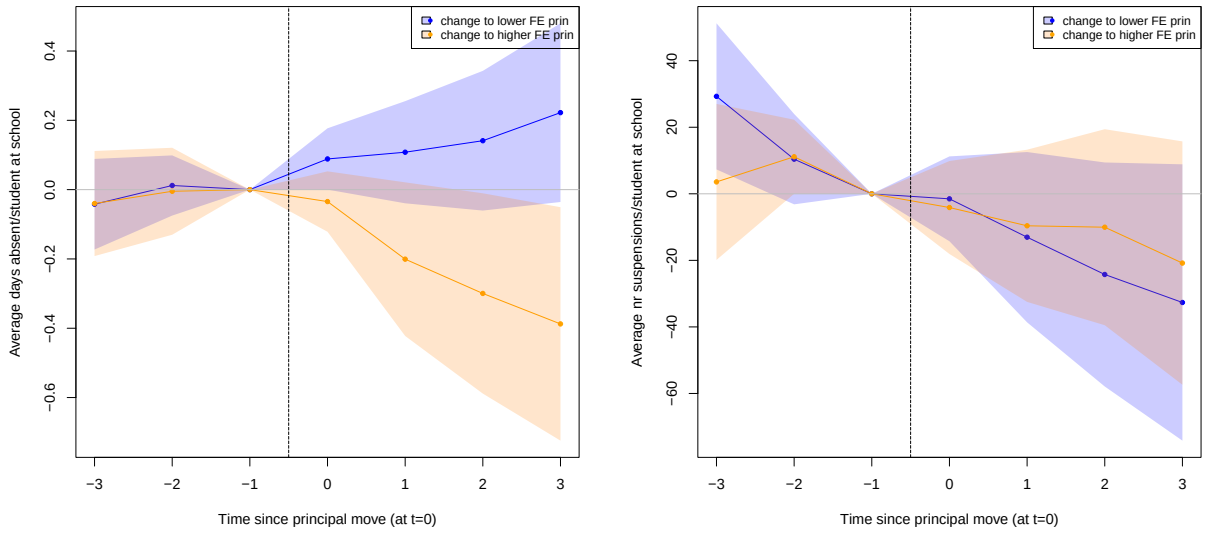
(a) Outcome: Avg. school reading scores (b) Outcome: Avg. school math scores

Note: This figure plots event study estimates of how average test scores change within a school around a principal change. The outcome variable in panel (a) is the average standardized reading score, while the outcome variable in panel (b) is the average standardized math score, for students in Grades 3-8. Principal changes are classified as going from to a higher or lower-FE principal, where the FE are associated with test scores. Data are from Texas schools.

Figures 3 and 4 illustrate an example of the results for one set of estimates: for all four student outcomes and two principal changes, classified based on the principal's productivity

defined by test scores.[10] The pre-trends in the outcomes of interest are relatively flat and do not significantly differ between the different types of principal changes. Schools that receive a principal with a higher test score fixed effect—that is, a principal who I identify as being more effective at raising test scores—experience increased math test scores within one year of the transition, weakly increased reading test scores within one to two years of the transition, and decreased student absenteeism within a year. In contrast, schools that receive a principal who is less effective at raising test scores experience sharply decreased math and reading scores within one year and weakly increased absenteeism within three years. In both cases, the average number of suspensions at schools that change principals weakly declines within three years.

Figure 4: Event study estimates of changes in non-cognitive student outcomes around principal change



(a) Outcome: Avg. nr. absences at school (b) Outcome: Avg. nr. suspensions at school

Note: These figures plots event study estimates of how the average number of behavioral incidents at a school change around a principal change. The outcome in panel (a) is the average number of absences at the school, while the outcome in panel (b) is the average number of in-school or out-of-school suspensions at the school. Principal changes are classified as going from to a higher or lower-FE principal, where the FE are associated with test scores. Data are from Texas schools.

[10] To avoid clutter, the graphs exclude trends associated with changing to similarly productive principals.

6 Principal Productivity and Teacher Outcomes

The results of the prior section indicate that more productive principals may be able to induce sizeable changes in student outcomes. How might principals be achieving these effects? Given teachers, not principals, are the staff who spend the most time with students, one possibility is that productive principals are more likely to hire new effective teachers or to implement school practices that improve the performance of existing teachers. This section examines whether personnel management appears to be an empirically important channel through which principals can affect student outcomes.

6.1 Teacher Hiring and Retention

Next, I use the same event study design to examine how the hiring and retention of teachers changes at schools that experience a principal change event. I define four main outcomes of interest at the school-year level. The first is the "rehire rate" in year t, which I define as the percentage of teachers who were employed at the school in $t-1$ who are rehired at the school in year t. The second is the "new hire rate" in year t, which I define as the percentage of teachers in year t who did not work at the school in $t-1$. The third is the "average leavers' value-add" in year t, which I define as the average teacher value-added (taken as the estimated teacher fixed effects) among teachers who are working in the school in year t but not in the following year $t+1$. Finally, the fourth outcome is the "average new hires' value-add" in year t, which I define as the average teacher value-added among teachers who are not working at the school in $t-1$ but are working at the school in year t.

There are two primary approaches that a principal could take to raise student achievement through their hiring and retention practices. One approach could be to let go of low value-add teachers and hire higher value-add teachers to take their place, which would correspond to a decrease in the average leavers' value-add and an increase in the average new hires' value add. However, it is not obvious that replacing teachers would be the most effective strategy at raising student test scores, as high turnover can decrease morale and productivity, and given that teacher experience is a key predictor of teacher value-add, teacher retention is also a key mechanism through which teacher effectiveness may be improved. Another

approach might therefore let go of low value-add teachers initially but focus on stabilizing their workforce as quickly as possible, which would correspond to an initial dip in the rehire rate but a quick recovery as well as a decreasing new hire rate.

