

The Benefits of Early and Unconstrained Hiring: Evidence from Teacher Labor Markets

Matthew A. Kraft
Brown University

John P. Papay
Brown University

Leigh Wedenoja
Rockefeller Institute

Nathan Jones
Boston University

April 2021

Abstract

Public schools face substantial constraints on when and whom they can hire. We study Boston Public Schools' (BPS) efforts to reduce these frictions by eliminating seniority privileges and forced placements as well as accelerating the open hiring timeline. Using a difference-in-difference framework where some BPS schools enjoyed hiring autonomies prior to the reforms, we find that mutual consent hiring moved up the median hire date by more than two months and increased the experience, diversity, effectiveness, and retention of both internal and external new hires. These findings illustrate the large returns to conducting early, open searches in the public teacher labor market.

JEL Codes: No. I2, J22, J24

Correspondence can be sent to Matthew Kraft; mkraft@brown.edu; P.O. Box, 1938, Brown University, Providence RI, 02912. We are grateful for the support of Ross Wilson, Emily Qazilbash, Jonathan Barrows and the Boston Public Schools as well as feedback from David Figlio, Kevin Lang, Dick Murnane, Marcus Winters and seminar participants at Boston University and the IZA workshop on the Economics of Education.

I. Introduction

Economic models of employee search and job matching predict that firms have the potential to realize substantial productivity gains by improving their hiring practices (Lazear and Oyer, 2009). The potential returns to improved search and screening processes are thought to be particularly large in the public sector where labor markets are often characterized by high degrees of friction and limited choice. However, there exists little empirical evidence on the actual return to improved hiring practices despite their perceived importance (Oyer and Schaefer, 2011).

In this paper, we examine efforts to reduce hiring frictions and improve the returns to search in the public education sector. Public school teachers are among the largest skilled occupation in the United States. With over 3.5 million K-12 public school teachers, policymakers and school administrators must expend considerable resources trying to staff schools effectively. These are consequential decisions for students, schools, and the economy. Teachers have large effects on students' short and long-run outcomes (Chetty, Friedman, & Rockoff, 2014). Tenure laws provide experienced teachers with substantial job protections in most states. Spending on teacher salaries and benefits in K12 public schools alone approaches \$345 billion annually. The teacher labor market also provides an advantageous context in which to study hiring given that productivity is heterogeneous across individuals (Hanushek and Rivkin, 2010) and worker-firm pairings (Jackson, 2013).

Hiring in many labor markets is characterized as a two-way matching process. Firms select candidates whom best meet their needs and workers select firms where they are most attracted to work. However, the teacher labor market often functions in ways that can hinder effective matching, especially in large urban districts. Collective bargaining agreements and

district policies often delay the hiring process and limit the flexibility of principals to hire the teachers they believe are the strongest, best-matched candidates. In fact, in many urban districts, one quarter to one third of all new teachers are hired after the school year starts, which imposes considerable costs on student achievement (Papay & Kraft, 2016). This constrained choice and delayed timing cause many new teachers to apply widely and accept the first offer they receive rather than risk waiting for their preferred match.

We study Boston Public Schools' (BPS) "mutual consent" hiring reforms to provide schools with the autonomy to conduct early, open, and intensive searches for vacant teaching positions and to hire their preferred candidate. For decades, BPS' hiring process was characterized by the "late, rushed and information poor" practices found in most urban school districts (Liu & Johnson 2006; Levin & Quinn, 2003). Teachers whose positions had been eliminated ("excessed teachers") could be administratively placed in other schools without principal approval, more senior teachers had substantial advantages in the internal transfer process, and schools could not hire external candidates until quite late. In the years prior to the reforms, 26% of new hires in traditional BPS schools were hired late and 44% left their schools after their first year, both signs of low match quality. Consistent with prior evidence (Papay & Kraft, 2016), we show that the practice hiring teachers late in the summer and after the beginning of school year imposes real costs on students by lowering their achievement.

Starting in 2014, BPS extended hiring autonomies to traditional district schools, a condition previously reserved for district turnaround schools and within-district charter schools. The district also eliminated administrative placements of excessed teachers and allowed schools to "open post" vacant positions for external candidates much earlier in the hiring cycle. We employ a difference-in-differences (DD) framework to examine the differential impact of these

reforms on traditional BPS schools compared to almost a third of the district schools that already operated with hiring autonomies. We examine how the reforms impacted the hiring time, characteristics, and retention of new hires that are new to the district or that taught previously in another BPS school. We then estimate the school-wide effects of the reforms on achievement as well as the direct effects on those students taught by new hires.

We find that the BPS hiring reforms indeed accelerated hiring timelines district-wide by over two months and reduced the prevalence of late hiring in traditional BPS schools by 50%. The reforms also had substantial positive impacts on teachers and students, increasing the diversity, experience, and retention of new hires. We estimate that the reforms substantially increased the effectiveness of new hired teachers, with effects on student achievement ranging between 0.14σ to 0.30σ for both external and internal new hires. Overall, we find the reforms raise student achievement in traditional BPS schools by 0.09σ and 0.07σ in math and English Language Arts (ELA). We conduct a range test for heterogeneous effects and find little evidence that the reforms benefitted specific schools or increased widen gaps in hiring outcomes across BPS schools.

We make several important contributions with this research. Our study informs the broader personnel and public economics literatures on hiring by providing a direct empirical test of the value of unconstrained, competitive search in the labor market (Oyer and Schaefer, 2011). We also contribute to the teacher effectiveness and diversity literatures. Prior studies have largely focused on efforts to improve hiring via the selection process and financial incentives.¹

¹ Research on the selection process includes Jacob, Rockoff, Taylor, Lindy, and Rosen (2018); Goldhaber, Grout, and Huntington-Klein (2017); Bruno and Strunk (2018). Research on financial incentives includes Glazerman et al. (2013); Steele, Murnane, and Willett (2010); Clotfelter, Glennie, Ladd, and Vigdor (2008); Feng and Sass (2018); Bueno and Sass (2019).

Recent research illustrating the benefits increasing teacher diversity points to the value of efforts to attract more teachers of color. We provide the first evidence on the impacts of mutual consent hiring reforms that have been adopted in several of the large districts including New York City, Chicago, Denver, and Baltimore and remain highly contested (Sawchuk, 2010). Our findings also extend the literature on employer-employee match quality with evidence of increased employee retention and improved effectiveness when both firms and workers are afforded more time, autonomy, and choice during the search process (Jackson, 2013). Finally, our analyses of the hiring reforms speak to the broader literature on the importance of managerial and organizational practices for both school and firm productivity (Bender et al., 2018; Fryer, 2014; Bloom et al., 2015; Angrist, Pathak, and Walters, 2013; Papay et al., 2020).

II. Teacher Hiring in U.S. Public Schools

The teacher labor market is characterized by a competitive search process with bilateral asymmetric information among both teachers and schools (Lazear and Oyer, 2009). Schools are attempting to maximize the general ability of new hires as well as the match quality between new hires and the school's organizational practices, the racial/ethnic composition of its students, and the specific position. Candidates for teaching positions are looking to maximize the utility they derive from a position and school context, conditional on their general ability. A successful hiring process is one that produces a teacher-school match that is optimally productive for both sides – the teacher finds a school where she wants to teach and can have success, while the school finds a teacher who can promote the organization's goals. However, there are a number of unique barriers to the optimal functioning of the job matching process in public schools.

Constrained Choice

Principals in most public schools face considerable restrictions on their ability to conduct an early and open search for job candidates. State tenure laws, which provide teachers with substantial job protections, create strong incentives for districts to conduct an “involuntary transfer” process. Typically, tenured teachers whose positions have been eliminated – due to school restructuring or closure, budget cuts, or reductions in student enrollment – bid on or are placed into open positions for which they are licensed (Boyd et al., 2011; Grissom, Loeb, & Nakashima, 2014; Levin, Mulhern, & Schunck, 2005). The imperative to provide job placements for excessed tenured teachers even leads some districts to “bump” non-tenured teachers from their position, regardless of their performance, and sometimes without the input of principals or school-based hiring committees.

District policy and collective bargaining agreements (CBAs) often constrain when different teacher candidates can be considered. In many districts, teachers who wish to transfer schools must be considered first during a “voluntary transfer” process before schools can hire external candidates (Levin, Mulhern, & Schunck, 2005). In some cases, principals have no input about which teachers transfer to their school, while other policies require principals to choose among qualified transfer applicants or interview and at least consider these candidates.

While teachers unions are often solely blamed for such policies, many districts in states that have repealed tenure laws or explicitly prohibit collective bargaining still operate with similar hiring constraints. In states where collective bargaining is illegal such as Georgia, Arizona, and Texas, many districts willingly honor informal agreements with local teacher associations and non-union professional organizations which secure important advantages for internal transfers. In other districts, central offices still maintain close control over the hiring

process. Schools often have limited autonomy and can still be subject to forced placements and constrained choice even in a state like North Carolina where both tenure and unions are illegal.

Late Hiring

The voluntary and involuntary transfer processes described above contribute to the widespread phenomenon of late teacher hiring. Complicated student enrollment projections, policies that allow teachers to announce their departure after the school year ends, and district budget processes that often rely on external political actors further delay job postings and hiring (Levin, Mulhern, & Schunck, 2005; Levin & Quinn, 2003). Recent estimates suggest that anywhere from 11 to 30 percent of newly hired teachers are hired after the start of the school year (Engel, 2012; Jones, Maier, & Grogan, 2011; Liu & Johnson, 2006; Papay & Kraft, 2016). These studies also show that late hiring is concentrated most in schools that serve large populations of low-income students and students of color.

Ultimately, late hiring is detrimental to student learning. When teachers are hired after the school year begins, they have no time to plan curriculum or prepare instructional materials. They also start without the opportunity to establish a classroom culture that is supportive to learning and where all students feel like they belong. In the absence of a permanent teacher, students are taught by substitutes, temporarily absorbed into other classes, or simply left under the supervision of a rotating cast of administrators and other school staff. Papay and Kraft (2016) find that late hiring directly reduced student achievement by between 0.03σ and 0.05σ . They also find teachers hired late leave their schools at higher rates than their peers hired on time, likely a result of lower-quality matches between teachers and schools.

