Journal of Public Economics 166 (2018) 81–97

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Teacher applicant hiring and teacher performance: Evidence from DC public schools☆,☆☆

Brian A. Jacob [a], Jonah E. Rockoff [b,*], Eric S. Taylor [c], Benjamin Lindy [d], Rachel Rosen [e]

[a] University of Michigan, 735 South State Street, Ann Arbor, MI 48109, United States of America
[b] Columbia Business School, 3022 Broadway #603, New York, NY 10027, United States of America
[c] Harvard Graduate School of Education, Gutman Library 469, 6 Appian Way, Cambridge, MA 02138, United States of America
[d] Teach For America, 1110 Main St., Cincinnati, OH 45202, United States of America
[e] MRDC, 19th Floor, 16 East 34 Street, New York, NY 10016, United States of America

ARTICLE INFO

Article history:
Received 1 December 2017
Received in revised form 22 August 2018
Accepted 22 August 2018
Available online xxxx

JEL classification:
I2
J2
M51

Keywords:
Teachers
Hiring
Screening
Job performance

ABSTRACT

Selecting more productive employees among a pool of job applicants can be a cost-effective means of improving organizational performance and may be particularly important in the public sector. We study the relationship among applicant characteristics, hiring outcomes, and job performance for teachers in the Washington DC Public Schools. Applicants' academic background (e.g., undergraduate GPA) is essentially uncorrelated with hiring. Screening measures (written assessments, interviews, and sample lessons) help applicants get jobs by placing them on a list of recommended candidates, but they are only weakly associated with the likelihood of being hired conditional on making the list. Yet both academic background and screening measures strongly predict teacher job performance, suggesting considerable scope for improving schools via the selection process.

> *"The best means of improving a school system is to improve its teachers. One of the most effective means of improving the teacher corps is by wise selection."*
>
> [Ervin Eugene Lewis, Superintendent of Schools, Flint, Michigan, 1925]

1. Introduction

The importance of employee selection is widely recognized by practitioners and researchers alike. While a large literature in psychology has explored the power of applicant screening to identify successful employees (see, for example, McDaniel et al., 1994), economists have paid far less attention to this issue. As noted by Oyer and Schaefer (2011), in contrast to economists' work on understanding employee incentives, "the literature has been less successful at explaining how firms can find the right employees in the first place." This has recently begun to change, with papers on the role of personal referrals (Schmutte, 2015; Burks et al., 2015; Brown et al., 2016), placement agencies (Stanton and Thomas, 2016), objective screening technologies (Hoffman et al., 2015), and recruitment messaging (Ashraf et al., 2016) in the hiring process.

There are several reasons why employee selection is particularly important in the field of public education. First, there is substantial variance in teacher effectiveness, and good teachers have positive impacts on long-term student outcomes (see Koedel et al., 2015 for a review). Second, there are substantial financial and political costs to

☆ We first thank the District of Columbia Public Schools, in particular Michael Gaskins, Anna Gregory, Brooke Miller, Jason Kamras, and Scott Thompson. Generous financial support was provided by the Smith Richardson Foundation, grant # 2011-8709. We received helpful comments and suggestions from seminar participants at Brown, Chicago, Clemson, Cornell, Delaware, Johns Hopkins, Kentucky, LSU, New York Fed, NYU, Northwestern University, Paris School of Economics, Princeton, Stanford, Texas A&M, UC Santa Barbara, Universidad Nacional La Plata, Universidad San Andres, University of Warwick, APPAM, and AEFP.

☆☆ The authors of this publication were consultants to the District of Columbia Public Schools. The terms of this relationship and this publication have been reviewed and found to be in accordance with the DCPS policy on objectivity in research by the Office of Talent and Culture and by the Office of Instructional Practice District of Columbia Public Schools.

