Background

In most school human resources policies across the world, teacher human capital characteristics  - such as education background and teaching experience - constantly play a central role. This is likely due to the fact that teacher characteristics are relatively accurate and straightforward measures, which are practical and computational cheaper than direct teacher quality measures such as value-added models, classroom observations, and student/parent surveys (Rice, 2013). Either explicitly or implicitly, however, these characteristics-based policies assume that education attainment and experience improve a teacher’s knowledge, skills, and productivity, which is not strongly supported by scientific research. In fact, research consistently shows that human capital measures frequently used in teacher evaluation and compensation explain little of a teacher's contribution to student academic growth (Aaronson et al., 2007).

The answer to whether teacher education attainment and experience have causal impact on student learning requires greater nuance and more accurate estimates. The relationships between teacher characteristics and student performance are rarely linear: whereas some researchers detected no relationship between teacher’s education attainment (Aaronson et al., 2007) or experience (Aaronson et al., 2007) and student performance, some researchers found evidence that teachers who obtained graduate degree (Guarino et al., 2013) improve student learning and the impact of experience is strongest in a teacher’s early career (Papay and Kraft, 2015). Moreover, research shows that the relations between teacher experience and student learning vary across levels of education (Harris and Sass, 2011) and subject areas (Clotfelter et al., 2007). Lastly and most importantly, the relationships observed between teacher characteristics and change in student learning are unstable because they are biased by the fact that higher performing and more motivated students often have more access to teachers with higher human capital profile.

Significance

Our study is among a rigorous literature body that contributes to the understanding of the causal impact of teacher education attainment and experience on student academic outcomes. We make at least three efforts to obtain accurate and precise estimates of teacher effects. First, we effectively address threats to internal validity by leveraging a random assignment of teachers to students that is enforced by a national policy and confirmed in a nationally representative, longitudinal, student-level data from China. Second, given random assignment, we are able to use between-teacher variation to estimate teacher effects without having to make strict assumptions like traditional teacher effect studies do when they rely on within-teacher variations (see Ladd and Sorensen, 2017; Papay and Kraft, 2015). Third, we model teacher education attainment and experience in different formats (continuous and categorical) to understand more beyond their linear effects on student outcomes. Fourth, we control for school fixed-effects to account for systematic difference across schools and a rich set of student-, homeroom-, and teacher-level covariates to improve estimation precision, here, the most important to note is the cubic polynomial functions of four scores (in three core content subjects and cognitive test) in prior year that effectively absorb noise from individual learning ability.

Research Questions

Specifically, we answer two research questions:

1. Whether and to what extent teacher education attainment impacts student academic performance and confidence?
2. Whether and to what extent teacher teaching experience impacts student academic performance and confidence?

Data and Measures

We drew our sample from China Education Panel Survey (CEPS), China’s first nationally representative, longitudinal data of middle school students (more details in Appendix B Data Description). We focus on the baseline 7th grade students who were successfully followed up in 2014-15 (follow-up rate 91.93%) and restrict the sample to be in 63 public schools that not only reported random assignment of teachers to students before 7th grade began but also did not allow student within-school sorting from 2013-14 to 2014-15. Note that we do not worry about across-school sorting because school fixed effects absorb any time-invariant factors that drive students sorting in or out of school.

We obtain three separate samples by matching students with their Chinese, English, and math teachers. Missing rate on all key variables was below 2% except for two variables, teacher experience and student age, which were missing at 2-3%. We dropped all observations that had any missing value on predictor and outcome variables and replaced missing values on other variables with leave-one-out mean within homeroom (for student variables) or school (for teacher variables). We are left with 4,754, 4,855, and 4,887 students in Chinese, English, and math sample, respectively. Summary statistics of key variables are presented in Table 1. The distributions of student observations by teacher education and experience are visualized in Figure 1 and 2.

***Predictor variables***. Teacher education attainment is measured by three variables: education in years, an indicator for graduate degree, and an indicator for major in Educational Studies. Teacher experience is also measured by three variables: experience in years, an indicator for novice teacher (<= 2 years of experience), and a categorical variable in which the year measure is collapsed into bins (0-1 year is omitted).

***Outcome variables***. In each subject, student academic outcomes are measured by two variables: student’s score on the school-administered mid-fall semester exam (obtained from school record) and confidence level (student self-report). Both are standardized to have a mean of zero and standard deviation of one within each school.