III. Hiring Reforms in Boston Public Schools

Pre-Reform

Prior to its Human Capital Initiative in 2014, hiring in most BPS schools was a highly-structured, multi-stage process that prioritized teacher seniority and was heavily controlled by the central office. In January, the district would conduct a district-wide staffing needs assessment based on programmatic changes, student enrollment trends, school closures, budget cuts and teacher licensing requirements (Phase 1). Typically, the district would notify around 400 of the district's 4,700 teachers that they would not have a position in their current school for the following year and would place them in the "excess pool."

In mid-April, the district would begin the voluntary and involuntary teacher transfer process (Phase 2). School principals posted vacant positions to an internal job board and tenured BPS teachers had ten days to apply. School leaders were required to interview all internal candidates who applied and had to hire one of them as long as there were at least two applicants. At the end of this process there often remained a significant number of tenured teachers in the excess pool either because they never applied to a position or were not among the internal transfer candidates who were selected. These teachers then bid on remaining jobs in their certification area in order of seniority (Phase 3). Once a position had three bids the central office would unilaterally assign the most senior teacher to the position. Any remaining teachers were administratively placed in open positions within their certification area. In some cases, the district "bumped" probationary teachers from their positions to place unattached tenured teachers in jobs. External hiring (Phase 4) typically began in June, pushing the hiring process well into the beginning of the school year.

In addition to hiring delays, the constrained hiring process in BPS also created a number of other unintended consequences. Many principals deliberately attempted to hide vacancies or

avoid having an excess pool teacher placed at their school. This was because principals often used the excess process as a quick and less contentious avenue for removing low-performing teachers from their school instead of the lengthy and an uncertain evaluation process. As one BPS principal described, “We all know of cases when a colleague closed out a position instead of evaluating out an ineffective teacher,” (Johnson & Suesse, 2005, p.10). In 2009, 52% of new positions were posted late in the summer after the voluntary and involuntary transfer process had ended (National Council for Teacher Quality, 2010). This practice left principals “scrambling in August to fill vacancies that we’d hidden to protect ourselves from a disruptive placement” (Boston Municipal Research Bureau, p. 6). Others would try keep excessed teachers from bidding on open position: “I did everything possible to discourage people from selecting my school. I told them that we had a shooting the previous year near the school yard, that we don’t have parking—anything to make the school seem as unappealing as possible” (National Council for Teacher Quality, 2010, p.10). Finally, some principals simply chose to stick with low-performing teachers in their building rather than risk having a vacancy filled from the excess pool. A principal explained, “The one you know is better than the one you don’t” and that he was hesitant “to see how low it can go” (Kraft & Gilmour, p.242).

The Growth of Autonomous Schools

Legislative changes over the past two decades enabled BPS to experiment with new school models that afforded greater autonomy over hiring and other school operations. In-district charter schools (called Pilot schools locally) first began in 1998 with autonomy over budgets, hiring and staffing decisions, school calendars, curriculum, professional development, and compensation approaches. In 2010, the Massachusetts state legislature authorized two new types of semi-autonomous public schools – Turnaround and Innovation schools. By 2014, almost one

out of every three BPS schools was operating with hiring autonomy that allowed them to bypass the internal transfer and administrative placement process and open post positions for external candidates early in the hiring process. These flexibilities created some tension in the district because autonomous schools could effectively remove tenured teachers from their schools by eliminating positions, but they were not required to hire excessed teachers or receive teachers who were administratively placed. The turnaround process placed additional pressure on the existing system by requiring some schools to dismiss more than half of their teachers, many of whom were tenured teachers who then had to be reassigned to traditional district schools.

Hiring Reforms

In early 2014, BPS rolled out a suite of hiring reforms built around two principles: accelerating the hiring timeline and ensuring that principals had the flexibility to hire any teacher they believed was the best matched candidate (mutual consent). The district aimed to “to attract and hire the most diverse, qualified and effective teachers as early as possible” by ending forced placements and allowing schools to open post positions at the beginning of the hiring cycle (Boston Public Schools, 2015, p.1). Importantly, the district worked within the structure of the existing collective bargaining agreement to achieve these goals. These reforms extended the flexibility and local control over hiring to all schools that had only previously been afforded to autonomous schools in the district.

The district unilaterally assumed the risk of paying the salaries of all teachers in the excess pool who did not obtain a job through the open search process. At the same time, the central office took a variety of measures to reduce the number of tenure teachers in the excess pool by providing job placement supports, subsidizing training for new certifications, offering early retirement incentives, and actively dismissing teachers with low evaluation ratings.

Teachers who failed to obtain a job offer were placed in a position of “suitable professional capacity,” (SPC) typically serving in supplemental instructional roles such as teachers’ aides, academic interventionists, and long-term substitutes. The intent was to place these teachers in roles in which they could grow professionally in hopes of improving their job candidacy for the following school year.

BPS also took advantage of a rarely used provision in the union contract that allowed schools to skip the internal transfer process and open post positions immediately by attaching a \$1,250 stipend to any new position that required additional duties. The district committed to paying this one-time stipend for every vacant position in traditional BPS schools. This effectively eliminated the internal transfer process and created a district-wide open posting system. Together, these reforms eliminated Phases 2 and 3 of the old hiring process, shifting external candidate hiring from mid to late summer to the spring.²

Theoretical Predictions

Hiring reforms such as those in BPS could improve the ability of new teacher hires, teacher retention, and ultimately student achievement through several theoretical channels. First, they allow principals to select the best candidate among a larger and higher-quality applicant pool instead of a small, potentially negatively selected pool of involuntary and voluntary transfers. Second, they allow districts to attract job candidates by eliminating the risk that even a high-performing early-career teacher could be “bumped” from their position. Third, they allow schools to better compete for job candidates with first-mover advantage in an uncertain hiring

² The district also undertook three related initiatives that stood to benefit both traditional and autonomous schools: merging the Office of Human Resources with the Office of Educator Effectiveness to create a unified Office of Human Capital; offering training and resources to aid principals in the hiring process, and expanding targeting recruitment efforts to attract more teacher candidates of color. Our estimates will also capture any differential effects of these reforms on traditional schools relative to autonomous schools.

market. Fourth, they provide time for more robust candidate screening efforts (Goldhaber, Grout, and Huntington-Klein, 2017; Jacob et al., 2018; Bruno & Strunk, 2018).

Mutual consent hiring reforms could also improve the productivity of new hires by enhancing match quality. The fit between a teacher and the culture at a school play an outsized role in the search process for teachers because wage schedules are fixed within districts and differ very little across districts. Jackson (2013) shows that teachers seek out better matches when transferring schools. Better matches may also result from pairings that maximize teachers' task-specific human capital across specific subjects (Fox, 2016; Goldhaber, Cowan, and Walch, 2013) and student populations (Aaronson, Barrow, and Sanders, 2007; Dee, 2004; Loeb, Soland, and Fox, 2014; Masters, Loeb, and Wyckoff, 2017).

Mutual consent hiring reforms would likely improve match quality through two primary channels. First, they allow principals to select candidates who share their schools' values and have skills that align well with the specific job opening. Second, they give external candidates a wider pool of job openings to consider and more time to search. Providing teachers and schools with more options during the search process can create gains through allocative efficiency. Improvements in match quality might also lead to a virtuous cycle of reduced teacher turnover allowing for more sustained on-the-job improvement. Papay and Kraft (2016) find that teachers hired after the school year starts leave their school at higher rates than on-time hires. On average, teachers who stay longer on the job become more effective by increasing their general and task-specific human capital (Papay & Kraft, 2015; Cook and Mansfield, 2016; Ost 2014).

IV. Data

BPS is the largest district in Massachusetts, serving approximately 54,000 students in prekindergarten through 12th grade and employing over 4,700 teachers. In Table 1 we provide a snapshot of student and teacher characteristics in BPS in 2013-14, the year prior to the reforms. BPS students come from diverse backgrounds: 39% of students were Hispanic, 36% Black, 14% White, and 9% Asian. Over 81% of all BPS students were from low-income households, while 23% of students are English language learners and 20% have special education needs. Teachers in BPS were predominantly white (61%) and female (74%) with Asian, Black, and Hispanic teachers constituting 6%, 22%, and 10% of the teacher labor force, respectively. 23% of teachers left their schools in BPS, a high but comparable rate as other large urban districts (Papay et al., 2017). Public school teachers earn tenure in Massachusetts after three years of service.

Table 1 also illustrates meaningful differences in population of students and teachers across traditional and autonomous schools in BPS. Students who attended traditional schools were, on average, substantially higher performing with an average difference in student achievement of 0.30σ in both math and ELA. This makes sense given that traditional schools include the districts test-in exam schools and that many autonomous schools were granted greater flexibility in an attempt to improve low performance. Traditional schools were also somewhat more likely to serve white students, students from families that were not low-income, and students who did not receive special education services. Teachers in traditional schools shared similar racial and ethnic backgrounds as their peers in autonomous schools, but were older, somewhat less likely to be novices and more likely to have a master's degree. Annual teacher turnover rates among autonomous schools was 32% compared to only 18% among traditional schools. This likely reflects both differential probabilities across schools types and the

fact that some autonomous schools were forced to dismiss up to 50% of their staff as part of a turnaround process.

We utilize a range of BPS administrative records on teachers, students, and schools across the 2010-11 to 2017-18 school years for which we have detailed hiring records. We construct two types of primary analytic datasets, a teacher-by-year-level panel and a student-by-year-level panel linked to students' math and ELA teachers in 4th through 8th grade. For our teacher-level analyses, we restrict our sample to teachers who both appear in the HR records with a job code as "teacher" and who are employed at a specific BPS school.³

We focus most of our teacher analyses on two types of new hires, with new hires defined as any teacher who is new to a BPS school. We distinguish between "external" and "internal" new hires. External new hires are in their first year of employment in BPS public schools and could be either novice teachers or experienced teachers who transferred to BPS. Internal new hires are teachers who previously taught in BPS and transferred across schools. BPS hires roughly 300 external new hires each year and 550 internal new hires.

