Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement

John P. Papay *, Matthew A. Kraft

Brown University, 340 Brook Street, Box 1938, Providence, RI 02912, USA

ABSTRACT

We present new evidence on the relationship between teacher productivity and job experience. Econometric challenges require identifying assumptions to model the within-teacher returns to experience with teacher fixed effects. We describe the identifying assumptions used in past models and in a new approach that we propose, and we demonstrate how violations of these assumptions can lead to substantial bias. Consistent with past research, we find that teachers experience rapid productivity improvement early in their careers. However, we also find evidence of returns to experience later in the career, indicating that teachers continue to build human capital beyond these first years.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Over the past decade, efforts to improve the elementary and secondary education system in the United States have focused on ensuring that all students have an effective teacher in their classroom. The debates over how to accomplish this goal have been increasingly informed by teacher effectiveness research that has blossomed in recent years with the availability of large-scale datasets that link teachers to students and test scores. These data have allowed researchers to examine central questions about the teacher labor market, including productivity dynamics—in other words, how do teachers improve their effectiveness over the course of their careers?

The extent to which teacher performance in the classroom changes with experience has both theoretical and practical implications. Better understanding this dynamic will shed light on the relationship between employee productivity and job experience, and also inform current education policy initiatives such as teacher pay, evaluation, retention, and tenure. Many analyses of the relationship between teacher experience and productivity have relied on cross-sectional data, comparing the effectiveness of teachers at different experience levels. However, this comparison does not provide a clear picture of how teachers improve over the course of their careers, largely because it ignores the issue of attrition. Even if teachers do improve with experience, we can find flat returns to experience in the cross-section if the most effective teachers leave. Thus, the extent of within-teacher returns to experience provides more relevant guidance to policymakers about teacher improvement throughout the career.

For much of the past decade, this question has been treated as settled (Rice, 2013; TNTP, 2012). Policymakers and researchers tend to believe that teachers improve rapidly during their initial years in the classroom, but that the returns to experience flatten out after the first few years of teaching. These results have become quite influential in the policy community. However, two recent papers in this journal find otherwise, providing evidence that teachers continue to improve over the course of their careers (Harris and Sass, 2011; Wiswall, 2013).[1]

In the first half of our paper, we reconcile these divergent results by laying out explicitly the identifying assumptions that researchers have used in estimating the within-teacher returns to experience (with teacher fixed effects), given the collinearity between experience and year for nearly all teachers. We demonstrate analytically and through simulation how violations of each assumption can bias estimates, sometimes substantially. We also propose a new approach that relies on a substantively different assumption and, thus, is subject to a different source of bias. In the second half, we use data from a large urban school district to present estimates of the within-teacher returns to experience from these different models. Examining estimates from models that rely on distinct identifying assumptions provides a clearer picture of the biases in each approach and enables us to present stronger evidence about the extent of later-career returns to experience.

* Corresponding author. Tel.: +1 401 863 5137.
   E-mail addresses: john_papay@brown.edu (J.P. Papay), mkraft@brown.edu (M.A. Kraft).

[1] Given that "tenure" and "seniority" have specific meanings in the field of education, we use the term "experience" to reflect the number of years a teacher has been in the profession.

Like past researchers, and consistent with theory, we find that teachers in the district improve most rapidly at the beginning of their careers. However, across models, we find that teachers continue to improve, albeit at lesser rates, past their first five years in the classroom. We also find suggestive evidence of continued returns to experience throughout the career, particularly in mathematics. These results make sense, as labor economists have long observed that employee wages continue to rise with job experience. Human capital theory supports this pattern, holding that workers build skills that translate to greater productivity (Becker, 1993). Taken together, our results suggest that the question of whether teachers continue to improve with experience is at least not settled and that policymakers should temper their policies to acknowledge this reality.

In the next section, we describe past efforts to estimate the productivity returns to teaching experience. In Section 3, we describe our dataset and measures. We then articulate the key assumptions that underlie existing approaches, propose an alternative method, and discuss the bias introduced by each approach. In Section 5, we present the estimated returns to teacher experience from each of these approaches in our data. We describe several threats to the validity of our inferences and our attempts to address them in Section 6. Finally, we conclude with a discussion of the economic and educational implications of this work.

2. Estimates of the returns to experience in teaching

The education sector is among the few industries for which direct estimates of worker productivity are available for much of the labor force. In recent years, education economists have produced a growing body of literature that examines the productivity returns to job experience among teachers, using estimated contributions to student test score gains as a proxy for productivity (see Todd and Wolpin, 2003; McCaffrey et al., 2004; Harris and Sass, 2006). We focus on all aspects of productivity improvement (as measured by student test scores) that accrue to teachers over their careers—in other words, we seek to estimate the overall effect of experience on productivity, rather than disentangling the reasons for these returns.[2] Thus, we include as "returns to experience" the effects of formal on-the-job training, informal on-the-job learning, out-of-work training (such as formal education) and any other factors that improve teacher effectiveness over time.

