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

July 2020

Abstract

Public schools face substantial constraints on when and whom they can hire. We study Boston Public Schools' (BPS) mutual consent hiring reforms, which eliminated the forced placement of tenured teachers without jobs and allowed schools to conduct early, open searches rather than prioritizing internal transfers. We estimate the effect of these hiring reforms in a difference-in-differences framework, where some BPS schools enjoyed hiring autonomies prior to the reforms. The reforms moved up the median hire date by more than two months, reduced late hiring, increased new-hire diversity and retention, and improved student achievement.

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 returns to improved search and screening processes are thought to be particularly large in sectors where productivity is heterogeneous across individuals and worker-firm pairings. However, there exists little empirical evidence on the actual return to improved search practices despite their perceived importance (Oyer and Schaefer, 2011).

In this paper, we provide direct empirical evidence of the benefits of hiring reforms in the public education sector intended to improve search practices on both the extensive and intensive margins. Hiring is particularly consequential in public education, where teacher performance is largely decoupled from wages, few teachers are removed for their performance, and tenure laws provide experienced teachers with substantial job protections. The teacher labor market also provides an advantageous context in which to study hiring for several reasons: 1) administrative data allow researchers to estimate direct measures of employee productivity, 2) there exists considerable variability in productivity across individual teachers (Hanushek and Rivkin, 2010), and 3) teacher-school match quality is an important driver of productivity (Jackson, 2013).

Ideally, hiring in the teacher labor market would be a two-way matching process. Schools would select candidates who best meet their needs and teachers would select schools where they could best see themselves teaching. 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. We replicate the findings of Papay and Kraft (2016) using data in BPS to show that late hiring lowers student achievement by 0.07 standard deviations (σ) in math and 0.06σ in reading.

Starting in 2014, BPS extended hiring autonomy 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. 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 find that the BPS hiring reforms indeed accelerated hiring timelines and reduced the prevalence of late hiring in traditional BPS schools by nearly 50 percent. The reforms also had substantial positive impacts on teachers and students – decreasing turnover among new hires, increasing the diversity of new hires, and increasing student achievement in traditional BPS schools by 0.07σ to 0.09σ. These achievement impacts largely operated through improved effectiveness of new hires.

This research provides rigorous empirical evidence on the value of competitive search and hiring practices in the labor market. We build on recent studies examining districts' efforts to improve teacher hiring via the selection process and financial incentives.[1] Specifically, 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. 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

[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).

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, the school's organizational practices, and the specific position. Candidates for teaching positions are looking to maximize the utility they derive from a position and school context as well, 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.

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 also 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 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 past the end of the school year. By this time, some candidates had already secured positions in other schools and school-based personnel were not regularly available to screen and interview new applicants.

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” given that “I haven't begun 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: ensuring that principals had the flexibility to hire any teacher they believed was the best matched candidate (mutual consent) and accelerating the hiring timeline. The district aimed to "to attract and hire the most diverse, qualified and effective teachers as early as possible" through two primary initiatives: ending forced placements and allowing schools to open post positions at the beginning of the hiring cycle (Boston Public Schools, 2015, p.1). 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.

BPS substantially expanded hiring autonomy by ending the practice of administratively placing excessed tenured teachers into vacant positions. Instead, the district 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, including providing job placement supports, subsidizing training for new certifications, offering early retirement incentives, and 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.

To accelerate the hiring timeline, the district worked within the structure of the existing collective bargaining agreement. It took advantage of a rarely used provision 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. Now all principals and school site committees in BPS had the autonomy to hire the candidate of their choice among a pool that included both internal and external job candidates. 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, 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 given first-mover advantage in an uncertain hiring 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 differ very little across districts

[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.

and do not vary within 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 twelfth grade and employing over 4,000 teachers. BPS students come from diverse backgrounds: 41% of students are Hispanic, 34% Black, 14% White, and 9% Asian. Over 80% of all BPS students come from low-income households, while 30% of students are English language learners and 20% have special education needs. On average, approximately 21% of teachers leave their schools each year in BPS, a comparable within-school turnover rate

as other large urban districts (Papay et al., 2017). Public school teachers earn tenure in Massachusetts after four years of service.

We utilize a range of BPS administrative records on teachers, students, and schools across the 2006-07 to 2017-18 school years. 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 English language arts 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. 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.