HR records allow us to identify teachers' start dates, the day teachers officially report to the district for work. New teachers are typically expected to report to the district for a New Teacher Institute on the Monday of the fourth week of August, two weeks before the start of the school year. All teachers report on the Tuesday after Labor Day in September and classes begin on the following day. We define a new teacher as a late hire if she is hired after the first day of new teacher orientation, but before December 1st when the district typically moves to long-term substitute contracts for any remaining hires. We combine these administrative data with detailed hiring records collected by the district starting for candidates applying for open positions in

³ This definition excludes specialists, substitute teachers, and administrators as well as district-wide curriculum specialists and teachers who provide special education services across multiple schools.

2010. These data allow us to identify the hire approval date, when a school’s decision to hire a new teacher has been officially approved by the district. This occurs when a school has interviewed the candidate, made an offer, the offer has been accepted, and the hire has been approved by the central office.

V. Econometric Methods

We leverage the context of the BPS hiring reforms as a unique natural experiment to estimate the direct impact of these reforms on the types of teachers hired and their retention. We then show how late hiring affects student achievement in BPS, documenting the potential for improved outcomes by moving up hiring timelines. These analyses replicate models detailed in Papay and Kraft (2015) by leveraging student and school-grade-year fixed effects to compare deviations in within student performance over time among students in the same school, grade and year that were assigned to either an on-time hired teacher or a late-hired teacher (see Appendix A for econometric details). Finally, we examine the direct effects of the hiring reform on student achievement for all students as well as specifically for those taught by new hired teachers.

Hiring Reform Effects on Teacher Characteristics

We analyze the sudden and largely unanticipated hiring policy reforms using a differences-in-differences (DD) approach, where traditional BPS schools serve as the treated group and autonomous BPS schools serve as the comparison group. Importantly, the hiring reforms likely had both district-wide effects and effects that were concentrated among traditional schools. Our DD approach primarily identifies the latter, namely the joint effect of ending the

involuntary and voluntary transfer process and allowing schools to immediately consider external candidates, flexibilities that autonomous BPS schools already enjoyed.

This approach has several important strengths as well as limitations. Using BPS schools that operated with autonomy prior to 2014-15 as the comparison group allows us to control for any district-specific shocks to funding, teacher labor supply, or student achievement such as changes in local funding which constitute approximately 40% of total expenditures. However, this approach does not capture the full district-wide effect of the reforms which likely benefited autonomous schools to some degree as well. The reforms might also have had negative spillover effects on autonomous schools if traditional schools were better able to compete for external hires. We observe evidence of some degree of overlapping labor pools among teachers applying to open positions in BPS. In the four post-reform years where we have detailed applicant data, 54% of job candidates submitted applications to both a traditional and autonomous school while 27% only applied to autonomous schools and 19% only applied to traditional schools. We explore the degree that the reforms affected autonomous schools by testing for changes in outcomes for new hires at autonomous schools pre and post reforms, on average.

Our primary estimation approach applies a simple two-group, two-period difference-in-differences model to our data. For teacher outcome we fit the following model:

$$Y_{jst} = \beta(TRAD_t * POST_t) + \bar{X}'_{st}\theta + \mu_s + \pi_t + \varepsilon_{jst} \quad (1)$$

where a given characteristic, Y , for teacher j in school s in year t is a function of a vector of time-varying student characteristics averaged at school-level, school fixed effects, μ_s , and year fixed effects, π_t . School fixed effects serve to remove any time-invariant differences across traditional

and autonomous schools while year fixed effect account for any district-wide shocks to the characteristics of new teachers. Here, β , the coefficient associated with the interaction of an indicator for being a traditional school (*TRAD*) during the post-reform period (*POST*), captures the average treatment effect of the hiring reforms. We present models both with and without our vector of average student characteristics to tests the sensitivity of our results and cluster our standard errors at the school level across all models.

Our primary identify assumptions for the model are that autonomous school provide a valid counterfactual for estimating the trends we would have observed in autonomous schools in the absence of treatment, and that no other relevant policy changes occurred concurrently with the hiring reforms that would have differentially benefitted traditional schools over autonomous schools. We directly examine the pre-trends assumption by plot time series of our outcomes of interest for traditional and autonomous schools over our panel in Appendix Figures 1 through 3. These times series provide strong visual evidence in support of the parallel trends assumption. Extensive conversations with BPS officials did not reveal other policy changes in the district that would pose a potential threat to our identification. The package of hiring reforms implemented in 2014-2015 was the districts' signature initiative.

Hiring Reform Effects on Teacher Retention

We examine the effect of hiring reforms on the retention of newly hired teachers by nesting our difference-in-differences model in equation (2) within a discrete-time hazard framework. We model the hazard (i.e. the conditional probability) of a new hire leaving the BPS school where they were first hired using logistic regression in a person-period dataset, as follows:

$$h(\textit{Turnover}_{jk}) = \Pr[T_j = k | T_j \geq k \textit{ and } \bar{X}_{st}, \mu_s, \pi_t] =$$

$$\frac{1}{1 + e^{-\left(\sum_{g=1}^{g=4} 1(k=g)\tau_k + \sum_{g=1}^{g=4} 1(k=g)*(TRAD_s*POST_t)\gamma_k + \bar{X}'_{st}\theta + \mu_s + \pi_t + \varepsilon_{ist}\right)}} \quad (2)$$

Here, the probability that teacher j leaves their school, s , in time t , conditional on having not left the district in previous years, is a function of a set of indicator variables for relative time, $\sum_{g=1}^{g=4} 1(k = g)$. Relative time is defined as the number of years since a new hire began working in BPS where $k \in [1,4]$. We interact this set of time indicators with our treatment indicator ($TRAD_{st} * POST_t$) and include the full set of fixed effects and controls described above. The γ_j 's provide treatment effect estimates on the conditional probability a new hire leaves their school after a given year.

Hiring Reform Effects on Student Achievement

We analyze how the BPS hiring reforms affected student achievement by fitting a difference-in-differences model analogous to equation (1) in a student-year panel dataset. For a given subject, we model a student i 's test score, A , on the MCAS as follows:

$$A_{it} = \vartheta A_{i,t-1} + \alpha(TRAD_s * POST_t) + X'_{it}\phi + \bar{X}'_{jt}\lambda + \mu_s + \pi_{gt} + \varepsilon_{it} \quad (3)$$

Here, current year achievement as function of prior performance in both math and ELA, $A_{i,t-1}$, as well as vectors of individual student characteristics, X_{it} , and the average characteristics of students taught by teacher j in year t , \bar{X}_{jt} . We also include our interaction term for traditional schools in the post period and school and grade-by-year fixed effects. The coefficient α captures

the average treatment effect of hiring reforms on student achievement across traditional BPS schools.

We start by fitting models in our full sample of students linked to all teachers to test for school-wide effects. These whole-school models capture both direct and indirect mechanisms through which the hiring reforms might have benefits schools and ultimately student achievement. We then restrict our samples to focus on external and internal new hires, separately, in their first year on the job to examine the direct effects on teacher effectiveness. Finally, we examine indirect spillover effects on veteran teachers who were never new hires in either the pre or post periods. We assess pretrends by plotting event studies of whole-school effects on student achievement.

VI. Findings

Effects of the Hiring Reforms on New Hires

Hiring records as well as conversations with BPS officials suggest that the hiring reforms had substantial effects on the timing of new external teacher hiring across the district. In Figure 1, we plot the distribution of hire approval dates in the four years prior to (2010-2013) and after the reforms (2014-2017). Prior to the reforms, hire approvals were almost non-existent before July 1st and were heavily concentrated in August and early September. The hiring reforms transformed the distribution of hire approval dates for school the district including autonomous schools, moving it up by several months and distributing it more evenly across the hiring cycle. In the post-reform period, over 54% of new hires were approved before July 1st. In this way, the reforms likely benefit both traditional and autonomous schools by creating formal expectations

of early hiring and a more service oriented to processing new hires rather than a compliance oriented focus.

Although the hiring reforms accelerated the timing of hiring across all BPS schools, descriptive patterns suggest that the compositional effects on new external hires were largely concentrated among traditional schools. In Table 2 Panel A, we examine average changes in the characteristics of newly hired teachers pre and post reforms to better understand the underlying dynamics of our difference-in-differences model. We observe a significant reduction in the number of external new hires in autonomous schools who were hired late, who were white, who were novices, and who left their schools after the first year. In contrast, we see no meaningful or significant changes over time among new external hires at autonomous schools suggesting the reforms did not appreciably change their hiring process along these observed dimensions. The total number of new external hires also expanded quite similarly across the two school types, increasing by roughly 11% and 13% among traditional and autonomous schools. Together, these patterns point to benefits of the hiring process accruing to traditional schools without negative spillovers to autonomous schools.

Among internal hires, we see a somewhat different pattern. Traditional schools again experience positive changes over time with the teacher workforce becoming more diverse and reductions in turnover. However, we see evidence that internal hires in autonomous schools were less likely to be Hispanic and more likely to leave their schools in the post period. We also find divergent patterns of overall hiring rates where the number of internal new hires declined by 30% among traditional schools but increased 55% among autonomous schools. These patterns suggest that the reforms appreciably changed the likelihood open positions were filled by internal transfers at traditional and autonomous schools and that any net effects of the reform

may reflect both positive benefits to traditional schools and negative spillovers autonomous schools.

Formal estimates of the effect of mutual consent hiring reforms from our DD models confirm the implied effects from the first and second differences described above. In Table 3, we present estimates from our teacher-level DD model both without and with controls for time-varying school characteristics. In our preferred model with controls, we find that the reforms reduced the frequency of late hiring for external new hires by 13 percentage points, a 49% reduction in late hires due to the hiring reforms. Estimates of the effect of hiring reforms on new hire diversity in Table 3 suggest that the reforms increased the share of Black or Hispanic external and internal new hires. We find that the reforms raise the percentage of hires that are Black or Hispanic by 10 percentage points for external hires and 11 percentage points for internal hires. These represent a 38% and 30% increase in the probability a new external and internal teacher was Black or Hispanic. We also find suggestive evidence that the reforms reduced the likelihood an external new hire was a novice, although these estimates are not statistically significant in our model with controls.

