

Minh Nguyen[†], Steven G. Rivkin, Lauren Sartain, and Jeffrey C. Schiman*

School district investments in general skills: The case of principal residency programs

Abstract

Districts are investing more in school leadership, including the implementation of residency training programs for aspiring principals. There is limited evidence about the district return on these investments, and none in the context in which principal hiring is decentralized at the school level. This paper develops a model highlighting the conditions under which districts benefit from investments in general leadership skills and examines the residency program in the Chicago Public Schools (CPS). The event-study analysis that addresses treatment effect heterogeneity finds that principals who complete a residency are significantly more effective at raising achievement. Although the leadership skills gained through the residency are likely to make a principal more valuable to districts outside of CPS, the large majority of residents remain in CPS despite the absence of salary premia for completion of a residency or high performance. These findings suggest the presence of transition costs that enable CPS to retain more effective, residency-trained principals without having to increase pay, thereby realizing some of the return on the investment in general skills.

Current version: February 6, 2023

Keywords: school principal training, principal labor markets, general skill investment

JEL-codes: I20, J45, M53

Corresponding author: Jeffrey C. Schiman
jschiman@georgiasouthern.edu

[†]Nguyen, Ball State University; Rivkin, University of Illinois at Chicago and NBER; Sartain, University of North Carolina at Chapel Hill and University of Chicago Consortium on School Research; and Schiman, Georgia Southern University. The authors are grateful to the Smith Richardson Foundation for financial support. They also thank the Chicago Public Schools for data access and information about the policy context. The research was approved by the Institutional Review Board.

1 Introduction

A growing emphasis on leadership as the focal point of school improvement has led states, districts, and school management organizations to adopt various strategies to elevate the effectiveness of school principals. Policy initiatives can be divided into two broad categories. The first includes those designed to raise the effectiveness of current principals including the introduction of rigorous and informative principal evaluations, heightened school accountability, or a closer association between compensation and performance (Dee and Jacob 2011, Ladd 1997, Hanushek and Raymond 2005, Neal and Schanzenbach 2010, Superville 2014), while the second set of initiatives focuses directly on improvements in the preparation and skills of incoming principals (Gates et al 2014, Goldring and Sims 2005, Hale and Moorman 2003, Herrington and Wills 2005).

Concerns about the misalignment of the content of training in traditional administrator preparation programs and the skills principals need on the job have led some charter school management organizations and traditional public-school districts to alter principal preparation by making a paid residency a key component of the program (Gates et al. 2014, Hess and Kelly 2005, Knechtel et al. 2015). Rather than acquiring administrator certification and service as an assistant principal as the stepping-stone to a principal position, residents have the chance to focus on the acquisition of leadership skills under the tutelage of a mentor principal without assistant principal responsibilities. Districts and charter management organizations, including KIPP Charter Schools that operate schools in urban areas across the U.S., have been particularly interested in residency programs and have invested substantial resources to this end (Knechtel et al. 2015, Braun, Billups, and Gable 2013).

From the district perspective, the return on investments in a residency program depend primarily upon three factors: 1) the value of the training as measured by its impact on leadership effectiveness; 2) the difference between the salary paid to a resident-trained principal and the value of the principal's work; and 3) timing and length of service in a leadership role in the district. Importantly, much, if not all, of the residency training likely augments general leadership skills that are valued by other districts and potentially other professions. Expanded opportunities at other districts could raise the probability of moving to a new position or push the district to increase compensation, either of which transfers at least some of the benefits of training from the district to the resident.

In this paper, we describe a model based on Becker (1962) that highlights the conditions under which districts realize benefits of investing in general leadership skills and then investigate the effects of a residency on principal effectiveness as measured by student achievement. The model shows that a residency program that augments skills likely elevates the effectiveness of a trainee in both the current and alternative districts, thereby elevating outside employer demand. This would be expected to increase turnover in the absence of salary increases or moving costs large enough to deter transitions to other employers. Therefore, the district return on investments in residents depends not only on the residency impact on quality but also on how long it takes trainees to ascend to a principal position, length of tenure as a principal in the district, and any competition-induced salary increases.

We use event-study methods that are robust to treatment effect heterogeneity by school and time to estimate residency program effects in the Chicago Public Schools (CPS) and find that the hiring of a resident-trained principal leads to higher annual achievement growth.

Because principal transitions are endogenous and may differ by residency status, we focus comparisons on achievement differences following the hiring of a new resident or nonresident principal regardless of how long they remain in the position; limiting the sample to schools in which new hires remain at least three years has little effect on the results. Importantly, selection of candidates into residency programs may be as important as training in the determination of program benefits if the program attracts and selects more skilled candidates than would have entered the principal pipeline in the absence of the program. There may even be another channel of selection if the program generates better information than work as an assistant principal on expected future effectiveness as a school leader. We acknowledge that selection may be part of what makes a residency program successful, and that we lack the data necessary to separate training and selection effects. Regardless, the higher average effectiveness of residents has important policy implications in the absence of other methods for attracting and selecting high-performing school leaders.

In terms of career trajectories following program completion, few residents enter a principalship at a CPS school directly from their training program, and only about half are in a CPS principalship four years following program completion. Despite the potential appeal to other districts, residency completion does not increase the probability a principal exits the CPS. Given the positive program impact, this suggests the presence of moving costs including a preference to live or work in Chicago that are substantial enough to deter transitions even in the presence of higher demand from outside employers.

Our study extends research on principal residency programs and preparation programs more broadly through the analysis of a large, urban, highly decentralized district, the use of methods that address complications introduced by heterogeneous treatment effects, the inclusion of school fixed effects to account for unobserved time-invariant differences among schools, and the steps taken to mitigate biases introduced by endogenous principal transitions. The positive program effects in the CPS contrast with the findings of little or no effect of the 2003 New York City school district alternative certification program for school leaders that included a residency component (Corcoran, et al, 2012). The New York City Aspiring Principals Program was developed and operated by the district and employed a curriculum that concentrated coursework during a six-week period prior to the 10-month residency. Following completion of the residency the aspiring principals participated in a transitional planning summer. This contrasts sharply with the structure of the CPS program where participating graduate programs including local institution and nationally known principal preparation programs integrate a residency component into an urban leadership curriculum that focuses on issues faced by leaders in urban schools and is aligned with the State of Illinois Principal Standards. Whether the differences in findings emerge from program differences, the advantages in CPS from having programs compete with one another for students, the use of empirical methods that account for heterogeneous treatment effects in this but not the NYC study, or other factors remains uncertain. The results in this study also contrast with a recent quasi-experimental empirical analysis on principal preparation program effects that does not find strong evidence that some preparation programs outperform others across a range of outcomes (see for example, Grissom et al. 2019).

The remainder of the paper has the following structure. Section 2 describes the CPS residency program including curricula, the program costs that are divided among the

district, participants and other revenue sources, and employment conditions. Section 3 describes the CPS data. Section 4 builds on Becker (1962) and Acemoglu and Pischke (1999a; 1999b) to develop a conceptual framework for consideration of returns to investment in general school leadership skills given the CPS institutional structure. Section 5 describes post-program career paths. Section 6 compares the effectiveness of residency participants with other newly hired principals, and Section 7 summarizes the findings and discusses implications for policy.

2 The Context of Chicago Public Schools

2.1 District-Sponsored Principal Training Programs

Due to concerns about the quality of school principals combined with turnover rates around 20 percent annually (Goldring and Taie 2018), CPS launched a principal residency program called the Chicago Leadership Collaborative (CLC) in November 2011 (Chicago Public School 2020). In the spirit of well-known teacher preparation programs like Teach for America, the CLC sought to attract candidates with proven track records of success into leadership roles in Chicago schools. Specifically, the goal of the CLC was to strengthen the pipeline of aspiring school leaders, starting with classroom-based principal training focused on the skills needed to work in CPS specifically and urban schools more broadly, followed by a school-based residency under the guidance of a mentor principal and a transition into a CPS principal position. According to CPS promotional materials, the hope was that the CLC would make Chicago a “destination of choice for urban school leaders.” The programs advertised the rigor of their admissions processes, suggesting that at least some of the value of these programs may be in the screening and selection of principal candidates. Requirements for admission often included prior classroom experience, as well as prior leadership positions such as department or grade-level chairs, teacher mentors, or leadership positions outside the field of education. We note that current principals are not eligible to participate in the CLC.

The CLC includes both classroom and residency components. Students enroll in a participating graduate program at either a local institution or in partnerships with nationally known principal preparation programs (e.g., New Leaders for New Schools), in contrast to the New York City Aspiring Principals program run by the district which compresses classroom instruction into a six-week summer course offered prior to the yearlong residency (Corcoran et al., 2012). The content of the CLC training focuses on issues faced by leaders in urban schools and is aligned with the State of Illinois Principal Standards. A key component of the CLC is the year-long residency that candidates complete in CPS schools under the tutelage of a “high-performing” principal. The CLC training programs also continue to support candidates once they attain principal positions, usually for 1-2 years following program completion. Upon completion of the program and depending on the program requirements, candidates receive either a masters in school administration (MSA) or a doctoral degree in education (EdD).