We focus most of our analyses on new hires to the district, any teacher who is in her first year teaching in BPS regardless of past experience in other school districts. We identify new hires using hire dates in the HR administrative records and crosscheck that these teachers have not been employed in BPS in prior years. In Table 1, we provide descriptive statistics for various samples of new hires. Overall, we identify 3,343 new hire teachers between the fall of 2006 and 2017, or approximately 280 per year on average. Almost three out of four new teachers in BPS are female, 65% are white, 15% are Black; 11% are Hispanic, and 7% are Asian. Most new hires are novice teacher (55%) and hold master's degrees (68%).

HR records allow us to identify teachers' start dates, the day teachers officially report to the district for work. New teachers are 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. In practice, late hire teachers are those hired in September, October, and November.[3] Thus, all new hires fall into one of three categories: on time hires (hired between June and August), late hires (hired between September and November), and other hires (hired between December and March).[4]

We combine these administrative data with detailed hiring records collected by the district starting for candidates applying for open positions in 2010. These data allow us to identity 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 estimate the impact of the BPS hiring reforms in several ways. First, we estimate the impact of late teacher hiring on student achievement in BPS, documenting the potential for improved outcomes by moving up hiring timelines. Second, 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, on attrition rates for hired teachers, and on student achievement.

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

[3] Teachers hired after December 1st are considered long term substitutes by BPS

[4] We do not include teachers whose formal hire date falls between December 1 and March 1 as late hires because these teachers are more likely mid-semester replacements for teachers who leave (e.g., for maternity leave) rather than true "late hires" who fill positions that were vacant at the start of the school year.

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*.[5]

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

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.

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.[6] We also fit models where we replace student

[5] 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.
[6] 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.

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.

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. 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 teacher labor supply or student achievement. Importantly, the hiring reforms likely had both district-wide effects and effects that were concentrated in 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. However, it does not capture the full district-wide effect of the reforms. We explore the degree that the reforms had district-wide effects by plotting event studies of our outcomes of interest separately for traditional and autonomous schools.

We describe the characteristics of traditional and autonomous BPS schools in Table 2. Overall, autonomous school enroll slightly higher percentages of low-income students and students of color. Students in autonomous schools also perform meaningfully lower on the state MCAS tests in math and reading given many of these schools were granted autonomies in response to poor student performance.

We fit the following model:

$$Y_{jst} = \beta(TRAD_t * POST_t) + \bar{X}'_{st}\theta + \mu_s + \pi_t + \varepsilon_{jst} \qquad (2)$$

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.

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:

$$h\left(Turnover_{jk}\right) = \Pr\left[T_j = k \middle| T_j \geq k \; and \; \bar{X}_{st}, \mu_s, \; \pi_t\right] =$$

$$\frac{1}{1 + e^{-(\tau_k \sum_{g=1}^{g=4} 1(k=g) + \gamma_k \sum_{g=1}^{g=4} 1(k=g) * (TRAD_s * POST_t) + \bar{X}'_{st}\theta + \mu_s + \pi_t + \varepsilon_{ist})}} \qquad (3)$$

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 (2) in a student-year dataset. For a given subject, we model a student *i*'s test score, *A*, on the MCAS as follows:

$$A_{it} = \vartheta A_{i,t-1} + \alpha(TRAD_s * POST_t) + X'_{it}\phi + \bar{X}'_{jt}\lambda + \mu_s + \pi_{gt} + \varepsilon_{it} \qquad (4)$$

Here we model achievement as function of prior performance in both math and reading, $A_{i,t-1}$, as well as vectors of individual student characteristics, X, 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.

VI. Findings

Late Hiring Reduces Student Achievement

Descriptive patterns suggest teachers hired late in BPS are less diverse, leave their schools at higher rates and are less effective than their peers who are hired on-time. As shown in Table 1, teachers who are hired late are less likely to be Hispanic and turnover at much higher rates after one year (45% vs. 34%). Late hired teachers are also more likely to be novices (64% vs. 54%) and less likely to hold master's degrees (59% vs. 69%) relative to on-time hires. While we find no difference in summative evaluation ratings that do not incorporate measures of student performance, late hire that teach in tested grades and subjects have lower value-added scores in ELA (-0.10 vs. -0.02) and math (-0.03 vs. -0.01).

Our formal estimates of the effect of late hires on student achievement confirm the negative effect of starting the school year without a full-time teacher and being taught by a late-hired teacher. As shown in Table 3, we estimate that late hiring reduces student achievement between -0.07σ to -0.11σ in math and -0.05σ to -0.10σ in reading. These estimates are robust across models that include student fixed effects (Column 1) or controls for prior achievement and student characteristics (Column 3). Our estimates also remain quite similar when we include controls for teacher characteristics (Columns 2 and 4) suggesting these effects are not driven by differences in the experience, education, race, or gender of teachers who are hired late versus on time.