We estimate that the hiring reforms had large and persistent effects on the probability a new external hire remained at their school. In Table 4, we show that the hiring reforms reduced the number of new hires who left their schools after their first year on the job by 9.8 percentage points, and by an additional 8.4 percentage points among teachers who returned for their second year. These effects on teacher retention are suggestive of better matches between new hires and schools caused by the reforms. However, they could also reflect the fact that the hiring reforms changed the composition of new hires – most notably reducing the number of novice new hires. We fit an additional model controlling for teacher experience in order to assess whether the

effects we find are driven primarily by the changing composition of new hires or other explanations including better teacher-school matches. As shown in column 3, our results remain largely unchanged after including a flexible set of indicators to control for teacher experience, suggesting match effects play an important role in reducing turnover rates.

Replicating the Effects of Late Hiring on Student Achievement

We next estimate the negative effects of late hiring on student achievement to illustrate one pathway through which the hiring reforms might have benefited student achievement. These replication analyses confirm prior evidence of the negative effect of starting the school year without a full-time teacher and being taught by a late-hired teacher. As shown in Appendix Table A1, we estimate that late hiring reduces student achievement between -0.07σ to -0.11σ in math and -0.05σ to -0.10σ in ELA. These estimates are robust across models that include student fixed effects or controls for prior achievement and student characteristics. Our estimates also remain quite similar when we include controls for teacher characteristics suggesting these effects are not driven by differences in the experience, education, race, or gender of teachers who are hired late versus on time. These findings serve to highlight one specific avenue through which more open, early hiring stands to benefit students.

Effects of the Hiring Reforms on Student Achievement

Estimates of the effect of hiring reforms on student achievement reveal the reforms led to meaningful improvements in achievement for BPS students. We begin by plotting event studies of school-wide effects on achievement in math and ELA in Figure 2 Panel A and B. These figures are suggestive of positive effect in the post period (solid dots) although estimates for individual year terms are imprecisely estimated. The figures also present little visual evidence of strong differential pretrends that might explain away our findings.

We estimate more precise school-wide effects by pooling across the pre and post periods. As shown in Table 5 Column 1, effects on academic achievement for all teachers with students in tested grades are 0.09σ in math and 0.07σ in ELA. We next examine the direct effects on the productivity of new hires by restricting our sample to students taught by new hires in their first year in their new position. As shown in column 2, the effect of the reforms on the productivity of new external hires as measured by contributions to academic achievement are large. We find effects of 0.20σ in math and 0.23σ in ELA. Results for internal new hires in column 3 are large as well, with effects of 0.30σ in math and 0.14σ in ELA. This suggests the reforms allowed traditional schools to hire teachers whose first-year performance was more than an entire standard deviation higher in the distribution of teacher effectiveness, likely because the schools were able to attract more competitive candidates in the market and because the hiring reforms promoted better teacher-school matches.

We also find some evidence that the hiring reforms served to raise the productivity of veteran teachers in traditional schools. Estimates for veteran teachers who are “never new hires” in the pre period compared to the post period are a significant 0.09σ in math and an insignificant 0.04σ in ELA. These improvements in the performance of veteran teachers could reflect positive spillovers from working with more effective peers (Jackson & Bruegman, 2009; Sun, Loeb, Grissom, 2017), changes in principals’ approaches to strategic retention, and/or the benefits of an improved school culture due to better teacher-school matches.

Falsification Tests

For our difference-in-differences models to provide valid impact estimates on school-wide student achievement, we must assume that the hiring reforms did not change students’ and families’ enrollment patterns across traditional and autonomous schools. Our analyses of effects

on new hires requires an additional assumption that student sorting within schools to new hires did not change because of the reforms. In Table 6, we present results from a series of unconditional DD models with school and year fixed effects where we replace our achievement measures with student characteristics and prior test scores, which should be unaffected by the hiring reforms.

As shown in Panel A, we find no evidence that the hiring reforms affected the types of students who attended traditional schools. All of our estimates are very small in magnitude and statistically insignificant. In Panel B, we find little evidence of differential sorting or assignment of students to newly hired external teachers caused by the reforms. Effects on all but one measure are insignificant. We do find that newly hired teachers were 10 percentage points more likely to be assigned to teach Hispanic students. In Panel C, we see more consistent evidence of dynamic sorting where internal new hires were assigned students who were 6 percentage points more likely to be Black and have substantially lower prior achievement in math (-0.19σ) and ELA (-0.28σ). This pattern of differential sorting would, if anything, attenuate the effects we find on student achievement.

Heterogeneous Effects

Understanding the distributional effects of the reforms across schools in BPS has important equity implications. Here we explore whether granting greater autonomy to all traditional schools particularly advantaged those schools serving more affluent, whiter, and higher achieving student populations which prior research reveals are correlated with teacher transfer patterns (Boyd et al., 2013). In Table 7 we report results from equation (1) where we allow for an interaction between Traditional*Post and four different school characteristics: % of student that are from low-income families, % of students that are Black or Hispanic, and average

prior achievement in math and ELA. We fail to find a single significant interaction across the 24 different coefficients we estimate. We also augment equation (3) with a parallel set of interaction terms to test for differential effects on student achievement. As shown in Table 8, we find no evidence of heterogeneous effects suggesting that the reforms appear to have been broadly beneficial for schools serving all different types of students.

VIII. Discussion and Conclusion

The estimates we present here provide compelling evidence that reducing frictions and increasing choice in the hiring process can lead to substantial productivity gains in the education sector. We find substantial school-wide returns to the BPS mutual consent hiring reforms at large scale—the reforms affected 84 traditional BPS schools. The magnitudes – 0.09σ in math and 0.07σ in ELA – are larger than the average effect sizes Fryer (2017) found across 72 school-based experimental interventions (0.07σ in math and 0.05σ in reading). They are also larger than 70% of the effect sizes from 181 large-scale randomized control trials of education interventions catalogued in Kraft (2020). Decomposing these school-wide estimates to better understand the primary mechanisms reveals that the reforms allowed schools to hire new teachers that were much more effective and/or better matched to their schools and specific positions. Effects on the productivity of new hires of approximately 0.2σ in student-level achievement equate to improving the productivity of new hires by a full standard deviation or more.

We find evidence that these benefits of mutual consent hiring reforms for student achievement likely operate through several mechanisms. First, the reforms substantially reduced the incidence of late teacher hiring which we show lowers achievement (Papay & Kraft, 2016). Second, the reforms increased the effectiveness of new hires with fewer novice teachers (Papay & Kraft, 2015) and better teacher-school matches (Jackson, 2013). Third, the reforms reduced

teacher turnover which, on average, is detrimental to student achievement (Ronfeldt, Loeb & Wyckoff, 2013). Fourth, schools were able to more successfully compete for teachers of color, hiring approximately 110 more Black and Hispanic teachers in the four years after the reforms than we predict would have occurred otherwise. This is particularly important given the large racial imbalance between BPS students (36% Black; 40% Hispanic) and teachers (16% Black; 11% Hispanic) and the growing literature demonstrating the academic and social-emotional benefits to students of color when they are taught by a teacher who shares their same race (Dee, 2004, 2007; Egalite et al., 2015; Lindsay & Hart, 2017; Holt & Gershenson, 2015; Gershenson et al., 2018).

These benefits of mutual consent hiring reforms came at a cost of \$10.5 million in the first year. This translates to approximately \$200 per student, or about 1% of BPS's billion dollar budget, a per-pupil cost that is quite low relative to other education interventions (Kraft, 2020). The \$1,250 stipends required by the CBA for open-posting positions early cost \$575,000 in the first year, though the stipend was eliminated in the new CBA between the district and the union in 2016. The vast majority of incurred costs came in the form of salaries and benefits for tenured teachers placed in the excess pool who were unable to secure jobs. Ending forced placements meant that the district now had to absorb the salaries and benefits of between 70 to 100 teachers placed in SPC positions. Encouragingly, the district was able to incrementally lower costs to \$8 million by 2018, or \$150 per student. The decreasing number of excess teachers who remained in SPC positions may be a natural result of improved teacher-school fit under mutual consent hiring. It also reflects increased voluntary separations and the districts efforts to place or terminate teachers in the excess pool.

Finally, our findings lend strong empirical support to economic theories that posit there are large unrealized gains to be had from opening the search process and improving job matching. In the education sector, public school districts have the potential to substantially improve teacher productivity and retention by providing schools with more autonomy and job applicants with more choices in the hiring process. This is particularly true for urban districts, which tend to hire later than their suburban peers. In general equilibrium framework, we would expect these districts to benefit from hiring reforms, while the suburban districts they compete with may lose. We cannot study these dynamics directly, although our evidence from autonomous schools in BPS suggests that they were not harmed by this competition. This may be because of the importance of teacher-school match in determining productivity.

References

- Boston Public Schools. 2015. "Superintendent's Circular." HRS-HS-7. https://bps.solutions/BPS/Superintendent_Circular_HRS-HS-7_Staffing_Reassignment_and_Hiring_2015-2016.pdf
- Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. "Teachers and student achievement in the Chicago public high schools." *Journal of Labor Economics* 25 (1): 95-135.
- Angrist, Joshua D., Parag A. Pathak, and Christopher R. Walters. 2013. "Explaining charter school effectiveness." *American Economic Journal: Applied Economics* 5 (4): 1-27.
- Bender, Stefan, Nicholas Bloom, David Card, John Van Reenen, and Stefanie Wolter. 2018. "Management Practices, Workforce Selection, and Productivity". *Journal of Labor Economics* 36 (1): 371-409.
- Bloom, Nicholas, Renata Lemos, Rafaella Sadun, and John Van Reenen. 2015. "Does Management Matter in Schools?" *The Economic Journal* 125 (584): 647-674.
- Boston Municipal Research Bureau. 2016. "Boston Public Schools Human Capital Initiative" Special Report 16-1. <https://bmr.org/wp-content/uploads/2014/01/SR16-1HCI.pdf>
- Boyd, Donald, Hamilton Lankford, Susanna Loeb, Matthew Ronfeldt, and James Wyckoff. 2011. "The role of teacher quality in retention and hiring: Using applications to transfer to uncover preferences of teachers and schools." *Journal of Policy Analysis and Management* 30: 88–110.
- Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2013. "Analyzing the determinants of the matching of public school teachers to jobs: Disentangling the preferences of teachers and employers." *Journal of Labor Economics* 31(1): 83-117.
- Bruno, Paul, and Katharine O. Strunk. 2018. "Making the Cut: The Effectiveness of Teacher Screening and Hiring in the Los Angeles Unified School District." *Educational Evaluation and Policy Analysis* 41 (4): 426-460.
- Bueno, Carycruz, and Tim R. Sass. 2019. "The Effects of Differential Pay on Teacher Recruitment and Retention. Working Paper No. 219-0519." *National Center for Analysis of Longitudinal Data in Education Research (CALDER)*.
- Clotfelter, Charles, Elizabeth Glennie, Helen Ladd, and Jacob Vigdor. 2008. "Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina." *Journal of Public Economics* 92(5-6): 1352-1370.
- Cook, Jason B., and Richard K. Mansfield. 2016. "Task-specific experience and task-specific talent: Decomposing the productivity of high school teachers." *Journal of Public Economics* 140: 51-72.