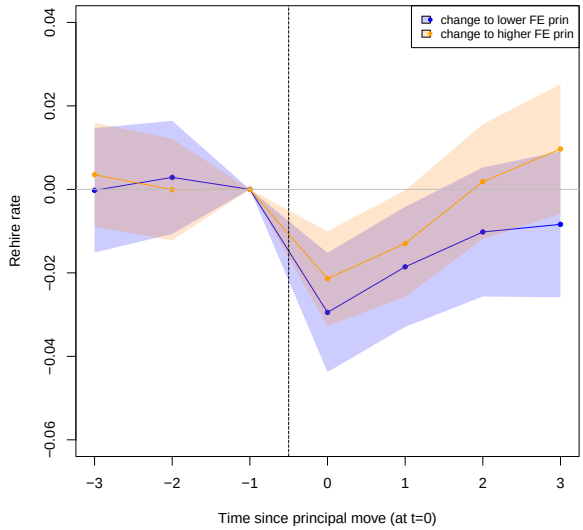
Figure 5 shows the event study estimates corresponding with each of the four outcomes related to hiring and retention. Regardless of the type of principal change that occurs, the average rehire rate falls by 2 p.p. in the first year of the principal change, while the average new hire rate remains steady. However, at schools that change to more productive principals, teacher retention increases quickly, with rehire rates bouncing back and rising above their original levels and new hire rates falling within two to three years. Furthermore, schools that receive more productive principals experience a weak decrease in average leavers' value-add and no change in average new hires' value-add, while schools that receive less productive principals experience a sharp increase in the average leavers' value-add and a corresponding increase in average new hires' value-add. These results suggest that principal who are productive at raising test scores tend to let go of less effective teachers initially but to quickly stabilize their workforce, while less productive principals are more likely to lose effective teachers and have more difficulty stabilizing their retention rates in the short run.

6.2 Productivity of Staying Teachers

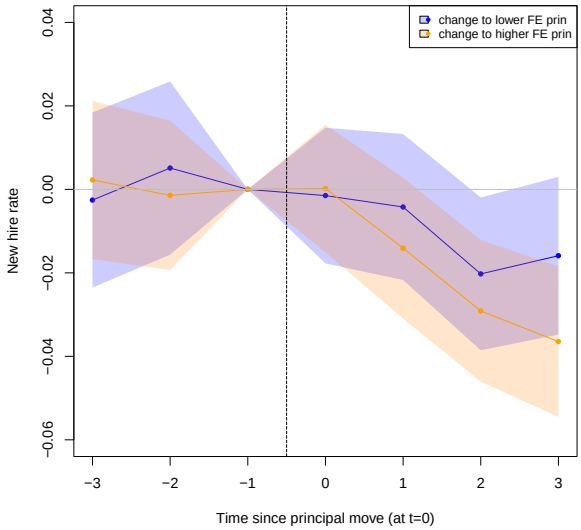
Personnel management entails not only bringing in and pushing out workers, but also training, monitoring, and motivating of the workers who stay. Another way then that productive principals might be able to improve student test scores through personnel management is through increasing the productivity of staying teachers, which could arise from implementing new training practices or improving the workplace climate.

Using the same event study design, I examine whether productivity appears to change among teachers who stay at schools that experience a principal change event. Because I estimate teacher value-add as a fixed effect that does not vary over time, I use raw test scores as a measure of teacher output and add controls for students' prior test scores and school wide test scores. Figure 6 shows the resulting estimates. Student test scores in math sharply increase among staying teachers at schools that receive a more productive principal. Remarkably, student test scores in reading, which tend to be stickier than math

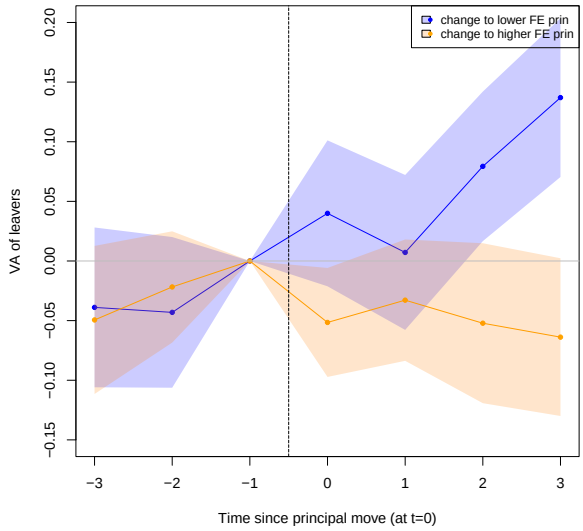
Figure 5: Event study estimates of changes in teacher outcomes around a principal change



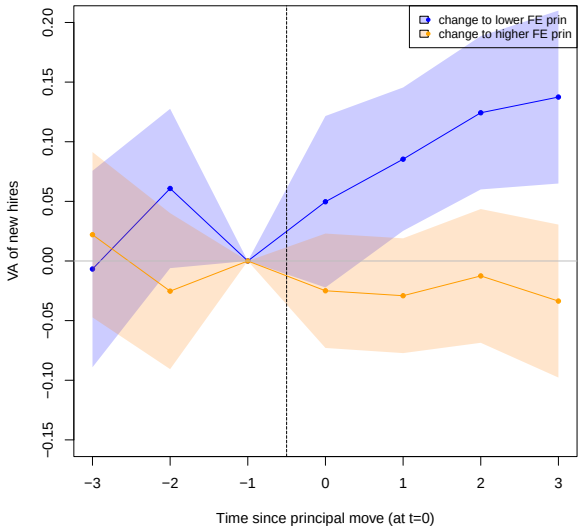
(a) Outcome: Avg. % of teachers from $t-1$ who are rehired in year t