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.

BPS Hiring Reforms Reduce Turnover, Improve Diversity, and Increase Achievement

Hiring Timing. Graphical evidence suggests that the hiring reforms had substantial effects on the timing and characteristics of new teacher hires. 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 September. Only 58% of new hires had hire approval dates before September 1st. The hiring reforms transformed the distribution of hire approval dates, 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 and 89% were approved before September 1st.

Figure 2 shows how the hiring reforms resulted in an immediate and dramatic shift in the timing of hire approvals for both traditional and autonomous schools. We plot the average hire date across our panel for both school types. The hiring reforms benefited all schools, moving the median hire date up for traditional schools from September 8 to June 26 and for autonomous schools from August 23 to June 29. This event study graph illustrates how our DD estimates of the effect of the reforms on the timing of hiring will substantially underestimate the full, district-wide effects because the comparison group, autonomous schools, also benefited from the reforms. It also foreshadows the impacts we will see, as traditional schools used to hire several weeks after autonomous schools, on average, but closed the gap entirely after the hiring reform.

The dramatic moving up of hire dates coincide with a substantial impact on late hiring in traditional schools. We depict trends in the proportion of new teachers hired late for traditional

and autonomous schools in Figure 3. We see a substantial and immediate reduction in the prevalence of late hiring among traditional BPS schools with little change in late hiring among autonomous school. The share of teachers hired late in traditional schools dropped from 26% in the pre-period to 15% in the post period, converging with the rate of late hiring in autonomous schools across the panel.

Formal estimates of the effect of mutual consent hiring reforms from our DD models confirm these effects. In Table 4, we present estimates from our DD model both without and with controls for time-varying school characteristics. We find that the reforms moved up the hire approval dates by 17 days more in traditional schools relative to autonomous schools and reduced the frequency of late hiring by 13 percentage points. This represents a 49% reduction in late hires due to the hiring reforms.

New Hire Characteristics. Both graphical evidence and our formal DD estimates indicate the hiring reforms also served to increase both the diversity of new hires and the share of new hires with prior teaching experience. Figure 4 depict trends in the proportion of new hires who are Black and Hispanic, suggesting traditional schools were able to hire relatively more teachers of color after the reforms, while, if anything, autonomous schools hired slightly fewer. The lack of an upward shift in the proportion of new teacher hires of color in autonomous schools suggests that the district's efforts to diversify the applicant pool did not have appreciable district-wide effects. Formal estimates of the effect of hiring reforms on new hire diversity in Table 4 suggest that the reforms increased the share of new hires of color by 9.4 percentage points, a 27% increase.

We also see a sudden and sustained decrease in the proportion of new hires who had no prior teaching experience among traditional schools relative to autonomous schools in Figure 5.

Novice new hire rates dropped from 68% in the pre-period to 51% in the post period among traditional schools, with little change over time among autonomous schools. Formal DD estimates suggest a 10 percentage point reduction in novice new hires, although these estimates are not statistically significant in our model with controls.

Teacher Turnover. We estimate that the hiring reforms had large and persistent effects on the probability a new hire remained at their school. In Table 5, 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%, and by an additional 8.4% in 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 Table 5, 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.

Student Achievement. Estimates of the effect of hiring reforms on student achievement reveal the reforms led to meaningful improvements in achievement for BPS students. As shown in Table 6 Column 1, school-wide effects on academic achievement for all teachers with students in tested grades are 0.09σ in math and 0.07σ in ELA.

We next decompose these school-wide effects by estimating separate models for students taught by teachers who were newly hired into their current school during our panel (2010-2017) and those hired prior to 2010. Estimates from the first sample reflect impacts due to an

improvement in the effectiveness of newly hired teachers. Estimates from our "never new hires" sample capture any whole-school effects of the hiring reforms or spillover effects from new hires. Finally, we restrict our "ever new hires" sample to only include teachers' first year on the job to further explore the degree to which the reforms allowed schools to hire teachers who were immediately more effective. Together, these analyses shed light on the degree to which any effects of hiring reforms result from the specific teachers schools are able to hire or broader school-wide effects.