- Dee, Thomas S. 2004. "Teachers, race, and student achievement in a randomized experiment." *Review of Economics and Statistics* 86 (1): 195-210.
- Dee, Thomas S. 2007. "Teachers and the gender gaps in student achievement." *Journal of Human Resources* 42 (3): 528-554.
- Egalite, Anna J., Brian Kisida, and Marcus A. Winters. 2015. "Representation in the classroom: The effect of own-race teachers on student achievement." *Economics of Education Review* 45: 44-52.
- Engel, Mimi. 2012. "The timing of teacher hires and teacher qualifications: Is there an association?" *Teachers College Record* 114: 1–29.
- Feng, Li, and Tim R. Sass. 2018. "The impact of incentives to recruit and retain teachers in "hard-to-staff" subjects." *Journal of Policy Analysis and Management* 37 (1): 112-135.
- Fox, Lindsay. 2016. "Playing to teachers' strengths: Using multiple measures of teacher effectiveness to improve teacher assignments." *Education Finance and Policy* 11 (1): 70-96.
- Fryer Jr, Roland G. 2014 "Injecting charter school best practices into traditional public schools: Evidence from field experiments." *The Quarterly Journal of Economics* 129, no. 3: 1355-1407.
- Fryer Jr, Roland G. 2017. "The production of human capital in developed countries: Evidence from 196 randomized field experiments." In *Handbook of economic field experiments 2*: 95-322. North-Holland.
- Gershenson, Seth, Cassandra Hart, Joshua Hyman, Constance Lindsay, and Nicholas W. Papageorge. 2018. "The Long-Run Impacts of Same-Race Teachers." NBER Working Paper 25254.
- Glazerman, Steven, Ali Protik, Bing-ru Teh, Julie Bruch, and Jeffrey Max. 2013. "Transfer Incentives for High-Performing Teachers: Final Results from a Multisite Randomized Experiment. NCEE 2014-4004." Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Grissom, Jason A., Susanna Loeb, and Nathaniel A. Nakashima. 2014. "Strategic involuntary teacher transfers and teacher performance: Examining equity and efficiency." *Journal of Policy Analysis and Management* 33: 112–140.
- Goldhaber, Dan, Cyrus Grout, and Nick Huntington-Klein. 2017. "Screen twice, cut once: Assessing the predictive validity of applicant selection tools." *Education Finance and Policy* 12 (2): 197-223.

- Goldhaber, Dan, James Cowan, and Joe Walch. 2013. "Is a good elementary teacher always good? Assessing teacher performance estimates across subjects." *Economics of Education Review* 36: 216-228.
- Hanushek, Eric A., and Steven G. Rivkin. 2010. "Generalizations about Using Value-Added Measures of Teacher Quality." *American Economic Review* 100 (2): 267–271.
- Holt, Stephen B., and Seth Gershenson. 2015. "The Impact of Teacher Demographic Representation on Student Attendance and Suspensions." IZA Discussion Paper 9554.
- Jackson, C. Kirabo. 2013. "Match quality, worker productivity, and worker mobility: Direct evidence from teachers." *Review of Economics and Statistics* 95: 1096–1116.
- Jackson, C. Kirabo, and Elias Bruegmann. 2009. "Teaching students and teaching each other: The importance of peer learning for teachers." *American Economic Journal: Applied Economics* 1: 85–108.
- Jacob, Brian A., Jonah E. Rockoff, Eric S. Taylor, Benjamin Lindy, and Rachel Rosen. 2018. "Teacher applicant hiring and teacher performance: Evidence from DC Public Schools". *Journal of Public Economics* 166: 81-97.
- Jacob, Brian A., Jonah E. Rockoff, Eric S. Taylor, Benjamin Lindy, and Rachel Rosen. 2018. "Teacher Applicant Hiring and Teacher Performance: Evidence from DC Public Schools." *Journal of Public Economics*, 166, 81-97.
- Jones, Nathan D., Adam Maier, and Erin Grogan. 2011. "The extent of late-hiring and its relationship with teacher turnover: Evidence from Michigan." Unpublished.
- Johnson, Susan. Moore, and Suesse, Jennifer. M. (2005). Staffing the Boston public schools. PELP Case Study. PEL-024.
- Kraft, Matthew A., and Allison F. Gilmour. 2017. "Revisiting the Widget Effect: Teacher evaluation reforms and the distribution of teacher effectiveness." *Educational Researcher* 46 (5): 234-249.
- Lazear, Edward P., and Paul Oyer. 2009. "Personnel Economics," in *Handbook of Organizational Economics*, ed. Robert Gibbons and John Roberts (Princeton University Press), 479-519.
- Levin, Jessica, Jennifer Mulhern, and Joan Schunck. 2005. "Unintended Consequences: The Case for Reforming the Staffing Rules in Urban Teachers Union Contracts." New York, NY: The New Teacher Project.
- Levin, Jessica, and Meredith Quinn. 2003. "Missed opportunities: How we keep high-quality teachers out of urban classrooms." New York, NY: The New Teacher Project.

- Lindsay, Constance A., and Cassandra M. D. Hart. 2017. "Exposure to Same-Race Teachers and Student Disciplinary Outcomes for Black Students in North Carolina." *Evaluation and Policy Analysis* 39 (3): 485-510.
- Liu, Edward, and Susan M. Johnson. 2006. "New teachers' experiences of hiring: Late, rushed, and information-poor." *Educational Administration Quarterly* 42: 324-360.
- Loeb, Susanna, James Soland, and Lindsay Fox. 2014. "Is a Good Teacher a Good Teacher for All? Comparing Value-Added of Teachers With Their English Learners and Non-English Learners." *Educational Evaluation and Policy Analysis* 36 (4): 457-475.
- Master, Benjamin, Susanna Loeb, and James Wyckoff. 2017. "More than content: The persistent cross-subject effects of English language arts teachers' instruction." *Educational Evaluation and Policy Analysis* 39 (3): 429-447.
- National Council on Teacher Quality. 2010 *Human Capital in Boston Public Schools: Rethinking How to Attract, Develop and Retain Effective Teachers*. ERIC Clearinghouse.
- Ost, Ben. 2014. "How do teachers improve? The relative importance of specific and general human capital." *American Economic Journal: Applied Economics* 6 (2): 127-51.
- Oyer, Paul and Scott Schaefer. 2011. "Personnel Economics: Hiring and Incentives." *Handbook of Labor Economics* 4: 1769-1823.
- Papay, John P., and Matthew A. Kraft. 2015. "Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement." *Journal of Public Economics* 130: 105-119.
- Papay, John P., and Matthew A. Kraft. 2016. "The productivity costs of inefficient hiring practices: Evidence from late teacher hiring." *Journal of Policy Analysis and Management* 35 (4): 791-817.
- Papay, John P., Andrew Bacher-Hicks, Lindsay C. Page, and William H. Marinell. 2017. "The challenge of teacher retention in urban schools: Evidence of variation from a cross-site analysis." *Educational Researcher* 46 (8): 434-448.
- Papay, John P., Eric S. Taylor, John H. Tyler, and Mary E. Laski. 2020. "Learning job skills from colleagues at work: Evidence from a field experiment using teacher performance data." *American Economic Journal: Economic Policy* 12(1): 359-88.
- Sawchuk, Stephen, "Mutual Consent' Teacher Placement Gains Ground," *Education Week*, July 6, 2010. https://www.edweek.org/ew/articles/2010/07/01/36placement_ep.h29.html.

- Steele, Jennifer L., Richard J. Murnane, and John B. Willett. 2010. "Do financial incentives help low-performing schools attract and keep academically talented teachers? Evidence from California." *Journal of Policy Analysis and Management* 29 (3): 451-478.
- Sun, Min, Susanna Loeb, and Jason A. Grissom. 2017. "Building Teacher Teams: Evidence of Positive Spillovers from More Effective Colleagues." *Education Evaluation and Policy Analysis* 39, no. 1: 104-125.