Methods

Notes on writing – teacher experience

* Introduction
  + Rice, 2013
* Methods
  + Method review
    - Papay and Kraft, 2015
  + Indicator variable model
    - Clotfelter et al., 2007
    - Harris and Sass, 2011
  + Direct using of experience in years
    - Aaronson et al., 2007
  + Novice teacher indicator
    - Aaronson et al., 2007
  + GAM and piecewise regression
    - Hu et al., 2017

Teacher effect in general

Student level data matters:

* (Aaronson et al., 2007) A brief sampling of other work on teacher effects includes Murnane (1975), Goldhaber and Brewer (1997), Angrist and Lavy (2001), Jepsen and Rivkin (2002), Rivers and Sanders (2002), Jacob and Lefgren (2004), Rockoff (2004), Kane and Staiger (2005), Rivkin, Hanushek, and Kain (2005), and Kane, Rockoff, and Staiger (2006). The earliest studies on teacher quality were hampered by data availability and thus often relied on state- or school-level variation. Ag- gregation and measurement error compounded by proxies such as student-teacher ratios and average teacher experience can introduce significant bias. More recent studies, such as Rockoff (2004), Kane and Staiger (2005), Rivkin et al. (2005), and Kane et al. (2006), use administrative data like ours to minimize these concerns.

Teacher education background/teacher preparation

Aaronson et al., 2007:

* Measures
  + Graduate degree
  + Major
  + And more
* Findings
  + No relationship with teacher value-added on math in Chicago public high schools

Teacher experience/return to teaching

Aaronson et al., 2007:

* Models and measures
  + Directly using years measure (age-education-6 then take average within teacher)
  + Also check the dummy variable years<=1
* Findings
  + No relationship with teacher value-added on math in Chicago public high schools

Clotfelter et al., 2007:

* Models and measures
  + “Indicator variable model”, “by using within-bin variation to estimate the year effects, the Indicator Variable Model relies on a similar functional form assumption. In this case, it assumes that teacher productivity does not change meaningfully within each of these experience bins” (Papay and Kraft, 2015)
    - Specify years of experience as a series of indicator variables (1-2, 3-5, 6-12, 13-20, 21-27, and 27+), with the base/left-out category being no experience
  + Student fixed-effects
  + 10 years of north Carolina data
* Findings (p.676)
  + Larger effects for math than for reading
  + On student score gains, all coefficients of teacher experience indicators are significant; size ranging from 0.072-0.118 in math and 0.043-0.096 in reading

Harris and Sass, 2011:

* Methods and measures
  + Indicator variable model
    - Compare teachers with 1-2, 3-4, 5-9, 10-14, 15-24, and 25+ years’ experience separately to new teachers
  + Panel data; leveraging student, teacher, and school fixed effects (see paper for the advanced methods getting around computational challenge)
* Findings (p.805)
  + For elementary and middle school teachers, experience effects are quantitatively substantial, ranging from 0.03-0.06 SD for the first 1-2 years of experience as much as 0.16 SD for 15-24 years of experience
  + High school teachers are the opposite, more experienced teachers are less productive than when they were new teachers (teacher fixed effects)

Hu et al., 2017:

* using generalized additive modeling (GAM), a nonlinear method for “identifying likely thresholds by estimating the relations between an independent variable and a dependent variable without making any assumptions about whether the relation is linear or non-linear
  + likely thresholds are identified by graphically identifying the regions where there appears to by systematic change in the relation between the two variables and both baseline and ceiling thresholds are possible
  + also a graphic representation of the non-linear regression trend between the two variables
  + it doesn’t provide exact cut-off points, rather, it provides visual guidance on the score ranges within which thresholds are likely to exist
* then using piecewise regression (spline regression approach) to validate the possible thresholds
  + meaning to test whether the regression slopes vary across the different regions thus provide evidence of the validity of the GAM-derived thresholds
  + for each outcome variable, the results from the simple linear regression model (assuming one constant slope) and the piece- wise regression model (assuming varying slopes as defined by the thresholds) could be compared statistically. e.g., statistical test, adjusted R2s from two models, and the Akaike information criterion (AIC; Akaike, 1974) are available as evidence for the viability of the thresholds.
  + Lastly, compare linear regression with piecewise regression based on F statistics (statistical significance), adjusted R2 (higher value indicating better fit), and AIC (lower value indicating better fit)

Graphical user interface, text, application, email

Description automatically generated

Papay and Kraft, 2015:

* Models
  + Censored growth model
  + Indicator variable model
  + Discontinuous career model
  + Two-stage model
* Findings
  + using “estimated contributions to student test score gains as a proxy for productivity”, 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” in a large southern United States district.