As shown in Column 2, the effect of the reforms on the productivity new hires as measured by contributions to academic achievement are large. We find effects of 0.17σ in math and 0.22σ in ELA among teachers who were "ever new hires" during this period. This suggests the reforms allowed traditional schools to hire teachers whose first-year performance was 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 evidence that the hiring reforms served to raise the productivity of veteran teachers hired before 2010. Estimate for these "never new hires" are a significant 0.11σ in math and an insignificant 0.03σ 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. Finally, in Column 4 we report estimates among our sample of ever new hires using only students taught by new hires in their first year on the job. The large positive effects of 0.20σ in math and 0.19σ in

ELA further confirm that the reforms resulted in new hires who were immediately more effective in their first year on the job compared to new hires prior to the reforms.

VII. Robustness Tests

We test the robustness of our estimated effects on student achievement to a range of potential threats to validity. First, dynamic student sorting and omitted variable bias could bias upwards our estimates. For our difference-in-differences models to provide valid impact estimates, we must assume that students are not endogenously sorting across school types because of the hiring reforms. In Table 7, we present results from a series of falsification tests that replace our achievement measures with student characteristics and prior test scores, both of which should be unaffected by the hiring reforms, and fitting a simplified DD model with school and year fixed effects. We report results for students of all teachers in the school (Panel A) as well as those taught by "ever new hire" teachers hired during our panel (Panel B).

We find no evidence that the hiring reforms affected the types of students who attended traditional schools. In Panel B, we find little evidence of differential sorting or assignment of students to newly hired teacher caused by the reforms. We do find that the hiring reform led newly hired teachers to be assigned fewer Black students and Asian students and more Hispanic students. However, the small magnitude and pattern of this differential sorting seem unlikely to drive the effects we find on student achievement. Most importantly, we find no evidence that our findings are driven by school-level or within-school sorting of students with higher levels of prior achievement. Estimated effects on students' prior achievement in both math and ELA are small and not statistically significant across both our full and "ever new hire" samples.

We further test these assumptions by refitting our DD models without controls for prior achievement, student characteristics, and classroom characteristics. Unlike value-added models that rely critically on prior achievement and other controls, our DD framework does not require these controls for identification. Dropping these controls helps to shed light on the degree to which our DD results might be biased due to dynamic sorting correlated to these observable student and classroom measures. As shown in Table 8, the magnitude and pattern of estimated effects on achievement without controls are quite similar to our preferred estimates, although, as expected, they are substantially less precise.

A second key assumption is that autonomous schools provide an appropriate estimate of the counterfactual trends in achievement for traditional schools in the absence of the hiring reforms. Here, the direction of any potential bias is unclear. Using autonomous schools to estimate a second difference could understate the district-wide effect of hiring reforms on student achievement if autonomous schools also benefitted from the hiring reforms. Event studies suggest this is at least the case for outcomes related to the timing of new hires. At the same time, autonomous schools may have been negatively affected by the reforms if they faced increased competition from traditional schools in the local labor market.

The tests described above provide some evidence that differences in student demographic or achievement trends in traditional and autonomous schools are not driving our results. We also address this concern directly by fitting an alternative specification of our DD model where we focus entirely on traditional schools. Here, we compare changes over time in the performance of new hires to teachers who were never new hires during our panel. This approach removes autonomous schools from the sample and instead uses veteran teachers within traditional schools to estimate the counterfactual trend. This provides a lower bound estimate of the effect of the

hiring reforms that guards against the potential that our effects are biased upwards due to negative treatment effects on autonomous schools caused by increased competition in the local labor market for new hires.

The effects from this alternative within-traditional-schools DD specification presented in Table 9 are positive and significant for both subjects. We find that the hiring reforms increased the achievement of students taught by "ever new hire" teachers relative to "never new hire" teachers in traditional schools by 0.07σ in math and 0.12σ in ELA, confirming that the reforms had a meaningful impact on student achievement. Together, this alternative DD approach combined with the falsification tests and results from baseline models without controls all support the conclusion that mutual consent hiring reforms in BPS had a positive causal effect on student achievement.

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 excessed 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://bmrb.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.