Figures

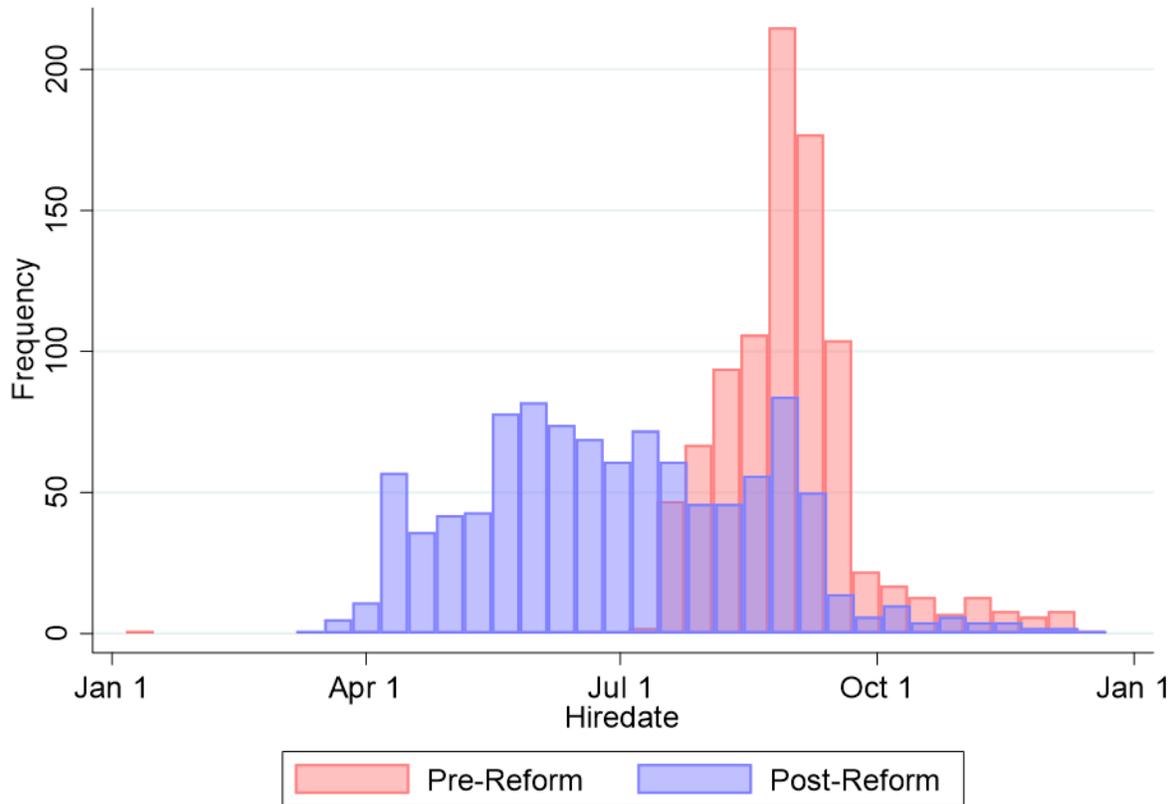
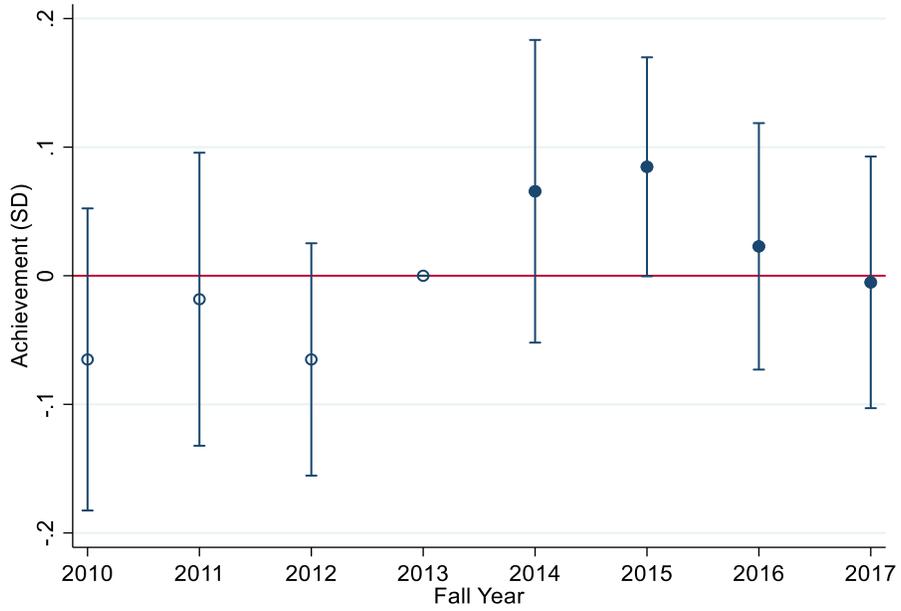


Figure 1. The distribution of hire approval dates prior to BPS mutual consent hiring reforms (2010-2013) and after the hiring reforms (2014-2017).

Panel A: Math



Panel B: English Language Arts

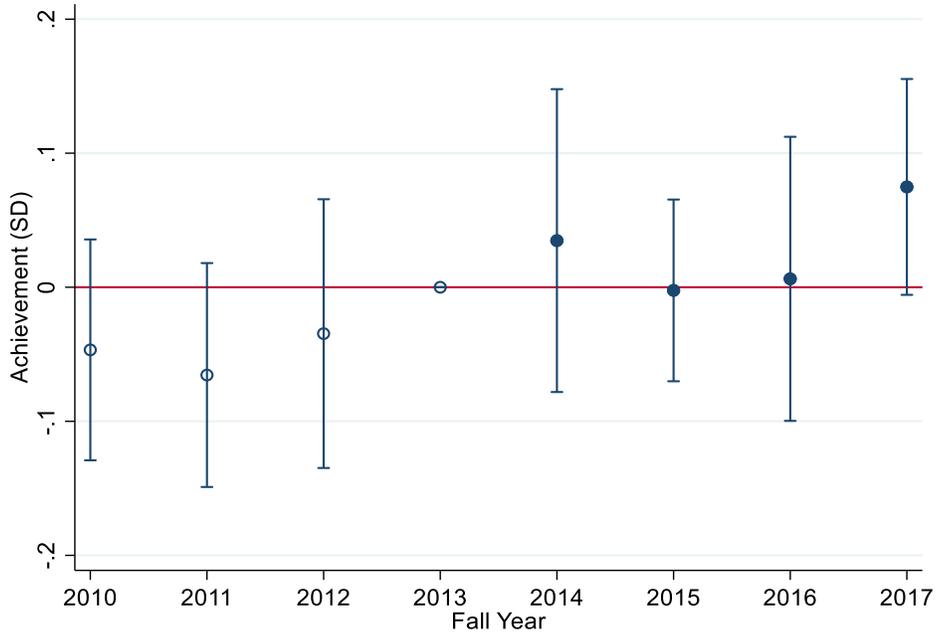


Figure 2: Event-study estimates of the effect of hiring reforms on student achievement in traditional BPS schools.

Tables

Table 1: Teacher and Student Characteristics in 2013-14 School Year

	All	Traditional	Autonomous	Difference (Trad- Auto)
Panel A: Teacher Characteristics				
Age	40.83	42.86	37.09	5.77***
Female	0.74	0.75	0.71	0.03**
Asian	0.06	0.06	0.06	0.01
Black	0.22	0.22	0.23	-0.02
Hispanic	0.10	0.10	0.10	0.00
White	0.61	0.62	0.60	0.02
Novice	0.04	0.02	0.07	-0.05***
Master's Degree	0.82	0.83	0.80	0.03**
Turnover	0.23	0.18	0.32	-0.13***
n (teachers)	4,388	2,843	1,545	
Panel B: Student Characteristics				
Special Education	0.20	0.18	0.24	-0.05**
Low Income	0.81	0.78	0.86	-0.07*
Asian	0.09	0.11	0.05	0.06**
Black	0.36	0.32	0.45	-0.12***
Hispanic	0.39	0.38	0.40	-0.02
White	0.14	0.16	0.08	0.08**
English language learner	0.23	0.24	0.23	0.01
Math Score (std)	0.05	0.13	-0.17	0.30***
Reading Score (std)	0.03	0.11	-0.19	0.30**
n (students)	47,108	33,586	13,825	
n (schools)	121	84	37	

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Significance test of the difference in student characteristics are based on coefficients from a regression model where standard errors are clustered at the school level.

Table 2: New Hire Characteristics Pre and Post Reforms

	Traditional Schools			Autonomous Schools		
	Pre	Post	Diff	Pre	Post	Diff
Panel A: External Hires						
Asian	0.09	0.08	-0.01	0.06	0.07	0.01
Black	0.15	0.17	0.02	0.18	0.17	-0.01
Hispanic	0.11	0.14	0.03	0.12	0.10	-0.02
White	0.65	0.56	-0.09**	0.63	0.60	-0.03
Novice	0.68	0.51	-0.17***	0.58	0.56	-0.02
Turnover	0.45	0.37	-0.07**	0.33	0.35	0.02
Late Hire	0.27	0.15	-0.12***	0.11	0.10	-0.01
n	366	415		607	678	
Panel B: Internal Hires						
Asian	0.05	0.05	0.00	0.05	0.06	0.01
Black	0.25	0.26	0.01	0.29	0.28	0.00
Hispanic	0.11	0.14	0.02*	0.14	0.09	-0.05***
White	0.58	0.55	-0.03	0.52	0.56	0.04
Turnover	0.28	0.24	-0.04**	0.33	0.39	0.07**
n	1,384	966		532	822	

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. P-values for the statistical significance of mean differences are estimated in a regression model where standard errors are clustered at the school level.

Table 3: The Effect of Hiring Reform on New Hired Teachers

	External Hires			Internal Hires		
	Traditional Pre-Reform Mean			Traditional Pre-Reform Mean		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Late Hire						
Traditional*Post	0.268	-0.137*** (0.044)	-0.134*** (0.045)			
n		2,052				
Panel B: Teacher of Color						
Traditional*Post	0.260	0.108** (0.049)	0.103** (0.049)	0.367	0.122*** (0.040)	0.113*** (0.042)
n		2,052			3,636	
Panel C: Novice Teachers						
Traditional*Post	0.679	-0.110* (0.064)	-0.087 (0.067)			
n		2,052				
School Characteristics		N	Y		N	Y

Note: * p<0.1, ** p<0.05, *** p<.01. Standard errors reported in parentheses are clustered at the school level.

Table 4: The Effect of Hiring Reforms on Turnover among New Hires

	(1)	(2)	(3)
Traditional*Post*($k=1$)	-0.124*** (0.039)	-0.098** (0.039)	-0.093** (0.041)
Traditional*Post*($k=2$)	-0.111** (0.049)	-0.084* (0.049)	-0.150*** (0.052)
Traditional*Post*($k=3$)	-0.070 (0.064)	-0.042 (0.063)	-0.029 (0.062)
Traditional*Post*($k=4$)	0.000 (0.076)	0.027 (0.075)	-0.076 (0.087)
n	4,244	4,235	4,240
School Characteristics	N	Y	Y
Teacher Experience	N	N	Y

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors reported in parentheses are clustered at the school level. Sample size drops when covariates are added due to a small number of observations being perfectly predicted. Notes sample sizes differ due to observations that are dropped because they are perfectly predicted.