(b) Outcome: Avg. % of teachers in year t who are new hires



(c) Outcome: Avg. teacher FE among teachers who leave at the end of year t

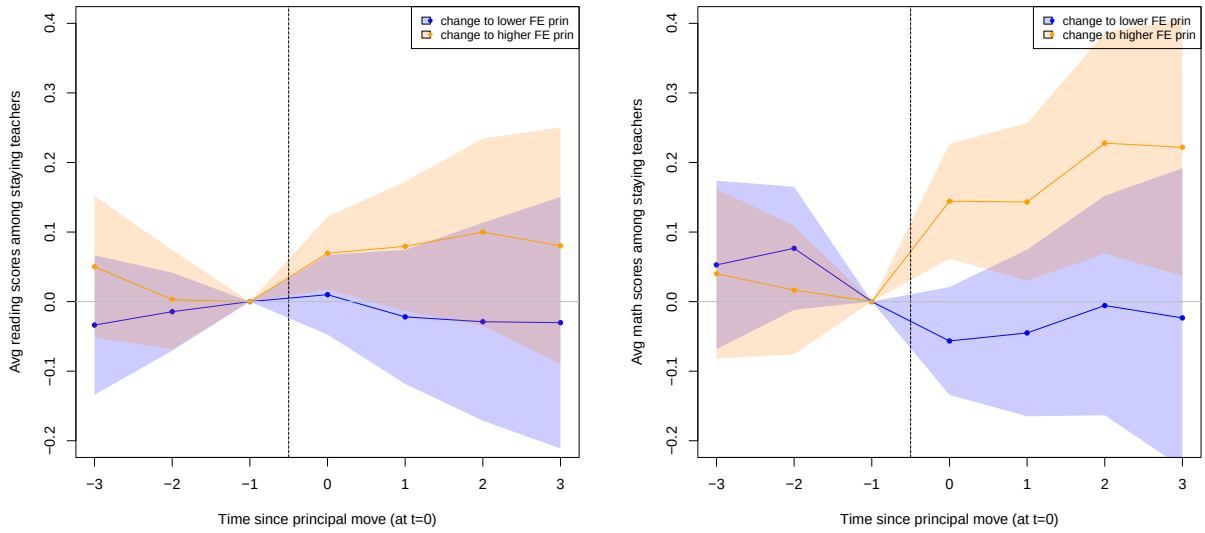


(d) Outcome: Avg. teacher FE among new hires who begin in year t

Note: These figures show event study estimates of how teacher retention, hiring rates, and teacher fixed effects change around the time of a principal change. The outcomes are: (a) the average percentage of teachers in year $t-1$ who are rehired (i.e., appear in the school as a teacher again) in year t, (b) the average percentage of teachers in year t who were not at the school in year $t-1$, (c) the average teacher fixed effect among teachers who are no longer observed at the school in year $t+1$, and (d) the average teacher fixed effect among teachers who are newly observed at the school in year t. Principal changes are classified as going from to a higher or lower-FE principal, where the FE are associated with test scores. Data are from Texas public schools. Data are from Texas schools.

test scores, also weakly increase. In contrast, math and reading scores remain unchanged among staying teachers at schools that receive a less productive principal. These results suggest that principals who raise student test scores not only manage the hiring and retention of their teachers differently from less productive principals, but also find ways to increase the productivity of the teachers they retain.

Figure 6: Event study estimates of how average test scores change for staying teachers around a principal change



(a) Outcome: Avg. reading scores among staying teachers

(b) Outcome: Avg. math scores among staying teachers

Note: These figures plot event study estimates of how average test scores at a school change among incumbent teachers who stay for the next 3 years in the school after a principal change event. Principal changes are classified as going from to a higher or lower-FE principal, where the FE are associated with test scores. Data are from Texas schools.

7 Teachers' Perceptions of Productive Principals

What might productive principals do differently around the school that could plausibly impact the productivity of staying teachers? Recent work in the management literature suggests that managers' "people management skills" as measured using employee surveys may decrease employee attrition (e.g. **?**). In this section, I leverage similar survey data

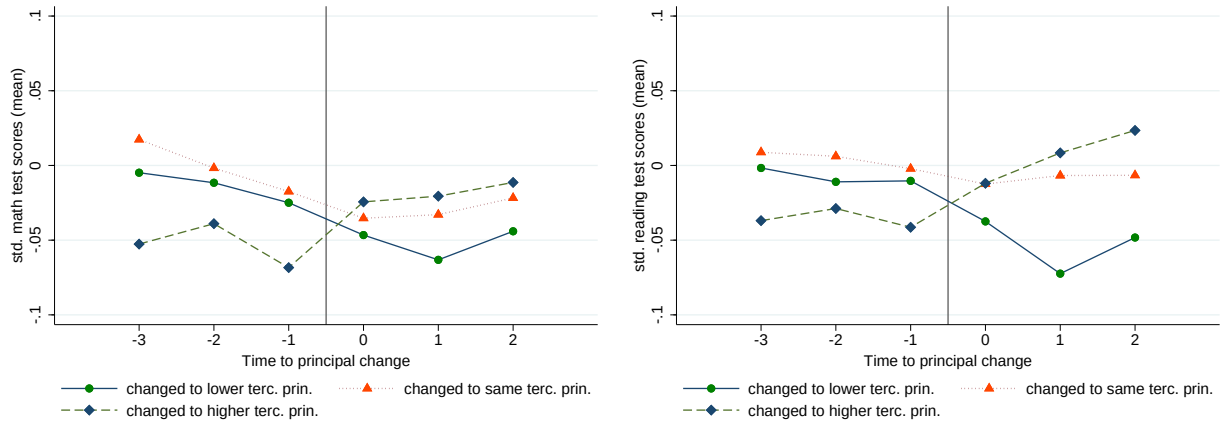
from Kentucky public schools, which asks teachers to rate various aspects of their working conditions including qualities of the school leadership, to examine whether more productive principals exhibit different specific management skills.

I first examine the relationship between teachers' ratings of their principals' leadership skills in a given year with raw student test scores. Using an approach similar to the heuristic checks for match and drift effects, I plot school-level raw means in student test scores in math and reading around the instance of a principal change. I classify principal changes based on whether the new principal receives a higher, similar, or lower rating than the old principal in terms of their overall leadership ability.[11] Figure 7 plots the resulting trends. Within schools, principals with higher leadership ratings are associated with higher math and reading test scores, while the opposite is true for principals with lower leadership ratings. While these patterns cannot be interpreted as causal, as it is possible that teachers are more likely to positively rate principals once test scores improve for other reasons, the patterns are nonetheless suggestive that principals' management skills, as experienced by teachers, may be an important mechanism through which principals can influence student outcomes.

Next, I leverage the fact that the survey asks what teachers what their immediate professional plans are in the following year—staying at the school, switching schools, or exiting teaching—to examine whether teachers in the Kentucky survey respond similarly to high-FE principals as teachers in the Texas administrative data do. Table 5 shows the estimated coefficients resulting from regressing the estimated principal FE on the share of teacher respondents who report intending to leave education entirely in the following year (Column (1)), and on the share of teacher respondents who report intending to stay in education in various forms in the following year (Column (2)). Teachers who work under higher-FE principals report being significantly less likely to leave education in the following year than the average teacher. The effect is entirely explained by teachers reporting being more likely to stay at their current school, rather than leave to teach at another location or to an administrative or non-administrative role in education. These results corroborate the event study findings in Texas that show that higher-FE principals achieve greater retention

[11] Note that, due to sparseness of the data, it is not possible to classify principal changes based on the pre-move rating that each principal receives.

Figure 7: Trends in test scores around principal changes, by leadership ratings



(a) Math scores, Kentucky

(b) Reading scores, Kentucky

Note: These figures plot trends in mean standardized test scores around the event of a principal change for three groups: schools that change to a principal with a higher tercile overall leadership rating in the TELL Survey (blue diamonds), schools that change to a similar tercile leadership rating principal (orange triangles), and schools that change to a lower tercile leadership rating principal (green circles). Data are from Kentucky schools.

in the short run, which suggest that the fixed effects share the same interpretation across the two states.

Finally, I examine the relationship between the estimated principal fixed effects from equation and teachers' ratings of a wide range of principal qualities and school working conditions. Because the survey data is only available every two years and survey responses are anonymous, it is not possible to conduct similar event study exercises as in the previous sections to isolate how individual teachers' ratings change within schools when new principals enter. Instead, I report OLS estimates of the extent to which various teachers' ratings predict the estimated principal fixed effects.