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

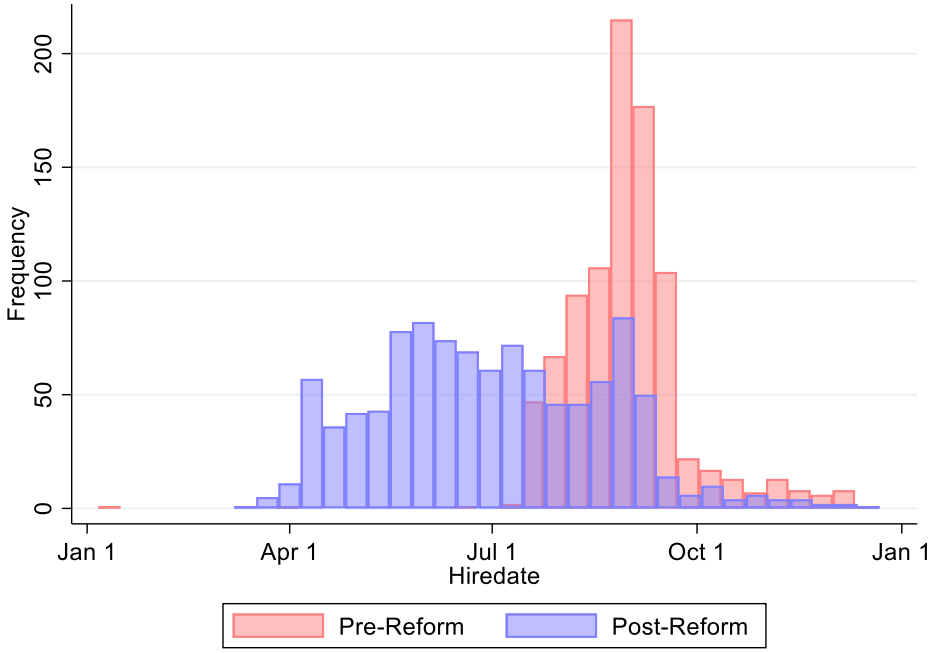


Figure 1. The distribution of hire approval dates prior to BPS mutual consent hiring reforms (2010-2013) and after the hiring reforms (2014-2017).

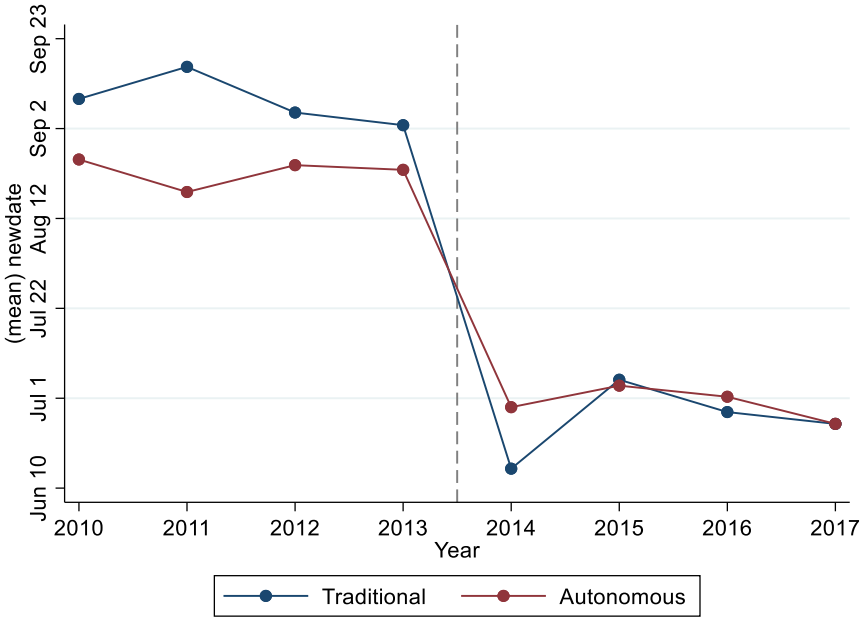
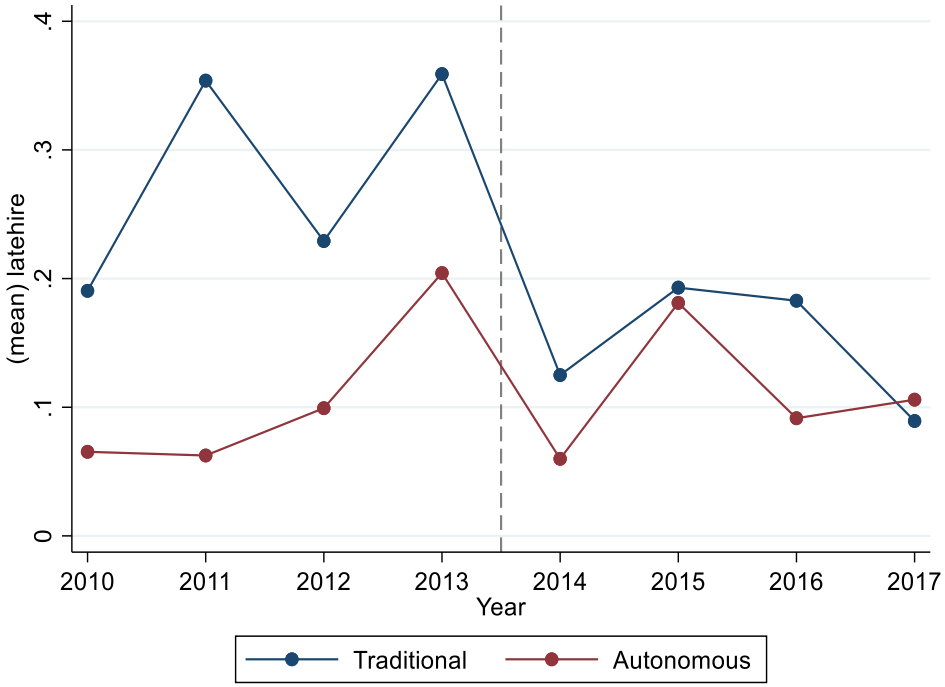