The district commits substantial funding to the CLC program. Between 2012 and 2017, the district has awarded subcontracts to the principal preparation programs totaling \$16,190,004. Per district contracts, that funding goes toward development of curriculum, recruitment of

candidates to the programs, and compensation to mentors for candidates who are placed as principals. Candidates participating in the residency receive compensation equivalent to \$80,000 per year or their annual salary if they were already in a CPS position, whichever is greater. They also receive the same benefits offered to CPS employees. Candidates must cover the costs of tuition and fees for the preparation programs. Recognizing that the training may make program participants more attractive to other districts, participants are contractually bound to repay the district up to \$7,500 if they decline an offer as a CPS administrator or they do not work for CPS for four years upon completion.

2.2 Contracts for and Expectations of CPS Principals

The Chicago School Reform Act of 1988 established four-year principal contracts along with a decentralized governance structure in which democratically elected “local school councils” (LSCs) comprised of teachers, parents, and community members maintain authority over principal hiring and evaluation unless the district intervenes for cause (Ryan et al. 1997; Schiman et al. 2018) the 1988 Chicago School Reform Act decentralized school governance by forming elected local school councils (LSCs). Consequently, while the district may influence principal placement, they do not control who is offered a principal contract in most schools.¹ Candidates to open principal positions are required to hold the state’s administrator license and complete the district’s eligibility process. The job positions are listed centrally on the CPS website, and central office staff process applications to make sure candidates meet the minimum requirements. Beyond that, LSCs drive all other aspects of the hiring process – they write the job postings and choose what skills to emphasize in the description, they review all qualified applicants, they decide which candidates to interview, and finally they determine which candidate to hire. In addition to awarding principal contracts, the LSCs are required to conduct annual evaluations of school principals, and principals are also evaluated by their supervisor in CPS. At the end of a principal’s four-year contract, the LSC can decide to renew or terminate the contract.

Upon entering a CPS principalship, the principal has multiple roles and expectations. Across school districts, and including CPS, the role of the principal is quite extensive. Principals are expected to lead the school operations, including setting the budget, managing personnel, and maintaining facilities. In CPS, school funding occurs through a student-based budgeting process, which means that dollars are allocated to schools based on enrollment. CPS principals then have discretion around allocating those funds to hire teachers and other school personnel when there are vacancies, as well as evaluating teachers and providing teachers with actionable feedback to improve their instructional practice and, thus, improve learning. CPS principals are also the school’s designated instructional leader, meaning that they are supposed to make key decisions as they relate to learning and teacher development. Thus, there are many mechanisms through which principals can ultimately affect student outcomes.

¹ The few exceptions are 1) schools that are persistently low performing on the district’s accountability system, and 2) schools that do not have full membership of their local school councils (e.g., there were not enough candidates who ran to sit on the LSC).

2.3 Principal Compensation

Another important consideration is how principals are compensated in CPS. Similar to other districts and to teacher pay, school administrators in CPS are awarded compensation packages of salaries and benefits as established by the school district and approved by the school board.² As of 2017, the salary range for CPS principals was \$125,000 to \$169,000. The actual salary earned is determined by enrollment, so principals at larger schools receive higher salaries. Notably, the salary range is quite compressed with little room for advancement. Further, principal performance is not a determinant of salary, though there have been pilot initiatives that implement a merit pay model (Schiman 2021).

3 Data

We combine CPS administrative data from 2007 to 2019 with personnel data on CPS principals and residency program participants that are available from 2013 to 2018. The administrative data include student demographics such as race/ethnicity and gender, academic information including test scores and special education status, and a unique identifier to link observations across years so long as the student remains in the CPS. We use CPS personnel data to build a panel of principals linked to the school in which they are employed and merge these with residency program participant data that include cohort year, program, and residency school. Merging the resident participant roster with the personnel data enables us to track careers in CPS following completion of the residency.

Test score data come from the Illinois Standards Achievement Test (ISAT) from 2007 to 2012 and the Northwest Educational Assessment (NWEA) test from 2013 to 2019, the tests used for accountability during this timeframe. We use achievement scores to generate school average annual achievement and achievement gain, the measures of principal effectiveness which we describe further in section VI. Achievement constitutes only one type of skill, and principals likely affect the acquisition of socio-emotional and behavioral skills as well (Jackson et al 2020). The absence of measures of these skills leads us to focus on achievement. Importantly, prior work on CPS principals shows a strong relationship between principal value added to achievement and teacher evaluations of principal leadership skills (Branch et al., 2020), suggesting that differences in achievement and particularly achievement growth capture important variation in principal effectiveness.

Table 1 reports the number of residents and first-year principals and assistant principals by year for our sample of elementary schools.³ In total, 265 residents were trained between 2013 and 2017, with the annual number declining by more than one third following 2015. Importantly, the number of program graduates exceeds the number of first-year principals in the

² Information on the CPS pay policy is available at the following link: <https://www.cps.edu/sites/cps-policy-rules/policies/300/302/302-8/> (last accessed 10 January 2023).

³ Our analysis focuses on elementary schools in the CPS, which run from kindergarten through 8th grade (primary schooling). In the district, there are roughly 478 elementary schools serving over 300,000 students. In any given year, rates of principal turnover in the CPS hover around 20%, suggesting roughly 96 principal vacancies at the elementary level. In our dataset, we are unable to determine the reason for principal departures and vacancies though evidence from Betteille, Kalogrides, and Loeb (2012) suggests that principals depart schools for a variety of reasons including district-level decisions or principal initiated moves, often gravitating towards higher achieving schools. In the Chicago context, the Local School Council (made up of parents, community members, teachers, and the principal themselves) are responsible for principal contract renewal decisions every four years for the majority of schools.

subsequent year, raising the possibility that a shortage of jobs slows entry into a leadership position. Although it may contribute to slower progress, fewer than 50 percent of open positions are filled by a residency-trained principal in any given year. This suggests that school-based hiring by local school councils, perhaps combined with resident application decisions, serves as the primary hindrance to becoming a principal.

The top panel of Table 2 describes the characteristics of schools where a resident first becomes principal (Column 1) and where a non-resident first becomes principal (Column 2). A comparison of Columns 1 and 2 shows that residents who assume a principal position tend

Table 1 Number of Residency-program Graduates by Cohort of Program Completion, and Number of First-year Elementary School Principals and Assistant Principals by Year and Residency-Program Participation Status

	Residency program graduates	First-year principals		First-year assistant principals	
		Residency program participants	Non-participants	Residency program participants	Non-participants
2013	63				
2014	65	5	36	18	70
2015	63	18	25	19	68
2016	39	18	23	17	40
2017	35	15	27	12	25
2018		14	15	6	28
Total	265	70	126	72	231

Notes: First-year principal or assistant principals are counted once in the first year they appear in a CPS school. The sample of first-year principals represents those that are included in our estimation sample.

Table 2 Descriptive Statistics for the Principal Sample

	Residents	Non-residents	Difference (Residents minus non-Residents)
<i>Student Variables</i>			
Math z-score	-0.076 [0.062]	-0.010 [0.046]	-0.065
Reading z-score	-0.075 [0.057]	-0.022 [0.038]	-0.053
Black	0.487 [0.050]	0.370 [0.040]	0.117*
Hispanic	0.367 [0.048]	0.452 [0.038]	-0.085
Other race	0.042 [0.010]	0.069 [0.017]	-0.027
Female	0.489 [0.004]	0.497 [0.003]	-0.007
Special Education	0.150 [0.005]	0.145 [0.004]	0.005
Grade Average Enrollment	61.599 [4.269]	70.345 [3.695]	-8.746

(continued)

Table 2 Continued

	Residents	Non-residents	Difference (Residents minus non-Residents)
<i>Principal variables</i>			
Asian	0.000 [0.000]	0.014 [0.014]	-0.014
Black	0.015 [0.015]	0.010 [0.010]	0.005
Hispanic	0.425 [0.062]	0.350 [0.044]	0.075
Multiple race	0.143 [0.044]	0.167 [0.041]	-0.024
Female	0.754 [0.052]	0.587 [0.050]	0.167**
Birth Year	1977.436 [0.717]	1971.143 [0.579]	6.293***
Proportion with the job title: acting	0.034 [0.020]	0.049 [0.021]	-0.015
Proportion with the job title: interim	0.524 [0.064]	0.671 [0.047]	-0.147*
Proportion with the job title: contract	0.441 [0.064]	0.280 [0.044]	0.161**
Number of observations	23,356	41,465	

Notes: Student characteristics are measured in the year prior to the entry of a the new principal or assistant principal. Standard deviation in square brackets. Statistical significance of the resident/non-resident difference is based on regressions with standard errors clustered by school. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

to work in schools that have lower test scores, higher Black enrollment, and lower Hispanic enrollment. Average mathematics achievement is roughly 0.065 standard deviations lower in the schools of program participants than those of non-participants, and the gap in reading achievement is slightly smaller at 0.053 standard deviations. The Black enrollment share is roughly 12 percentage points higher in schools led by a program participant, and share Hispanic is 8.5 percentage points lower. Differences in demographic composition must be considered in any examination of the effectiveness of resident-trained principals in comparison to other new principals. However, note that differences in the Black enrollment share constitute the only statistically significant differences by resident participation.