Table 5: The Effect of Hiring Reforms on Student Achievement

	All Teachers	External Hires in First Year	Internal Hires in First Year	Veterans
	(1)	(2)	(3)	(4)
Panel A: Math				
Traditional*Post	0.090*** (0.034)	0.204** (0.086)	0.289*** (0.060)	0.085* (0.047)
n	127,899	5,575	12,234	85,250
Panel B: ELA				
Traditional*Post	0.068** (0.028)	0.228*** (0.077)	0.139** (0.063)	0.039 (0.045)
n	139,659	6,869	12,730	91,995
School FE	Y	Y	Y	Y
Student Characteristics	Y	Y	Y	Y
Prior Achievement	Y	Y	Y	Y
Classroom Characteristics	Y	Y	Y	Y

Notes: * p<0.1, ** p<0.05, *** p<.01. Standard errors reported in parentheses are clustered at the school level. All models include grade-by-year fixed effects.

Table 6: Falsification Tests of the Effect of Hiring Reforms on Student Characteristics

	Special Education	Low Income	Asian	Black	Hispanic	White	English Language Learner	Prior Math Scores	Prior ELA scores
Panel A: All Teachers									
Traditional*Post	0.003 (0.014)	0.013 (0.016)	-0.001 (0.007)	-0.007 (0.017)	0.022 (0.019)	-0.015 (0.011)	0.016 (0.019)	0.011 (0.053)	-0.028 (0.042)
n	139,659	139,659	139,659	139,659	139,659	139,659	139,659	138,738	139,659
Panel B: External Hires in First Year									
Traditional*Post	0.061 (0.058)	0.015 (0.041)	-0.030 (0.020)	-0.034 (0.048)	0.099* (0.052)	-0.019 (0.032)	-0.065 (0.067)	-0.157 (0.139)	-0.151 (0.153)
n	6,869	6,869	6,869	6,869	6,869	6,869	6,869	6,757	6,869
Panel C: Internal Hires in First Year									
Traditional*Post	0.011 (0.040)	0.023 (0.026)	-0.011 (0.011)	0.058* (0.031)	-0.047 (0.031)	-0.007 (0.014)	-0.011 (0.038)	-0.185* (0.094)	-0.277*** (0.082)
n	12,730	12,730	12,730	12,730	12,730	12,730	12,730	12,665	12,730

Notes: * p<0.1, ** p<0.05, *** p<.01. Standard errors reported in parentheses are clustered at the school level.

Table 7: Testing for Heterogeneous Effects of Hiring Reform on New Hired Teachers

	External Hires				Internal Hires			
	% Low Income	% Black or Hispanic	Prior Math Achievement	Prior ELA Achievement	% Low Income	% Black or Hispanic	Prior Math Achievement	Prior ELA Achievement
Panel A: Late Hire								
Traditional*Post	-0.246 (0.178)	-0.262* (0.145)	-0.139*** (0.047)	-0.139*** (0.047)				
Traditional*Post*School Characteristic	0.015 (0.023)	0.016 (0.017)	-0.041 (0.031)	-0.049 (0.030)				
n	2,052	2,052	1,631	1,628				
Panel B: Teacher of Color								
Traditional*Post	0.328** (0.138)	0.251** (0.103)	0.102* (0.055)	0.102* (0.055)	-0.010 (0.143)	0.001 (0.103)	0.106** (0.045)	0.106** (0.044)
Traditional*Post*School Characteristic	-0.030 (0.019)	-0.019 (0.015)	0.015 (0.032)	0.015 (0.031)	0.016 (0.017)	0.014 (0.012)	-0.036 (0.025)	-0.022 (0.027)
n	2,052	2,052	1,631	1,628	3,636	3,636	3,291	3,284
Panel C: Novice Teachers								
Traditional*Post	-0.143 (0.172)	0.000 (0.135)	-0.068 (0.078)	-0.068 (0.078)				
Traditional*Post*School Characteristic	0.007 (0.022)	-0.011 (0.016)	-0.015 (0.032)	-0.007 (0.031)				
n	2,052	2,052	1,631	1,628				
Panel D: Turnover								
Traditional*Post	0.003 (0.135)	-0.053 (0.146)	-0.070 (0.046)	-0.071 (0.046)	-0.276** (0.109)	-0.264*** (0.082)	-0.129*** (0.048)	-0.128*** (0.048)
Traditional*Post*School Characteristic	-0.015 (0.018)	-0.008 (0.017)	0.006 (0.027)	0.022 (0.023)	0.021 (0.014)	0.019* (0.010)	-0.035 (0.022)	-0.027 (0.022)
n	2,052	2,052	1,631	1,628	3,636	3,636	3,291	3,284
School Characteristics	Y	Y	Y	Y	Y	Y	Y	Y

Note: * p<0.1, ** p<0.05, *** p<.01. Standard errors reported in parentheses are clustered at the school level. School-level measures percent low income and percent Black and Hispanic are scaled so that a one unit change is equivalent to a 10 percentage point change. Average school-level achievement in the prior year is standardized at the school level so a one unit change is a one standard deviation change in the distribution of school-level mean differences.

Table 8: Testing for Heterogeneous Effects of Hiring Reforms on Student Achievement

	All Teachers			
	% Low Income	% Black or Hispanic	Prior Math Achievement	Prior ELA Achievement
	(1)	(2)	(3)	(4)
	Panel A: Math			
Traditional*Post	0.122**	0.127***	0.091**	0.091**
	(0.047)	(0.042)	(0.037)	(0.037)
Traditional*Post*School Characteristic	-0.043	-0.047	0.021	0.018
	(0.057)	(0.048)	(0.027)	(0.030)
n	127,899	127,899	127,899	127,899
	Panel B: ELA			
Traditional*Post	0.054	0.045	0.068**	0.068**
	(0.068)	(0.050)	(0.031)	(0.031)
Traditional*Post*School Characteristic	0.019	0.025	-0.020	-0.039
	(0.085)	(0.061)	(0.036)	(0.041)
n	139,659	139,659	139,659	139,659
School FE	Y	Y	Y	Y
Student Characteristics	Y	Y	Y	Y
Prior Achievement	Y	Y	Y	Y
Classroom Characteristics	Y	Y	Y	Y

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors reported in parentheses are clustered at the school level. All models include grade-by-year fixed effects. School-level measures percent low income and percent Black and Hispanic are scaled so that a one unit change is equivalent to a 10 percentage point change. Average school-level achievement in the prior year is standardized at the school level so a one unit change is a one standard deviation change in the distribution of school-level mean differences.

Appendix A

Late Hiring Effects on Student Achievement

We replicate the primary analytic approach outlined in Papay and Kraft (2016) using a model that exploits variation within-students over time and across teachers within each grade in a given school and year. We code a time-invariant indicator ($EVER_LATE_j$) to indicate whether a teacher was initially hired into the district late and a predictor ($NEWHIRE_{jt}$) to indicate whether the teacher was newly hired in the district in a given year. We fit different specifications of the following model for student i with teacher j in grade g , school s , and year t .⁴

$$Y_{it} = \beta_1 EVER_{LATE_j} + \beta_2 NEWHIRE_{jt} + \beta_3 EVER_{LATE} * NEWHIRE_{jt} + \gamma f[(EXPER_{jt})] + \bar{X}'_{jt} \xi + \delta_i + \theta_{sgt} + \varepsilon_{it} \quad (1)$$

Parameters β_1 , β_3 , and their linear combination identify our three primary quantities of interest. The parameter sum ($\beta_1 + \beta_3$) represents the differential effect on student achievement of being assigned to a late-hired teacher's classroom in her first year relative to an on-time-hired teacher in her first year. We decompose this total effect into a disruption effect (β_3), the effect on student achievement that *only* occurs in the year a teacher was hired late, and a labor market effect (β_1), the average permanent effect of late hiring across all years a teacher is in the district, including the teacher's first year.

⁴ In all cases we cluster our standard errors at the school-grade-year level to account for the fact that students in the same school and grade are likely to share common unmeasured influences on their achievement.

Student fixed effects (δ_i) account for any time-invariant differences across students who are assigned to teachers hired late or on-time. School-by-grade-by-year fixed effects (θ_{sgt}) control for the non-random sorting of students or teachers to schools and any cohort effects – focusing our estimates on schools that had an on-time new hire and late new hire in the same grade and year. We also include a vector of teacher-year-level averages (\bar{X}_{jt}) of students' prior-year test scores and other demographic characteristics to account for classroom composition effects not captured by the student fixed effects.⁵ We also fit models where we replace student fixed effects with student-level covariates, including a cubic polynomial of students' prior-year achievement test scores in both mathematics and reading, to instead control for sorting on time-varying student measures. For both models, we report results with and without a flexible set of indicators for teacher experience to examine the degree to which difference in general teaching experience between late and on-time hires are driving any effects we find.

Additional Discussion of the Results

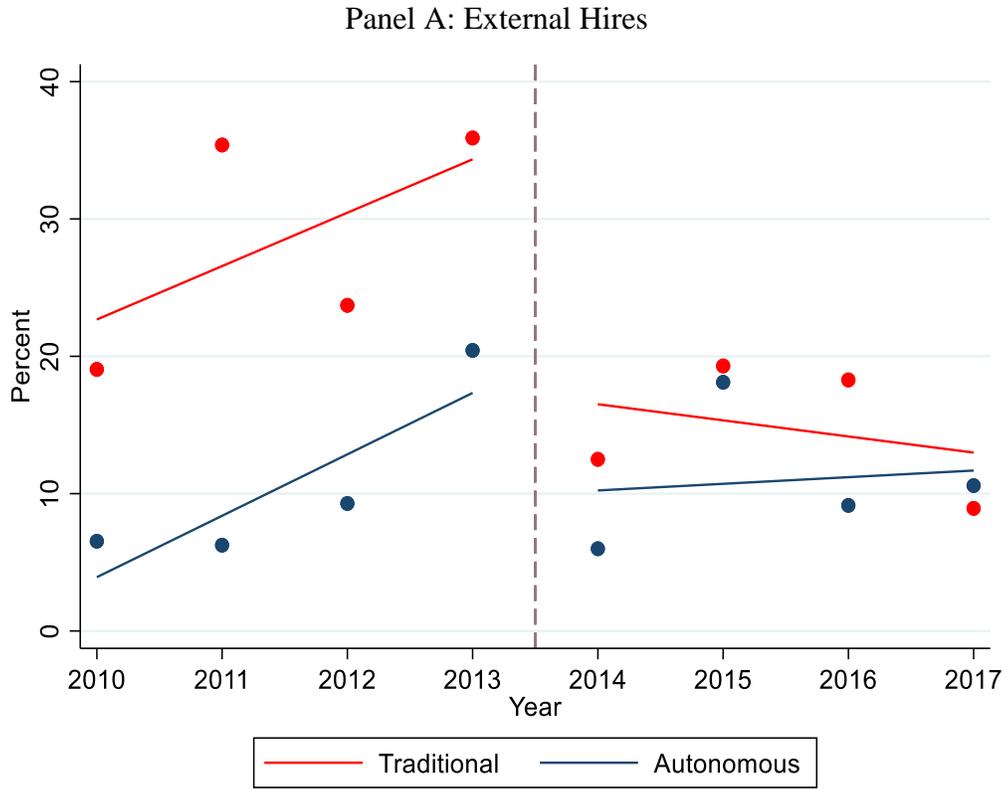
When we decompose these late hire effects into a first-year “disruption” effect and persistent “labor market” effects, we find that the negative effects are almost entirely concentrated in the first year. This suggests that in the Boston labor market teachers who are hired late suffer from temporary disadvantages caused by starting the school year late rather than being negatively selected. In prior research, Papay and Kraft (2016) found that similar negative effects in reading were also explained by a temporary disruption but that negative effects in math were the result of both disruption and negative selection in the labor market.