Figure 2. Trends in hire approval dates over time for traditional and autonomous BPS schools.



Figures 3. Trends over time in the percentage of new hires that are hired late for traditional and autonomous BPS schools.

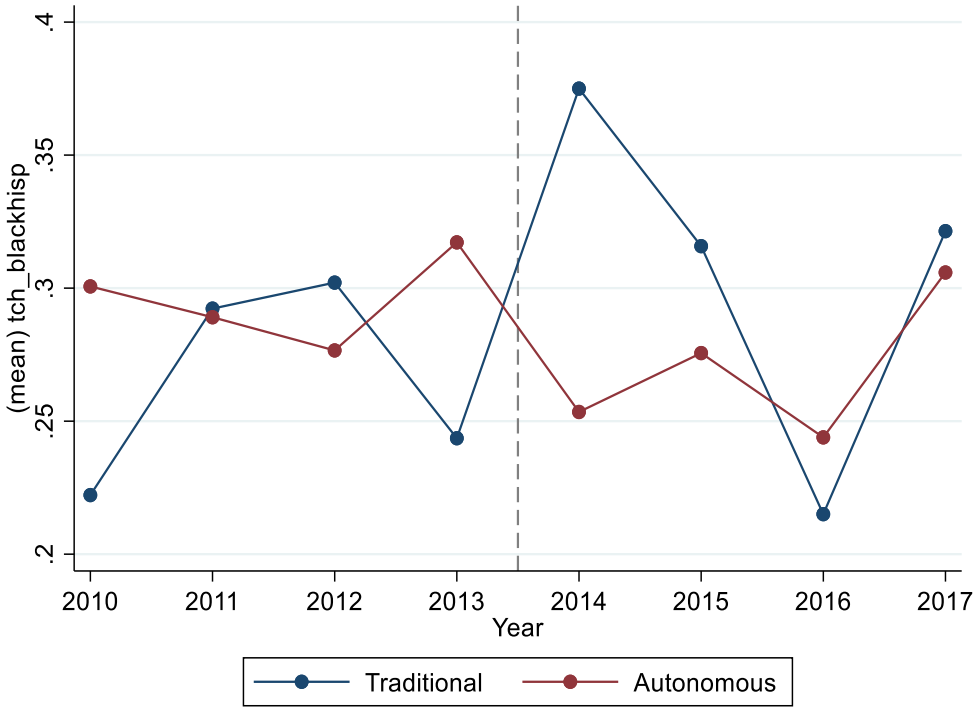


Figure 4. Trends over time in the probability a new hire is either Black or Hispanic for traditional and autonomous BPS schools