The bottom panel reports the characteristics of resident-trained and non-resident trained principals. Resident principals are more likely to be female (75% versus 59%), younger (born in 1977 on average versus 1971 on average), and to be hired into a contract rather than an interim principal position. Gender and age differences likely come at least in part from residency training program efforts to recruit more diverse classes. Roughly 52% of residents become interim principals and 44% become contract principals, while roughly 67% of non-residents become interim principals and 28% become contract principals. Importantly, the distinction between the classifications of contract and interim primarily reflects the hiring authority rather than the expected length of service. Interim principals

are appointed by the district CEO rather than contracted by the Local School Council, either because the Council no longer has hiring authority due to consistently low school performance (here the interim appointment lasts indeterminately until the school leaves probation) or failure of the Council to approve a principal by voting (here the appointment lasts up to one year).⁴

4 Conceptual Framework

The impact of the residency program on principal effectiveness is a primary but not the only determinant of value of the program to CPS. First, the magnitude of the impact may depend on the subsequent career trajectory, as the benefit may be smaller for those serving as an assistant principal than those serving as a principal, and it may depreciate the longer the gap between program completion and ascension to a principal position. Second, the training likely augments general leadership skills or at least those relevant to urban settings as opposed to only CPS. Consequently, residency completion may lead to more attractive job offers from outside employers who recognize the value of CPS' training program. This would be expected to increase transitions out of CPS despite the small financial penalty for leaving the district. The increase in outside employer demand would erode the return on the district investment by either shortening careers in CPS or forcing CPS to raise salaries of resident principals to deter exits.

The canonical Becker (1962) analysis of job training shows that firms or organizations realize little or no benefit from general training that raises productivity outside the current organization because others will bid away trained employees unless the organization raises salary by the full amount of the productivity increase. Becker's results raise questions about the benefits of principal residency training programs to the district, given that these programs likely enhance principal effectiveness in other settings. However, deviations from a perfectly competitive, frictionless labor market for principals may essentially transform the general training to be quasi-firm specific.⁵ We note that this discussion refers to the skills acquired during the residency, but selection based on pre-existing skills is another channel through which the residency program raise leadership quality. Importantly, the conceptual framework holds regardless of whether any higher productivity of residents comes from skills acquisition, selection, or some combination of the two.

We first present a model of investment in general skills within a perfectly competitive, frictionless labor market and then extend the basic framework to incorporate the institutional structure of CPS and the principal labor market. Specifically, we allow delay in ascension to a principal position to depreciate the value of the training, incorporate the rigid CPS principal salary structure, and recognize that there may be fixed costs to transitioning to a position outside of CPS. Taken together, these deviations from the perfectly competitive, frictionless environment have an ambiguous effect on district returns to funding a principal residency program, though the decentralized governance structure might present an impediment to the realization of a positive return on program investment.

⁴ https://www.cpsboe.org/content/documents/chapter_iv_august_2013.pdf (last accessed 7 March 2021)

⁵ Components of the training may also focus on CPS and therefore be district specific.

4.1 Baseline Model

Consider the investment decision for an employer who operates in perfectly competitive and frictionless output and labor markets, where there are no hiring costs, and no geography so that workers bear zero cost to change jobs. In the presence of training costs, firms consider both current and future costs and revenues in making hiring decisions. To simplify the presentation, assume that all training costs occur in period zero, the increase in marginal revenue product (MRP) remains constant following training and firms hire until the point that the expected present value of the marginal revenue products over $(y+1)$ years equals the expected present value of wage and training costs for $(y+1)$ years

$$(1) \text{ Expected present value of MRP for } (y+1) \text{ years} = (1-\gamma) * MRP_0 + \sum_{y=1}^Y \frac{(MRP_0 + \Delta)_y}{(1+r)^y}$$

$$(2) \text{ Expected present value of wage and training cost for } (y+1) \text{ years} = w_0 + k + \sum_{y=1}^Y \frac{w_y}{(1+r)^y}$$

where w is wage, k is the financial cost of training, Δ is the training-induced increase in MRP , MRP_0 is marginal revenue product prior to training, and γ is the share of productive time lost to training. Assume the firm incurs all training costs in year 0.

In equilibrium, firms hire until (1) and (2) are equal. Given the assumptions of a perfectly competitive and frictionless labor market with zero hiring costs, outside firms will be willing to pay a wage equal to $MRP_0 + \Delta$. Therefore, to retain a trained worker the firm must raise the wage equal to $MRP_0 + \Delta$ in all subsequent years. This produces the familiar result that $MRP_0 + \Delta = w_y$ in all periods, the summation terms in equations (1) and (2) exactly cancel out, and firms receive no benefit from the increase in productivity. Consequently, profit maximizing firms would not be willing to bear any training costs. As equation (3) shows, financial (k) and time (γMRP_0) costs of training are borne by the worker in the form of lower wages in period zero.

$$(3) \quad w_0 = MRP_0 - (\gamma MRP_0 + k)$$

4.2 Deviations from Profit Maximization and Competitive Markets

Of course, CPS and other large urban districts do not function as profit maximizing firms, and the labor market for principals diverges from perfect competition in myriad ways including rigidities in district salary structures, the treatment of experience gained in other districts, the geographic specificity of schools, licensing requirements, and imperfect information about productivity. This highlights the importance of incorporating relevant institutional features into the analysis. Three particularly important factors for CPS would appear to be 1) delayed hiring as a principal resulting from local school autonomy over personnel decisions rather than centralized district assignment of candidates to schools; 2) salary compression within CPS which is typical for large urban districts; and 3) the fixed costs of switching districts.⁶

As mentioned previously, principal contracts in CPS are managed in a decentralized way, relying on school-based governance, which means that central office does not determine who is

⁶ Dinerstein, Megalokonomou, and Yannelis (2019) find that a longer delay between teacher training and hiring leads to skill depreciation, and Autor et al. (2015) finds that employment delay leads to human capital deterioration.

hired to be a school principal. Therefore, residents must compete for positions school by school, and that likely decreases the probability a resident is hired simply based on participation in the program. Additionally, an LSC may prefer a candidate over a resident for reasons other than leadership skills. These circumstances would result in increased time between program completion and entrance into the principalship with most candidates entering assistant principal or other positions immediately following completion of the CLC program rather than going straight to the principalship. Regardless of the reason, to the extent that the CLC training develops skills that are more applicable to a principal's role and responsibilities than an assistant principal's job, the delays in procuring a principalship likely lower the training benefit to the district.

Spending additional years as an assistant principal likely also degrades the value of the training. In the standard career path, a teacher transitions to an assistant principal position and subsequently to a principalship. A primary rationale for residencies is that assistant principal responsibilities hinder the acquisition of leadership skills and that residents will learn much more from the mentor principal who is selected based on performance. However, if the resident remains an assistant principal following program completion, some of the knowledge acquired during the residency will be lost or potentially irrelevant as the district changes. A longer delay likely amplifies this skill depreciation. Even if the training enhances productivity as an assistant principal, this is unlikely to compensate for the negative effects of the delay. Note that such skill degradation becomes far less relevant if selection on pre-existing skills rather than skill acquisition during the residency determines the higher productivity of residents.

Regardless of whether program participants serve as principals or assistant principals, the largely fixed CPS salary structure based on years of experience and size of school introduces a divergence between productivity and salary. Higher pay does not accompany program completion nor is pay typically determined by performance, so if residencies elevate productivity the salary in CPS will not reflect the increase in effectiveness. As Acemoglu and Pischke (1999a; 1999b) emphasize, such salary compression with respect to productivity potentially redistributes some of the returns to the acquisition of general skills from the worker to the firm, or in this case from the program participant to the district.

Equations (4) and (5) incorporate the delayed ascension to the principal position and CPS salary compression into equations (1) and (2)

(4) Expected present value of MRP for $y+1$ years working in CPS =

$$(1-\gamma)^*MRP_0 + \sum_{y=1}^A \frac{(MRP^a + \Delta^a)_y}{(1+r)^y} + \sum_{y=A+1}^Y \frac{(MRP^p + \lambda^A \Delta^p)_y}{(1+r)^y}$$

(5) Expected present value wage and training cost for $y+1$ years =

$$w_0 + k + \sum_{y=1}^A \frac{\bar{w}_y^a}{(1+r)^y} + \sum_{y=A+1}^Y \frac{\bar{w}_y^p}{(1+r)^y}$$

where the benefits of training (Δ) and the wage depend upon whether the participant is an assistant principal (a) or principal (p), the wage is determined by salary schedule and not productivity ($\bar{w}$) and the value of the training to productivity as a principal depreciates geometrically with the length of delay (A) at the rate $1-\lambda$. We assume $\Delta p > \Delta a$, i.e. that the benefit of the training is larger when working as a principal than as an assistant principal.

Equations (4) and (5) illustrate that the fixed salary structure and delayed employment as a principal have offsetting effects on the benefits of the program to CPS. On the one hand, the longer is the delay and the larger is the annual delay-related knowledge depreciation the lower is the training benefit. On the other hand, the wage compression for both assistant principals and principals redistributes at least a portion of the benefit to the district.

Of course, the duration of employment in CPS also affects the net return of the residency investment to the district. Differences in total compensation between a position in CPS and in an alternative district depend upon both pecuniary and non-pecuniary factors, and any weight placed on the opportunity to lead an urban school raises the probability of turning down a competing offer outside of CPS. In addition, commuting costs for those who have a strong preference for living in Chicago may be lower for a position within as opposed to outside of CPS. Finally, residency participants agree to pay CPS a nominal amount if they leave the district prior to an agreed upon number of years of employment, though the value of the promise likely exceeds the small monetary penalty.