Table 4 shows the estimates resulting from regressing principal fixed effects on the average teacher ratings from each main section of the survey: school leadership, time use, student conduct, teacher leadership, instructional practices, professional development, and community support.[12] I construct the average ratings by taking a simple average of teachers' responses to the set of all Likert-scale questions asked in each category. Importantly, I include

[12]A section on facilities and resources was added in later years, and is thus excluded from this analysis to maintain consistency across years.

teachers' answers to questions that pertain to areas that principals may not have direct influence over, such as whether parents in the community are very invested in the school operations, to control for as many factors that may affect teachers' ratings as possible.

Table 4: Relationship between Estimated Principal Fixed Effects and Teachers' Ratings of Principal's School

	Principal FE	
	(1)	(2)
Average leadership score	-0.0605*	-0.0612*
	(0.0334)	(0.0333)
Average time use score	-0.0180	-0.0186
	(0.0150)	(0.0150)
Average student conduct score	0.0264**	0.0265**
	(0.0131)	(0.0131)
Average teacher leadership score	-0.0029	-0.0018
	(0.0298)	(0.0298)
Average instructional practices score	0.0925***	0.0856***
	(0.0279)	(0.0282)
Average professional development score	-0.0128	-0.0122
	(0.0238)	(0.0238)
Average community support score	0.0348*	0.0409**
	(0.0185)	(0.0189)
Controls for respondent experience level	No	Yes
Observations	2,134	2,134
R^2	0.07897	0.08010
Adjusted R^2	0.07288	0.07358

Note: Data are at the school-year level, using data in years when the TELL Survey was run (every 2 years from 2011 to 2017).

Column (2) shows the estimates from our preferred specification, which includes controls for the average number of years of teaching experience the survey respondents have. Teachers who work with higher-FE principals tend to report below average school leadership qualities, above average student conduct and community support, and significantly above average instructional practices at their place of work compared to the average teacher.

To unpack the drivers of these patterns, Column (1) of Tables B.3 through B.6 show the estimates from regressing the principal fixed effects on all of the individual questions used to construct the average ratings. On issues of school leadership, teachers who work under higher-FE principals are less likely to agree that there is an atmosphere of trust and mutual

Table 5: Relationship between Estimated Principal Fixed Effects and Teachers' Reported Immediate Professional Plans

	Principal FE	
	(1)	(2)
Intend to leave education	-0.1384*	
	(0.0722)	
Intend to stay at school		0.1200*
		(0.0723)
Intend to teach at another school/district		0.0540
		(0.1661)
Intend to move to admin		-0.0374
		(0.0961)
Intend to move to non-admin role		0.0867
		(0.1228)
Controls for respondent experience level	Yes	Yes
Observations	2,134	2,134
R^2	0.06833	0.07421
Adjusted R^2	0.06438	0.06897

Note: Regressions are at the school-year level, in years when the TELL Survey was run (every 2 years from 2011 to 2017).

respect at their school, are less likely to report that their school leadership makes a sustained effort to address teachers' concerns about leadership issues, and are less likely to agree that they are held to high professional standards or objectively assessed. That said, they are also more likely to agree that they have a shared vision with the school leadership, that they feel comfortable raising issues and concerns, and that the school leadership facilities using data to improve student learning. On issues of instructional practices, teachers who work under higher-FE principals are significantly more likely to report using assessment data to inform their instruction. Taken together, these results are consistent with high-FE principals emphasizing instructional practices that can help to raise test scores, but potentially at the expense of establishing a trusting relationship with their employees.

The estimates also provide suggestive evidence that principal effectiveness, as measured by principal fixed effects, can be bolstered by working at a school where the community actively supports the school and students respond well to student conduct rules. On issues of student conduct, teachers who work under higher-FE principals are more likely to agree that the students at their school understand expectations for their conduct, but less likely to

agree that school administrators are consistent in their enforcement of student conduct rules. While in theory principals could help improve student conduct directly by e.g. supporting teachers' efforts to maintain discipline in the classroom, doing so does not appear to be a significant predictor of principal fixed effects. Furthermore, on issues of community support, teachers who work under higher-FE principals are more likely to agree that community members and parents/guardians support teachers, although less likely to agree that they as teachers provide adequate information about student learning to parents/guardians.

8 Conclusion

Despite the growing literature documenting the importance of managers, little is known about the extent to which a critical group of public sector managers—public school principals—contribute to student learning. Furthermore, data limitations make it difficult to assess how effective managers achieve better outcomes through their managerial practices.

This paper addresses both gaps. Using administrative data that links principal, teachers, and student outcomes in two U.S. states, I leverage moves of principals across teachers to estimate principal and teacher fixed effects. I find that principals explain less of the variance in student outcomes than teachers do, but that more effective principals appear to be associated with different management practices than less effective ones. Linking the estimated effects back to the administrative data and exploiting principal changes at schools, I find that more effective principals are associated with better student outcomes as well as greater teacher retention and increased productivity within teachers. Linking the administrative data with survey data from Kentucky teachers on their perceptions of their working conditions, I find that teachers who work under more effective principals report better overall school leadership, instructional practices, and student conduct, but less trust and respect in their school.

One caveat to the analysis is that, by separating the principal and teacher effects additively in the main model, the principal effect in the model is understated. Indeed, if principals with higher individual fixed effects are also principals who select effective teachers and can even improve teacher effectiveness, the principal's effect is really the sum total of

their mean effect and the effectiveness of the teachers they attract and improve. Adapting the model framework to account for the compounding effect of effective managers with effective teachers is left for future work.

Overall, my findings suggest that while public school principals may show limited ability to affect student outcomes, a primary channel through which such effects can arise is through personnel management, both through selecting and retaining effective teachers *and* improving the productivity of staying teachers. From a policy perspective, these results suggest that offering personnel management training could have large productivity gains without having to re-allocate principals across schools.

References

J. M. Abowd, F. Kramarz, and D. N. Margolis. High wage workers and high wage firms. *Econometrica*, 67(2):251–333, 1999.

M. J. Andrews, L. Gill, T. Schank, and R. Upward. High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society Series A: Statistics in Society*, 171(3):673–697, 2008.

O. Bandiera, A. Prat, S. Hansen, and R. Sadun. Ceo behavior and firm performance. *Journal of Political Economy*, 128(4):1325–1369, 2020.

B. Bartanen, A. N. Husain, and D. D. Liebowitz. Rethinking principal effects on student outcomes. *Journal of Public Economics*, 234:105115, 2024.

S. Bender, N. Bloom, D. Card, J. Van Reenen, and S. Wolter. Management practices, workforce selection, and productivity. *Journal of Labor Economics*, 36(S1):S371–S409, 2018.