Rice 2013:

Teacher experience has long been a central pillar of teacher workforce policies in U.S. school systems. The underlying assumption behind many of these policies is that experience promotes effectiveness, but is this really the case? What does existing evidence tell us about how, why, and for whom teacher experience matters? This policy brief distills the research on teacher experience into four general findings: (1) the effect of experience is most evident during the first few years of teaching; (2) the early-career experience effect varies by level of ed- ucation and subject area; (3) inexperienced teachers are most likely to teach in high-poverty schools; and (4) the impact of experience differs for teachers in high- versus low-poverty schools. The brief concludes by discussing the implications of these findings for several key policy measures including teacher compensation, support and professional development, and the unequal distribution of teachers across schools.

* Methods and measures
* Findings
  + Effect of experience is most evident during the first few years of teaching
  + The early-career experience effect varies by
    - Level of education
    - Subject area

Causal evidence

References

Aaronson, Barrow, L., & Sander, W. (2007). Teachers and Student Achievement in the Chicago Public High Schools. Journal of Labor Economics, 25(1), 95–135. <https://doi.org/10.1086/508733>

Clotfelter, Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. Economics of Education Review, 26(6), 673–682. <https://doi.org/10.1016/j.econedurev.2007.10.002>

Guarino, Dieterle, S. G., Bargagliotti, A. E., & Mason, W. M. (2013). What Can We Learn About Effective Early Mathematics Teaching? A Framework for Estimating Causal Effects Using Longitudinal Survey Data. Journal of Research on Educational Effectiveness, 6(2), 164–198. <https://doi.org/10.1080/19345747.2012.706695>

Harris, & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. Journal of Public Economics, 95(7), 798–812. <https://doi.org/10.1016/j.jpubeco.2010.11.009>

Ladd, & Sorensen, L. C. (2017). Returns to Teacher Experience: Student Achievement and Motivation in Middle School. Education Finance and Policy, 12(2), 241–279. <https://doi.org/10.1162/EDFP_a_00194>

Lim, & Meer, J. (2017). The impact of teacher-student gender matches: Random assignment evidence from South Korea. The Journal of Human Resources, 52(4), 979–997. <https://doi.org/10.3368/jhr.52.4.1215-7585R1>

Marioni, Freguglia, R. D. S., & Menezes-Filho, N. A. (2020). The impacts of teacher working conditions and human capital on student achievement: evidence from brazilian longitudinal data. Applied Economics, 52(6), 568–582. <https://doi.org/10.1080/00036846.2019.1650885>

Papay, & Kraft, M. A. (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. <https://doi.org/10.1016/j.jpubeco.2015.02.008>

Rice. (2013). Learning from Experience? Evidence on the Impact and Distribution of Teacher Experience and the Implications for Teacher Policy. Education Finance and Policy, 8(3), 332–348. <https://doi.org/10.1162/EDFP_a_00099>

Rockoff. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. The American Economic Review, 94(2), 247–252. <https://doi.org/10.1257/0002828041302244>

Sansone. (2017). Why does teacher gender matter? Economics of Education Review, 61(December), 9–18. <https://doi.org/10.1016/j.econedurev.2017.09.004>

Sansone. (2019). Teacher Characteristics, Student Beliefs, and the Gender Gap in STEM Fields. Educational Evaluation and Policy Analysis, 41(2), 127–144. <https://doi.org/10.3102/0162373718819830>

Wayne, & Youngs, P. (2003). Teacher characteristics and student achievement gains: a review. Review of Educational Research, 73(1), 89–122. https://doi.org/10.3102/00346543073001089

Winters, Haight, R. C., Swaim, T. T., & Pickering, K. A. (2013). The effect of same-gender teacher assignment on student achievement in the elementary and secondary grades: Evidence from panel data. Economics of Education Review, 34, 69–75. <https://doi.org/10.1016/j.econedurev.2013.01.007>

Wiswall. (2013). The dynamics of teacher quality. Journal of Public Economics, 100, 61–78. <https://doi.org/10.1016/j.jpubeco.2013.01.006>

Xu, & Li, Q. (2018). Gender achievement gaps among Chinese middle school students and the role of teachers’ gender. Economics of Education Review, 67, 82–93. <https://doi.org/10.1016/j.econedurev.2018.10.002>

Model specifications

Graphical user interface, text, application

Description automatically generated