Taken together, job and location preferences and the agreement to work in the district following program completion may introduce a large wedge between the CPS salary and the reservation offer that would induce a transition to a position outside of CPS even if the residency program substantially raises productivity. Therefore, the district may realize a substantial net benefit from investments in principal residency training programs that exceeds the net benefit of the next best use of education funds even if the program primarily augments general leadership skills. A more effective principal elevates the skill acquisition of the entire school rather than only a single classroom as is the case with a more effective teacher. Therefore, the influence on leadership quality would not have to be large to generate a positive return on the district investment.

5 Principal Career Paths

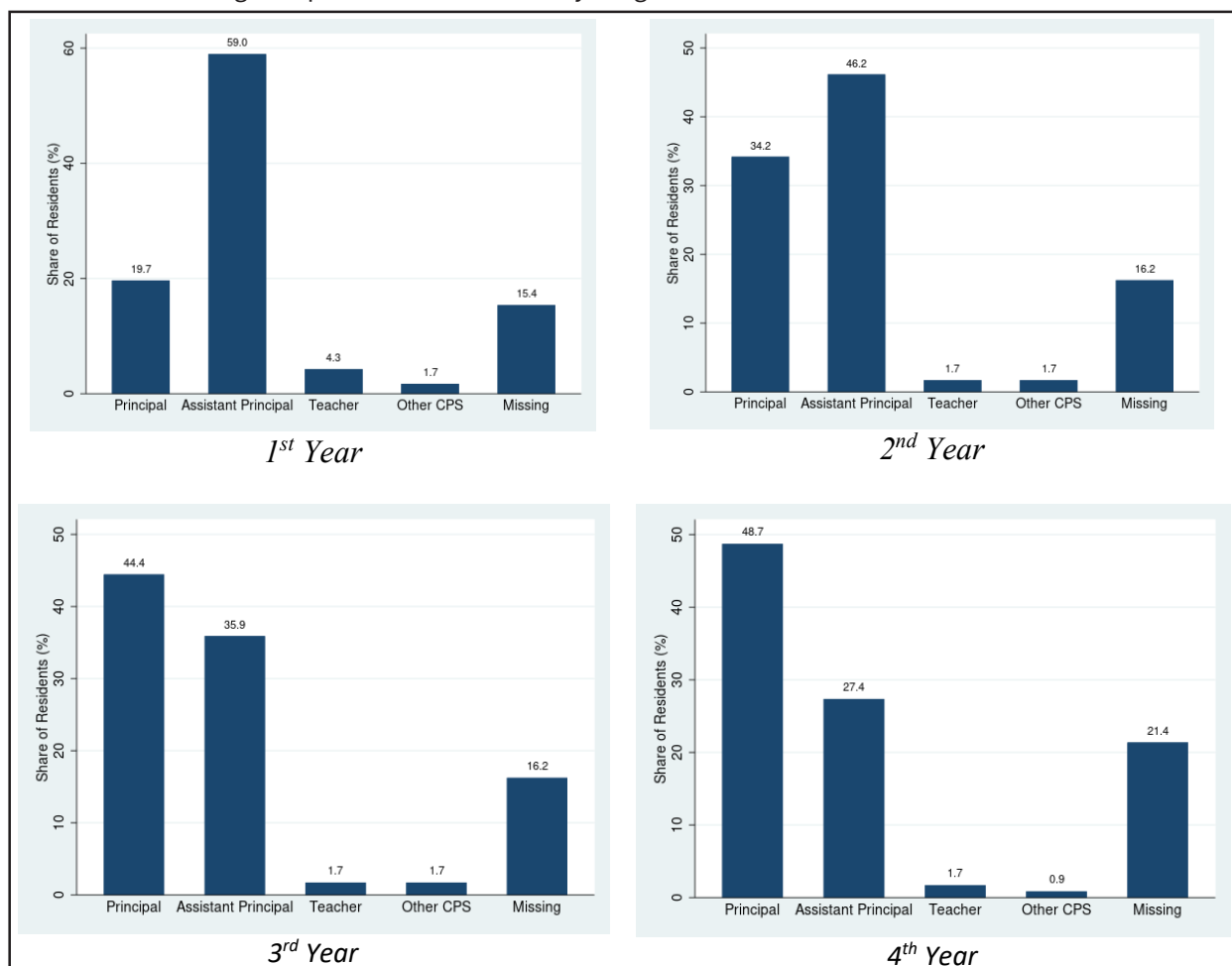
In the first component of our empirical analysis, we consider the career paths of residency program graduates, including progress toward, into, and out of principal positions. Figure 1 illustrates the employment distribution in each of the four years following program completion for the 2013 and 2014 cohorts and shows a steady increase in the fraction of principals, rising from less than one fifth in the first year following the residency to roughly 49 percent in the fourth year. The declines in the shares in instruction and other administrative positions (mostly assistant principals) roughly offset the increases in the principal share. Finally, the share of program completers who leave CPS increases by 6 percentage points between the first and fourth years. Missing residents include those who have exited the CPS sample for another district and those who exit the education sector.

To gain a better understanding of the relationship between residency participation and subsequent turnover, we focus on the following model,

$$(6) \text{ TURNOVER}_{jst} = \alpha + \beta \text{RESIDENT}_j + \gamma \mathbf{S}_{st} + \delta \mathbf{P}_j + \delta_t + \epsilon_{jst}$$

where turnover for the principal j serving in school s in year t is a function of an indicator for residency program training (RESIDENT_j), a vector of student demographic controls ($\mathbf{S}_{st}$) including school-year averages of racial composition, enrollment, share in special education, and share female, a vector of principal controls ($\mathbf{P}_j$) including race, female, birth year, and

Figure 1 Position Distribution of the 2013 and 2014 Residency Program Graduates, by Number of Years Following Completion of the Residency Program.



job title (interim principal, acting principal, or principal), year indicators (δ_t), and a random error term (ϵ_{jst}). In the first specification, $TURNOVER_{jst}$ is an indicator equal to one if it is the principal's last year leading a school and zero otherwise. In a second version of equation (6), we separate turnover into transitions to another position within CPS and transitions out of CPS entirely. Equation (6) is estimated using a sample of 184 principals hired between 2014 and 2017.⁷ The left panel estimates the probability a principal leaves her position during the sample period, and the right panel estimates the probability she leaves after her first year. Although year hired determines the number of years a principal is at risk of a transition, the entry year fixed effects ensure that residents are being compared with nonresidents whose terms begins in the same year.

The linear probability estimates in the top left panel of Table 3 suggest that turnover is higher among resident principals than among nonresidents, conditional on principal and school characteristics, but the multinomial logit estimates in the lower panels show that this reflects a higher rate of within- as opposed to between- district transitions. The estimates of the effects of residency on the relative probability of moving to another CPS school, $\beta_{within\ CPS}$,

⁷ Since the personnel data spans 2013 to 2018 and turnover is measured as a departure between t and $t+1$, we cannot calculate turnover in 2018. Residents begin entering the principal positions starting in 2014. Thus, the turnover analysis focuses on the period 2014 to 2017.

Table 3 Resident- Versus Non-Resident-Principal Transition Rates, by Destination and Period Examined

	Transition in any year (2014-2017)			Transition following the 1st year		
<i>Any destination</i>						
Resident Principal	0.130*	0.110	0.071	-0.00017	0.0018	-0.028
	(0.068)	(0.070)	(0.083)	(0.050)	(0.052)	(0.060)
Mean Turnover	0.30	0.30	0.30	0.10	0.10	0.10
Sample Size	184	184	184	184	184	184
<i>Multinomial Logit Estimates (relative to no transition)</i>						
Transition within CPS						
Resident Principal	2.65***	2.65***	3.30**	2.15*	2.40**	2.20
	(0.77)	(0.77)	(1.38)	(1.18)	(1.18)	(1.80)
Transition out of CPS						
Resident Principal	-0.10	-0.16	-0.55	-1.33	-1.36	-2.39**
	(0.42)	(0.44)	(0.60)	(0.86)	(0.96)	(1.18)
Sample Size	184	184	184	184	184	184
Student Controls	N	Y	Y	N	Y	Y
Principal Controls	N	N	Y	N	N	Y

Notes: We focus on newly entering principals between 2014 and 2017. The sample includes 184 newly entering principals. The left side of the table focuses on transitions in any year from 2014 to 2017. The right focuses on transitions occurring after the end of the new principal's first year. Student controls include school-year averages of racial composition, enrollment, share in special education, and share female. Principal controls include indicators for race, female, and job title (interim principal, acting principal, or principal) and birth year. All regressions control for year dummy variables. Robust standard errors are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

is positive and significant in all specifications. In contrast, estimates of the effects of residency on the relative probability of leaving CPS, $\beta_{leave\ CPS}$, is negative though not significant in the estimates that focus on all years. Note that when focusing on transitions out of the district after the first year (right panels), estimates from the full specification reported in Column 6 show that residents are significantly less likely to exit CPS and more likely to transition within CPS, though the latter effect is not significant.

The figure and table illustrate two patterns highlighted in the conceptual model that affect the return to residency program investment. On the one hand, the slow ascension to a principal position likely dampens the return due to both the delay in using skills learned during the residency and accompanying depreciation of their value. On the other hand, only one of five residents leaves the CPS even after four years, consistent with the notion that labor-market imperfections enable the CPS to extract some of the benefit of general training.