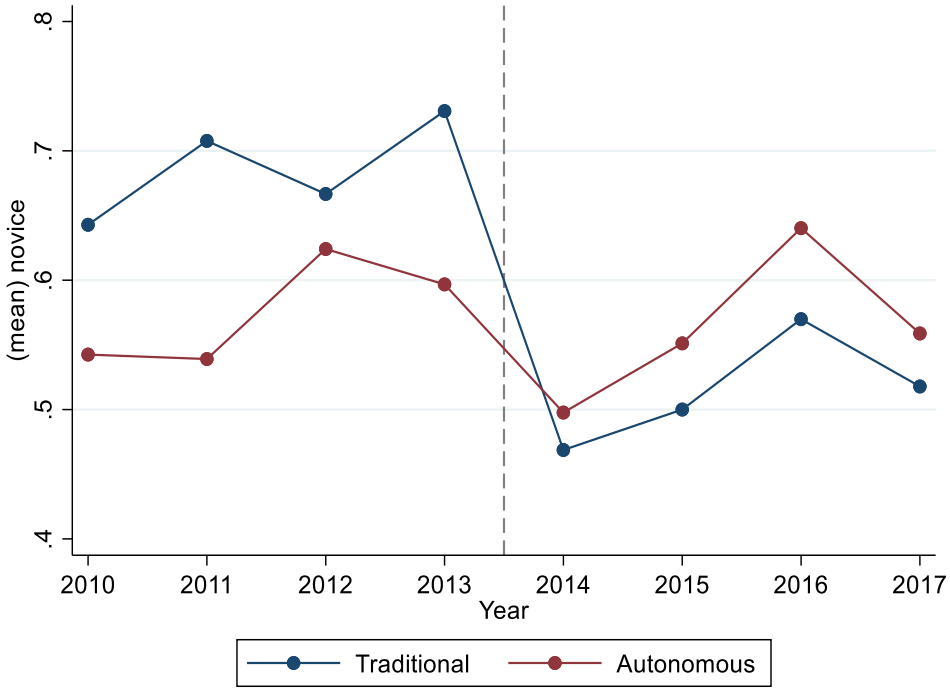


Figure 5. Trends over time in the probability a new hire is a novice teacher for traditional and autonomous BPS schools.

Tables

Table 1. New Hire Teacher Characteristics by Timing of Hire

	Late Hire Analytic Sample	Hiring Reforms Analytic Sample	On-time Hire	Late Hire (sep, oct, nov)	Difference (Late - On-time)
	2006-2017	2010-2017	2006-2017	2006-2017	
Age	30.31	29.87	29.85	32.97	3.12 ***
Female	0.72	0.73	0.73	0.70	-0.02
Black	0.15	0.16	0.14	0.16	0.02
Asian	0.07	0.07	0.07	0.08	0.00
Hispanic	0.11	0.11	0.11	0.08	-0.04 ***
White	0.65	0.61	0.64	0.68	0.03
Novice	0.55	0.57	0.54	0.64	0.11 ***
Masters Degree	0.68	0.67	0.69	0.59	-0.10 ***
Late Hire	0.15	0.16	0.00	1.00	
Turnover after first year	0.36	0.37	0.34	0.46	-0.11 ***
Evaluation Score	2.91	2.91	2.91	2.90	-0.02
Math VAM	-0.02	0.01	-0.01	-0.03	-0.02
ELA VAM	-0.03	-0.02	-0.02	-0.10	-0.08 **
n	3,343	2,182	2,856	487	

Notes: * p<0.1, ** p<0.05, *** p<.01. Sample sizes for hire appoval date, evluation score, and value-added estimates are smaller due to missing data.

Table 2: Student Characteristics Across Traditional and Autonomous Schools in 2013

	All	Traditional	Autonomous	Difference (Trad-Auto)
	(1)	(2)	(3)	(4)
Special Education	0.20	0.19	0.23	-0.05 *
Low Income	0.81	0.79	0.86	-0.07 *
African American	0.36	0.32	0.45	-0.13 ***
Asian	0.09	0.11	0.05	0.06 **
Hispanic	0.40	0.39	0.40	-0.01
White	0.13	0.15	0.08	0.07 **
English language learner	0.25	0.26	0.24	0.01
Math Score (std)	0.01	0.09	-0.19	0.28 **
Reading Score (std)	-0.01	0.07	-0.21	0.28 **
n (students)	47,108	33,586	13,825	
n (schools)	121	84	37	

Notes: * p<0.1, ** p<0.05, *** p<.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 3: 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*	-0.072**	-0.106**	-0.108**
		(0.035)	(0.034)	(0.046)	(0.046)
"Disruption Effect"	β_3	-0.075**	-0.074**	-0.091*	-0.087*
		(0.037)	(0.036)	(0.049)	(0.049)
"Labor Market Effect"	β_1	0.007	0.001	-0.016	-0.021
		(0.012)	(0.012)	(0.014)	(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	-0.059*	-0.094***	-0.099***
		(0.032)	(0.033)	(0.036)	(0.038)
"Disruption Effect"	β_3	-0.062*	-0.069**	-0.103***	-0.105***
		(0.033)	(0.034)	(0.038)	(0.039)
"Labor Market Effect"	β_1	0.013	0.010	0.009	0.006
		(0.009)	(0.009)	(0.011)	(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<.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.