* Corresponding author.
E-mail addresses: bajacob@umich.edu (B.A. Jacob), jonah.rockoff@columbia.edu (J.E. Rockoff), eric_taylor@harvard.edu (E.S. Taylor), rachel.rosen@mdrc.org (R. Rosen).

https://doi.org/10.1016/j.jpubeco.2018.08.011
0047-2727/

removing teachers who perform poorly on the job.[1] Third, in most parts of the country and in most subject areas there is an abundance of *potential* teachers (see Greenberg et al., 2013), yet most prior research suggests that school systems are not very good at selecting the individuals most likely to become successful teachers (see Ballou, 1996; Kane and Staiger, 2005; Harris et al., 2014; Hinrichs, 2014).

Only one study, in addition to ours, provides solid evidence on whether pre-employment measures from an actual job application can predict successful teaching.[2] Goldhaber et al. (2017) examine teacher applicants in Spokane Washington, where applications contain background information (education, experience, licensure), recommendation letters, and narrative statements. District personnel do an initial screen, then school-specific personnel fill out more detailed evaluations on candidates of interest. They find that teachers with higher rated applications (at either stage) have lower attrition rates, and they find a significant positive relationship between the school-based evaluation of teacher candidates and subsequent value-added in math (but not English).[3]

While these results are encouraging, the importance of the issue merits much further study. To that end, we use uncommonly detailed data on applications, job offers, employment, and performance of teachers in the Washington, DC Public Schools (hereafter DCPS) to make several contributions to the literature on employee selection in public schools.

First, we present an analysis of the hiring process. Previous work on labor demand for teachers has been limited, in contrast to the large literature on teacher labor supply, where it is often assumed (sometimes implicitly) that employment outcomes stem from teachers' choices, not those of school or district administration. We provide evidence on the extent to which schools use the information collected centrally by the district, and investigate whether applicants' characteristics are related to the types of schools where they receive job offers and are hired. Using data on both job offers and hiring outcomes provides additional support for the idea that our analysis is capturing demand-side factors.

Second, our analysis focuses on a broader measure of teacher performance than previous work. Our performance metric is based largely on classroom observations of teacher instruction and interaction with students, although it also incorporates a variety of other inputs such as supervisor ratings, student work, and (when available) value-added to students' standardized test scores. We also examine performance measured by test-score value-added alone, but our statistical power is limited.

The prevalence and high-stakes use of observation-based measures alone make them an important subject of inquiry. Over the past decade, new rubric-based classroom observation evaluations have been introduced by nearly all states (at least 46) and the nation's 25 largest districts (Steinberg and Donaldson, 2016). These new observations are structured and scored using detailed rubrics, and generate considerable variation.[4]

Beyond their growing use in educational policy, classroom observation scores can capture important variation in teacher job performance, making them an informative measure of teacher quality. While value-added can be calculated for only about one-third of teachers in certain grades and subjects, all teachers can be measured with observations, and there is growing evidence that classroom observations do predict student achievement gains. Several studies report positive correlations (about 0.2) between observation scores and student test scores or teacher value-added (see, for example, Milanowski, 2004, Kane et al., 2011, Kane and Staiger, 2011, and Grossman et al., 2014) and we find similar correlations in our DCPS data.[5] Recent studies that randomly assign teachers to classes can rule out explanations due to within-school student sorting (see, for example, Kane et al., 2013, Garret and Steinberg, 2015, Araujo et al., 2016, Bacher-Hicks et al., 2017).[6]

A teacher's job includes responsibilities beyond those reflected in test scores, and classroom observation scores also partly reflect non-test score student outcomes. Blazar and Kraft (2017) show that observation measures predict students' self-assessment of their math ability, happiness in class, and behavior in class. Similarly, Araujo et al. (2016) find observation scores predict executive functioning skills among kindergarten students.[7] In this sense our observation-based outcome is broader in the scope of teaching job responsibilities measured, although it is not a direct measure of student learning.