⁵ We also include indicators for “other” hires to ensure that we include all possible students in the district. We omit these coefficients from our tables and discussion for simplicity.

Table A1: The Effect of Late Hiring on Student Achievement

Comparison	Parameter	(1)	(2)	(3)	(4)
Panel A. Mathematics					
Late new hire vs. on-time new hire	$\beta_1 + \beta_3$	-0.068* (0.035)	-0.072** (0.034)	-0.106** (0.046)	-0.108** (0.046)
"Disruption Effect"	β_3	-0.075** (0.037)	-0.074** (0.036)	-0.091* (0.049)	-0.087* (0.049)
"Labor Market Effect"	β_1	0.007 (0.012)	0.001 (0.012)	-0.016 (0.014)	-0.021 (0.015)
n		214,555	214,555	207,697	207,697
Panel B. Reading					
Late new hire vs. on-time new hire	$\beta_1 + \beta_3$	-0.049 (0.032)	-0.059* (0.033)	-0.094*** (0.036)	-0.099*** (0.038)
"Disruption Effect"	β_3	-0.062* (0.033)	-0.069** (0.034)	-0.103*** (0.038)	-0.105*** (0.039)
"Labor Market Effect"	β_1	0.013 (0.009)	0.010 (0.009)	0.009 (0.011)	0.006 (0.011)
n		230,929	230,929	221,153	221,153
Student fixed effects		Y	Y	N	N
School-grade-year fixed effects		Y	Y	Y	Y
Student characteristics		N	N	Y	Y
Teacher-year-level averages		Y	Y	Y	Y
Teacher characteristics		N	Y	N	Y

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered by school-by-grade-by-year reported in parentheses. Teacher controls include age, a flexible function of experience, and indicators for gender, race, having a Master's degree.



Figures A1. Trends over time in the percentage of new hires that are hired late for traditional and autonomous BPS schools.

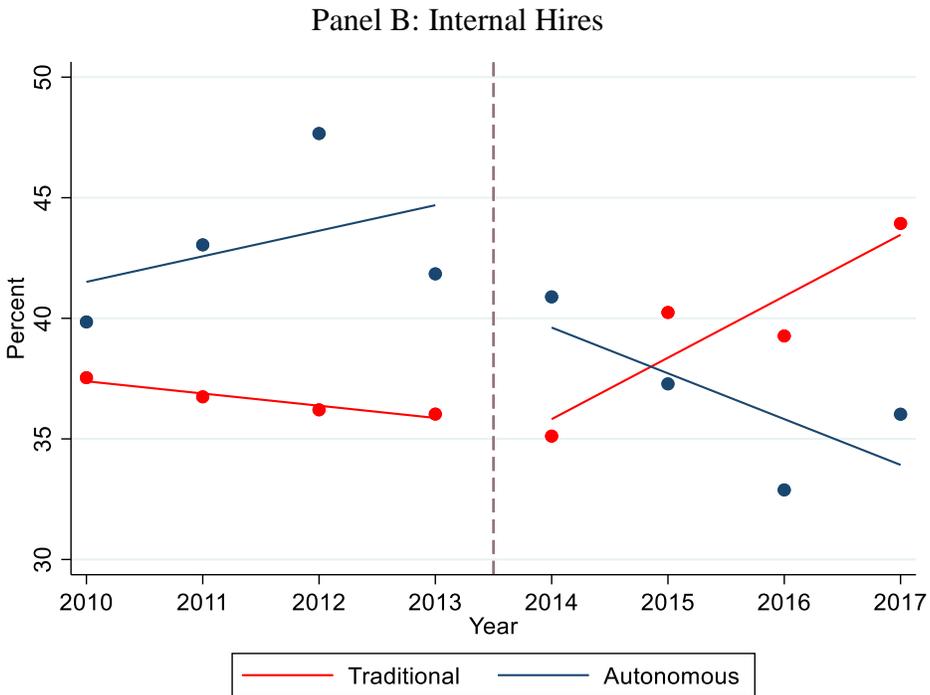
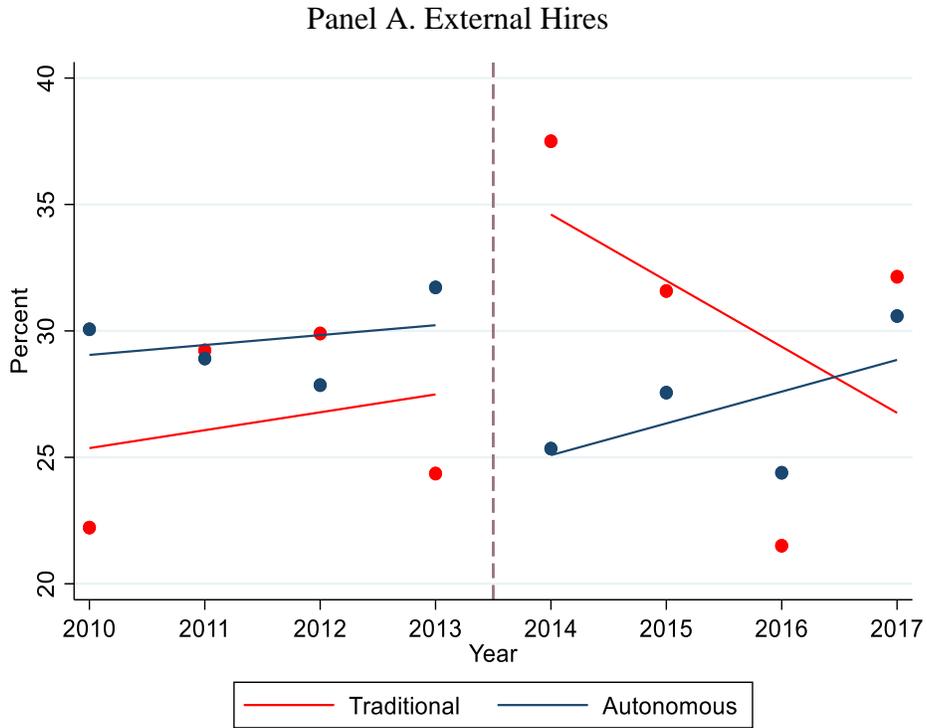


Figure A2. Trends over time in the probability a new hire is either Black or Hispanic for traditional and autonomous BPS schools

Panel A: External Hires

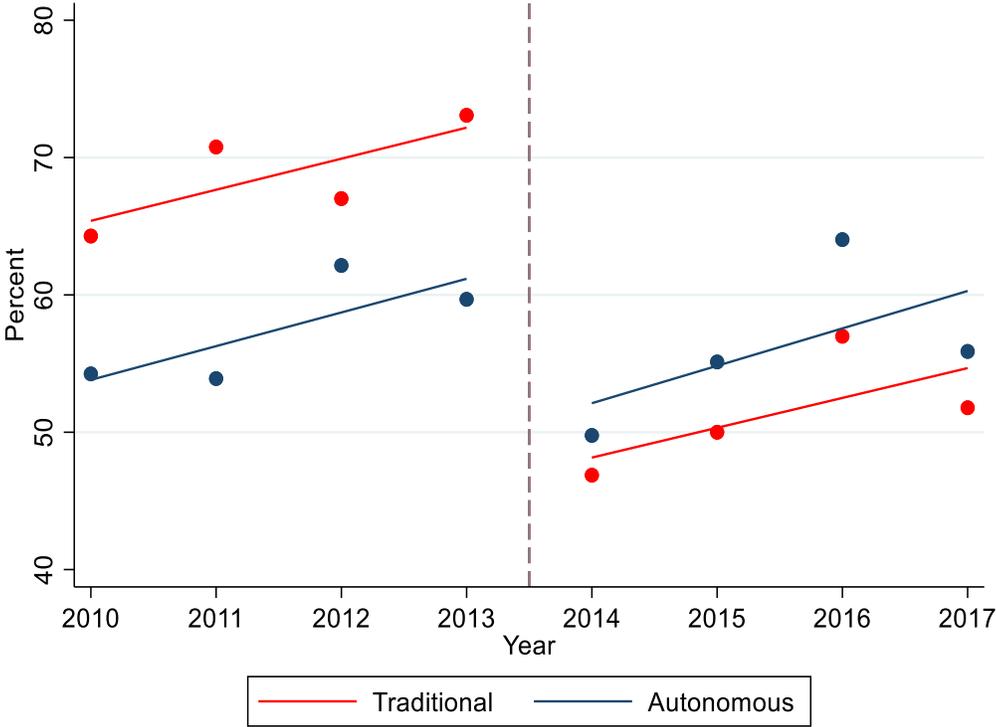


Figure A3. Trends over time in the probability a new hire is a novice teacher for traditional and autonomous BPS schools.