6 Productivity Effects of Residency Participation

This section uses an event-study framework to estimate differences in achievement between schools that hired a residency participant and those that hired a principal or assistant principal who did not complete a residency. After describing the empirical framework, we present the estimated effects of hiring a resident principal on achievement and annual achievement growth.

6.1 Empirical Framework

A growing body of research highlights the difficulties of identifying principal effectiveness as measured by value added to achievement or other outcomes Branch et al (2020), (Chiang, Lipscomb, and Gill (2016), Grissom, Kalogrides, and Loeb (2015), and Miller (2013)). The separation of principal effects from time-invariant and time-varying school effects including student composition presents a primary challenge, and we adopt an event-study framework with schools that hire non-resident principals during the same time period as controls. Restricting the controls to schools with principal transitions accounts for the dynamics identified by Miller (2013), which finds substantial average decreases in achievement just prior to principal transitions. Information on pretrends and changes in demographic composition illuminate divergent trends in unobserved school factors and student characteristics that could introduce bias.

Because of the endogeneity of transitions, we do not restrict the sample to the years in which the new principal remains in a school. Rather, the parameters capture the intent to treat and reflect differences in both productivity and transition costs resulting from turnover. Alternative estimates from samples that exclude principals who either leave within three years or are hired too recently to have three years of data are quite similar. The hiring date is determined by the first principal transition observed during the 2014 to 2018 sample frame, so a subsequent transition does not alter the initial classification based on residency status.

Equation 7 presents the basic event-study specification:

$$(7) \quad A_{ist} = \alpha + \sum_{j=-5}^{-2} \rho_j (\text{lag } j)_{st} + \sum_{k=1}^2 \tau_k (\text{lead } k)_{st} + \theta_s + \delta_t + \epsilon_{ist}$$

where achievement for student i in school s in year t (A_{ist}) is a function of school and time fixed effects, a random error and indicators for the periods leading up to and following the hiring of a resident principal. Period $j = -1$, the year prior to the new hire, is the omitted period, and $k=0$, the first year of a new principal spell, is dropped because of concerns about a recovery from a dip in the final year of the previous principal and the more limited influence over staffing, climate and other school operations in the first year. This two-way fixed effects (TWFE) framework controls for fixed differences between schools and time periods, and the parameters on the lags and leads capture the time-pattern of achievement around the hiring of a resident principal.

The key estimates of interest, ρ_j and τ_k , represent the periods j and k differences in achievement between schools that hired a new resident and the controls. The omitted period in the regression is the year before the new resident was hired (period $t = -1$), meaning that ρ_j and τ_k are benchmarked against that period. The pre-period estimates of ρ_j provide evidence on the common-trends assumption, and the estimates of τ_k capture resident-nonresident differences in achievement *after* the hiring of the new principal. If the hiring of a resident does not improve leadership quality and raise achievement, the estimates of τ_k would not be expected to be significantly different from zero. However, if residents are associated with improvements to test scores, we would expect to see a positive divergence from the pre-trend and positive and significant estimates of τ_k .

Because the effectiveness of resident principals may vary by program cohort, year hired, and school, concerns that TWFE models do not produce easily interpretable estimates of treatment effects certainly arise in this context.⁸ Consequently, we adopt the two-stage estimation approach developed in Gardner (2022) to identify average treatment effects. The first stage uses untreated observations to estimate school and period effects, and these effects are then removed prior to comparing outcomes for treated and untreated observations. Unfortunately, we lack the power to produce separate estimates by treatment cohort and year hired to gain a better understanding of the character of any heterogeneity.

Todd and Wolpin (2003) model achievement as a cumulative function of family, school and neighborhood factors, leading to a focus on achievement gain rather than level in the estimation of school, teacher or principal effects. We therefore replace the level of achievement with achievement gain in our preferred specifications. Differencing out the prior-year score accounts for unobserved differences in academic preparation owing to the cumulative effects of student, family, school and community factors. In this specification, differences in principal effectiveness appear as differences in learning and annual achievement growth. Importantly, the gain measure equals $(A_{it} - 0.7 * A_{i,t-1})$, thereby allowing for knowledge depreciation of 30 percent. Estimates are similar under the simple gain model equal to the difference in test scores between the current and previous year, though the simple gain model imposes the assumption of no knowledge depreciation which is not consistent with the evidence.⁹ The alternative approach of including lagged test score as a control is not in line with the Gardner (2022) approach that does not include control variables and, moreover, introduces complications related to the inclusion of a lagged endogenous variable.

Even in a gains specification, endogenous changes in student composition in response to the hiring of a new resident could introduce bias if they are related to achievement growth. Importantly, pretrends would be unaffected by such responses, and we therefore utilize an event-study framework to investigate changes in observed student characteristics following the hiring of a resident.

Similar to Grissom, Mitani, and Woo (2019), our results reflect changes in performance that occur in the first 2 to 3 years after hiring a new principal. It is possible that the benefits of a residency do not appear in the first three years in the leadership position. However, Branch et al. (2020) show evidence of substantial variation in principal value added to achievement within the first three years that is significantly related to teacher survey responses on leadership effectiveness. Moreover, roughly half of principal spells are four years or shorter, suggesting that the impact of a successful intervention would almost certainly appear within the first three years. Because we have only two years of data for principals hired in 2018, the set of schools that contribute to the third-year effect differs from the set that contributes to the second-year effect. Therefore, some estimates come from samples restricted to schools that contribute three treatment years. Note also that the main estimates ignore any subsequent transitions, meaning that the estimates capture the effect of hiring a resident on outcomes two or three years later regardless of whether the resident remains or transitions out of a school.

8 A growing number of papers address this issue including Borusyak et al. (2021), Callaway and Sant'Anna (2021), De Chaisemartin and d'Haultfoeuille (2020), Goodman-Bacon (2021) and Sun and Abraham (2021).

9 Meghir and Rivkin (2011) discuss the assumptions imposed by the achievement gain specification.

Alternative estimates come from samples limited to new principals who remain in their positions for at least three years.

6.2 Residency effects on achievement level and gain

This section reports the residency effects on principal productivity as measured by achievement level and gain. Table 4 reports treatment effects for math and reading achievement level and gain based on the Equation 7 event-study specifications and the Gardner (2022) two-stage method. Subsequent tables report estimates based on restricted samples that either exclude schools with fewer than three years of outcomes following the principal hire (Table 5) or principal spells lasting fewer than three years (Table 6). Standard errors are clustered by school. Appendix Tables a1 to a3 present TWFE estimates that do not account for heterogeneous treatment effects for the respective samples.

The pre-period coefficients in Table 4 (period=-5 to period=-2) for schools that subsequently hire a resident participant reveal modest fluctuations in the level of achievement in

Table 4 Gardner (2022) two-stage event-study estimates

		Levels		Gains	
		Math	Reading	Math	Reading
Pre-period	Resident*(t=-5)	-0.0048 (0.0155)	-0.0116 (0.0106)	0.0101 (0.0086)	0.0001 (0.0057)
	Resident*(t=-4)	0.0108 (0.0142)	0.0109 (0.0090)	0.0210* (0.0108)	0.0183** (0.0085)
	Resident*(t=-3)	-0.00004 (0.0102)	0.0061 (0.0083)	-0.0021 (0.0104)	-0.0033 (0.0074)
	Resident*(t=-2)	0.0070 (0.0131)	0.0084 (0.0091)	-0.0013 (0.0099)	0.0048 (0.0083)
	Reference period (t=-1)				
Post-period	Resident*(t=0) (dropped)				
	Resident*(t=1)	0.0451 (0.0305)	0.0374 (0.0240)	0.0516*** (0.0153)	0.0321*** (0.0122)
	Resident*(t=2)	0.0714* (0.0427)	0.0481* (0.0277)	0.0503** (0.0195)	0.0165 (0.0111)
	Sample Size	440968	440968	352705	351864

Notes: The samples include students in schools that hire a new principal (resident trained or non-resident trained) in the years 2014 to 2018. Schools have at least 5 years of data prior to the entry and at least 2 years following the entry of the new principal. Note that schools where a new principal entered in 2018 will only have 2 “post-period” years.

Estimates are based on the two-stage approach in Gardner (2022), where the first stage uses untreated observations to estimate school and period effects, and these effects are then removed prior to comparing outcomes for treated and untreated observations. Regressions include school fixed effects and year fixed effects. Regressions also control for indicators for grade. Period t=0 which is the first year of the new principal’s tenure is dropped from the sample.

Standard errors clustered by school are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

periods -5 to -3, none of which are significant at conventional levels. However, the combination of modest negative effects in period -5 and positive effects in period -4 produces significant and somewhat larger jumps in achievement gain for both math and reading in period -4. The magnitude of the coefficient in period -4 is roughly 40 percent as large as the treatment effects in math and even larger in reading. In contrast, the pre-period coefficients in periods -3 and -2 are roughly one tenth as large in math and one fourth as large in reading and insignificant. Thus, math and reading achievement and achievement gains appear to be moving in parallel in treatment and control schools in the three years leading up to the hiring of a resident principal following a modest jump in period -4. In fact, none of the math coefficients are larger than 0.007 in the levels specification or larger than 0.002 in magnitude in the gain specification in the years following period -4; the magnitudes in the reading specifications are only slightly larger.