M. Bertrand and A. Schoar. Managing with style: The effect of managers on firm policies. *The Quarterly journal of economics*, 118(4):1169–1208, 2003.

M. C. Best, J. Hjort, and D. Szakonyi. Individuals and organizations as sources of state effectiveness. *American Economic Review*, 113(8):2121–2167, 2023.

B. Biasi, C. Fu, and J. Stromme. Equilibrium in the market for public school teachers: District wage strategies and teacher comparative advantage. Technical report, National Bureau of Economic Research, 2021.

N. Bloom and J. Van Reenen. Measuring and explaining management practices across firms and countries. *The quarterly journal of Economics*, 122(4):1351–1408, 2007.

N. Bloom, B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts. Does management matter? evidence from india. *The Quarterly journal of economics*, 128(1):1–51, 2013.

N. Bloom, R. Lemos, R. Sadun, and J. Van Reenen. Does management matter in schools? *The Economic Journal*, 125(584):647–674, 2015.

S. Bonhomme, T. Lamadon, and E. Manresa. A distributional framework for matched employer employee data. *Econometrica*, 87(3):699–739, 2019.

G. F. Branch, E. A. Hanushek, and S. G. Rivkin. Estimating the effect of leaders on public sector productivity: The case of school principals. Technical report, National Bureau of Economic Research, 2012.

A. Bryson, L. Stokes, and D. Wilkinson. Is pupil attainment higher in well-managed schools? *Education Economics*, 31(1):129–144, 2023.

J. B. Bushnell and C. Wolfram. The guy at the controls: Labor quality and power plant efficiency. In *International differences in the business practices and productivity of firms*, pages 79–102. University of Chicago Press, 2009.

D. Card, J. Heining, and P. Kline. Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics*, 128(3):967–1015, 2013.

R. Chetty, J. N. Friedman, and J. E. Rockoff. Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American economic review*, 104(9):2593–2632, 2014.

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

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

M. Engel, M. Cannata, and F. C. Curran. Principal influence in teacher hiring: Documenting decentralization over time. *Journal of Educational Administration*, 56(3):277–296, 2018.

A. Fenizia. Managers and productivity in the public sector. *Econometrica*, 90(3):1063–1084, 2022.

R. G. Fryer et al. Management and student achievement: Evidence from a randomized field experiment. Technical report, National Bureau of Economic Research, 2017.

M. Gilraine and N. G. Pope. Making teaching last: Long-run value-added. Technical report, National Bureau of Economic Research, 2021.

J. A. Grissom and B. Bartanen. Strategic retention: Principal effectiveness and teacher turnover in multiple-measure teacher evaluation systems. *American Educational Research Journal*, 56(2):514–555, 2019.

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

J. A. Grissom, D. Kalogrides, and S. Loeb. Using student test scores to measure principal performance. *Educational evaluation and policy analysis*, 37(1):3–28, 2015.

E. A. Hanushek, A. J. Morgan, S. G. Rivkin, J. C. Schiman, A. Shakeel, and L. Sartain. The lasting impacts of middle school principals. Technical report, National Bureau of Economic Research, 2024.

M. Hoffman and S. Tadelis. People management skills, employee attrition, and manager rewards: An empirical analysis. *Journal of political economy*, 129(1):243–285, 2021.

C. Ichniowski and K. Shaw. Beyond incentive pay: Insiders' estimates of the value of complementary human resource management practices. *Journal of Economic Perspectives*, 17 (1):155–180, 2003.

C. Ichniowski, K. Shaw, and G. Prennushi. The effects of human resource management practices on productivity: A study of steel finishing lines. *American Economic Review*, 87 (3):291–313, 1997.

R. M. Ingersoll, P. Sirinides, and P. Dougherty. School leadership, teachers' roles in school decisionmaking, and student achievement. working paper. wp# 2017-2. *Consortium for Policy Research in Education*, 2017.

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

B. A. Jacob. Do principals fire the worst teachers? *Educational Evaluation and Policy Analysis*, 33(4):403–434, 2011.

T. J. Kane and D. O. Staiger. Estimating teacher impacts on student achievement: An experimental evaluation. Technical report, National Bureau of Economic Research, 2008.

E. P. Lazear, K. L. Shaw, and C. T. Stanton. The value of bosses. *Journal of Labor Economics*, 33(4):823–861, 2015.

D. D. Liebowitz and L. Porter. 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, 2019.

R. D. Metcalfe, A. B. Sollaci, and C. Syverson. Managers and productivity in retail. Technical report, National Bureau of Economic Research, 2023.

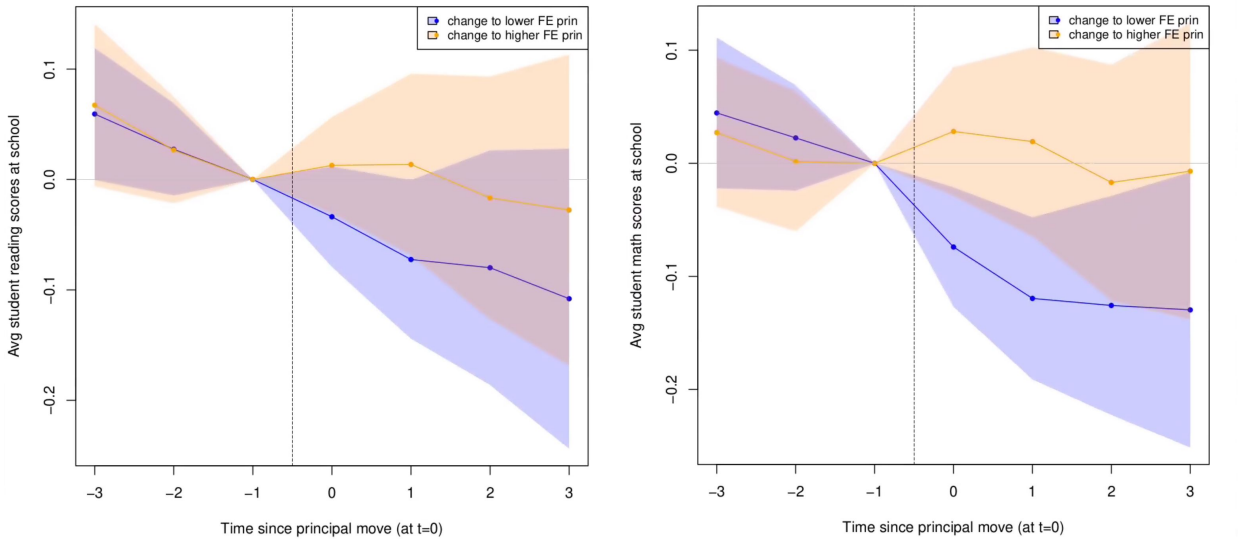
A. Miller. Principal turnover and student achievement. *Economics of Education Review*, 36: 60–72, 2013.