In our study, we examine the relationship between the teacher performance measures in DCPS and applicant characteristics. These include "traditional" measures of applicant quality (e.g., SAT score, college GPA) as well as measures based on candidates' writing, interviews, and auditions. To address potential bias from selection into hire, we exploit idiosyncratic features of the DCPS hiring process that create discontinuities between applicant scores and hiring/job offers, and test the robustness of our findings to non-random sorting of teachers using specifications with school fixed effects.

Several interesting findings emerge. First, the district's less-traditional screening measures are strong predictors of teacher performance. Second, we find that several academic background characteristics (e.g., undergraduate GPA) also strongly predict subsequent teacher performance. Pooling all of these measures to create an index of predicted performance, we find the actual performance of "top quartile" hires is more than two-thirds of a standard deviation (0.71σ) higher than those from the bottom quartile.[8]

[1] Barnes et al. (2007) estimate that turnover costs districts roughly $10,000 per teacher. Staiger and Rockoff (2010) illustrate the academic cost of exposing students to newly hired but ineffective teachers, and Rothstein (2015) highlights the cost of compensating teachers for increasing the risk of job separation. In addition, collection of performance measures on teachers (e.g. standardized student testing, classroom observation, portfolios of student work) requires significant public resources and often entails difficult labor negotiations (e.g., Baker and Santora, 2013), while schools and school districts have wide freedom in deciding what information to collect from applicants as part of the hiring process. Issues over teacher removal have also been the subject of major lawsuits (Treu, 2014).

[2] Other studies that do speak to the issue of teacher selection suffer from important shortcomings stemming from the use of measures collected in low-stakes research surveys (Rockoff et al., 2011) or administrative data unavailable to schools and school districts (Boyd et al., 2008). Moreover, these studies only examine data on teachers who are already hired, rather than data on a pool of applicants, preventing them from addressing issues of selection. There is also a large literature outside economics that has studied the teacher hiring process, but this work is often qualitative in nature or relies on small samples with limited measures of teacher performance (see Appendix B for a more comprehensive discussion of this work).

[3] The estimated relationship between district screening scores and math value added is considerably smaller than the estimate for math and school-based scores, and is not statistically significant. Note that the authors also address potential biases from selective hiring by relying on arithmetic errors made in computing applicant scores.

[4] Weisberg et al. (2009) provide evidence that "older generation" evaluations based on classroom observation that were more holistic often offer limited variation, with large fractions of teachers earning top scores. While final scalar evaluation scores (in DC and elsewhere) are also somewhat coarse, the raw data we use are collected in multiple observations, by multiple observers, scoring multiple practices, and provide meaningful variation.

[5] The correlation between a teacher's value-added score in year t and her classroom observation score in year t is 0.27. The correlation between value-added in t and mean observation score in years other than t is 0.25.

[6] As further evidence against sorting bias, Bacher-Hicks et al. (2017) find that a teacher's observation score when students are assigned naturally is an unbiased predictor of the teacher's score when students are assigned randomly. White (2018) finds a similar result, as does Kane et al. (2013). Related evidence comes from Cantrell et al. (2008), who experimentally test the link between test scores and teachers' ratings by the National Board for Professional Teaching Standards; these ratings are based on a portfolio of teacher work including recorded lessons and tests of content knowledge. Additionally, there is some evidence that the process of being evaluated with classroom observations can itself improve teacher value-added to student test scores (Taylor and Tyler, 2012).

[7] Other work has shown that teachers have effects on various other non-cognitive skills (Jackson, 2016; Petek and Pope, 2017). Given the evidence cited above, it seems likely that observation-based measures will predict these outcomes as well.

[8] To underscore the large magnitude of this finding, one can compare it to the average on-the-job improvement exhibited by DCPS teachers over their first three years working at DCPS, a period in teachers' careers when performance has been consistently shown to improve rapidly (e.g., Rice, 2013; Ost, 2014; Papay and Kraft, 2015). Among the new teachers in our sample who remain in DCPS for three years, their average three-year growth in performance is 0.37 standard deviations, roughly half of the difference in performance between top- and bottom-quartile applicants entering DCPS.