The resident interaction-term coefficients for periods 1 and 2 showing the treatment effects in period 1 (second year) and period 2 (third year) reveal positive effects of hiring a resident principal that are larger and more significant in math. The coefficients of 0.0516 (period 1) and 0.0503 (period 2) in the preferred gains specification are highly significant and large: they exceed in magnitude estimates of a one standard deviation difference in principal effectiveness for the Chicago Public Schools (Branch et al., 2022). The reading coefficient in Column 4

Table 5 Gardner (2022) two-stage event-study estimates (school has 3 post-period years)

		Levels		Gains	
		Math	Reading	Math	Reading
Pre-period	Resident*(t=-5)	-0.0035 (0.0153)	-0.0031 (0.0105)	0.0046 (0.0091)	0.0009 (0.0062)
	Resident*(t=-4)	0.0021 (0.0129)	0.0057 (0.0089)	0.0131 (0.0091)	0.0114 (0.0073)
	Resident*(t=-3)	-0.00007 (0.0115)	0.0028 (0.0093)	0.0041 (0.0108)	-0.0006 (0.0076)
	Resident*(t=-2)	0.0072 (0.0141)	0.0053 (0.0101)	-0.0007 (0.0108)	0.0047 (0.0098)
	Reference period (t=-1)				
Post-period	Resident*(t=0) (dropped)				
	Resident*(t=1)	0.0417 (0.0363)	0.0502* (0.0286)	0.0437** (0.0183)	0.0416*** (0.0147)
	Resident*(t=2)	0.0721* (0.0438)	0.0494* (0.0291)	0.0439** (0.0197)	0.0124 (0.0116)
	Sample Size	385322	385322	308109	307332

Notes: The samples include students in schools that hire a new principal (resident trained or non-resident trained) in the years 2014 to 2017. Schools have at least 5 years of data prior to the entry and 3 years following the entry of the new principal. Note that schools where a new principal entered in 2018 will not contribute to these estimates.

Estimates are based on the two-stage approach in Gardner (2022), where the first stage uses untreated observations to estimate school and period effects, and these effects are then removed prior to comparing outcomes for treated and untreated observations. Regressions include school fixed effects and year fixed effects. Regressions also control for indicators for grade. Period t=0 which is the first year of the new principal's tenure is dropped from the sample.

Standard errors clustered by school are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

is roughly 40 percent smaller though highly significant in period 2 and half as large and not significant in period 3.

The unbalanced panels used to produce the Table 4 estimates include principals hired in 2018 who do not contribute to the estimates of period 2 effects, meaning that the differences between period 1 and 2 effects come from variation in composition and tenure. Table 5 reports estimates for the same specifications estimated over balanced panels of schools that all have three years of post-period observations. The estimates are quite similar to those for the unbalanced panel in Table 4: only slightly smaller but still significant effects on math achievement gain in both periods, a highly significant effect on reading achievement gain in period 1, and a smaller and not significant effect on reading achievement gain in period 2. The strength of the resident effects on math achievement gain supports the belief that residents significantly improve math instruction and achievement growth.

The estimates in Tables 4 and 5 come from samples that include observations following the departure of a resident or nonresident principal, meaning that they capture the effect of hiring a resident, not of the effectiveness differences between residents and nonresidents. It is informative to produce estimates for the restricted sample of schools where the newly hired principal remains at least three years, recognizing that these principals are not randomly

Table 6 Gardner (2022) two-stage event-study estimates (principal remains there 3 years)

		Levels		Gains	
		Math	Reading	Math	Reading
Pre-period	Resident*(t=-5)	0.0076 (0.0198)	0.0011 (0.0126)	0.0117 (0.0119)	0.0005 (0.0069)
	Resident*(t=-4)	0.0074 (0.0157)	0.0019 (0.0098)	0.0116 (0.0106)	0.0049 (0.0067)
	Resident*(t=-3)	-0.0085 (0.0139)	0.0038 (0.0119)	-0.0011 (0.0129)	0.0053 (0.0096)
	Resident*(t=-2)	0.0071 (0.0183)	0.0031 (0.0129)	-0.0031 (0.0139)	0.0025 (0.0127)
	Reference period (t=-1)				
Post-period	Resident*(t=0) (dropped)				
	Resident*(t=1)	0.0466 (0.0520)	0.0616 (0.0413)	0.0519* (0.0269)	0.0541*** (0.0210)
	Resident*(t=2)	0.1007 (0.0660)	0.0787* (0.0467)	0.0654** (0.0276)	0.0247 (0.0165)
	Sample Size	229918	229918	183850	183468

Notes: The samples include students in schools that hire a new principal (resident trained or non-resident trained) in the years 2014 to 2016. To be included in the sample, principals must remain in their school at least 3 years. Thus, principals from entry cohorts in 2017 and 2018 will not contribute to these estimates.

Estimates are based on the two-stage approach in Gardner (2022), where the first stage uses untreated observations to estimate school and period effects, and these effects are then removed prior to comparing outcomes for treated and untreated observations. Regressions include school fixed effects and year fixed effects. Regressions also control for indicators for grade. Period t=0 which is the first year of the new principal's tenure is dropped from the sample,

Standard errors clustered by school are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

drawn from all new hires. The slightly larger estimates, particularly in period 2, are consistent with the notion that turnover imposes transaction costs that adversely affect learning and achievement, though the differences may also reflect the positive selection of resident principals who remain in their positions for at least three years.

A comparison between the estimates in Table 4 and Appendix Table a1 reveals that the failure to account for heterogeneous treatment effects leads to much larger coefficients in the gains specifications in both the pre and treatment periods; the same pattern holds for the more restrictive samples used to produce the estimates shown in Tables 5 and 6. These differences highlight the importance of accounting for heterogeneous treatment effects in this context.

A final sensitivity test examines the importance of accounting for any achievement dip in the year immediately prior to the hiring of a nonresident principal. Appendix Table a4 shows that estimates are slightly smaller but quite similar in magnitude and significance to those in Table 4 that do not include any controls.

6.3 Changes in student composition

Changes in student composition following the hiring of a resident could potentially introduce bias, and Table 7 reproduces the event-study analysis substituting student characteristics in

Table 7 Gardner (2022) two-stage event-study estimates (school covariates)

	Enrollment	Black	Hispanic	Other	Female	Special Education
Pre-period	Resident*(t=-5)	-0.2652 (0.7526)	0.0093*** (0.0031)	-0.0054* (0.0033)	0.0015 (0.0010)	0.0022 0.0022 0.0019
	Resident*(t=-4)	-0.0998 (0.4900)	0.0039** (0.0019)	-0.0034 (0.0023)	0.0017 (0.0011)	-0.0004 0.0016 0.0017
	Resident*(t=-3)	-0.1110 (0.3483)	-0.0006 (0.0011)	-0.0015 (0.0011)	0.0006 (0.0007)	0.0015 0.0017 0.0013
	Resident*(t=-2)	0.2523 (0.4461)	-0.0038* (0.0022)	0.0029 (0.0023)	-0.0009 (0.0006)	0.0012 0.0018 0.0015
	Reference period (t=-1)					
Post-period	Resident*(t=0) (dropped)					
	Resident*(t=1)	-1.1356 (1.2912)	-0.0155** (0.0058)	0.0131** (0.0066)	-0.0039 (0.0029)	0.0019 0.0045 0.0041
	Resident*(t=2)	-2.1136 (1.1582)	-0.0152* (0.0081)	0.0185** (0.0091)	-0.0051 (0.0037)	0.0027 0.0055 0.0048
	Sample Size	440968	440968	440968	440968	440968

Notes: Similar to Table 4, the sample here include students in schools that hire a new principal (resident trained or non-resident trained) in the years 2014 to 2018. Schools have at least 5 years of data prior to the entry and at least 2 years following the entry of the new principal. Note that schools where a new principal entered in 2018 will only have 2 “post-period” years.

Estimates are based on the two-stage approach in Gardner (2022), where the first stage uses untreated observations to estimate school and period effects, and these effects are then removed prior to comparing outcomes for treated and untreated observations. Regressions include school fixed effects and year fixed effects. Regressions also control for indicators for grade. Period t=0 which is the first year of the new principal’s tenure is dropped from the sample,

Standard errors clustered by school are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

place of achievement as outcome variables. The estimates suggest a decline in share Black and roughly corresponding increase in share Hispanic of approximately 1.5 percentage points following the hiring of a resident principal. Estimates of the relationship between achievement and achievement gain on the one hand and student characteristics on the other from TWFE models that include covariates reported in Appendix Table a5 indicate that such a compositional change is likely to increase achievement. However, the coefficients for Black and Hispanic students in the math achievement gain specifications of -0.18 and -0.11, respectively, suggest the compositional change effect is quite small: multiplying the difference in coefficients by 0.015 indicates that the change in race-ethnic composition increases achievement gain by 0.001 standard deviations, which is less than 2 percent of the resident effect. Even in the presence of a positive change in orthogonal unobserved student characteristics that has an effect that is twice as large as that introduced by the change in racial composition, compositional changes would account for less than 10 percent of the resident effect for math and an even smaller amount for reading given the smaller difference between coefficients for Black and Hispanic students reported in Appendix Table a5.