Table 4: Difference-in-Difference Estimates of The Effect of Hiring Reform on New Hired Teachers

	Traditional School Pre-Reform Mean	(1)	(2)
		Hire Approval Date	
Traditional*Post	September 8th	-20.607***	-16.718***
		(5.381)	(5.028)
n		1826	
		Late Hire	
Traditional*Post	0.266	-0.133***	-0.130***
		(0.044)	(0.048)
n		2052	
		Teacher of Color	
Traditional*Post	0.352	0.105**	0.094*
		(0.049)	(0.052)
n		1984	
		Novice Teachers	
Traditional*Post	0.679	-0.110*	-0.099
		(0.064)	(0.066)
n		2052	
School Characteristics		N	Y

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

Table 5: Hazard Difference-in-Difference Estimates of The Effect of Hiring Reforms on Teacher Turnover

	(1)	(2)	(3)
Traditional*Post*(k=1)	-0.124***	-0.098**	-0.093**
	(0.039)	(0.039)	(0.041)
Traditional*Post*(k=2)	-0.111**	-0.084*	-0.150***
	(0.049)	(0.049)	(0.052)
Traditional*Post*(k=3)	-0.070	-0.042	-0.029
	(0.064)	(0.063)	(0.062)
Traditional*Post*(k=4)	0.000	0.027	-0.076
	(0.076)	(0.075)	(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<.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 preferly predicted.

Table 6: The Effect of Hiring Reforms on Student Achievement

	All Teachers	Ever New Hires	Never New Hires	New Hires in First Year
	(1)	(2)	(3)	(4)
			Panel A: Math	
Traditional*Post	0.090***	0.174***	0.111**	0.204**
	(0.034)	(0.047)	(0.042)	(0.086)
n	127,899	25,806	102,092	5,575
			Panel B: ELA	
Traditional*Post	0.068**	0.218***	0.034	0.191**
	(0.028)	(0.044)	(0.036)	(0.079)
n	139,659	31,285	108,373	7,219
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 7: Falsification Tests of the Effect of Hiring Reforms on Student Characteristics

	Special Education	Low Income	African American	Asian	Hispanic	White	LEP	Prior Math Scores	Prior ELA scores
					Panel A: All Teachers				
Traditional*Post	0.005	-0.002	-0.010	0.001	0.012	-0.003	0.002	0.002	-0.010
	(0.007)	(0.010)	(0.010)	(0.003)	(0.012)	(0.008)	(0.013)	(0.035)	(0.032)
n	823,370	823,370	822,365	822,365	822,365	822,365	822,280	339,818	336,040
					Panel B: Ever New Hires				
Traditional*Post	-0.009	0.011	-0.044*	-0.024*	0.065**	0.000	-0.035	-0.039	-0.005
	(0.036)	(0.020)	(0.023)	(0.013)	(0.029)	(0.013)	(0.033)	(0.089)	(0.069)
n	35,493	35,493	35,453	35,453	35,453	35,453	35,452	15,763	15,628

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

Table 8: The Effect of Hiring Reforms on Student Achievement without Controls

	All Teachers	Ever New Hires	Never New Hires	New Hires in First Year
		Panel A: Math		
Traditional*Post	0.093	0.292***	0.114**	0.305**
	(0.062)	(0.087)	(0.057)	(0.150)
n	128,054	25,824	102,229	5,578
		Panel B: ELA		
Traditional*Post	0.046	0.259***	0.059	0.071
	(0.053)	(0.063)	(0.074)	(0.122)
n	139,722	31,288	108,433	7,220
School FE	Y	Y	Y	Y
Student Characteristics	N	N	N	N
Prior Achievement	N	N	N	N
Classroom Characteristics	N	N	N	N

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 9: The Effect of BPS Hiring Reforms on Student Achievement using only Traditional Schools

	Traditional School
	Panel A: Math
Traditional*Post	0.067**
	(0.033)
n	96574
	Panel B: ELA
Traditional*Post	0.117***
	(0.025)
n	106025
School FE	Y
Student Characteristics	Y
Prior Achievement	Y
Classroom Characteristics	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.