P. Muñoz and M. Prem. Managers' productivity and recruitment in the public sector. *American Economic Journal: Economic Policy*, 2024.

C. Syverson. What determines productivity? *Journal of Economic literature*, 49(2):326–365, 2011.

Appendices

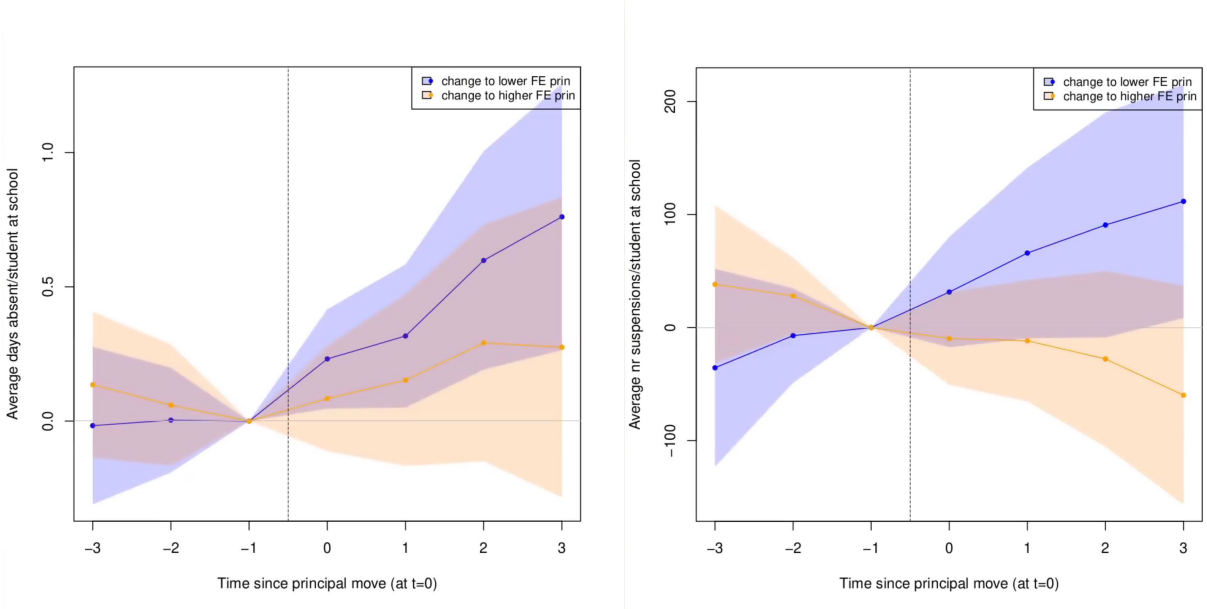
A Additional Figures



(a) Outcome: Avg. school reading scores

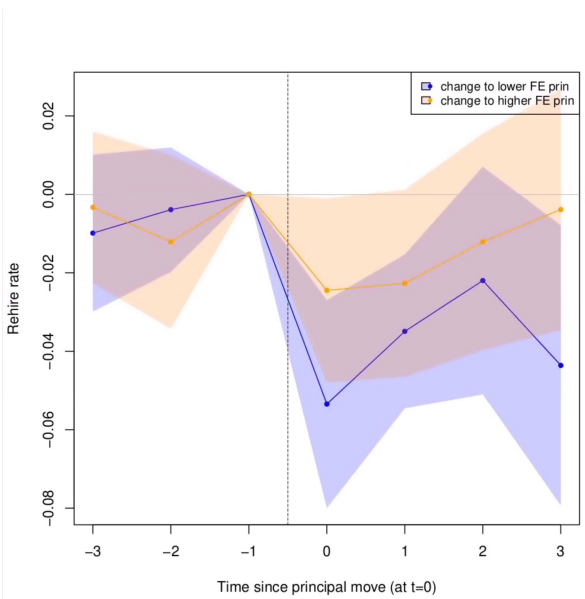
(b) Outcome: Avg. school math scores

Figure A.1: Event study estimates of how average test scores change within a school around a principal change. Principal changes are classified as going from to a higher or lower-FE principal, where the FE are associated with test scores. Fix axes

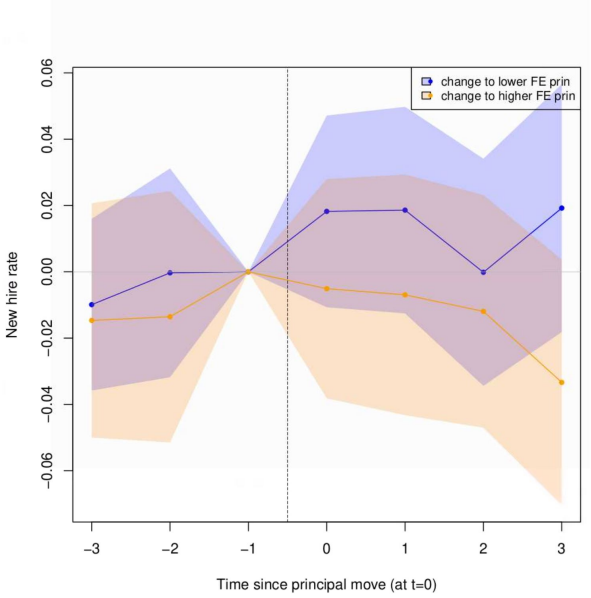


(a) Outcome: Avg. nr. absences at school (b) Outcome: Avg. nr. suspensions at school

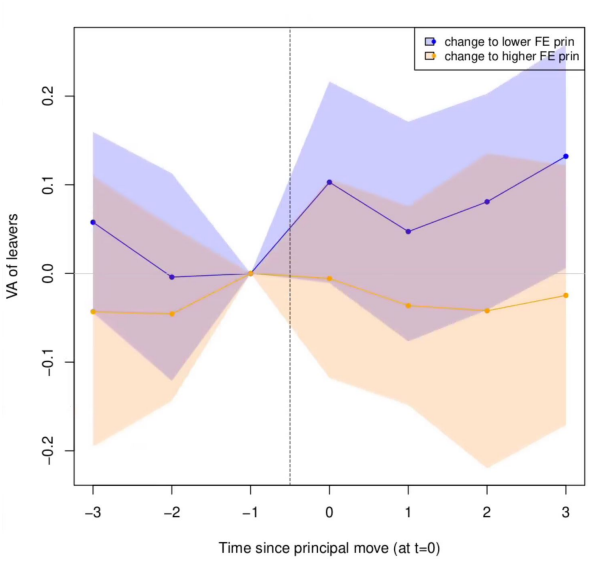
Figure A.2: Event study estimates of how the average number of behavioral incidents at a school change around a principal change. Principal changes are classified as going from to a higher or lower-FE principal, where the FE are associated with test scores.