7 Conclusion

The confluence of a growing emphasis on school leadership and long-held concerns about the training of school administrators has led to a search for better approaches to training and identifying effective school principals. The Chicago Public Schools have devoted substantial resources to residency programs designed to raise the effectiveness of school leaders. Although appealing, there are potential pitfalls that could dampen the return on this investment in addition to any shortfalls in program efficacy. These include the loss of trained principals to other districts that value their augmented leadership skills and a delay in obtaining a CPS principal position that may depreciate the skills produced by the residency. Consequently, the district return on the investment depends on both the program impact on principal effectiveness and principal career paths. Given the scarcity of resources in CPS, district investments in principal training that raise productivity but end up benefiting primarily the residents through higher salaries or other districts to which residents may transition would be unlikely to have positive returns; such resources could be used to reduce class size, provide additional academic support, give bonuses to high-performing educators or in myriad other uses.

Evidence suggests that that residency programs increase principal effectiveness at raising reading and especially math achievement and that only 20 percent of principals leave CPS within four years of program completion. The limited loss of resident-trained principals in combination with positive effects on productivity suggest that market imperfections or the commitment to remain in the district enables CPS to capture some of the benefits of the training program. The long time-lag between program completion and ascension to a principal position likely dampens the return on the investment given the absence of positive effects on achievement in schools with resident-trained assistant principals and potential depreciation of leadership skills produced by the program. Comparisons between hiring processes run by LSCs and those run by the district suggest that the decentralization of governance contributes to the slow pace at which residents become principals.

An assessment of the district return to investment in a residency program requires a comprehensive assessment of program costs and benefits. On the benefit side, the focus on achievement for reasons of data availability may fail to capture resident effects on the development of noncognitive skills. Nevertheless, a comparison between residency effects and the effects of an equally costly reduction in class size on achievement provides information on the value of this investment in terms of raising cognitive skills relative to a widely used alternative use of resources. Program costs per resident principal are determined by the sum of the stipends paid to all residents plus district support for the principal training programs, dividing by the number of residents who become principals. As of 2017 slightly more than one fourth of residents had become principals, though that fraction would be expected to increase given the experiences of the earlier cohorts. Dividing the 16.2 million dollars among the 70 residents who become principals and dividing the \$80,000 average stipend by the share that become principals ($\$80,000/0.264$) sums to a total cost of roughly \$535,000 per resident.

The magnitude of the class-size reduction generated by spending an additional \$535,000 depends crucially over how many years the money is spread. Because residency costs do not vary with the number of years a resident serves as principal, the length of service determines the appropriate number of years over which to divide the money. Evidence to date suggests that three years is likely to be an underestimate of the average spell length of residents who become principals in CPS. This would leave approximately \$175,000 per year for the costs of salaries and benefits for additional teachers, allowing schools to add roughly two teachers. An additional two teachers would constitute around a 10% average increase in the number of classroom teachers, leading to a decline of around 3 students per class. Whether the benefit of the residency program exceeds the benefit of a 3 student per class reduction in class size depends crucially on the benefits of the smaller classes. If the effects of a 10-student reduction exceed 0.18 standard deviations as in Krueger (1999), the answer is more likely to be no. However, if the benefits are smaller as argued in Hanushek (2002), the answer is more likely to be yes.

Of course, these calculations assume that only one fourth of residents become principals, that the average length of service is only three years, and that the district realizes no benefits from residents who work as assistant principals or in other capacities. Based on the career paths of residents in the early cohorts it is likely that the share who become principals will actually approach 50 percent. Over time the average spell lengths would also be expected to increase. Therefore, it appears that this program that augments general skills is likely to generate a positive return to the district.

Bibliography

- Acemoglu, Daron, and Jorn-Steffen Pischke. 1999a. "Beyond Becker: Training in Imperfect Labour Markets." *The Economic Journal* 109 (453): 112–42. <https://doi.org/10.1111/1468-0297.00405>.
- Acemoglu, Daron, and Jörn-Steffen Pischke. 1999b. "The Structure of Wages and Investment in General Training." *Journal of Political Economy* 107 (3): 539–72. <https://doi.org/10.1086/250071>.
- Autor, David, Nicole Maestas, Kathleen J. Mullen, and Alexander Strand. 2015. "Does delay cause decay? The effect of administrative decision time on the labor force participation and earnings of disability applicants." *NBER Working Paper No. 20840*. National Bureau of Economic Research.
- Becker, Gary S. 1962. "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy* 70 (5): 9–49.
- Béteille Tara, Demetra Kalogrides & Susanna Loeb, 2012. "Stepping Stones: Principal Career Paths and School Outcomes," July 2012. *Social Science Research* 41(4):904-19

- Borusyak, K., Jaravel, X., & Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. arXiv preprint arXiv:2108.12419.
- Branch, Gregory F., Eric Hanushek, Steven Rivkin, and Jeffrey C. Schiman. 2020. "How Much Does Leadership Matter? Evidence from Public Schools." *Unpublished Manuscript*.
- Braun, Donna, Felice D. Billups, and Robert K. Gable. 2013. "Transforming equity-oriented leaders: Principal Residency Network program evaluation." *K-12 Education*. Center for Research and Evaluation.
- Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225 (2), 200-230.
- Chiang, Hanley, Stephen Lipscomb, and Brian Gill. 2016. "Is School Value Added Indicative of Principal Quality?" *Education Finance and Policy* 11 (3): 283-309. https://doi.org/10.1162/EDFP_a_00184.
- Chicago Public School. 2020. *Chicago Leadership Collaborative*. Retrieved at <https://cps.edu/PrincipalQuality/Pages/Pipeline.aspx> (accessed February 20, 2020).
- Corcoran, Sean P., Amy Ellen Schwartz, and Meryle Weinstein. 2012. "Training Your Own: The Impact of New York City's Aspiring Principals Program on Student Achievement." *Educational Evaluation and Policy Analysis* 34 (2): 232-53. <https://doi.org/10.3102/0162373712437206>.
- De Chaisemartin, C., & d'Haultfoeulle, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9), 2964-96.
- Dee, Thomas S. and Brian Jacob. 2011. "The impact of No Child Left Behind on student achievement." *Journal of Policy Analysis and Management* 30 (3): 418-446.
- Dinerstein, Michael, Rigissa Megalokonomou, and Constantine Yannelis. 2019. "Human Capital Depreciation." *Unpublished Manuscript*.
- Gardner, John. 2022. "Two-stage differences in differences," Papers 2207.05943, arXiv.org.
- Gates, Susan M., Laura S. Hamilton, Paco Martorell, Susan Burkhauser, Paul Heaton, Ashley Pierson, Matthew Baird, Mirka Vuollo, Jennifer J. Li, Diana C. Lavery, Melody Harvey, Kun Gu. *Preparing Principals to Raise Students Achievement: Implementation and Effects of the New Leaders Program in Ten Districts*. Santa Monica, CA: Rand Corporation.
- Goldring, Ellen and Pearl Sims. 2005. "Modeling creative and courageous school leadership through district-community-university partnerships." *Educational Policy* 19 (1): 223-249.
- Goldring, R., & Taie, S. (2018). Principal attrition and mobility: results from the 2016-17 principal follow-up survey first look (NCES 2018-066). Washington DC: U.S. Department of Education National Center for Education Statistics.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254-277.
- 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. <https://doi.org/10.3102/0162373714523831>.
- Grissom, J. A., Mitani, H., & Woo, D. S. (2019). Principal Preparation Programs and Principal Outcomes. *Educational Administration Quarterly*, 55(1), 73-115. <https://doi.org/10.1177/0013161X18785865>
- Hale, Elizabeth L, and Hunter N. Moorman. 2003. *Preparing School Principals: A National Perspective on Policy and Program Innovations*. Institute for Educational Leadership, Washington, DC and Illinois Education Research Council, Edwardsville, IL.
- Hanushek, Eric. 2002. "Evidence, Politics, and the Class Size Debate." (Mishel and Rothstein eds). *The Class Size Debate*, Economic Policy Institute, Washington, DC: 37-65.
- Hanushek, Eric and Margaret E. Raymond. 2005. "Does school accountability lead to improved student performance?" *Journal of Policy Analysis and Management* 24 (2): 297-327.
- Herrington, Carolyn D. and Barbara K. Wills. 2005. "Decertifying the principalship: The politics and administrator preparation in Florida." *Educational Policy* 19 (1): 181-200.
- Hess, Frederick M. and Andrew P. Kelly. 2005. "An innovative look, a recalcitrant reality: The politics of principal preparation reform." *Educational Policy* 19 (1): 155-180.
- Jackson, Kirabo, Shanette C. Porter, John Q. Easton, Alyssa Blanchard, and Sebastian Kiguel. 2020. "School effects on socioemotional development, school-based arrests, and educational attainment." *American Economic Review: Insights* 2 (4): 491-508.
- Knechtel, Virginia, Mary Anne Anderson, Alyson Burnett, Thomas Coen, Margaret Sullivan, Christina Clark Tuttle, and Philip Gleason. 2015. *Understanding the Effect of KIPP as It Scales: Volume II, Leadership Practices at KIPP*. Mathematica Policy Research. Retrieved at <https://files.eric.ed.gov/fulltext/ED581456.pdf> (accessed February 20, 2020).
- Krueger, Alan. 1999. "Experimental Estimates of Education Production Functions." *The Quarterly Journal of Economics* Vol. 114, No. 2: 497-532.