Most research suggests that teachers improve a great deal at the beginning of their careers (e.g., Rockoff, 2004). Fast early-career improvement in productivity is not surprising, given that theory implies more rapid human capital development and greater investment earlier in the career (Becker, 1993). This pattern mirrors theories of the teacher career arc, where novice teachers are often characterized as simply trying to survive in the classroom as they build key classroom management skills, learn the curriculum, and add to their instructional abilities (Johnson and the Project on the Next Generation of Teachers, 2004). Many factors contribute to the extent of early-career productivity growth, including the availability of effective colleagues (Jackson and Bruegmann, 2009), consistency in teaching assignments (Ost, 2014), and supportive work environments (Kraft and Papay, 2014).

However, there is less agreement about the nature of returns to experience after these early years. On one hand, shirking models suggest that teachers, who face minimal oversight and enjoy strong job protections, may stop improving once they become established in their schools (Hansen, 2009). On the other, some theories of teacher career development suggest that, beyond their first few years, teachers may continue to refine their practice and gain the relationships and time to collaborate with colleagues about instruction (Huberman, 1992). Recent evidence suggests that veteran teachers can improve their instructional effectiveness if they participate in a rigorous teacher evaluation program (Taylor and Tyler, 2012), find more productive school matches (Jackson, 2013), or engage in effective on-the-job training (e.g., Matsumura et al., 2010; Neuman and Cunningham, 2009; Powell et al., 2010; Allen et al., 2011).

As Murnane and Phillips (1981) made clear, cross-sectional estimates cannot fully distinguish between true individual returns to job experience and vintage effects (i.e., average differences in quality across teacher cohorts) or selection effects (i.e., differential attrition). We focus on this question by estimating the within-teacher returns to experience using longitudinal data with teacher fixed effects. This line of work builds on Rockoff's (2004) analysis of data from two school districts in New Jersey. Rockoff finds substantial early-career returns to teaching experience, particularly on reading test scores, but the returns to experience on all but reading comprehension scores diminish rapidly after the first few years in the classroom. More recently, Boyd et al. (2008) have applied Rockoff's general approach to examine data in New York City and North Carolina, respectively, finding qualitatively similar results.

These cross-sectional and longitudinal findings have been widely interpreted as evidence that teachers do not improve their performance beyond their first few years in the classroom (Rivkin et al., 2005). This interpretation has had a profound effect on education policy. For example, Bill Gates (2009) asserted that "once somebody has taught for three years, their teaching quality does not change thereafter." However, recent evidence suggests that teachers may improve throughout their careers. Using data from Florida, Harris and Sass (2011) find that while the largest gains in experience accrue in the first few years, there are "continuing gains beyond the first five years of a teacher's career" (p. 1). Using data on 5th grade teachers in North Carolina, Wiswall finds that "teaching experience has a substantial and statistically significant impact on mathematics achievement, even beyond the first few years of teaching" (2013, p. 62), although he finds no such returns in reading. We seek to resolve this divergent evidence by examining these approaches in more detail.

3. Dataset and measures

3.1. Dataset

In order to examine within-teacher returns to experience, we use a comprehensive administrative dataset from a large, urban school district in the southern United States that includes student, teacher, and test records from the 2000–01 to the 2008–09 school years. This district has over 100,000 students and nearly 9000 teachers. Student data include demographic information, teacher–student links, and annual state test results in reading and mathematics. We standardize these test scores to interpret our estimates as standard deviation differences in student performance.[3] Because appropriate estimation of the education production function requires both baseline and outcome test data, we focus on teachers in grades four through eight. We exclude any students in atypically small classes or substantially separate special

---

[2] There are both substantive and practical reasons for this. Substantively, we are interested in understanding how teachers improve over the course of their careers on average. Different teachers may take different paths to such improvement. Practically, many of these elements are notoriously difficult to measure. For example, in-school professional development can take many forms, only some of which are recorded. Formal education can be captured in aggregate, such as whether teachers earn a masters' degree, but we cannot distinguish finer-grained course-taking. As such, we focus on the broader question of whether teachers improve their productivity throughout their career. We find nearly identical returns to experience when we condition on teachers' formal education.

---

[3] Note that this standardization does not make the scales comparable from year to year because of differences in tested material and changes in the distribution of student ability from year-to-year. However, the test measure we use does not have a vertical scale that enables inferences about student growth from year-to-year.