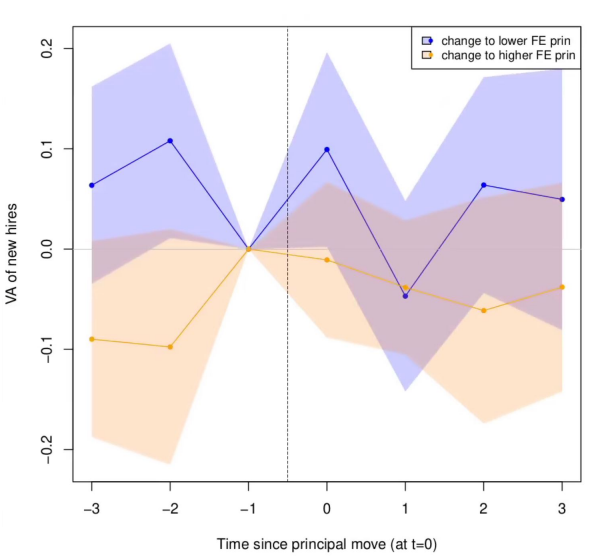
(a) Outcome: Avg. % of teachers from $t-1$ who are rehired in year t



(b) Outcome: Avg. % of teachers in year t who are new hires

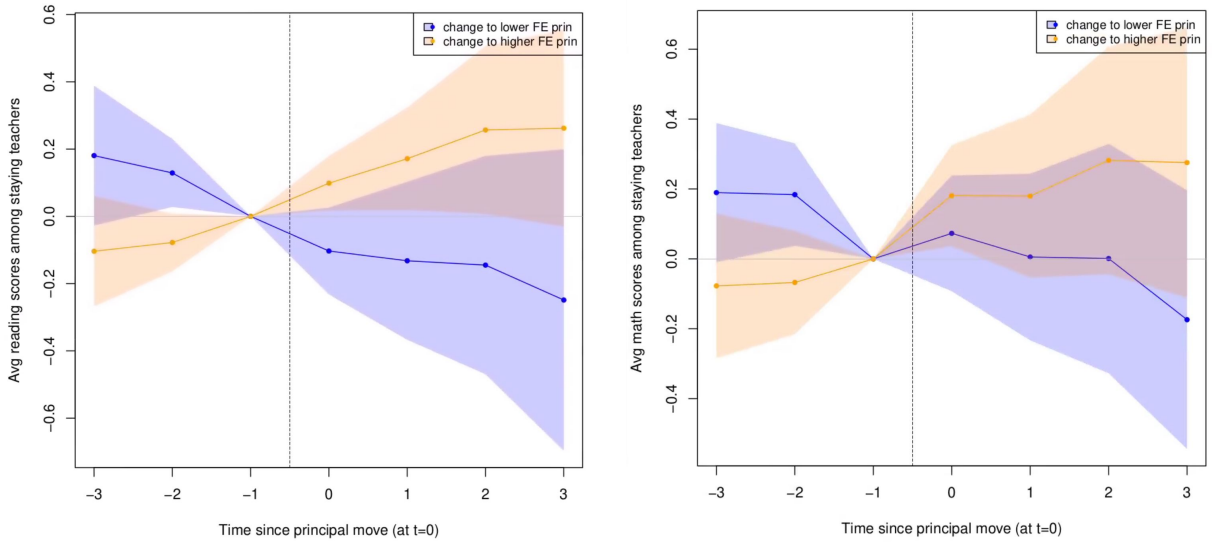


(c) Outcome: Avg. teacher FE among teachers who leave at the end of year t



(d) Outcome: Avg. teacher FE among new hires who begin in year t

Figure A.3: Event study estimates of how teacher retention, hiring rates, and teacher fixed effects change around the time of a principal change. Principal changes are classified as going from to a higher or lower-FE principal, where the FE are associated with test scores. Data are from Texas public schools.



(a) Outcome: Avg. reading scores among staying teachers

(b) Outcome: Avg. math scores among staying teachers

Figure A.4: Event study estimates of how average test scores change for **staying teachers**, i.e. incumbent teachers who stay for the next 3 years in the school, around a principal change. Principal changes are classified as going from to a higher or lower-FE principal, where the FE are associated with test scores.

B Additional Tables

Table B.1: Comparison between value-add estimates in the literature change to principal FE only

		Teacher FE only	Teacher and Principal FE	
	Data	s.d.(teacher FE) (1)	s.d.(teacher FE) (2)	s.d.(principal FE) (3)
My estimates	KY	0.15	0.31	0.13
From the literature				
Chetty et al., 2013	U.S.	0.10-0.14		
Kane and Staiger, 2008	LAUSD	0.11-0.18		
Branch et al., 2013	TX			0.24
Munoz and Prem, 2021	Peru			0.29

Note: Own estimates use standardized reading scores in elementary and middle schools in KY.

Table B.2: Comparison of relationship between principal and school fixed effects with teachers' average ratings of the school

	Principal FE	School FE
	(1)	(2)
Average leadership score	-0.0612*	-0.0488
	(0.0333)	(0.0510)
Average time use score	-0.0186	-0.0491**
	(0.0150)	(0.0245)
Average student conduct score	0.0265**	0.0517**
	(0.0131)	(0.0200)
Average teacher leadership score	-0.0018	0.0943**
	(0.0298)	(0.0452)
Average instructional practices score	0.0856***	-0.0833*
	(0.0282)	(0.0463)
Average professional development score	-0.0122	-0.0228
	(0.0238)	(0.0391)
Average community support score	0.0409**	0.0046
	(0.0189)	(0.0310)
Observations	2,134	1,478
R^2	0.08010	0.49213
Adjusted R^2	0.07358	0.48692

Note: Regressions are at the school-year level, in years when the TELL Survey was run (every 2 years from 2011 to 2017).