- Ladd, Helen F. 1999. "The Dallas school accountability and incentive program: an evaluation of its impact on student outcomes." *Economics of Education Review* 18 (1): 1-16.
- Miller, Ashley. 2013. "Principal turnover and student achievement." *Economics of Education Review* 36: 60-72.
- Meghir, Costas and Steven Rivkin. 2011. "Econometric Methods for Research in Education." in (Eric Hanushek, Stephen Machin and Ludger Woessman, eds.), *Handbook of the Economics of Education*, vol 3, Elsevier.
- Neal, Derek and Whitmore Schanzenbach. 2010. "Left behind by design: Proficiency counts and test-based accountability." *Review of Economics and Statistics* 92 (2): 263-283.
- Ryan, Susan, Anthony S. Bryk, Gudelia Lopez, Kimberly P. Williams, Kathleen Hall, and Stuart Luppescu. 1997. "Charting Reform: LSCs—Local Leadership at Work." Consortium on Chicago School Research.
- Schiman, Jeffrey C., Derek Laing, Steven G. Rivkin, and Jason Ward. 2018. "Decentralized Governance and the Quality of School Leadership." Unpublished Manuscript.
- Schiman, Jeffrey C., 2021. "The Information Content of Employee Awards," *Economics of Education Review*, Elsevier, vol. 84.
- Sun, L., & Abraham, S.** (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175-199.

Appendix Table A1 TWFE event-study estimates

		Levels		Gains	
		Math	Reading	Math	Reading
Pre-period	Resident*(t=-5)	0.0087 (0.0431)	0.0034 (0.0274)	0.0477* (0.0270)	0.0285 (0.0175)
	Resident*(t=-4)	0.0258 (0.0404)	0.0286 (0.0246)	0.0576** (0.0284)	0.0459** (0.0206)
	Resident*(t=-3)	0.0154 (0.0213)	0.0244 (0.0217)	0.0322 (0.0266)	0.0230 (0.0171)
	Resident*(t=-2)	0.0212 (0.0267)	0.0236 (0.0169)	0.0298 (0.0199)	0.0270 (0.0183)
	Reference period (t=-1)				
	Resident*(t=0) (dropped)				
Post-period	Resident*(t=1)	0.0541** (0.0267)	0.0501** (0.0201)	0.0779*** (0.0215)	0.0528*** (0.0172)
	Resident*(t=2)	0.0782* (0.0411)	0.0584** (0.0241)	0.0759*** (0.0255)	0.0357*** (0.0154)
	Sample Size	440968	440968	352705	351864

Notes: The samples include students in schools that hire a new principal (resident trained or non-resident trained) in the years 2014 to 2018. Schools have at least 5 years of data prior to the entry and at least 2 years following the entry of the new principal. Note that schools where a new principal entered in 2018 will only have 2 “post-period” years.

Estimates are based on a two-way fixed effect regression with school fixed effect, year fixed effects, and grade indicators. Period t=0 which is the first year of the new principal’s tenure is dropped from the sample.

Standard errors clustered by school are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table A2 TWFE event-study estimates (school has 3 post-period years)

		Levels		Gains	
		Math	Reading	Math	Reading
Pre-period	Resident*(t=-5)	-0.0047 (0.0434)	0.0043 (0.0303)	0.0297 (0.0268)	0.0190 (0.0202)
	Resident*(t=-4)	0.0028 (0.0399)	0.0151 (0.0269)	0.0386 (0.0262)	0.0304 (0.0201)
	Resident*(t=-3)	0.0019 (0.0318)	0.0129 (0.0245)	0.0282 (0.0249)	0.0176 (0.0200)
	Resident*(t=-2)	0.0090 (0.0241)	0.0148 (0.0186)	0.0207 (0.0242)	0.0216 (0.0231)
	Reference period (t=-1)				
Post-period	Resident*(t=0) (dropped)				
	Resident*(t=1)	0.0347 (0.0304)	0.0490** (0.0239)	0.0606** (0.0236)	0.0528** (0.0212)
	Resident*(t=2)	0.0667* (0.0392)	0.0540** (0.0243)	0.0630** (0.0248)	0.0284* (0.0168)
	Sample Size	385322	385322	308109	307332

Notes: The samples include students in schools that hire a new principal (resident trained or non-resident trained) in the years 2014 to 2017. Schools have at least 5 years of data prior to the entry and 3 years following the entry of the new principal. Note that schools where a new principal entered in 2018 will not contribute to these estimates.

Estimates are based on a two-way fixed effect regression with school fixed effect, year fixed effects, and grade indicators. Period t=0 which is the first year of the new principal's tenure is dropped from the sample. Standard errors clustered by school are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table A3 TWFE event-study estimates (principal remains there 3 years)

		Levels		Gains	
		Math	Reading	Math	Reading
Pre-period	Resident*(t=-5)	0.0178 (0.0602)	0.0055 (0.0395)	0.0436 (0.0396)	0.0113 (0.0277)
	Resident*(t=-4)	0.0181 (0.0549)	0.0085 (0.0344)	0.0436 (0.0391)	0.0176 (0.0264)
	Resident*(t=-3)	0.0021 (0.0446)	0.0114 (0.0341)	0.0291 (0.0365)	0.0185 (0.0303)
	Resident*(t=-2)	0.0163 (0.0372)	0.0099 (0.0283)	0.0286 (0.0368)	0.0140 (0.0319)
	Reference period (t=-1)				
	Resident*(t=0) (dropped)				
Post-period	Resident*(t=1)	0.0526 (0.0426)	0.0622* (0.0341)	0.0767** (0.0350)	0.0614** (0.0292)
	Resident*(t=2)	0.1004* (0.0561)	0.0777* (0.0394)	0.0861** (0.0331)	0.0331 (0.0246)
	Sample Size	229918	229918	183850	183468

Notes: The samples include students in schools that hire a new principal (resident trained or non-resident trained) in the years 2014 to 2016. To be included in the sample, principals must remain in their school at least 3 years. Thus, principals from entry cohorts in 2017 and 2018 will not contribute to these estimates.

Estimates are based on a two-way fixed effect regression with school fixed effect, year fixed effects, and grade indicators. Period t=0 which is the first year of the new principal's tenure is dropped from the sample.

Standard errors clustered by school are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table A4 Gardner (2022) two-stage event-study estimates (controlling for period before)

		Levels		Gains	
		Math	Reading	Math	Reading
Pre-period	Resident*(t=-5)	-0.0052 (0.0154)	-0.0121 (0.0105)	0.0090 (0.0086)	0.0005 (0.0057)
	Resident*(t=-4)	0.0098 (0.0135)	0.0094 (0.0087)	0.0179* (0.0105)	0.0159* (0.0073)
	Resident*(t=-3)	-0.0017 (0.0097)	0.0034 (0.0097)	-0.0075 (0.0104)	-0.0074 (0.0073)
	Resident*(t=-2)	0.0045 (0.0147)	0.0043 (0.0097)	-0.0095 (0.0101)	-0.0014 (0.0078)
	Reference period (t=-1)				
Post-period	Resident*(t=0) (dropped)				
	Resident*(t=1)	0.0435 (0.0321)	0.0333 (0.0252)	0.0433*** (0.0157)	0.0258** (0.0126)
	Resident*(t=2)	0.0687 (0.0436)	0.0438 (0.0287)	0.0418** (0.0197)	0.0097 (0.0114)
	Sample Size	440968	440968	352705	351864

Notes: Similar to Table 4, the sample here include students in schools that hire a new principal (resident trained or non-resident trained) in the years 2014 to 2018. Schools have at least 5 years of data prior to the entry and at least 2 years following the entry of the new principal. Note that schools where a new principal entered in 2018 will only have 2 “post-period” years.

Estimates are based on the two-stage approach in Gardner (2022), where the first stage uses untreated observations to estimate school and period effects, and these effects are then removed prior to comparing outcomes for treated and untreated observations. Regressions include school fixed effects and year fixed effects. Regressions also control for indicators for grade and an indicator for the period before the transition (t=-1). Period t=0 which is the first year of the new principal’s tenure is dropped from the sample,

Standard errors clustered by school are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

Appendix Table A5 TWFE control coefficients

	Mathematics		Reading	
	Levels	Gains	Levels	Gains
Student Black	-0.5523*** 0.0346	-0.1834*** 0.0112	-0.3968*** 0.0368	-0.1482*** 0.0111
Student Hispanic	0.3420*** 0.0261	-0.1112*** 0.0083	-0.327*** 0.0297	-0.1047*** 0.0094
Student Other Race	0.0989* 0.0569	0.0532** 0.0210	-0.0816*** 0.0279	-0.0061 0.0113
Student Female	-0.0434*** 0.0051	-0.0061*** 0.0019	0.1095*** 0.0046	0.0312*** 0.0019
Student Special Education	-0.0975 0.0175	-0.2951*** 0.0088	-1.2028*** 0.0145	-0.3671*** 0.0076
Sample Size	440968	352705	440968	351864

Notes: Similar to Table 4, the sample here include students in schools that hire a new principal (resident trained or non-resident trained) in the years 2014 to 2018. Schools have at least 5 years of data prior to the entry and at least 2 years following the entry of the new principal. Note that schools where a new principal entered in 2018 will only have 2 “post-period” years.

The control coefficient estimates come from regressions that include school fixed effects and year fixed effects. Regressions also control for indicators for grade. Period $t=0$ which is the first year of the new principal’s tenure is dropped from the sample,

Standard errors clustered by school are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.