Table B.3: Comparison of relationship between principal and school fixed effects and teachers' ratings of school leadership

	Principal FE	School FE
	(1)	(2)
The faculty and leadership have a shared vision.	0.0730***	-0.0012
	(0.0235)	(0.0373)
There is an atmosphere of trust and mutual respect.	-0.1045***	-0.0155
	(0.0238)	(0.0359)
Teachers feel comfortable raising issues and concerns.	0.0370*	0.0766**
	(0.0215)	(0.0341)
The school leadership consistently supports teachers.	0.0294	0.0298
	(0.0277)	(0.0435)
Teachers are held to high professional standards for delivering instruction.	-0.0413*	0.0331
	(0.0240)	(0.0347)
The school leadership facilitates using data to improve student learning.	0.0497**	0.1272***
	(0.0245)	(0.0355)
Teacher performance is assessed objectively.	-0.0717**	-0.0130
	(0.0306)	(0.0466)
Teachers receive feedback that can help them improve teaching.	0.0315	-0.0399
	(0.0262)	(0.0383)
The procedures for teacher evaluation are consistent.	0.0192	-0.1730***
	(0.0284)	(0.0416)
The school improvement team provides effective leadership at this school.	0.0041	-0.0872*
	(0.0274)	(0.0445)
The faculty are recognized for accomplishments.	0.0252*	0.0402*
	(0.0144)	(0.0221)
The school leadership makes a sustained effort to address teacher concerns about...		
Leadership issues.	-0.0515*	0.0713
	(0.0294)	(0.0438)
Facilities/resources.	0.0285	-0.0265
	(0.0249)	(0.0400)
The use of time in my school.	-0.0203	-0.0870**
	(0.0240)	(0.0367)
Prof. learning.	0.0077	0.1270***
	(0.0216)	(0.0334)
Teacher leadership.	0.0069	-0.0494
	(0.0370)	(0.0541)
Community support and involvement.	0.0071	-0.0456
	(0.0255)	(0.0381)
Managing student conduct.	-0.0040	0.0381
	(0.0163)	(0.0243)
Instruc. practices and support.	0.0355	-0.0672
	(0.0345)	(0.0494)
New teacher support.	-0.0221	0.0346
	(0.0201)	(0.0262)
Observations	2,134	1,478
R^2	0.09096	0.52096
Adjusted R^2	0.07887	0.51171

Note: Regressions are at the school-year level, in years when the TELL Survey was run (every 2 years from 2011 to 2017).

Table B.4: Comparison of relationship between principal and school fixed effects with teachers' ratings of student conduct

	Principal FE	School FE
	(1)	(2)
Students at this school understand expectations for their conduct.	0.0689***	0.0181
	(0.0228)	(0.0328)
Students at this school follow rules of conduct.	0.0267	0.0101
	(0.0176)	(0.0260)
Policies and procedures about student conduct are clearly understood by the faculty.	-0.0140	-0.0645**
	(0.0206)	(0.0323)
School administrators consistently enforce rules for student conduct.	-0.0590**	0.1406***
	(0.0253)	(0.0375)
School administrators support teachers' efforts to maintain discipline in the classroom.	0.0318	-0.1390***
	(0.0254)	(0.0381)
Teachers consistently enforce rules for student conduct.	-0.0038	0.0532**
	(0.0158)	(0.0251)
The faculty work in an environment that is safe.	-0.0123	-0.0273
	(0.0183)	(0.0273)
Observations	2,134	1,478
R^2	0.08136	0.49504
Adjusted R^2	0.07486	0.48986

Note: Regressions are at the school-year level, in years when the TELL Survey was run (every 2 years from 2011 to 2017).

Table B.5: Comparison of relationship between principal and school fixed effects with teachers' ratings of instructional practices

	Principal FE	School FE
	(1)	(2)
Teachers use assessment data to inform their instruction.	0.0656***	0.1567***
	(0.0210)	(0.0334)
Teachers work in prof. learning communities to develop/align practices.	-0.0069	0.0902***
	(0.0174)	(0.0266)
Provided supports translate to improvements in instruc. practices.	-0.0141	-0.0628*
	(0.0241)	(0.0377)
Teachers are encouraged to try new things to improve instruction.	0.0310	-0.0905***
	(0.0238)	(0.0346)
Teachers are assigned classes that maximize their likelihood of success with students.	-0.0129	-0.0463*
	(0.0160)	(0.0252)
Teachers have autonomy to make decisions about instruc. delivery.	0.0068	-0.0226
	(0.0131)	(0.0205)
Observations	2,134	1,478
R^2	0.07683	0.50901
Adjusted R^2	0.07073	0.50431

Note: Regressions are at the school-year level, in years when the TELL Survey was run (every 2 years from 2011 to 2017).

Table B.6: Comparison of relationship between principal and school fixed effects with teachers' ratings of community support at school

	Principal FE	School FE
	(1)	(2)
Parents/guardians are influential decision makers in this school.	-0.0128	-0.0459**
	(0.0143)	(0.0226)
This school maintains clear, two-way comm. with the community.	-0.0682***	-0.1684***
	(0.0255)	(0.0373)
This school does a good job of encouraging parent/guardian involvement.	0.0206	0.2112***
	(0.0239)	(0.0340)
Teachers provide parents/guardians with useful info about student learning.	-0.0679***	-0.0489
	(0.0254)	(0.0391)
Parents/guardians know what is going on in this school.	0.0545**	0.1086***
	(0.0267)	(0.0407)
Parents/guardians support teachers, contributing to their success with students.	0.0366*	-0.0304
	(0.0195)	(0.0302)
Community members support teachers, contributing to their success with students.	0.0493*	0.0277
	(0.0282)	(0.0403)
The community we serve is supportive of this school.	0.0019	-0.0894**
	(0.0257)	(0.0378)
Observations	2,134	1,478
R^2	0.08716	0.50866
Adjusted R^2	0.08026	0.50327

Note: Regressions are at the school-year level, in years when the TELL Survey was run (every 2 years from 2011 to 2017).

Table B.7: Comparison of relationship between principal and school fixed effects with teachers' reported immediate plans: Exit outcomes

	Principal FE	School FE
	(1)	(2)
Intend to leave education	-0.1384*	-0.0684
	(0.0722)	(0.1346)
Observations	2,134	1,478
R^2	0.06833	0.48348
Adjusted R^2	0.06438	0.48031

Note: Regressions are at the school-year level, in years when the TELL Survey was run (every 2 years from 2011 to 2017).

Table B.8: Comparison of relationship between principal and school fixed effects with teachers' reported immediate plans: Staying outcomes

	Principal FE	School FE
	(1)	(2)
Intend to stay at school	0.1200*	0.0564
	(0.0723)	(0.1343)
Intend to teach at another school/district	0.0540	-0.1245
	(0.1661)	(0.2902)
Intend to move to admin	-0.0374	-0.4398***
	(0.0961)	(0.1669)
Intend to move to non-admin role	0.0867	0.2105
	(0.1228)	(0.2089)
Observations	2,134	1,478
R^2	0.07421	0.49483
Adjusted R^2	0.06897	0.49069

Note: Regressions are at the school-year level, in years when the TELL Survey was run (every 2 years from 2011 to 2017).